Work Toward a Theory of Brain Function

A selection of scientific papers and a commentary submitted to the University of Otago in fulfilment of requirements for the degree of Doctor of Science

by

J. J. Wright

April 2015
For my wonderful Adrienne,
who made my world,
4th January 1963 to 25th July 2011, and forever.
From your Jim
Contents

Introduction 7
Part I: The split brain and the brain stem 8
Part II: The electrocorticogram 120
Part III: The embryogenesis of the cortex 469
Conclusion: Towards a theory of brain function 522
Appendix: Curriculum Vitae 529
Introduction

This essay and collection of scientific papers describes my efforts to determine as much as I can about fundamental mechanisms of brain information processing.

The papers are selected from all those I have published, to show the continuity and development of the research. A complete list, including other, more redundant, papers, those to which I made no very great contribution although I was a listed author, or on unrelated subjects, is given in the Appendix.

The selected papers are in roughly chronological order of publication, and more particularly in the order of development of the ideas they contain. I have tried to describe the institutional context and the change of ideas current at each stage. Complete referencing in the essay would be impracticable, so I refer to influential scientists by name, or to entire bodies of work by a descriptive field label, except in a few footnoted instances where exact reference is germane. In this way, I hope to give a sense of the intellectual atmosphere of this research field through changing times, from my perspective. In that spirit, I have not hesitated to include references to family events and their impact, and the political climate, rather than adhering to a convention of detachment.

The research developed in three main stages, and these are reflected in three principal parts.

The first part arises from unilateral electrical stimulation of motivational/reward pathways in the lateral hypothalamus and brain stem of “split-brain” cats, in which the great cerebral commissures were surgically divided. This showed that motivation systems in split-brain animals exert joint influence upon learning in both of the divided cerebral hemispheres, in contrast to the separation of cognitive functions produced by commissurotomy. However, attempts to identify separate signatures of electrocortical activity associated with the diffuse motivational/alerting effects and those of the cortically lateralised processes failed to achieve this goal, and showed that an adequate model of cerebral information processing was lacking.

The second part describes how this recognition of inadequacy led into computer simulations of large populations of cortical neurons – work which slowly led my colleagues and me to successful explanations of mechanisms for cortical synchrony and oscillation, and of evoked potentials and the global EEG. These results complemented the work of overseas groups led by Nunez, by Freeman, by Lopes da Silva and others, but also differed from the directions taken by these workers in certain important respects. It became possible to conceive of information transfer in the active cortex as a series of punctuated equilibria of signal exchange among cortical neurons – equilibria reached repeatedly, with sequential perturbations of the neural activity away from equilibrium caused by exogenous inputs and endogenous pulse-bursting, thus forming a basis for cognitive sequences.

In the third part, I report on how our explanation of synchrony gave rise to a new theory of the regulation of embryonic cortical growth and the emergence of mature functional connections. This work is based upon very different assumptions, and reaches very different conclusions, to that of pioneers of the field such as Hubel and Wiesel, whose ideas have dominated cortical physiology for more than fifty years.

In conclusion, I link findings from all the stages of this research together, to show they provide a sketch of the working brain, fitting within and helping to unify wider contemporary concepts of brain function.
Part I. The split brain and the brain stem

In 1971, I found myself, to my great good fortune, in Caltech, in the laboratory of Roger Sperry.

Caltech’s atmosphere was heady. The Jet Propulsion Laboratory was at the centre of the Space Program. Caltech quantum physicists led the world. There was optimism that big problems were there to be solved, not respected. And, thus, discovery of the mechanism of function of the brain was on the list of things to be done. Perhaps because of Caltech’s great strength in physics, and in biology, in the work of Pauling, and of Delbrück, it was abstract, unifying principles that were sought – not accumulation of specifics. For my own part, I had read Sherrington’s great work on reflex analysis, and Hebb’s great speculative work on the organization of neural systems. While at Caltech I learned of McCulloch and Pitts’ proof that networks of simplified neurons could be arranged to perform all logical functions and I saw a simple electro-mechanical perceptron simulating visual recognition, itself based on the theories of Hubel and Wiesel, arising from their study of stimulus-specific responses of cortical neurons.

Roger Sperry was already famous, particularly for his work on hemispheric specialization. His early work tested (to destruction) concepts advanced by Karl Lashley on the hotly contested issue of whether the cortex worked in some holistic manner, or whether its function could be reduced to specific, separable circuitry components. Sperry’s work favoured the latter. He was also regarded as a significant mental philosopher, who, although he considered the function of the brain to be analysable in physical terms, believed the brain/mind problem to be a concern of a different type and therefore the brain, as mechanism, was the simpler side of the problem.

Down the hallway was the laboratory of James Olds, the discoverer of intracranial self-stimulation (ICSS), then known as “pleasure centre stimulation”. This work had exerted huge influence on contemporary physiological psychology, creating the impression that seemingly impossibly complex mechanisms of motivation and emotion might be within grasp, since they could be triggered by electrical excitation of specific tracts in the brain stem. In a similar way, the earlier work of Moruzzi and Magoun on the brain-stem reticular activating system had excited belief that the very mechanisms of consciousness were coming into experimental range.

Both Sperry and Olds were most encouraging to me, when I set about the work that I would submit two years later for my MD (Otago) dissertation. This work was on the unilateral electrical stimulation of

---

motivational/reward pathways in the lateral hypothalamus and brain stem of “split-brain” cats, in which the great cerebral commissures were surgically divided.

The immediate motivation for the work came from the wider goal of determining which functions could be clearly ascribed to either the left or the right cerebral hemisphere, and which functions required joint mediation via subcortical systems. There was a suggestion, based on clinical reports of an odd effect in surgically split-brain humans, that emotional states were shared by both cerebral hemispheres. Humorous stimuli shown to the right hemisphere made the reporting, verbal, left hemisphere amused, but unable to give a spoken account of the source of the humour. Some light might be cast on the matter by stimulating ICSS sites on one side of the brain in split-brain animals, and seeing if the stimulation could influence learning separately, in both hemispheres. This proved to be the case. My first paper


showed that motivation systems in split-brain animals exert joint influence upon learning in both of the divided cerebral hemispheres, in contrast to the separation of cognitive functions produced by commissurotomy. The result strongly suggested that the neural pathways involved exerted some diffuse or spreading effect – presumably through neurons with wide fields of termination – on other neurons in the telencephalon that had been specifically involved in sensory-motor or cognitive processes in the short period prior to the delivery of the ICSS stimulus, and had thus consolidated learning of immediately prior neural sequences.

After less than two years at Caltech, Adrienne and I, and our two little children, moved on to London. We had greatly enjoyed both Caltech and California. My Post-Doctoral Fellowship remained open, but I needed to get on with my professional training. In London, in the course of clinical training, I was also able to continue research in the laboratory of Professor Giles Brindley, at the Institute of Psychiatry, King’s College, London. There I met and began work with my friend and colleague, Michael Craggs.

Working now on rhesus monkeys I was able to extend the commissurotomy surgery beyond the great cerebral commissures and the optic chiasm, to include the massa intermedia of the thalamus. Instead of exploring the behavioural effects of unilateral ICSS on learning in the separated hemispheres, using implanted electrocortical recording electrodes we showed that the bilateral effects of the stimulus were associated with joint electrocortical activation of both hemispheres, as reported in


The same animals showed separation of electrophysiological function in their separated hemispheres related to ongoing cognition and attention. Changes in the spectra of their electrocorticograms produced by eye cover were also exhibited in spontaneous shifts of electrocortical state between hemispheres, moment to moment. However, these shifts were similar to those induced by the unilateral brain-stem stimulations, and so did not distinguish particular cognitive activity from general activation. See

Again, after three years this time, home and career called. Shooed on our way by a house-fire in which our new baby son almost died, our little family returned to New Zealand, and Auckland Medical School. Mike Craggs obtained an MRC Travelling Fellowship, and for another year we were able to set up a laboratory and continue our work in the Department of Psychiatry.

Since we had provisionally established that there were bilateral electrocortical effects associated with the brain-stem stimulation, in contrast to unilateral differences accompanying cognition, or perception, our goal now was to attempt to identify separate signatures of electrocortical activity associated with the diffuse motivational/alerting effects and those of the cortically lateralisd cognitive processes. We now mixed unilateral lesion of the same pathways that produced the motivation, reward and arousal, with stimulation of the same pathways on the “intact” side. The unilateral lesion produced, as had been described by others, a contralateral sensorimotor neglect syndrome. We showed the neglect syndrome was associated with attenuated contralateral cortical activation produced by the ipsilateral stimulus, as we reported in:


Still, these asymmetries seemed only relative shifts on a continuum of activation. In company now with colleagues Rob Kydd (then MRC Training Fellow), Gordon Lees (a neurochemist), and John West and Nick Hawthorne in technical roles, we set out to establish quantitative differences attributable to these manipulations of the brain-stem inputs to the cortex, and to establish the neurochemical nature of the pathways involved. To attempt this we were fortunate in having within the Department a PDP-8 computer and a student colleague, Alex Sergejew, who could program it. We also needed a theoretical framework within which to express our results.

By this time I was conscious of how inadequate was contemporary understanding of cortical function and its relationship with the electrocortical signal (or the clinical EEG). The dramatic functional role of the reticular, lateral hypothalamic and cortico-thalamic pathways in motivation and alerting, and electrocortical activation, was established, but there was no agreement about the physiological mechanisms. Theories of the logical operations of networks of neurons, begun by McCullough and Pitts and followed by Hebb’s famous learning rule in the 1940s, had generated a developing field, but no theory of any sort was available which linked the logical operations of neurons to the EEG – at that time the only directly observable physical measure of global brain operation. Controversy over its functional significance could be dated from the earliest days of the EEG. Although it was early recognised that the EEG signal was sensitive to psychological events, its relevance was often denigrated, and it was popular for detractors to say that it was “only the hum of the engine”, not a reflection of fundamental aspects of cortical function, basing this comment on the failure of attempts to correlate firing of cortical neurons to the EEG signal. Although it was known that the signal mainly arose from the dendrites of the cerebral cortex, under modulation by, and
interaction with, subcortical systems, the source of particular components of the EEG signal within the brain was widely disputed.

In response to this background controversy and uncertainty, attempts were being made elsewhere to reach more clearly testable theoretical accounts of the EEG. Mathematical models of the biophysical processes leading to the generation of the EEG signals were being formulated, and the three leading models were those of Nunez, Lopes da Silva, and Freeman. All could offer a match to limited aspects of experimental data, but fundamental differences were apparent, and it was not clear whether their work could be reconciled. Nunez believed the EEG rhythmic activity arose from wholly cortical global resonant modes of the brain, Lopes da Silva’s group focused on more restricted cortico-thalamic interactions, and Freeman, whose work on the olfactory bulb carried wide implications elsewhere in the brain, favoured a cortical origin quite different to that supposed by Nunez. The most fundamental difference of all was between the classes of mathematics that were applicable. Nunez and Lopes da Silva favoured linear wave mechanics – that is, an assumed analogy of the waves of electrical activity spreading through a field of mutually exciting neurons, to the wave-superposition principles applicable to waves of light or sound. Freeman thought this naïve at best, pointing out that individual neurons exhibit sharp threshold behaviour, and their interactions and summations must, therefore, be highly nonlinear. That view was taken up enthusiastically in some quarters, despite the fact that no general means of analysing highly nonlinear systems of such complexity was in sight – nor is today. The distinction was not as cut and dried as the high nonlinearity of neurons suggested. The deeper issue was whether the activity of populations of neurons could be smoothed over into average properties of the neural field, or whether the interaction of the neurons was critically sensitive to all of their states at once – that is, whether brain activity was not, or was, chaotic. Means of distinguishing between the alternatives were not apparent – so since nobody else could agree, I thought it might be best to assume linearity as the best practical method, and to see if evidence for wave linearity could be established. A first try in this direction was published, with the help of a mathematician colleague, as


From that start, and with the help of further generous tutorials from members of the Mathematics and Engineering Departments (particularly Dr Ah Chung Tsoi), I was able to advance an initial, and crude, treatment of the EEG as arising from linked oscillators, and to propose tests of whether or not the EEG signal had properties fairly typical of linear waves. The initial theory was given in


and the tests for the theory reported in

---


The experiments used linear model fitting to power spectra of the electrocorticogram from left and right hemispheres, before and after unilateral lesion of the lateral hypothalamus, in rats. This experimental preparation offered very good control of before and after effects, with one hemisphere acting as control for the other, so that relative changes in the amplitude of each frequency component could be compared between the directly affected, and the control, hemisphere.

Our results showed that there appeared to be constant frequencies of rhythmic activity in the hemisphere, and laws for the relative speed of transmission of waves at all frequencies, that were independent of manipulation of the controlling input pathways – with the brain-stem input regulating both the rate of damping of the resonances, and the strength of noise-like input signals driving the resonant activity in the cortical circuits. We later showed that the concurrent changes in relative amplitude and phase of electrocortical signals were consistent with linear wave theory. With Gordon Lees joining us in the enterprise, bringing the techniques of neuronal histoflourescence, we showed that this controlling effect was mediated by catecholaminergic neurons – the group of neurons implicated in ICSS, and of increasing importance as the substrate of action of antipsychotic and antidepressant drugs. These findings we reported in


A start had been made. Some of the neurochemistry of the joint control of electrocortical activity and reward had been defined, and this was interpretable within linear systems theory, in a physically valid way. It was clear that a much better formulated theory would be needed, and much more detailed analysis of the cortical surface signal. We would try to do both.
Unilateral Pleasure-Center Stimulation in Split-Brain Cats

J. J. Wright

Division of Biology, California Institute of Technology, Pasadena, California 91109

Received December 28, 1972

Unilateral bipolar stimulation at seven positively reinforcing sites ranging from the ventral mesencephalic reticular formation to the preoptic region, in five split-brain cats, was shown to reinforce pattern discrimination learning in both hemispheres. The training method used permitted simultaneous presentation of discriminanda to both hemispheres of the split-brain animal. The separated hemispheres were subsequently tested individually for establishment of learning, and it was found that learning had occurred in both hemispheres, in most cases. Monocular training showed reinforcement of behavior can be produced in both hemispheres at every electrode site tested. It is concluded that unilateral pleasure-center stimulation may exert a bilateral central effect via direct and crossed projections from the brain stem.

INTRODUCTION

It has been amply demonstrated that after division of the intercortical commissures of the mammalian brain, the divided hemispheres have separate perceptual, learning, and memory processes (5). This separation has been shown to include some aspects of limbic system function. Webster and Yoneida (8) have shown that unilateral ablation of the hippocampus produces a unilateral deficit in reversal learning in split brain cats, and Doty, Yamaga, and Negrao (1) have reported a lateralized loss of visual fear for human beings, in split-brain macaques with unilateral amygdalecmy. To what extent separation of limbic system functions includes the motivational and affective mechanisms in the split-brain animal is uncertain. Commisurotomy as usually performed includes the anterior and hippocampal commissures so that direct interhemispheric links between the higher limbic structures are severed. There exists no evidence for or against the

1 This work was supported by United States Public Health Service Grant MH03372 to Professor R. W. Sperry to whom the author extends added thanks for advice throughout the study. Present address: Department of Psychiatry, St. George’s Hospital Medical School, University of London, London, England.

Copyright © 1973 by Academic Press, Inc.
All rights of reproduction in any form reserved.
pleasure center

possibility of lateralized interaction between the cerebral hemispheres and lower brainstem structures involved in motivation and affect. Studies of intracranial self-stimulation use behaviors of the whole animal in response to unilateral pleasure-center stimulation. With the intercortical commissures intact, it follows that learning reinforced by pleasure-center interactions with one hemisphere would rapidly be transferred to the other hemisphere, thus masking the lateralized nature of the interaction. Alternatively there may be no such lateralized interaction, and reward-system activity may exert a more diffuse bilateral action upon the higher centers. The present study was aimed at separating these alternatives, using split-brain cats and unilateral pleasure-center stimulation at a variety of sites.

METHODS

The subjects were five previously untrained adult cats (two male, three female). Surgery was performed in three stages, with a convalescent period of 3–6 weeks after each procedure. First, the optic chiasm was divided in the midline, using the transbuccal approach (3). Second, the corpus callosum and anterior commissure were divided, using the method of Sperry (6). Third, three chronically implanted bipolar stimulating electrodes were placed stereotaxically in each cat, using sites selected from the atlas of the sites of intracranial self-stimulation in cats (9). The bipolar stimulating electrodes were made from 0.01-in. diameter lacquered stainless steel with a tip separation of 0.5 mm. The shorter electrode tip was oriented laterally in situ.

After convalescence, the reinforcing character of each implanted electrode was tested by the method of O’Donahue and Hageman (4), using 60-Hz square waves of various current strengths and train duration, delivered by a Nuclear Chicago 7150 constant-current stimulator. Seven of the electrode sites were found to exert a motivating effect sufficient to readily motivate the cat to cross a 3-ft-square open space to receive further stimulation.

Pattern-discrimination training was then performed in a training box 30 in. long × 21 in. wide × 18 in. deep. Each end of the box was fitted with two IEE one-plane readout display panels, placed 9 in. apart. An electromechanical control system permitted the pattern stimuli to be displayed first at one end of the box, and then after left/right alternation in a pseudorandom sequence, at the other end of the training box. The cat in training was thus required to run back and forth in the training box, attempting a two-choice pattern discrimination each time he reached the end of the box with the panels illuminated. A correct choice resulted in the animal’s receiving a unilateral intracranial stimulus (of the current and duration shown in Table 1), when the panel was touched with nose or forepaw. An incorrect choice resulted in the animal receiving no stimulation. All the animals were
pretrained to run in the box for intracranial stimulation using a simple absolute brightness cue on the display panels. Each animal ran between 20 and 60 trials per day. Training with one electrode and one discrimination problem was completed before training with the next electrode and another problem, in the same cat.

*Use of Binocular and Monocular Forms of Training.* Occlusion of vision to one eye, and hence to one hemisphere, was produced by use of black plastic individually fitted eye occluders. Training of the animals was performed in the following way:

1. With both eyes allowed free vision, the cat was trained to a criterion of 19 out of 20 responses correct. Then the eye occluder was introduced, and each eye was allowed sole vision for 40 trials. This was done in blocks of 10 trials with each eye. Thus, a measure was obtained of the degree to which memory had become consolidated in each of the separated hemispheres, after identical visual experience of the task. After this, the animal was again allowed uninterrupted binocular vision and was overtrained for 240 trials. Then performance of each hemisphere was again tested by the monocular method, to again assess the degree to which consolidation had occurred in each hemisphere.

2. The above form of training and testing permitted comparison of levels of performance in the separated hemispheres after identical opportunity to form the critical stimulus reward association. It was assumed that a unilateral rewarding effect of unilateral pleasure-center stimulation would produce unilateral consolidation of learning. In the event of both hemispheres showing evidence of learning the discrimination, it might be objected that one hemisphere had learnt the discrimination by repeatedly cooperating with its active partner, which may have alone been subject to the positive reinforcement of the pleasure-center stimulus. To control for this difficulty, a second training task was performed for each electrode. The animals were trained monocularly with alternate 20 trial blocks being performed using each eye. The hemisphere contralateral to the stimulating electrode was trained in the reversal of the previously established discrimination, while the homolateral hemisphere continued to receive reward for the previous discrimination response. The capacity to reverse learning in one hemisphere while maintaining performance in the other was taken as evidence that the hemispheres were separated for the perception of patterns and the learning of operant responses and that the intracranial stimulation produced a reinforcement effect which was specifically associated with the rewarded pattern in each hemisphere. If this were not so, either reversal learning would not occur in the contralateral hemisphere, or performance would gradually deteriorate to chance in the homolateral hemisphere.
Histological Technique. After the completion of testing, the cats were given an overdose of barbiturates. Electrode tips were marked by the electrolytic deposition of tin using a direct current of 100 μA for 30 sec. Cerebral perfusion was carried out by the transcardiac method, and the brain was perfused with 1 liter of heparinized physiological saline solution, 500 ml of 10% formol–saline and 500 ml of 1% potassium ferricyanide solution. The removed brain was sectioned at 35 μm and sections were stained with either hematoxylin and eosin with an excess of eosin, against which the blue electrode tip depositions stood out, or with the Weil stain to confirm adequacy of the commissuromyotomy and optic chiasm section. Electrode tip locations were recorded by reference to the atlas of Jasper and Ajmone-Marsan (2).

RESULTS

At the conclusion of the experiment gross and histological examination confirmed complete section of the corpus callosum, anterior commissure, and optic chiasm in all subjects. The seven electrodes which were found to be sufficiently motivating to maintain performance during training were located as shown in Table 1.

Establishment of Learning in Each Hemisphere after Binocular Presentation of Discriminanda With Unilateral Stimulation As Reward. Results of this part of the study are shown in Table 2, and in graphical form in Figs. 1 and 2. There is a trend toward superior performance of the learned response by the contralateral hemisphere, but this trend does not reach statistical significance (by t test) either after training to criterion or after overtraining. Monocular performance was often not significantly better.

<table>
<thead>
<tr>
<th>Cat</th>
<th>Electrode</th>
<th>Stereotaxic site</th>
<th>Corresponding anatomical site</th>
<th>Reward-stimulus strength (60 cps square waves)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. W. F.</td>
<td>1</td>
<td>FR 11.5 H-3.5 RL 3.5</td>
<td>Lateral hypothalamus</td>
<td>200 μA, 1 sec</td>
</tr>
<tr>
<td>B. W. F.</td>
<td>2</td>
<td>FR 2.0 H-5.0 RL 3.0</td>
<td>Ventral mesencephalic reticular formation</td>
<td>150 μA, 1 sec</td>
</tr>
<tr>
<td>L. M.</td>
<td>3</td>
<td>FR 10.5 H-3.5 RL 3.5</td>
<td>Lateral hypothalamus</td>
<td>200 μA, 1 sec</td>
</tr>
<tr>
<td>B. V.</td>
<td>4</td>
<td>FR 11.5 H-4.5 RL 3.0</td>
<td>Lateral hypothalamus</td>
<td>200 μA, 1 sec</td>
</tr>
<tr>
<td>R. L. K.</td>
<td>5</td>
<td>FR 5.0 H-3.5 RL 2.5</td>
<td>Lateral magnocellular red nucleus</td>
<td>150 μA, 1 sec</td>
</tr>
<tr>
<td>G. O. P.</td>
<td>6</td>
<td>FR 14.0 H-3.0 RL 3.0</td>
<td>Preoptic region</td>
<td>200 μA, 1 sec</td>
</tr>
<tr>
<td>G. O. P.</td>
<td>7</td>
<td>FR 10.5 H-3.0 RL 2.0</td>
<td>Between lateral and dorsal hypothalamus</td>
<td>75 μA, 1 sec</td>
</tr>
</tbody>
</table>
than chance (by chi-square test) after binocular training to criterion, but in most instances improved to a highly significant level after binocular overtraining. The homolateral hemisphere trained with reward from electrodes 1 and 4 failed to establish significant evidence of consolidation of learning in this method of training.

![Graph showing performance comparison](image-url)

**Fig. 1.** Comparison of performance of each hemisphere after binocular training to criterion. Open bars: hemisphere on the side of the electrode; shaded bars: hemisphere contralateral to the electrode.
Fig. 2. Comparison of performance of each hemisphere after binocular overtraining Bars as in Fig. 1.

Fig. 3. Concurrent reversal learning in the hemisphere contralateral to the electrode (unbroken line) and maintenance of performance of previously trained discrimination in the hemisphere homolateral to the electrode (broken line). "Trials correct" refers to the correct reversal response.
Control Test to Establish Independence of Cognitive Processes in the Separated Hemispheres, and Specificity of Reinforcement in Each Hemisphere.

The learning curves for this control experiment are shown in Fig. 3. In each case reversal learning proceeded satisfactorily to a criterion of 19/20 in the contralateral hemisphere, while the homolateral hemisphere maintained or improved its performance to the 19/20 level.

DISCUSSION

The foregoing results demonstrate that unilateral pleasure-center stimulation at a variety of sites in the split-brain cat can reinforce pattern-discrimination learning in the contralateral and the ipsilateral hemisphere. The results of the test for establishment of learning in the two hemispheres after simultaneous training resembles results obtained in split-brain monkeys, using a conventional food reward (7). The failure of consolidation of learning in the homolateral hemisphere when training with electrodes 1 and 4 did not persist when monocular training was undertaken. The significance of this failure of consolidation is uncertain, but whatever the mechanism, it would appear that bilateral reinforcement was possible from these electrode sites.

The most probable interpretation of these experimental results is that unilateral pleasure-center stimulation produces a central effect that is transferred by commissural and other fiber systems in the brain stem to produce reinforcement in both hemispheres. The possibility that the associated somatic and autonomic activity produced by the unilateral pleasure-center stimulus was positively reinforcing to one or both hemispheres cannot be excluded. Similar studies on animals subjected to brain bisection to a more caudal level, and studies of more rostral and diverse electrode placements in split-brain animals may help to clarify understanding of the central pathways involved in reinforcement.

REFERENCES


Arousal and Intracranial Self-Stimulation in Split-Brain Monkeys

J. J. Wright and M. D. Cracco

Institute of Psychiatry, London SE5 8AF, United Kingdom

Received August 12, 1966; revision received November 18, 1966

Split-brain monkeys were delivered noncontingent unilateral stimulation at sites in the vicinity of the lateral hypothalamus and the internal capsule. At self-stimulation sites, this produced suppression of alpha rhythm in both hemispheres of animals with either telencephalic or diencephalic commissurotomies. The time course of the alpha suppression in each hemisphere was identical and no habituation was seen with prolonged repetition of stimulation. Bilateral nonhabituating arousal was still obtained after lesion of the hypothalamus, including the medial forebrain bundle, contralateral to the self-stimulation site. Non-self-stimulation sites showed habituation. Bilateral nonhabituating cortical arousal may be involved in the bilateral reinforcement of learning produced by unilateral intracranial self-stimulation in split-brain animals.

INTRODUCTION

Split-brain animals show not only hemispheric independence of perception and learning (18) but also a high degree of electroencephalogram (EEG) dissociation between the hemispheres. Their EEGs at similar points on the two hemispheres are asynchronous, their alpha spindles vary independently in both phase and amplitude (5), and independent changes occur in the EEG power spectrum related to visual processing (24).

In contrast, the reinforcement effect of unilateral intracranial self-stimulation is not confined to the ipsilateral hemisphere (23). With sites of stimulation as far rostral as the orbitofrontal cortex in monkeys with commissurotomy extended to the midbrain level, lateralization of respond-

1 This work was supported financially by the Mental Health Trust and Research Fund and the Medical Research Council, (U.K.).
2 Address correspondence to Dr. J. J. Wright, Department of Psychiatry, Medical School, University of Auckland, Auckland, New Zealand.
3 Abbreviations: EEG electroencephalogram; ECoG electrocorticogram.
ing for intracranial self-stimulation is not consistently sustained (10). This suggests the mediation of reward through common neural systems in the brain stem.

Stimulation of lateral hypothalamic self-stimulation sites is associated with widespread activation of single units in the brain stem and with the production of cortical arousal (8, 12, 13), although the systems mediating reward and arousal, respectively, are unclear.

Scheibel and Scheibel summarize evidence which suggests that brain stem reticular neurons ascending through the interpeduncular, hypothalamic, and septal areas mediate cortical desynchronization (15), and pharmacological and neuroanatomical analyses of self-stimulation sites support the belief that catecholaminergic neurons (particularly those arising from the locus coeruleus and the ventral mesencephalic area) are involved in rewarding effects produced by stimulation of brain stem sites (7). These reward-mediating neurons ascend largely unilaterally through the medial forebrain bundle including the lateral hypothalamus (1). Descending systems of axons passing through the bundle may also play a role, but their role is somewhat unclear.

In the experiment to be described we have assessed the time course and habituation characteristics of cortical arousal in split-brain monkeys during unilateral stimulation of intracranial self-stimulation sites to determine whether or not consistent arousal changes occur in each hemisphere as a correlate of rewarding stimulation. We have also made a preliminary attempt to determine whether or not cortical arousal associated with stimulation of intracranial self-stimulation sites is mediated directly by neurons projecting rostrally via the lateral hypothalamus or via other pathways.

METHODS

Five adolescent rhesus monkeys were used. Under sodium pentobarbital anesthesia (40 mg/kg), each was subjected to a midline transection of the telencephalic and tectal commissures, and in two cases the massa intermedia was also divided. In four cases the optic chiasma including the suprachiasmatic commissures was transected, and in one case (in which the massa intermedia was intact) the right optic tract. Surgical techniques were similar to those described by Sperry (19). Hypothermia was used to diminish postoperative cerebral edema in surgery of the massa intermedia as described by Brinkman and Kuypers (3).

At a later stage, multipolar depth electrodes for stimulation were implanted near various known intracranial self-stimulation sites in and around the lateral hypothalamus and internal capsule, near the A11 plane
in the case of most lateral hypothalamic sites and the A14 plane in the
case of internal capsule sites and two of the lateral hypothalamic sites,
to obtain a variety of self-stimulation placements and adjacent non-self-
stimulation placements for comparison (2, 16). Depth electrodes were
insulated 0.075-mm platinum wires with the tips bared for 0.5 mm and
separated by 1 mm. Bilateral occipitoparietal extradural platinum record-
ing electrodes were also implanted, with an indifferent platinum electrode
attached to the skull.

Two weeks after recovery from surgery, self-stimulation rates were
determined for all depth electrode sites, using 100-Hz, 200-ms sine-wave
trains. A current was determined for each intracranial site, at which the
rate of self-stimulation (about 60/min) was consistently greater than
control rates (about 10/min) during repeated 10-min sessions. All stimu-
lation currents were in the range 400 to 800 μA rms (2 to 4 V rms); comparable currents were applied to sites found not to be self-stimulating.

While restrained in a primate chair, each monkey then underwent
repeated noncontingent stimulation of each site at a rate of about five/min,
while the behavioral response to each stimulus was observed and the
electrocochleogram (ECO2) recorded. By using a low rate of noncontingent stimulation it was possible both to maintain the animal in a rela-
tively low state of arousal and to observe the recovery from alpha sup-
pression between stimuli. Cortical signals from each hemisphere were
led to differential amplifiers (Telequipment DM6.3/VS) through a flexi-
ble low-noise cable and monitored on an oscilloscope. A bandpass filter
(24 db/octave) was used to select the 8 to 12-Hz range (alpha rhythm in
young monkeys) previously shown to reflect changes in cortical activation
in these animals (23). The alpha rhythm changes associated with depth
electrode stimulation could be (a) monitored and retained for inspection
on a storage oscilloscope and (b) squared and smoothed (time constant,
60 ms, 10 to 90%) to give the mean power, then computer averaged (CAT
400 c) for a block of 40 stimulations. Two to four such blocks of stimu-
lation were delivered consecutively, and each site of stimulation was
studied every 2 to 3 days, during a period of 2 to 4 weeks. Twenty-two
sites were studied in all, nine of which supported self-stimulation. Of
these self-stimulation sites, seven in the lateral hypothalamic were dis-
tributed in animals both with and without the massa intermedia divided.
The two internal capsule sites were in one "deep" split-brain animal. The
accuracy of commissurotomy and sites of electrode placement was con-
firmed histologically after frozen sectioning of the formalin-fixed brain
at 100-μm thickness and staining of selected sections with luxol fast blue
and cresyl violet.
RESULTS

Effect of Self-Stimulation and Non Self-Stimulation Sites on Arousal. At all self-stimulation sites behavioral alerting was consistently obtained in response to each noncontingent stimulus. No habituation was seen during the blocks of stimuli or between each block of trials for a given site. In association with behavioral alerting, the filtered ECoG sometimes showed a clear suppression of alpha activity in both hemispheres with each stimulus (Fig. 1). However, the onset of suppression was not always well defined, particularly when alpha activity was low, and the onset was best seen on the computer plot of the mean power change within the alpha band for the block of 40 stimuli. The signals averaged by the computer showed a simultaneous suppression of power in each hemisphere (Fig. 2). The average latency of onset of suppression was about 140 ms, after allowing for a 60-ms delay in the recording system. Maximal suppression occurred about 350 ms after the start of stimulation in each hemisphere, and the time course of recovery to prestimulation levels was similar in the two hemispheres. These features were the same at all self-stimulation sites, irrespective of whether the massa intermedia was intact or not.

Sites at which no self-stimulation could be obtained did not show this pattern; behavioral habituation occurred either within a few trials or over several blocks of trials, and ECoG evidence of change of arousal showed an accompanying diminution of effect, as habituation proceeded. At sites where habituation was slow enough to permit the gathering of adequate

Fig. 1. Bilateral suppression of alpha rhythm in a split-brain monkey following a single 100-Hz, 200-ms stimulus to an intracranial self-stimulation site in the left lateral hypothalamus. Top and middle traces: occipitoparietal ECoG, bandpass filtered 8 to 12 Hz. Bottom trace: stimulus marker.
averaged data, arousal in the early stage was bilateral and had a time course similar to that at self-stimulation sites.

**Effect on Self-Stimulation Arousal of Unilateral Hypothalamic Lesions.** The role of the contralateral lateral hypothalamus in propagating arousal influences from a self-stimulation site to the contralateral hemisphere was investigated in one animal whose bilateral lateral hypothalamus electrodes

---

**Fig. 2.** Mean alpha power in each hemisphere of the same monkey as shown in Fig. 1, averaged by computer over 40 stimuli delivered to the same site. Bottom trace: stimulus marker.

**Fig. 3.** Extent of unilateral lesion in the area of the right hypothalamus which failed to impair bilateral nonhabituated alpha suppression evoked by stimulation of a left hypothalamic self-stimulation site. Section taken at the A11 plane of Horsley and Clarke, through the median eminence. The electrode track is shown diagrammatically by the double dashed line. Dor A—dorsal hypothalamic area; DM—dorsomedial hypothalamus; F—fornix; Lat A—lateral hypothalamic area; Opt—optic tract; Ped C—cerebral peduncle; Teg A—area tegmentalis; VM—ventromedial hypothalamus.
were suitably placed at self-stimulation sites in the A11 plane. In this case, all telencephalic and tectal commissures and the optic chiasma were transected, but the massa intermedia was intact. Under barbiturate anesthesia, a large unilateral electrolytic lesion was made through adjacent self-stimulation electrodes in one hypothalamus, in two stages 12 h apart, using a direct current of 5 mA for 30 s at each adjacent self-stimulation electrode in sequence, with the depth electrodes anodal.

Following this large unilateral lesion, intracranial self-stimulation and arousal changes were abolished from that site but self-stimulation and typical bilateral nonhabituating ECoG changes and alerting reactions could still be elicited from the intact lateral hypothalamus. These observations were made 12 h after completion of the lesion and again 2 weeks later. The extent of the lesion was confirmed by histology and is shown diagrammatically in Fig. 3.

DISCUSSION

Our results show that stimulation in split-brain monkeys at lateral hypothalamus and internal capsule self-stimulation sites produces bilateral cortical arousal, as measured by alpha suppression, which does not habituate with repeated stimulation. Bilateral arousal was produced from these sites even in “deep” split-brain animals in which the most rostral intact commissure was the supramammillary commissure, so it would seem that neuronal systems in the midbrain region, or below, were involved. Nonhabituating cortical arousal is considered the hallmark of activation of the reticular arousal system, in contrast to other sites at which arousal responses undergo habituation (21, 22). The close parallel of onset and offset of arousal in the separated hemispheres, at all intracranial self-stimulation sites studied, strongly suggests the activation of a common arousal mechanism in the brain stem. Furthermore, our split-brain subject with a unilateral hypothalamic lesion indicates that neurons ascending toward the cortex through the lateral hypothalamus are not themselves solely responsible for nonhabituating arousal, because bilateral arousal was not impaired when contralateral self-stimulation sites were stimulated, either acutely or chronically following the lesion. Therefore activation of the cortex ipsilateral to the lesion must have occurred by the spread of arousal impulses through pathways in the brain stem and/or the massa intermedia, other than those of the medial forebrain bundle.

From these results, and the previous studies on intracranial self-stimulation reward in split-brain animals (10, 23), it seems that brain stem mechanisms regulating both reward and arousal are conjointly activated from unilateral self-stimulation sites, and these effects diffuse bilaterally
to higher centers. Such an association might result from the close anatomic contiguity of separate neural systems arising from or traversing the sites of stimulation and, thus, might be an experimental artifact. On the other hand, because cortical arousal is related to the acquisition of learning (11, 17), the relationship between intracranial self-stimulation and arousal may be of functional significance in correlating learning with reward under physiologic conditions. A functional relationship is further suggested by the coincidence of self-stimulation and arousal at sites in the internal capsule as well as the lateral hypothalamus.

A further possibility is that descending neural fibers running through the sites of stimulation toward the brain stem were activated, and reward and arousal were generated by their various projections. However, the sites studied, in the lateral hypothalamus and internal capsule, were all on the pathways of projection to the telencephalon of noradrenergic and/or dopaminergic neurons which have been implicated in the mediation of intracranial self-stimulation reward in previous studies. In the lateral hypothalamus, noradrenergic neurons from the locus coeruleus and dopaminergic neurons from the interpeduncular area are implicated (1, 7). In the internal capsule, dopaminergic neurons ascend from the substantia nigra which has also been found to mediate intracranial self-stimulation (6). These neurons project to the telencephalon in a predominantly unilateral manner, and in a split-brain animal their rostral projections are further unilaterally constrained by the commissurotomies. For them to be involved in the bilateral mediation of both reward and arousal in split-brain animals, their antidromic activation and diffuse interaction with arousal neurons and with contralateral reward neurons in the brain stem must be postulated. Such interactions have been shown to occur among some of the classes of neurons involved. Short axons from the cells of origin of catecholamine intracranial self-stimulation neurons passing into the reticular formation have been described (20). Self-stimulation via stimulation of the lateral hypothalamus produces a widespread bilateral activation of pontine and midbrain reticular neurons by antidromic and transynaptic mechanisms, as cited above (8, 12, 13), and coeruleus neurons are bilaterally activated from medial forebrain bundle self-stimulation sites (9).

If the activation of some or all of the catecholaminergic neurons mediating reward is closely associated with cortical arousal (as our results suggest) then these neurons might be expected to show a response in firing rate only in relation to rewards which result in cortical arousal. The locus coeruleus does not appear to be activated by natural food rewards (9), in contrast to units in the lateral hypothalamus and substantia innominata, which respond to the sight of food (4).
Because self-stimulation at the nucleus accumbens and some other rostral sites does not appear to be associated with arousal, Rolls has concluded that there is no evidence that the brain stem arousal system is directly involved in reward effects produced by self-stimulation, and he rejects the role of arousal to effects on the rate of self-stimulation, the "priming" effect, and locomotor activity (14). However, at the more caudal sites so far studied it seems possible that effects on cortical arousal may be involved in the establishment of learned responses in both hemispheres of split brain animals receiving unilateral self-stimulation reward. The essential pathways involved in the transfer of reinforcing effects through the brain stem might be determined by further studies on consolidation of learning in split-brain animals subjected to unilateral self-stimulation and selective lesion of self-stimulation pathways on the contralateral side.

REFERENCES

AROUSAL AND SELF-STIMULATION


generic counts for invertebrates. This work was supported by the NSF.

KEITH STEWART THOMSON

Department of Biology and Peabody Museum of Natural History,
Yale University,
New Haven, Connecticut 06520

Received March 15; accepted April 20, 1976.


Visual attention in split-brain monkeys

How the separated hemispheres of a split-brain animal avoid conflict with each other remains uncertain. In some circumstances both hemispheres seem to be able to attend to separate stimuli simultaneously14. Yet, when a split-brain animal has received separate training in each hemisphere on visual discrimination tasks which require opposite solutions, and is then placed in a situation where either hemisphere could respond, conflict does not occur. Instead, one hemisphere dominates behaviour for a variable time, then the other assumes control1. Independent variation of degree of attention in each hemisphere has been hypothesised to account for this4. We have sought electroencephalographic evidence of such hemispheric independence of visual attention in the split-brain animal, since a change in the power spectrum of the EEG accompanies increased visual attention—that is, an increase in θ (4–7 Hz) and β (13–25 Hz) frequency power, and a decrease in α (8–12 Hz) frequency power4. Power spectrum analysis has given evidence of alternate dominance of visual attention in the separated hemispheres. This finding has implications for the absence of hemispheric conflict in such animals, and the development of hemispheric specialisation in man.

Our subjects were three rhesus monkeys, which underwent surgical section of the telencephalic and tectal commissures, the optic chiasma, and in one case, the massa intermedia. During recording sessions the animals were seated in a restraint chair, allowed to gaze about them, but not required to make operant responses. Electroencephalograms from each occipito-parietal area were recorded on f.m. tape, and later computer analysed to give average power spectra over successive 2-min epochs. Variations in the power spectra of each hemisphere were compared with those of its partner.

Usually the power spectra from each hemisphere covaried with the animals’ general state of arousal and in those conditions the difference between successive 2-min spectra had a strong tendency to be closely similar in the two hemispheres. A relative increase in power in the θ and β bands was associated with a fall in the α band. On some occasions, however, the difference spectra in the two hemispheres would become almost mirror images suggesting a shift in the dominance of attention from both to one or other of the hemispheres. To determine whether these spectral shifts, which seemed to occur spontaneously, were related to the degree of visual processing in each hemisphere, we then checked the effect of covering each eye of the subject alternately, thus restricting visual input to one hemisphere at a time. Now difference spectra that were mirror images could invariably be produced by changing the cover from one eye to the other every 2 min (Fig. 1).

These observations suggest independent variations in the degree of visual attention in each hemisphere of the split-brain monkey. It seems that an increase in visual attention in one hemisphere may occur at the same time as a decrease in attention in the other hemisphere. Such a phenomenon perhaps involves a "switching" mechanism located at the midbrain level, or below, since in one case the midline bisection included the massa intermedia, as well as the tectal and telenchephalic commissures, optic chiasma and associated supraoptic commissures.

A mechanism of this type would permit one hemisphere or the other to dominate behaviour in the presence of stimuli which might otherwise result in interhemispheric conflict. The factors determining which hemisphere achieves dominance at any given time remain unclear.

In the special case of split-brain humans, asymmetries of the EEG between hemispheres during performance of verbal or spatial tasks have been described9. A neural mechanism controlling hemispheric dominance of attention could be an important influence leading to hemispheric specialisation in man.

![Fig. 1](image) The effect on the EEG power spectra of changing the cover from one eye to the other every 2 min. Difference spectra for the two hemispheres are mirror images, a relative increase in power in θ (< 7 Hz) and β (> 13 Hz) is associated with a fall in α (8–12 Hz) for the hemisphere visually attending. These results represent the average from three split-brain rhesus monkeys.
Tubulin synthesis in developing rat visual cortex

In some species, restriction of sensory experience during early postnatal life can affect functional connectivity within the brain. Yet, in these circumstances, little is known of the changes that occur to the internal organisation of brain cells. It is also significant that certain membrane components of brain cells are continuously turning over and some of these are responsible for imparting to the membrane and cells their specialised functions. Thus plastic changes may sometimes result from a local alteration in the number or nature of molecules available for incorporation into the membrane, and it then becomes relevant to consider those systems which regulate the differential distribution of materials needed for maintenance and growth of cellular processes and synaptic junctions. Consequently, both here and in previous papers we consider the possibility that the turnover of microtubular protein (tubulin) may be involved in some of the plastic changes in the brain that result from learning or early experience. We report here significant changes in tubulin synthesis that are correlated with eye opening and the duration of a critical period.

Norwegian black hooded rats were reared under normal fluorescent lighting with a 12-h light/dark cycle (background illumination was approximately 100 lx m⁻²). At various times after birth the rats were killed and a plug of tissue removed from area 17 of the visual cortex in each hemisphere. Area 17 was located using the maps of Adams and Forrester and Krieg. Plugs of diameter 2.0 mm and weight 6.0 ± 1.5 mg were removed from each hemisphere in animals up to 1 week old. Plugs of diameter 2.5 mm and weight 10.0 ± 2.0 mg were removed from animals aged 1 to 2 weeks, and plugs of diameter 3.5 mm and weight 15.0 ± 2.5 mg were removed from animals more than 2 weeks old. Care was taken to remove the underlying white matter.

Fig. 1 DNA concentration (determined by diphenylamine method) per mg wet weight visual cortex, as a function of postnatal age. EO, Eye opening.

Fig. 2 a, Concentration of soluble tubulin in visual cortex. Ordinate shows c.p.m. [³H]-colchicine (Radiochemical Centre, Amersham) bound to tubulin that has been precipitated with vinblastine sulphate (Ely Lilly & Co. Ltd). EO, Eye opening. b, Rate of soluble tubulin synthesis in visual cortex. Ordinate shows counts per minute [¹⁴C]-labelled L-leucine (Radiochemical Centre, Amersham) in tubulin that has been precipitated with vinblastine sulphate. c, Protein concentration (determined by Folin-Lowry method) per (DNA per mg wet weight visual cortex). Each point represents the mean value of at least four determinations. Bars indicate standard deviations.

Both the concentration of soluble tubulin and its rate of synthesis were determined by a double labelling technique adapted from a method described by Feit et al. Tritiated colchicine (200 mCi mmol⁻¹) was used to assay tubulin and [¹⁴C]-labelled L-leucine (348 mCi mmol⁻¹) injected into the brain ventricle 2 h before death was used to study rate of tubulin synthesis at different times postnatally. Because cortical cells increase their volume significantly in early
Changed cortical activation and the lateral hypothalamic syndrome: a study in the split-brain cat

J. J. WRIGHT and M. D. CRAGGS*

Department of Psychiatry, University of Auckland, School of Medicine, Auckland, (New Zealand)

(Accepted February 23rd, 1978)

Following earlier work by Von Hess and Flynn, the work of Marshall et al. demonstrated that a unilateral lesion in the lateral hypothalamus results in a syndrome they called sensorimotor neglect. That is, neglect for sensory stimuli in the contralateral perceptual field, as well as motor impairment. Bilateral lesions result in aphagia, adipsia and motor paresis. Recent work has focused on the identification of the neural systems involved in the lateral hypothalamic syndrome, and it has been demonstrated that damage to dopaminergic fibres in the nigrostriatal system can account for this syndrome. The mechanism of this impairment remains unclear. One important question is how the neglect for contralateral stimuli in all sensory modes is brought about. Behavioural experiments have suggested that the deficit is neither a sensory or motor impairment alone, but impaired responding occurs only when a contralateral response to a contralateral stimulus is required. Ungerstedt has suggested that derervation of the dopaminergic input to the caudate nucleus is responsible, and points out that the caudate is involved in sensori-motor functions.

The polymodal neglect seen in the lateral hypothalamic syndrome seems similar to that following some cortical lesions — classically in the parietal lobe, but also in the frontal lobe, and cingulate gyrus. Neglect syndromes following damage at cortical sites are associated with slow wave activity in the affected hemisphere and interference with a cortico-limbic-reticular system of connections has been postulated to explain this relationship. Unilateral lesion of the mesencephalic reticular formation produces ipsilateral hemispheric slow wave activity and contralateral sensory neglect, supporting the view that a causal relationship is present.

However, impaired function of the non-specific reticular activating system has not been demonstrated in the case of the lateral hypothalamic syndrome. Feldman and Waller showed that in cats with bilateral posterolateral hypothalamic lesions, behavioural somnolence occurred despite the integrity of the reticular activating system, as demonstrated by normal cortical EEG activation in response to reticular formation.

stimulation. Conversely, mesencephalic reticular formation lesions, while producing behavioural sluggishness and EEG slow waves, did not as severely impair behavioural activity as did the posterior hypothalamic lesions. While demonstrating the lack of essential dependence between cortical-electrographic activation and behavioural arousal, this does not prove that the neural systems involved are not functionally interrelated, nor that a change in cortical arousal may not be associated with sensory neglect in the lateral hypothalamic syndrome.

To determine whether such an interrelationship exists, we have studied changes in the occipitoparietal EEG power spectrum, induced by lateral hypothalamic lesions, in 6 split-brain cats. Split-brain animals were used because of their greater hemispheric electroencephalographic independence compared to normal animals. It was thought that this might enhance any unilateral electroencephalographic changes, and eliminate the possibility that negative results might be due to obscuration of unilateral changes by transcortical effects upon the EEG.

All 6 animals were subjected to mid-line section of the corpus callosum, anterior commissure and optic chiasm, using the method of Sperry. After a recovery period of at least two weeks, monopolar stainless steel chronic depth electrodes were then implanted bilaterally into the lateral hypothalamus, at sites at or between the A10 and A13 planes (from the atlas of Jasper and Ajmone-Marsan). Bipolar platinum ball recording electrodes were placed extradurally over the occipitoparietal cortex of each hemisphere, with the balls about 10 mm apart rostrocaudally and 7.5 mm from the mid-line. Skull screws provided ground electrodes. The bilateral insertion of chronic depth electrodes was intended to counterbalance any effects of passage of the depth electrode, and to allow surgical stabilisation to occur before the lesion was placed. Following a further two week recovery period baseline EEG data was collected for power spectral analysis. The cats sat quietly in a box while 2 sec samples of artifact-free EEG were collected about every 10 sec, in blocks of 90 samples. These samples were computer stored (Digital PDP8/e), for off-line power spectral analysis. All 90 power spectra in each block were then averaged for that day. After a week of such adaptation to the experimental situation, the cats were anaesthetised with intravenous Alphaxalone/alphadolone mixture (7 mg/kg), and a unilateral lesion was placed through one depth electrode using an anodal current of 5 mA for 1 min. The side of the lesion was counterbalanced between subjects.

Following the lesion, all animals showed a neglect for visual, auditory and cutaneous stimuli contralateral to the side of the lesion. Visually guided placing of the contralateral forepaw was impaired, and there was some poverty of movement in the contralateral forepaw. In an open field situation, they showed a marked tendency to circle to the side of the lesion. No gross motor paresis was evident. This syndrome was marked during the first 4 days, but much diminished or absent (2 cats) by the seventh day. Ratings of neglect were made independently by the experimenters, and without knowledge of the accompanying EEG manifestations. Further blocks of 90 EEG samples were collected from each cat on the first, second and seventh days after lesion, and in two cases on the fourth and tenth days. The animals were then sacrificed and perfused with formol-saline by the transcerebral route. Extent of lesion, electrode place-
Fig. 1. Sagittal plane through the brain stem at 3.3 mm from mid-line. The extent of the electrolytic lesion in each cat is shown encircled in black. Abbreviations: HLA, lateral hypothalamic area; OT, optic tract; SUB, subthalamic nucleus; FF, nucleus of the fields of forel; CM, nucleus centrum medianum; VM, ventromedial complex of the thalamus; MT, mamillothalamic tract.

ment, and completeness of commissurotomy was then confirmed histologically. Commissurotomy was found to be complete in all cases. The optic chiasm was incompletely divided in two cases. Electrode tip placement was as intended. The sites of the lesions are shown schematically in Fig. 1.

Changes in the averaged power spectrum produced by the lesion were then assessed for the lesioned and control hemispheres respectively, during both the neglect and recovery stages. This was done by deriving difference spectra to show the change in logarithmic units of power at each frequency, compared to the power before lesion. (Spectral samples from the 2 days before lesion were averaged to ascertain this baseline for each cat.)

It was found that while the control hemispheres showed various shifts in spectral power following lesion, the lesioned hemispheres consistently showed a relative increase in power compared to the changes in the control, in all frequencies up to about 20 Hz, but predominantly in the lower frequencies (see Fig. 2). When the animals had recovered, or almost recovered, from the neglect syndrome, this EEG disparity was correspondingly absent, or diminished.

It is concluded that a disparity of cortical activation in the affected hemisphere is an integral part of the lateral hypothalamic syndrome, and that on recovery from the acute effects of the lesion, compensatory processes can restore regional cortical arousal back towards parity. The relationship of these findings to the changes in EEG activation
associated with neglect syndromes following cortical and reticular lesions is uncertain. Our results show a similar predominant increase in lower frequencies (i.e., a tendency toward cortical synchrony) but as power spectral analysis was not applied in the earlier studies, comparability in the higher frequencies is uncertain. It is possible that the lateral hypothalamic area contains connections important to the cortico-limbic-reticular control of attention postulated by Heilman and Valenstein\textsuperscript{6} and Watson et al.\textsuperscript{17}.

The close covariance of behavioural neglect and lowered occipitoparietal cortical activation suggests that the same neural systems may have been responsible for both, and the nigrostriatal dopaminergic neurones are therefore most likely those involved. All the lesions in the present study were placed along the course of the nigrostriatal pathway as it has been described in the cat\textsuperscript{11}.

Whatever the fibre systems involved, it is likely that they are the same as those concerned with intracranial self-stimulation (ICSS). Since our lesions induced increased EEG synchrony, they presumably included fibres which would give rise to cortical arousal when stimulated. Within the lateral hypothalamus there is a correspondence between sites which produce cortical arousal when stimulated, and those that support ICSS\textsuperscript{18,20}. The present finding is therefore indirect evidence for a physiological role of hypothalamic ICSS neurons in the regulation of sensory association cortex arousal and attention, even in the quietly resting state.

This investigation was supported by a New Zealand Government grant to Dr Wright, and a Medical Research Council (U.K.) Travelling Fellowship to Dr Craggs.
Intracranial Self-Stimulation, Cortical Arousal, and the Sensorimotor Neglect Syndrome

J. J. Wright and M. D. Craggs

Department of Psychiatry, University of Auckland, School of Medicine, Auckland, New Zealand

Received July 31, 1978; revision received January 23, 1979

Split-brain cats were prepared with bilateral lateral hypothalamic depth electrodes, at intracranial self-stimulation sites. The occipitoparietal EEG was subjected to power spectral analysis, which showed that stimulation through either depth electrode induced bilateral cortical activation. One electrode was then used to produce a hypothalamic lesion, so that a unilateral sensorimotor neglect syndrome ensued. After the hypothalamic lesion, both hemispheres retained the capacity for bilateral cortical arousal from unilateral stimulation on the nonlesion side, but the neglectful hemisphere was now relatively synchronized at all stages of generalized arousal. This increased synchrony was evident for a wide frequency range of the EEG spectrum and varied with the degree of generalized arousal. These results show that neurons traversing, or near, hypothalamic intracranial self-stimulation sites exert interacting general and lateralized effects on cortical activity. It is inferred that neurons supporting intracranial self-stimulation may exert effects on the general and regional rate of information processing in the cortex.

INTRODUCTION

Lesion of the lateral hypothalamus at the level of the ventromedial nucleus gives rise to the transient syndrome of sensorimotor neglect, an important part of which is neglect for stimuli in all sensory modes in the contralateral perceptual field (6). The mechanism of this inattention is uncertain. Behavioral studies indicated the defect is neither motor nor sensory alone (18). Other studies demonstrated that a change in cortical activation was an integral part of the syndrome. With bilateral lesions,

Abbreviations: ICNS—intracranial self-stimulation, EEG—electroencephalogram, FET—field effect transistors.
1 This investigation was supported by a New Zealand Government grant to Dr. Wright, and a Medical Research Council (U.K.) Travelling Fellowship to Dr. Craggs.

42
SENSORIMOTOR NEGLECT

43

cortical slow waves and impaired sensory responsiveness were marked in the early days after lesion (4, 12, 14). Such a change in cortical synchrony was not necessarily generalized, or nonspecific, as recent work with split-brain cats showed that sensorimotor neglect was associated with a relative increase in cortical synchrony on the side of a unilateral hypothalamic lesion (20). Spectral analysis showed this increase in synchrony to involve increased power for a wide frequency band from 0 Hz to about 20 Hz, although the effect was predominant in the lower frequencies. In showing increased slow-wave activity, the sensorimotor neglect syndrome was therefore similar to contralateral neglect syndromes arising from certain localized cortical and subcortical lesions (3, 11).

Although the lateral hypothalamus contains many fiber systems, it is those sites in the lateral hypothalamic region supporting intracranial self-stimulation (ICSS) which give rise to cortical desynchronization when stimulated (13, 19), so it may be that disruption of ICSS neurons causes disturbance of cortical arousal, and hence of behavioral responsiveness. A further reason for supposing ICSS neurons to be involved is the finding that a sensorimotor neglect syndrome appeared attributable to damage of the dopaminergic neurons of the nigrostriatal bundle (5, 7). Other neurons, such as those from the locus ceruleus, may also play a part, because lesions of this nucleus also cause a rotation syndrome similar in some respects to that of nigrostriatal lesion, but whether or not this rotation is associated with lateralized sensory neglect has yet to be determined (9). The nigrostriatal bundle and the locus ceruleus are known to be ICSS sites, and together with certain other dopaminergic and noradrenergic neurons, appeared to be principal mediators of ICSS at this level in the brain (1, 2). It was also observed that stimulation of ICSS sites produces unilateral inattention in the period following stimulation (8). This may be interpreted as arising from disruption of the normal function of the ICSS neurons during the poststimulation period.

Whatever the specific neural systems involved in sensorimotor neglect and their relationship to cortical activation, further uncertainty arises, for it appears that some neurons traversing the lateral hypothalamus exert at least two different effects on cortical activation, one generalized and one lateralized. The unilateral increase in cortical synchrony found to accompany contralateral sensorimotor neglect in split-brain animals suggests that a sharply lateralized arousal activity is usually mediated by the involved neurons. In contrast, unilateral stimulation of hypothalamic ICSS sites in split-brain monkeys produced a diffuse bilateral cortical activation, despite the absence of the telencephalic commissures, and even with the addition of a large lesion in the corresponding part of the contralateral hypothalamic region (20). From this it was concluded that cortical arousal induced by ICSS site stimulation is diffused along extrahypothalamic sub-
cortical pathways. The spectral properties of this arousal response are
unknown, because the above study considered only the time course of sup-
pression of alpha activity in each hemisphere. An immediate question
arises about how the systems producing these lateralized and generalized
hypothalamic arousal effects are related and interact.

The present experiments were therefore done to gain further insight into
the nature of the changed cortical activation in unilateral sensorimotor
neglect, by studying the response of each hemisphere to ICSS site stimula-
tion before and after lesion of a contralateral ICSS site. Specifically it was
sought (a) to determine the shift in the electroencephalogram (EEG) power
spectrum in each hemisphere associated with stimulation at lateral hypo-
thalamic ICSS sites, so that this could be contrasted with the unilateral
spectral change accompanying contralateral sensorimotor neglect, and (b)
to determine whether or not this unilateral cortical synchrony persisted in
the presence of activation from the intact hypothalamic ICSS sites.

METHODS

Subjects

The subjects were four adult domestic cats, also included as subjects in
the previously reported study (20) on changes in the EEG power spectrum
taken at rest, during sensorimotor neglect. On each day of study, the rest-
ing EEG data were collected before the data reported here.

Surgery

Each animal was subjected, under routine general anesthesia, to midline
transection of the corpus callosum, anterior commissure, and optic
chiasm, using the method of Sperry (17). After a recovery of at least 2
weeks, monopolar, stainless-steel, implanted depth electrodes with a
tip exposure of 0.5 mm were then placed bilaterally in the lateral hypo-
thalamus. In each animal one electrode tip lay at the medial forebrain
bundle in the A13 plane, whereas the contralateral electrode was placed at
a corresponding site in two cats and at the A10 plane in the others. During
the same operation bipolar, platinum ball, recording electrodes (0.5 mm
diameter) were placed extradurally over the occipitoparietal cortex of each
hemisphere, with the balls about 10 mm apart rostrocaudally and 7.5 mm
from the midline. Skull screws were used as ground electrodes. A further
2-week recovery period elapsed before experiments commenced.

Intracranial Self-Stimulation Ratings for the Depth Electrode Sites

The ICSS properties for each depth electrode were tested using the
method of O’Donahue and Hagaman (8). Stimulation with a 100-Hz sine
wave for 0.2 s was delivered at constant voltage, while current was monitored on an oscilloscope. This method permits grading of ICSS on a 0–5 V scale, and in every case the depth electrodes yielded grades between III–IV at 300 to 800 μA (RMS). The same current was then used appropriately for each electrode on subsequent testing.

*Electroencephalogram Recordings and Power Spectrum Computation*

The EEG was recorded via a socket fixed to the skull of each cat. The connecting plug contained matched-pair unity gain field effect transistors (FET) powered remotely by batteries. The FET output was lead to differential amplifiers (Tektronix AM 502) and then filtered by Rockland 1022F filters, to pass a frequency band of 1 to 30 Hz, cut at 24 dB per octave at the high-frequency end and 6 dB per octave at the low-frequency end. The filtered signal was monitored on an oscilloscope, and then converted to 10-ms interval digital signals for storage by a small general-purpose computer (Digital PDP 8/e).

Samples of the EEG were collected for 2 s immediately before the forced delivery of an appropriate stimulus to the lateral hypothalamus. A further 2-s EEG sample beginning 0.5 s after the stimulation was then collected. A block of 30 such prestimulus and poststimulus samples was obtained for each electrode in each cat at two different rates of stimulation:

(i) *Slow Stimulation.* Stimuli were delivered while the cat sat quietly. Each stimulus elicited a mild behavioral alerting reaction. The animal was then allowed to settle down to quiet rest before delivery of the next stimulus. This took between several seconds and a few minutes.

(ii) *Fast Stimulation.* Stimuli were delivered every 4.5 s; i.e., as fast as consistent with the recording of pre- and poststimulation EEG data. During this condition, the cats were highly alert, and displayed seeking behavior throughout.

Each 2-s EEG data block was subjected to power spectral analysis offline on the PDP 8/e computer. A variant of the Autocorrelation/Fourier transformation algorithm was utilized (16) to calculate power at 0.5-Hz intervals from 1 to 30 Hz, for each 2-s EEG data sample. These 2-s power spectra were then averaged within each 30 trial block, into pre- and poststimulus average power spectra, for the two rates of stimulation at each electrode in each cat.

*Recordings before and after Unilateral Lesion of the Lateral Hypothalamus*

Such average power spectra were collected from each cat, utilizing seven of the depth electrode sites. From one of the seven electrodes, fast
stimulation data were not collected. After this, each cat was anesthetised with intravenous Alphaxalone–alphadolone mixture (Saffan) in a dose of 7 mg/kg. Then a unilateral lesion was placed through one depth electrode using an anodal DC current of 5 mA for 1 min. The side and site (A10 or A13 plane) of the lesion were varied in each of the cats, but the contralateral site left intact to support the ICSS subsequently, was always at the A13 plane.

Two days after the lesion, all animals showed a marked contralateral sensorimotor neglect as previously reported (20). Average EEG power spectral data associated with stimulation of the intact ICSS site were then obtained. The efficacy of these sites in sustaining ICSS was again checked, as was the abolition of ICSS from the lesion sites. The EEG power spectra before and after hypothalamic lesion could then be compared.

Additional Calculation of Results

The PDP 8/e computer was again used to obtain the averaged power spectra, which were used to calculate two types of difference spectra.

(i) Computation of the Spectral Change in Each Hemisphere Associated with Unilateral ICSS Site Stimulation. The averaged spectra obtained at the slow rate of stimulation before lesion were used to calculate difference spectra by subtracting the average power spectrum before stimulation (expressed as $\log_2 V^2$ at each 0.5-Hz interval) from the corresponding averaged poststimulation spectrum. The mean difference spectra for the hemispheres contralateral and ipsilateral to stimulation were then calculated by averaging the differences obtained by the use of six electrodes in the three cats for which complete left and right side stimulation data were available. The mean spectral changes ipsilateral and contralateral to stimulation were thus determined under precisely matched conditions.

(ii) Estimation of Changes in the EEG Spectrum after Unilateral Lesion. Additional difference spectra were calculated to determine ipsilateral changes in the EEG spectrum accompanying contralateral sensorimotor neglect for all states of stimulation. For each class of averaged power spectrum (pre/poststimulation, fast or slow stimulation, ipsilateral or contralateral to lesion) the spectrum before lesion (again expressed as $\log_2 V^2$ at each 0.5-Hz interval) was subtracted from the corresponding postlesion spectrum. Then, in each case the difference spectrum for the hemisphere on the side of stimulation (opposite the lesion) was treated as a baseline (i.e., a control change) and subtracted (again in $\log_2 V^2$ form) from the change in the spectrum on the side of the lesion. This yielded corrected difference spectra showing the relative change in spectral power between hemispheres, consequent to the lesion, for each state of activation.
Histological Verification

Ten days after lesion the animals were killed and perfused with formal saline by the transcardiac route. Extent of the lesion, electrode placement, and completeness of commissurotomy were then confirmed histologically.

RESULTS

Power Spectral Changes

(i) The Spectral Change in Each Hemisphere Associated with Unilateral ICSS Site Stimulation before Lesion. In all cases ICSS site stimulation produced marked cortical arousal from the resting state, in both hemispheres. Figure 1 shows the difference spectra associated with this arousal response at the slow rate of stimulation, calculated as described under Methods. Fast stimulation produced a sustained level of arousal in both hemispheres evident in both the pre- and poststimulation spectra, and generally similar to the poststimulation spectra gathered at the slow rate of stimulation.

It is seen that the suppression of low frequencies and the enhancement of the high frequencies (20 Hz or more) was greater on the side of stimulation. Applying a t-test to the mean and variance of these pooled data, at each 0.5-Hz interval to test the significance of this difference between hemispheres, we found highly significant differences at almost every point. As a guide, the significance levels at 5-Hz intervals are shown in Table 1.

![Graph showing power spectral changes](image_url)

**Fig. 1.** The effect of slow monosynaptic stimulation at a hypothalamic intracranial self-stimulation site on the changes in spectral power of the EEG. This is the mean of six experiments, two from each of three split-brain cats with bilateral hypothalamic electrodes. Difference spectra were calculated by subtracting the prestimulus from the poststimulus power at every frequency interval of 0.5 Hz. Large, connected filled circles—hemispheres ipsilateral to site of stimulation, small, unconnected filled circles—hemispheres contralateral to site of stimulation.
The Spectral Change in Each Hemisphere Associated with Unilateral ICSS Site Stimulation after Lesion. In all cases, both hemispheres remained responsive to unilateral stimulation, as is shown in a specific case in Fig. 2. This was true at slow and fast rates of stimulation, as before the lesion. The asymmetries present after the lesion complicate direct comparison of the degree of suppression in each hemisphere with that before lesion, so this was not attempted. However, comparison of the relative degree of suppression in each hemisphere is implicit in the following section.

(ii) Estimation of Changes in the EEG Spectrum after Unilateral Le-
tion. The neglectful hemispheres showed a relative increase in spectral power, at all states of activation. The average for all cases is shown in Fig. 3. A relative increase at all states of activation was evident in every individual case. Also apparent in the individual cases was the trend toward greater relative increase in power at the lower frequencies at the slow rate of stimulation, and to increased relative power at the high frequencies in each poststimulation state. Although consistent from case to case, mathematical difficulties implicit in the Autocorrelation/Fourier transformation method do not permit rigorous application of confidence limits to these trends, because of the unknown degree to which side-bands of the calculated power at each frequency interact (10).

**Histological Data**

Commissurotomy was found to be complete in all cases, but the optic chiasm was incompletely divided in two cases. Electrode tip placement was as intended. The sites of the lesions (which surrounded the tip of the electrode used to produce them) are shown schematically in Fig. 4.

![Graphs showing spectral power changes](image-url)

**Fig. 3.** Relative change in spectral power of neglectful hemispheres after lesion, and the change of this with stimulation of the site of intracranial self-stimulation. The change of spectral power in the nonlesion hemisphere is treated as baseline. These graphs show the mean of the relative increases, for all cats studied, of power in the neglectful hemisphere before and after stimulation at both rates.
FIG. 4. Sagittal plane through the brain stem at 3.3 mm from midline. The extent of the electrolytic lesion in each cat is encircled. Left and right lesions have been projected onto the same sagittal plane. Abbreviations: HLA—lateral hypothalamic area, OT—optic tract, SUB—subthalamic nucleus, FF—nucleus of the fields of forel, CM—nucleus centrum medianum, VM—ventromedial complex of the thalamus, MT—mamillothalamic tract. (Based on a sagittal plane from Alvin L. Berman, 1968, The Brain-Stem of the Cat. Univ. Wisconsin Press, Madison.)

DISCUSSION

These data show that unilateral lesion of the lateral hypothalamus at an ICSS site gives rise to a contralateral neglect syndrome and ipsilateral EEG synchrony over a wide frequency range. This increased synchrony relative to the intact hemisphere is shown to persist even when both hemispheres are raised to a high state of activation by stimulation of the intact side of the hypothalamus.
In contrast to this unilateral effect of a lesion at an ICSS site, unilateral stimulation of similar sites produces a bilateral cortical activation, with or without a lesion on the other side. Because the animals were split-brain this must have been via subcortical pathways. The spectral form of this bilateral cortical desynchronization was spread over a wide spectral range, from 0 to 20 Hz or more, and is thus like the mirror image of the unilateral change at rest, produced by lesion. Both lateralized and generalized modes of activation are therefore involved, and both desynchronize cortical activity over a wide spectrum, though by different pathways. It was argued previously that the bilateral arousal is probably due to the widespread activation of units of the reticular activating system of the brain stem, possibly by brain stem projections of ICSS neurons (19), whereas unilateral change in cortical activation accompanying sensorimotor neglect reflects interruption of some lateralized projection of the arousal system. The ascending projections of the dopaminergic fibers of the nigrostriatal bundle or related neurons appear likely candidates, as mentioned in the Introduction, an action presumably exerted through the caudate nucleus, as suggested by Ljungberg and Ungerstedt (5).

We assume the EEG as recorded in these experiments to arise from continuously active, spatially distributed information processing elements in the cortex. Information theory indicates that all information processing systems carry their maximum information when their activity is most unpredictable (15). Synchronization of the EEG means increased orderliness and predictability of the cortical elements' activity and implies a lowered rate of processing of information. In the present case, the increased synchrony after a hypothalamic lesion is not merely due to the appearance of fixed cortical slow-wave activity, but has a variable spectrum with increased synchrony of activity at higher frequencies as well. This might arise from a generally increased redundancy of activation of cortical elements. Support for this comes from a recent finding demonstrating relative increase in voltage in late components of the visually evoked response in the neglectful hemisphere (21).

Although generalized and lateralized effects upon cortical activation are both mediated by fibers traversing through or near ICSS sites, the definite identification of ICSS neurons with these functions requires further investigation. Until this is determined it may be hypothesized that such neurons play a physiological role in regulation, both general and regional, of flow of information in the cerebrum.

REFERENCES

2. CROW, T. J. 1971. The relation between electrical self-stimulation sites and catechol-
4. KOLB, B., AND I. Q. WHISHAW. 1977. Effects of brain lesions and atropine on hip-
5. LINDBERG, T., AND U. UNGEREDET. 1976. Sensory inattention produced by 6-
hydroxydopamine-induced degeneration of ascending dopaminergic neurons in the
6. MARSHALL, J. F., TURNER, B. M., AND P. TEITELBAUM. 1971. Sensory neglect pro-
damage and the lateral hypothalamic syndrome. *J. Comp. Physiol. Psychol.* **87**: 808–830.
producing self-stimulation and unilateral inattention. *Brain Res.* **5**: 289–305.
London.
physiol.* **30**: 83–86.
on voluntary behaviour and hippocampal electroencephalograms in the rat. *J. Comp.
Physiol. Psychol.* **86**: 768–786.
13. ROLLS, E. T. 1971. Involvement of brain-stem units in medial forebrain bundle self-
**27**: 379–423.
Wiley, New York.
*Harvey Lect.* **62**: 293–323.
18. TURNER, B. H. 1973. Sensorimotor syndrome produced by lesions of the amygdala and
lateral hypothalamicus. *J. Comp. Physiol. Psychol.* **82**: 37–47.
Visual-Evoked Response in Lateral Hypothalamic Neglect

J. J. Wright, M. D. Craggs, and A. A. Sergejew

Department of Psychiatry, University of Auckland School of Medicine,
Auckland, New Zealand

Received January 3, 1978

Visual-evoked responses were recorded bilaterally from the occipitoparietal region of six split-brain cats, before and 2 days after a unilateral hypothalamic lesion had produced a contralateral sensorimotor neglect syndrome. No consistent change in each visual-evoked response was detected, but there was an average increase in voltage in slow components on the side of the lesion compared to changes on the intact side. It is concluded that sensorimotor neglect is not associated with reduction of sensory input, and that the changed cortical synchrony accompanying this syndrome therefore reflects disruption of a nonsensory mechanism.

INTRODUCTION

Unilateral lesion of the lateral hypothalamus or elsewhere along the nigrostriatal pathway gives rise to the sensorimotor neglect syndrome (8, 9). This is characterized by neglect for stimuli in the sensory field contralateral to the lesion.

Similar neglect syndromes can arise from neural damage at a variety of sites other than the lateral hypothalamus—notably from damage to the parietal, frontal, or cingulate cortex, and also the reticular formation (1, 5, 13, 14). These syndromes have in common the occurrence of increased electroencephalograph (EEG) synchrony on the side of the lesion (6, 10, 14) and this was recently shown to be true for lateral hypothalamic sensorimotor neglect as well (3, 17, 18). The comparability of the increased cortical synchrony in these disparate syndromes, as regards the spectra character of the synchrony, is in doubt (17, 18). In the syndromes of cortica

1 This investigation was supported by a New Zealand Government grant to Dr. Wright and Medical Research Council (UK) Travelling Fellowship to Dr. Craggs. Dr. Craggs' permanent address is: Institute of Psychiatry, University of London, London, England.

Abbreviations: VER—visual-evoked response; EEG—electroencephalograph.

0014-4886/79/070178-08$02.00/0
Copyright © 1979 by Academic Press, Inc.
All rights of reproduction in any form reserved.
origin, there has been controversy as to whether neglect reflects impairment of sensory input to the cortex, or the derangement of central arousal mechanisms (15). In the case of the frontal-arcuate neglect syndrome, work on the somatosensory evoked potential strongly favors an arousal, rather than sensory mechanism, as only the later components of the evoked response are affected (15). On the other hand, it is noted that sensory pathway damage may indeed produce a similar behavioral syndrome (4, 11, 12).

The experiments described in this paper were an attempt to resolve this uncertainty in the case of the lateral hypothalamic syndrome. Although it seems unlikely that direct damage to classical sensory pathways might arise from a hypothalamic lesion, disruption of a centripetal mechanism regulating sensory input cannot be excluded. As Ljungberg and Ungerstedt (7) suggested, nigrostriatal tract damage may produce sensorimotor neglect by interference with caudate nucleus regulation of sensorimotor integration. High-frequency stimulation of this nucleus exerted marked effects upon cortical evoked responses (2). Thus sensorimotor neglect might involve changes in the cortical evoked potential.

Vision was chosen as the representative sensory mode for the present work because (a) sensorimotor neglect appears to include visual neglect along with other modes (7, 8, 16) and (b) because the anatomic proximity of the visual pathways to the lateral hypothalamus might make them prone to disruption by mechanical and thermal effects and ensuing cerebral edema arising from the lesion.

METHODS

Animals. These were six adult split-brain cats with bilateral hypothalamic depth electrodes and bipolar occipitoparietal extradural recording electrodes. The methods of their surgical preparation and placement of the hypothalamic lesions were described elsewhere (17, 18). Commissurotomy was intended to include the corpus callosum, anterior commissure, and optic chiasm in each case. Split-brain animals were used to minimize possible interhemispheric interactions of the evoked response, and to permit accurate within-subject control of nonspecific changes during the course of the experiment.

Recording Technique. Averaged visual-evoked responses (VER) were recorded as follows: The visual stimulus used was a white light flash of 1-μs duration, produced by a Grass Ps-3 photostimulator. Photic stimuli were delivered directly toward the midline of the cat's face from a distance of 91 cm (3 ft.), at intervals between 2 and 10 s, varied randomly, while the cat sat facing forward in a narrow three-sided box. Bipolar electroencephalograph (EEG) signals were recorded via an electrical connector fixed to the skull.
After unity-gain buffering by field-effect transistors, the signals were led to differential amplifiers, filtered to the 0- to 30-Hz band and stored on disk on a PDP8/e computer after analog to digital conversation at 2-ms intervals. Five-hundred-millisecond segments of EEG were captured time-locked to the photic stimuli, in groups of 30 to 120. These segments were then averaged off-line on the same computer and the average voltage and its standard deviation for each 2-ms interval was thus determined for each group of evoked responses. Such comparatively short recording times were used to avoid significant habituation during each test, which might have obscured early responses to stimuli.

*Demonstration of the Independence of the VER in Each Hemisphere of the Split-Brain Animal (before Hypothalamic Lesion).* Because evoked responses may be recorded at considerable distances from their origin, we sought to demonstrate that the VER recorded from each hemisphere of these subjects was independent of the other. In this we took advantage of the effect of optic chiasm transection, which confines visual input from each eye via the classical visual pathway to the ipsilateral hemisphere. By covering one eye at a time with a black mask we checked for corresponding independence of the VER. In addition to determining whether cross-interference by volume conduction was a factor, this experiment was intended to demonstrate any effects on the VER produced by transmission to the opposite hemisphere along nonclassical pathways.

*Measurement of the VER before and after Lateral Hypothalamic Lesion.* Averaged VERs were recorded from each animal, without eye cover, at separate daily sessions for 3 to 10 days before lesion to obtain baseline measurements under stable conditions immediately before the lesion was made. The cats were then anesthetized with intravenous alphaxalone/alphadolone mixture (7 mg/kg) and a unilateral lesion was placed through one depth electrode using an anodal current of 5 mA for 1 min. The side of the lesion was counterbalanced between subjects.

Averaged VERs were recorded again by the same technique on the 2nd day after the lesion. At that time all animals showed a neglect for visual, auditory, and cutaneous stimuli contralateral to the side of the lesion. Visually guided placing of the contralateral forepaw was impaired and there was some poverty of movement in the contralateral forepaw. In an open field they showed a marked tendency to circle to the side of the lesion. No gross motor paresis was evident. The subsequent resolution of these changes was consistent with previous reports (16).

*Histological Procedures.* At termination of the experiment the animals were killed and underwent cerebral perfusion with formol-saline by the transcardiac route. Commissurotomy, electrode position, and lesion extent were then checked using Nissl- and ferricyanide-stained sections.
RESULTS

**Independence of the VER in Each Hemisphere before Lesion.** In the four subjects with complete optic chiasm section, this was successfully demonstrated. An example is shown in Fig. 1. It will be noted that although the ipsilateral VER was eliminated by eye cover, entrainment of the EEG was not. The other two cats showed partial persistence of the ipsilateral VER with eye cover.
The VER before and after Unilateral Lateral Hypothalamic Lesion. The VERs recorded immediately before lesion were compared with those recorded on the 2nd day after lesion for each subject. One example is shown in Fig. 2. Considerable change in the VER was noted after the lesion, in both the neglectful and control hemispheres, to the extent that identification of individual prelesion peaks was uncertain in the postlesion VER in some cases. These variations in the VER were not, however, consistent from subject to subject either for voltage or latency variations of discernable peaks, in either hemisphere. To determine if a group trend was present, data from four subjects were pooled to obtain the group average and mean standard deviations of voltage at each 2-ms interval, for neglectful and control hemispheres, pre- and postlesion. The four subjects whose data were used in this way were those with VERs based on similar numbers (120 ± 3) of individual evoked potentials before and after lesion, so that equivalent conditions pertained for all cases. These group-averaged VERs showed four peak values discernible before and after lesion, with an initial positive deflection. These peaks were therefore denoted \( P_1, N_1, P_2, \)

![Diagrams showing VER changes before and after lesion](image-url)
TABLE 1
Means and Standard Errors of Amplitude of Major Wave Forms
in the Average Evoked Response

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pref lesion</th>
<th>Post lesion</th>
<th>Pref lesion</th>
<th>Post lesion</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak time (ms)</td>
<td>82</td>
<td>88</td>
<td>80</td>
<td>76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu V$</td>
<td>14.06 ± 0.86</td>
<td>15.36 ± 0.77</td>
<td>15.65 ± 0.68</td>
<td>12.78 ± 0.57</td>
<td>-2.87</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>$N_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak time (ms)</td>
<td>112</td>
<td>112</td>
<td>108</td>
<td>105</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu V$</td>
<td>-2.85 ± 0.87</td>
<td>0.41 ± 0.73</td>
<td>-0.61 ± 0.86</td>
<td>-1.15 ± 0.64</td>
<td>-2.44</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>$P_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak time (ms)</td>
<td>132</td>
<td>132</td>
<td>128</td>
<td>134</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu V$</td>
<td>5.29 ± 0.87</td>
<td>10.00 ± 0.82</td>
<td>5.00 ± 0.82</td>
<td>7.43 ± 0.67</td>
<td>-1.43</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>$N_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak time (ms)</td>
<td>230</td>
<td>236</td>
<td>238</td>
<td>230</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu V$</td>
<td>-10.94 ± 0.81</td>
<td>-13.82 ± 0.72</td>
<td>-8.20 ± 0.62</td>
<td>-12.48 ± 0.60</td>
<td>-1.01</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>479</td>
<td>480</td>
<td>479</td>
<td>480</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

and $N_2$ in order of increasing latency. Changes in their voltages were compared between hemispheres using a $t$-test of differences before and after lesion (see Table 1). Since direct measures of variation of peak latency are not available because of the system of data capture used, no estimation of confidence can be applied to the shifts in peak time.

Histological Results. These were previously described (17). In all animals commissurotomy was complete. Optic chiasm transection was complete in four cases and partial in two. Electrode situation and lesions were as intended in the lateral hypothalamus and have also been reported previously (17). The sites and extent of the lesions are shown in Fig. 3.

CONCLUSION

The absence of consistent change in the evoked potential during sensorimotor neglect compared to that produced by eye cover appears to rule out the possibility that any gross reduction of visual input was a factor in the neglect for visual stimuli. If changes in centrietal control of visual input are of importance to sensorimotor neglect, then they are likewise not of gross magnitude as measured by the VER. It would appear, rather, that the balance between hemispheres has been in some way changed. Caution is still warranted on this issue as (a) these results are limited to the visual mode, (b) a very gross visual stimulus was utilized, and (c) only evoked components in the 0- to 30-Hz band were considered. Data were collected from sampling runs too short for significant habituation to occur [a factor which might have differentiated these results from those of Watson et al.].
(15), who showed significant changes in the late components of the somatic evoked response in frontal-arcuate neglect. Yet, although inconsistent changes occurred in individual VERs, on average a significant increase in voltage in components of about 100-ms latency was shown in the present results. This is therefore comparable to the changes shown by Watson et al. Interpretation must therefore favor a nonsensory mechanism for the behavioral neglect, and the associated increased cortical synchrony.

We suggested elsewhere that the changed cortical synchrony associated with sensorimotor neglect reflects impaired modulation of central information processing and tentatively linked this function to those hypothalamic neurons involved in the intracranial self-stimulation phenomenon (17, 18). Such an interpretation therefore still stands.
REFERENCES


A PRELIMINARY MATHEMATICAL MODEL FOR LATERAL HYPOthalAMIC REGULATION OF ELECTROCORTICAL ACTIVITY

J.J. WRIGHT and G.R. IHAKA

(J.J.W.) Department of Psychiatry, University of Auckland School of Medicine, Auckland, and (G.R.I.) Department of Mathematics, University of Auckland, Auckland (New Zealand)

(Accepted for publication: April 10, 1981)

The electroencephalogram (EEG) arising from a particular domain of cortical tissue appears to arise from wave-like activity in the dendrites of cerebral neurones, which may result from postsynaptic potentials, dendro-dendritic interactions, activity in the surrounding ionic plasma, and/or mechanisms yet unknown (Pribram 1971; Elul 1972; Schmitt et al. 1976; Adey 1977). Since the EEG varies markedly in a variety of behavioural states, it is thus suspected of reflecting cortical information processing in some way (e.g., Walter et al. 1967). But electrocortical activity bears only a limited relationship to cross-correlation of firing in closely situated cortical neurones (Noda and Adey 1970) and in some circumstances is dissociated from neurone firing (Elul 1972, 1974). Such findings have led attempts at theoretical modelling of cortical information processing away from early nerve-net models (Hebb 1949) toward models at least equally concerned with electrotonic processing (Pribram 1971; Adey 1977). One requirement of these theories, if they are to provide complete models of information handling, is a mathematical description of the modulation of EEG activity by subcortical systems.

Perhaps the most important of these subcortical systems is the lateral hypothalamus, which exerts major 'desynchro-

nising' effects upon EEG activity (Scheibel and Scheibel 1967); effects associated with powerful alerting and other behavioural responses mediated by this pathway (Aden and Lindsley 1959; Olds and Peretz 1960; Glickman and Feldman 1961; Adey et al. 1962; Kolb and Whishaw 1977; De Ryck and Teitelbaum 1978; Wright and Craggs 1979). It has recently been shown that lesion of the lateral hypothalamus interrupts a mechanism exerting specific effects upon the spectral density of EEG power in resting conditions (Wright 1981). This indicates that lateral hypothalamic effects upon electrocortical activity might be susceptible to formulation within a mathematical model. The development of such a model may be dependent on whether linear means of analysis can be applied, at least in some defined preliminary conditions, since there are no general solutions for the analysis of non-linear systems.

The present paper advances a provisional model to this end. While this is a simple, linear model, constrained to stimulation and lesion data gathered in defined circumstances, it suggests that one of a family of higher order models of greater explanatory power may be found applicable outside these constraints.

Description of the model

The broad anatomical framework within which the model is set is diagrammed in Fig.
1 and arises from the following considerations. Cerebral electrocortical activity, in general terms, arises as the result of: (a) intracortical activities; (b) cortical-subcortical interactions, taking place over pathways which to some extent form cross-exciting and/or self-exciting loops; (c) signals input along sensory channels, and from elsewhere in the brain. The EEG arises largely from the outer cortical layers, and while the signal varies with position and depth in the cortex, each set of electrode positions gives rise to a time series derived from a weighted summation of activity over the extended surface.

Our initial presumption will be that the lateral hypothalamus exerts a linear filtering action upon this EEG time series (by means later hypothesised), and that this action is time-invariant so long as conditions remain stable, over epochs up to several minutes.

Our second presumption is that this description may be held to apply identically for opposite hemispheres of the brain, under conditions of symmetry — i.e., the input signals to each hemisphere may be regarded as approximately equal, and subsequent processes as following a mirror-image symmetry, especially in the absence of hemispheric specialisation. EEG signals arising from symmetrically opposite sites will thus be closely similar. While short-term asymmetries in the EEG occur and may be related to signal input (Wright and Craggs 1976), the left and right EEG spectrums covary in different behavioural states (Walter et al. 1967). This simplifying assumption therefore seems fairly applicable to average conditions.

Where minor asymmetries in electrode positioning occur, the effect on the symmetry of the EEG will be provisionally accounted for in this model by a linear transfer function, as will other time-invariant asymmetries. This is only a rough approximation, as a linear transfer description is only readily applicable to waves of surface voltage moving unchanged through the domain of recording. A fully general treatment would require consideration of attenuation of waves, differential rates of propagation, etc. These more involved processes may introduce non-linearities. However, averaging of data collected in similar conditions from several subjects will tend to randomise these effects.

Under these assumptions, let the left and right EEG signals be \( X(t) \) and \( Y(t) \). Then their complex spectra are related by

\[
f_{yx}(\omega) = A(\omega)f_{xx}(\omega)
\]

where \( \omega \) is frequency, \( f_{yx}(\omega) \) is the cross-spectrum of \( Y(t) \) and \( X(t) \), \( A(\omega) \) is the transfer function due to minor electrode asymmetry, and \( f_{xx}(\omega) \) is the autospectrum of \( X(t) \).

Now if separate epochs of \( X(t) \) and \( Y(t) \) are recorded in stable conditions before and after unilateral lesion of the lateral hypothalamus, then

\[
f_{yx}^{\text{pre}}(\omega) = A(\omega)f_{xx}^{\text{pre}}(\omega)
\]

\[
f_{yx}^{\text{post}}(\omega) = B(\omega)A(\omega)f_{xx}^{\text{post}}(\omega)
\]

where \( B(\omega) \) represents the new transfer function describing the change in filtering on the side of lesion, and the super-scripts 'pre' and 'post' indicate the pre- and post-lesion epoch of recording.

Therefore

\[
B(\omega) = \frac{f_{yx}^{\text{post}}}{f_{yx}^{\text{pre}}} \frac{f_{xx}^{\text{pre}}}{f_{xx}^{\text{post}}}(\omega)
\]
This leads by decomposition of $B(\omega)$ into gain and phase aspects $G(\omega)$ and $\Phi(\omega)$, to

$G(\omega) = |B(\omega)|$

and

$\Phi(\omega) = \arg(B(\omega))$

By definition, $G(\omega)$ may also be expressed as

$$G^2(\omega) = \frac{V_Y^{post} V_Y^{pre}}{V_X^{post} V_X^{pre}} (\omega)$$

where $G^2(\omega)$ is the relative squared gain attributable to the effect of the lesion, $V^2(\omega)$ is the power spectrum of the EEG, and the subscripts $Y$ and $X$ indicate the side of the lesion and the non-lesion side respectively.

These considerations enable examination of a mathematical model of the filtering action attributed to the lateral hypothalamus.

Our principle hypothesis is that the commonly seen peaks of activity in the EEG power spectrum represent natural resonant frequencies of EEG activity, which are damped by activity mediated via the lateral hypothalamus. The total system may thus be viewed as a group of interlinked damped harmonic oscillators driven by input signals. We make no hypothesis regarding the mechanism of oscillation, nor the means by which damping is effected, and leave out of immediate consideration the mathematical nature of the interlinkages.

In later fitting this model to experimental data obtained by computing $G^2(\omega)$ from EEG records made before and after unilateral lesion of the lateral hypothalamus, we make the following provisions:

(a) Since there is no reason to assume that electrocortical activity is generated in a constant variance medium, allowance must be made for a change after lesion in relative amplification of power, $K$, between hemispheres, as the change in damped activity may be expected to influence the signal variance, thus influencing total EEG power.

(b) If the experiment is conducted in steady-state conditions, at such a level of cortical activation that a single cerebral rhythm is dominant, then consideration can be constrained to the release from damping of driven oscillation at the dominant frequency.

A single oscillator of simple harmonic type, driven by an input signal and subject to damping forces is described by the relation

$$\frac{d^2 \hat{Y}(t)}{dt^2} + D \frac{d\hat{Y}(t)}{dt} + N^2 \hat{Y}(t) = \hat{X}(t)$$

where $\hat{Y}(t)$ is the output signal, $\hat{X}(t)$ is the driving signal, $D$ is the damping coefficient, and $N$ is the natural frequency of oscillation. Such a system acts as a linear filter, converting a generalized input signal, with a Fourier representation

$$\hat{X}(t) = \int_{-\infty}^{\infty} g(\omega) \cos(\omega t) \cdot d\omega$$

into

$$\hat{Y}(t) = \int_{-\infty}^{\infty} \hat{G}(\omega) g(\omega) \cos(\omega t + \Phi(\omega)) \cdot d\omega$$

where

$$\hat{G}(\omega) = \{(N^2 - \omega^2)^2 + D^2 \omega^4\}^{-1/2}$$

and

$$\Phi(\omega) = \arctan\left(\frac{D\omega}{N^2 - \omega^2}\right)$$

Under the provisions made above, we shall expect $G^2(\omega)$ found experimentally to be equal to

$$K \left[\frac{\hat{G}^2(\omega)^{post}}{\hat{G}^2(\omega)^{pre}}\right]$$

on side of lesion

$$\left[\frac{\hat{G}^2(\omega)^{post}}{\hat{G}^2(\omega)^{pre}}\right]$$

on control side

Under the condition of pre-lesion symmetry, the pre-lesion values of $\hat{G}^2(\omega)$ will be equal, i.e., $G^2(\omega)$ found experimentally should be fitted by the function

$$G^2(\omega) = K\left[\frac{(N^2 - \omega^2)^2 + D^2 \omega^4}{(N^2 - \omega^2)^2 + D^2 \omega^4}\right]$$
where $N$ will have a value equal to the frequency of the dominant oscillator, and $D_c$ and $D_L$ will be coefficients of damping on the control and lesion sides during the post-lesion epoch, of such values that $D_L$ will be less than $D_c$.

By similar reasoning, electrical stimulation of lateral hypothalamic sites should increase the damping effect, and if the damping system is capable of differential influence on different parts of the cortical system, unilateral stimulation should produce similar effects differing in degree on each side of the brain. If conditions are otherwise similar to those pertaining in the lesion studies, then using EEG data recorded in short epochs before and after unilateral stimulation of a lateral hypothalamic site, the averaged power spectra obtained from each recording site, before and after stimulation, should be related in such a way that where the subscripts $Y$ and $X$ now represent the side of stimulation and the opposite side, and the subscripts $R$ and $S$ represent values of the pre- and post-stimulation epochs respectively, then $V_{Y,R}/V_{Y,R}(\omega)$ and $V_{X,S}/V_{X,R}(\omega)$ will both be fitted by the expression

$$K \left[ \frac{(N^2 - \omega^2)^2 + D_R^2 \omega^2}{(N^2 - \omega^2)^2 + D_S^2 \omega^2} \right]$$

where $N$ will again be the frequency of the dominant oscillator. $D_S$ will be expected to be greater than $D_R$.

It will be noticed that the proposed model also makes predictions regarding phase shifts. For reasons given later, these predictions cannot be as easily constrained to consideration of a single oscillator.

### Methods

Detailed accounts of the experimental methods used in these studies are given in earlier papers (Wright and Craggs 1978, 1979; Wright 1981). The data utilised in the present paper were all obtained from animals with split brains, as an early concern was that transcommisural activity might obscure lateralising effects. Subsequent lesion studies on animals without commissurotomy have shown that essentially similar results are obtained (Wright 1981).

The experimental animals were adult domestic cats. They underwent section of the corpus callosum, anterior commissure and optic chiasm, by the method of Sperry (1968). After recovery, bilateral depth electrodes were inserted to the lateral hypothalamus at the level of the ventromedial nucleus, or the posterior hypothalamus, and bipolar extradural recording electrodes with 0.5 mm platinum ball ends were inserted over the posterior suprasylvian area, about 10 mm apart rostrocaudally, and about 7.5 mm from the midline. Skull screws provided ground electrodes, and all wiring was connected to sockets buried in acrylic on the animals skull. A 2 week (or longer) recovery period ensued before experiments. Preliminary stimulation studies showed that all depth electrodes supported intracranial self-stimulation (ICSS) or marked aversive behaviours.

### Lesion studies

While the animals sat in a state of quiet rest, 180 sec of EEG was recorded simultaneously from each channel, in 2 sec data blocks. After differential amplification and filtering to the 1–30 Hz band the signals were digitised at 10 msec intervals, and each 2 sec block was subjected to power spectral analysis off-line on a PDP 8/e computer, using the autocorrelation/Fourier transformation algorithm. These were then averaged to produce power spectral estimates at 0.5 Hz intervals for the 180 sec epoch. In these circumstances of quiet rest, theta rhythm was the dominant rhythm. The animals were then given a short acting anaesthetic and underwent unilateral electrolytic lesion via one of the depth electrodes. A contralateral sensorimotor neglect syndrome ensued. Two days after lesion, EEG recordings were again made in similar circumstances, and analysed as before.
The spectral estimates were then used to obtain the values of
\[ G^2(\omega) = \frac{V_x^{\text{post}}/V_x^{\text{pre}}}{V_y^{\text{post}}/V_y^{\text{pre}}} \] (\omega) for each animal.

Estimates of \( G^2(\omega) \) were then averaged from data obtained from 6 cats. Criteria for choosing data of comparable nature before and after lesion have been given in an earlier paper (Wright 1981).

**Stimulation studies**

These data were obtained from cats prepared as above, before they entered lesion studies similar to those reported above.

While the animals remained in a state of quiet rest (so theta rhythm was again the dominant rhythm), slow, non-contingent stimuli were delivered to the tip of one depth electrode. Stimulation was with a 100 Hz sine wave of constant voltage, and 0.2 sec duration, delivering an amperage varying between 300 and 800 \( \mu \)A at different sites — a stimulus previously found to support ICSS at each site. In the two second period immediately before stimulation, EEG samples were collected from each channel. A second pair of samples were collected in a 2 sec block beginning 0.5 sec after the stimulus was delivered. This was repeated 30 times in each session. The 2 sec blocks were then spectrally analysed, and the pre- and post-stimulation spectra averaged respectively, for each channel. This enabled \( V_x^{\text{R,s}}/V_x^{\text{R}}(\omega) \) and \( V_x^{\text{R,s}}/V_x^{\text{L}}(\omega) \) to be calculated for data obtained from each electrode site. These estimates of relative squared gain were obtained from the stimulation of 6 electrodes, two in each of 3 cats, and averaged over the 6 cases. This permitted comparison of effects ipsilateral and contralateral to stimulation in well matched conditions.

At the conclusion of the experiments, the position of the electrode tips was confirmed. All lesions were found to be centred in the lateral hypothalamus in the A10 or A13 planes, and were about 2—3 mm broad and 3 mm deep.

The various values of relative squared gain obtained were then fitted to the predicted functions using a Burroughs 6700 computer. Fitting was performed using least absolute deviations rather than least squares, since robustness was felt preferable to efficiency in these circumstances. Conventional estimates of ‘goodness of fit’ were not made, as they have little meaning in the present conditions.

**Results**

Fig. 2 shows a representative power spectrum from one cat, quietly at rest in the pre-lesion recording period. The dominance of power in the theta band is seen here. This was generally the case, although power peaks at other frequencies intruded to a variable extent in individual cases. The pre-stimulation spectra obtained in the stimulation studies revealed rather more delta power, suggesting that the animals had been at a lower state of arousal overall, in those circumstances. This apparent paradox may be attributed to the experimenters’ excessive zeal in awaiting a completely resting state before each stimulus was delivered.

Fig. 3 shows the fit of the predicted function to \( G^2(\omega) \) for the lesion data, and

![Fig. 2. The distribution of power at 0.5 Hz intervals in the EEG power spectrum of a cat quietly at rest, prior to unilateral lesion (original recording from the right posterior suprasylvian area).](image)
Fig. 3. Relative gain in the EEG power spectrum following unilateral lesion of the lateral hypothalamus. Experimental data fitted by the theoretical model, with parameters: N (natural frequency) = 5.16 Hz; control hemisphere damping = 28.76 cycles²/sec; experimental hemisphere damping = 16.20 cycles²/sec; K (amplification factor) = 0.62.

Fig. 4 shows the corresponding fits to

\[ V_{Y,S}^2 / V_{Y,R}^2(\omega) \text{ and } V_{X,S}^2 / V_{X,R}^2(\omega). \]

It will be seen that the experimental data are approximately fitted in each case, with values of N about the theta range. The values of the damping coefficients are of the expected relative magnitudes, and those of the stimulation data, compared to the lesion data, are consistent with relatively lower levels of arousal pertaining in the stimulation experiments. The value of K, the scaling parameter is of opposite sense in the two conditions, and the overall effect of stimulation is seen to be greater on the side of stimulation. The statistical significance of this lateralising effect has been demonstrated previously (Wright and Craggs 1979).

Conclusions

The present findings lend some support to the notion that the lateral hypothalamus exerts a filtering action upon electrocortical activity, and they suggest that this action might be provisionally analysed by linear techniques. Attempts to produce a more general model must address the following issues: (a) provision must be made for oscillatory activity at natural frequencies apart from the dominant oscillator, (b) the nature of the interlinkages between oscillators must be found, and integral to this question, the effects upon phase must be specified, (c) consideration must be given to conditions other than a steady state of quiet rest, (d) the spatial effects upon electrocortical activity must be considered — this amounts to analysing the selective action of the filtering process in different cortical regions, and its effects upon propagating and non-propagating electrocortical activity. While rather formidable tasks, these problems can all be overcome by standard techniques, as discussed below.

The intrusion of effects upon the relative squared gain related to oscillators at other
Fig. 4. The change in gain of the EEG power spectrum in each hemisphere, induced by unilateral stimulation of lateral hypothalamic ICSS sites. Data fitted by the theoretical model with parameters: Upper figure—side of stimulation: N (natural frequency) = 7.21 Hz; damping before stimulation = 9.27 cycles/sec; damping after stimulation = 23.89 cycles/sec; K (amplification factor) = 1.73. Lower figure—side opposite stimulation; N (natural frequency) = 7.84 Hz; damping before stimulation = 9.77 cycles/sec; damping after stimulation = 21.32 cycles/sec; K (amplification factor) = 1.56.
natural frequencies may be suspected in the present data. Fig. 3 shows the possible presence of spikes in the delta, alpha, beta and gamma bands. The suppression of activity in the delta band may be seen in both graphs in Fig. 4, as well as the intrusion of high frequency activity on the side of stimulation. Effects of a similar nature are seen in data published previously (Wright and Craggs 1979; Wright 1981). A general model would require allowance for those oscillators by specifying a family of models covering all possible parallel, series and feedback relations between oscillators, and testing these by optimisation methods applied to cross-spectral data obtained by similar methods to the above experiments. This would enable full utilisation of phase information – a task avoided in the present experiments, since the subdominant oscillators may be expected to exert considerable effects upon phase, so this problem is best deferred until a more general model is attempted.

The departure from steady-state experimental conditions would imply a significant test for the model, as it is then that the actual filtering influence of the lateral hypothalamus might be revealed as non-linear, thus greatly increasing modelling difficulties. However, a most fundamental attribute of linear analysis techniques is that they yield transfer functions having roots to their characteristic equations which are stable in varying conditions (Kuo 1962). The corresponding roots of the present model, and its proposed multiple oscillator general form, are the values of N, the natural frequencies, and D, the damping coefficients. The frequencies of the major cerebral rhythms are well known to be generally stable in individual cases. This suggests that the introduction of time-variant coefficients in D would suffice to describe time-variant conditions, and incidentally, to provide a means for describing how the filtering action could be regulated by other influences in the central nervous system.

The spatial effects of the proposed filtering process would be subject to analysis by utilisation of a rectangular grid of cortical recording electrodes in experiments similar to the present ones. Spatial Fourier transformation of the pattern of activity on the lesion and control hemispheres would then be possible.

The above considerations support the view advanced in the introduction that this sort of approach may prove heuristically useful in defining the actions of control mechanisms of electrocortical activity, and perhaps thus prove helpful in the understanding of the mechanisms of information processing.

Summary

The assumption that transhypothalamic neurones exert damping effects on electrocortical oscillatory activity enables a preliminary model to be advanced, in which the action of the lateral hypothalamus upon the EEG is that of a linear filter. Predictions of the model were tested on data obtained from lateral hypothalamic lesion and stimulation studies.

Résumé

Modèle mathématique préliminaire de la régulation de l'activité électrocorticale par l'hypothalamus latéral

L'affirmation suivant laquelle les neurones transhypothalamiques exercent des effets d'amortissement sur l'activité corticale oscillatoire permet d'avancer un modèle préliminaire, dans lequel l'action de l'hypothalamus latéral sur l'EEG est celle d'un filtre linéaire. Les prédictions du modèle sont testées sur les données obtenues à partir de lésions de l'hypothalamus latéral et d'études de stimulation.

This work was supported by the Medical Research Council of New Zealand. The contributions of M.D.
REGULATION OF ELECTROCORTICAL ACTIVITY

Craggs and Ah Chung Tsoi are gratefully acknowledged.

References

A Linear Theory for Global Electro cortical Activity and Its Control by the Lateral Hypothalamus

J. J. Wright and R. R. Kydd
Department of Psychiatry, University of Auckland School of Medicine, Auckland, New Zealand

Abstract. A linear model for electro cortical waves and their control by the lateral hypothalamus is proposed. It is argued that such a linear model is not in contradiction to non-linearity of neural elements on the microscopic scale. Telencephalic structures are treated as a mass of linked oscillators generating activity with a number of resonant modes. The lateral hypothalamus is regarded as controlling damping of activity in the telencephalic mass, and therefore exerting a specific parametric control over all signal processing in the cortical networks. An initial test is proposed to assess the constancy of telencephalic natural frequencies, with variation in lateral hypothalamic damping.

1 Introduction

An important aspect of the EEG is its close correlation with cognition and behaviour (e.g. Walter et al., 1967; Berkhout et al., 1969; Gevins et al., 1979). MacKay (1960) drew attention to the significance of electro cortical waves in the context of network theory. If cortical cells are viewed as information processing networks, then electro cortical waves might be viewed as statistical shifts in average local depolarisation, influencing the transition probability of local network elements, thus biasing the logical properties of the network. Determination of the laws of motion of these waves and their internal mechanisms of control is therefore a required step toward an overall understanding of brain information processing.

A major difficulty in determining these laws is seen to lie in the complexity and extreme non-linearity implicit in neural interactions (MacKay, 1960; MacKay and McCulloch, 1952, 1953). Successive neural models relating formal neurone properties to the origin of the EEG rhythms have been advanced (Pringle, 1951; Beurle, 1956; Wilson and Cowan, 1973; Karwahara, 1980; Lopes da Silva et al., 1974; Nakagawa and Ohashi, 1980; Nunez, 1981). These models have in common the attribution of the EEG waves to resonance among elements, but differ greatly in the kinds of couplings envisaged. This seems inevitable in the face of continuing physiological uncertainty about detailed cell-to-cell interactions. While the biophysial origin of the EEG signal from cortical dendritic trees appears established (Elul, 1972) the relative contribution of axosynaptic, dendrodendritic and other mechanisms remains uncertain (Noda and Aday, 1970; Schmitt et al., 1976; Aday, 1979, 1981).

In the face of this uncertainty, the theoretical models seem to provide only partial accounts of mechanism, and are therefore very difficult to put to experimental test.

We wish to propose that the properties of the gross electro cortical waves be clearly distinguished from the microscopic and non-linear interactions which underlie them. We will show that the gross waves may be represented as linear, and subject to a specific form of control by certain sub-cortical pathways. By linear waves we mean all of the following: that the waves are subject to the superposition law; that with given boundary conditions the gross electro cortical activity will exhibit resonant modes of fixed natural frequency, and that at each point in the system there is a specific dispersion relation, giving phase velocity for each wavelength. These are interdependent properties for time-invariant linear systems (Kuo, 1982; Feynman et al., 1963). We will show that the control of wave motion we ascribe to pathways passing through the lateral hypothalamus does not effect the linearity of the system, but instead introduces time-variation in the damping parameters of the linear differential equations describing the system.

To develop critical tests for the validity of the model, we propose indirect tests involving repeated measures comparisons of electro cortical activity in
each hemisphere before and after unilateral hypothalamic lesion. These indirect means are necessary because:

(a) Optimization in the time-domain of a noise-driven linear model, of uncertain order, with time-varying parameters, offers so many degrees of freedom that any signal with a degree of stationarity could be accounted for. Indeed it is already known that EEG waves can be fitted to sets of linear functions such as damped sinusoids, or other orthogonal functions (Nunez, 1981).

(b) Direct matching of input and output relations is not practicable, since input signals will undergo marked non-linear transformations from the activity of the microscopic elements, regardless of whether or not the gross waves exhibit linear properties.

(c) The recorded electrocortical signal bears a relationship to the surface signal which is not yet totally defined.

Instead, we aim to optimise an approximate version of our model in the frequency domain, and to test whether the resulting parameters yield results in further calculations which are internally consistent with those expected in a linear wave system.

Our motivation in formulating the theory in its present form came from earlier studies of an empirical nature, in which the effects of the lateral hypothalamus upon the EEG and behaviour were jointly studied (Wright and Craggs, 1976, 1979; Wright, 1981) and from a preliminary test of a very simple form of the model to be described (Wright and Ihaka, 1981).

2 The Role of the Lateral Hypothalamus

The lateral hypothalamus is chosen as the focal point in these studies because:

2.1

It is one of a number of fibre pathways well placed to exert controlling influences in the brain, being made up of numerous pathways ascending and descending from telencephalon to brainstem (Olds, 1956a; Nauta and Karten, 1970).

2.2

Lesion of this pathway leads to the sensorimotor neglect syndrome – an inattention syndrome not accounted for by sensory loss, or motor paralysis (Marshall et al., 1971, 1974; Turner, 1973; Wright et al., 1979).

2.3

Stimulation of the lateral hypothalamus exerts strong effects on motivation and attention, as manifested by the intracranial self-stimulation phenomena (Olds, 1956b, 1961; Olds and Percetz, 1960; Valenstein et al., 1970; Rolls, 1971, 1974; Wauquier and Rolls, 1976).

2.4

The electrocortical power spectra obtained in sensorimotor neglect and intracranial self-stimulation appear to show reciprocal changes in power at each frequency, when each is compared to a resting control state. (Wright and Craggs, 1979). The converse nature of these effects reduces the chance that either represents immediate, or later developing, artefactual events, and suggests that opposite shifts along a continuum are induced by these manipulations.

2.5

Sensorimotor neglect and intracranial self-stimulation appear to depend upon lesion, or stimulation respectively, of all, or sub-groups of cells within the mesotelencephalic dopaminergic group (Ljungberg and Ungerstedt, 1976; Wise, 1978; Wise and Bozarth, 1981).

The collective effect of Sects. 2.2–2.5 above is to indicate that the lateral hypothalamus exerts important control functions upon telencephalic activity, consistent with the structural considerations given in Sect. 2.1.

3 Description of the Model

3.1 Transformation of the Surface Electrotonic Activity in the Electrocorticogram

Essential features are diagrammed in Fig. 1. The output variable of our model is the recorded electrocorticogram (EEG). We treat this as a transform of electrotonic activity in cortical dendrites, following the findings of Elul (1972). He shows that the gross recording appears to arise from summation of "neuronal waves" detectable in the cortical dendritic trees. The neuronal waves in small areas of cortex are largely asynchronous and summation of the small degree of synchronous activity predominates in the gross recording. Elul treats the summation process simply. We note that at least the following factors will influence the sample obtained and its relation to the electrocortical spatial average: electrode position, coherence of the synchronous activity, attenuations and conductivity in the physical medium, and the spatial organisation and movement of the synchronous activities (Nunez, 1981). For the meantime, we treat the complicated spatial and temporal summations involved as an invariant linear transform from the local spatial average for each electrode site. Since this transform is undefined, we will
seek its elimination, along with other complicating factors, in later calculations, and eventually attempt its estimation. For the meantime we give the power spectrum of the recorded EEG, $V^2(\omega)$ as

$$V^2(\omega) = |A(\omega)|^2 \cdot |\Sigma x(\omega)|^2,$$

where $A(\omega)$ is the unknown transform and $|\Sigma x(\omega)|$ is the sum of the various electrocortical components of temporal-frequency $\omega$.

### 3.2 Assumptions and Initial Development

Consider the cerebral cortex as a mesh of highly interconnected voltage (or current) sources, equivalent to segments of dendritic trees. Note that while the interactions between elements are highly non-linear, they take place by processes involving fairly constant conduction times for each connection. Neuronal elements are ordered into many closed loops (within cortical columns, by cortico-cortical fibres, by cortico-subcortical pathways, etc.). These closed loops are also coupled by numerous and disparate interconnections, exhibiting both anatomical orderliness and non-linearity (e.g., Mountcastle, 1978). Thus the telencephalon constitutes a comparatively closed, highly interlinked system. The entire system may be regarded as driven both by input sensory signals, and active neuronal firing. Diffuse inhibitory pathways arising from the brain-stem exert wide-spread hyperpolarizing effects upon certain telencephalic neurones.

Let the potential of a segment of dendritic tree, $x_i(t)$, be represented by

$$\dot{x}_i + D_i(t)x_i + N_i^2(t)x_i = 0$$

noting both that the equation is inhomogeneous and can therefore be fitted to any oscillatory source when $D_i(t)$ and $N_i(t)$ are free parameters, and that these parameters are analogous to a damping coefficient and a natural frequency.

A mass of unit sources coupled to each other may be similarly represented by

$$\dot{x}_1 + D_1(t)x_1 + N_1^2(t)x_1 = K_1^2(t)x_2 + K_3^2(t)x_3 + \ldots + K_n^2(t)x_n$$

$$\dot{x}_2 + D_2(t)x_2 + N_2^2(t)x_2 = K_1^2(t)x_1 + K_3^2(t)x_3 + \ldots + K_n^2(t)x_n$$

to

$$\dot{x}_n + D_n(t)x_n + N_n^2(t)x_n = K_1^2(t)x_1 + K_2^2(t)x_2 + \ldots + K_{n-1}^2(t)x_{n-1},$$

where $D_i(t), N_i(t), K_i^2(t)$ are again free parameters. The parameters of this generalized description may be given physiological meaning under the following assumptions.

#### 3.2.1 Assumptions

3.2.1.1 All $N_i(t)$ have a finite variance $\sigma_{N_i(0)}$ about a mean $\overline{N_i}$ representing perturbation about a dominant cycle time, created by closed cycles of activity, along pathways of fixed conduction velocity, but subject to outside interference and non-linearities.

3.2.1.2 All $D_i(t)$ have a mean $\overline{D_i}$ and variance $\sigma_{D_i(0)}$ both partially determined by the level of inhibitory influences exerted from the brain-stem, either upon the cortical signal source; or upon participant components in the closed loops.

3.2.1.3 All $K_i^2(t)$ have a mean $\overline{K_i^2}$ and variance $\sigma_{K_i^2(t)}$ representing the strength of coupling between sources over numerous, disparate and interacting non-linear pathways, which link together self-exciting systems. (No particular type of distribution for $N_i(t), D_i(t)$ or $K_i^2(t)$ is assumed).

3.2.1.4 All $D_i(t), N_i(t), K_i^2(t)$ are stochastically independent, as each represents processes being perturbed
by very complicated non-linearities in the interactions of the linked oscillatory sources, with diverse input signals.

3.2.2 Development of State Transition Matrix

Let

\[ z_1 = x_1 \]
\[ z_2 = \dot{x}_1 = \dot{z}_1 \]
\[ z_3 = x_2 \]
\[ z_4 = \dot{x}_2 = \dot{z}_3 \]

so that

\[ z_{m-1} = x_n \]
\[ z_m = x_n = \dot{z}_{m-1} \quad m = 2n \]

So

\[ \dot{z}_2 = -D_2 z_2 - N_2^2 z_4 + K_2^1 z_3 + K_2^2 \dot{z}_3 + \ldots + K_2^m \dot{z}_{m-1} \]
\[ \dot{z}_4 = -D_2 z_4 - N_2^2 z_2 + K_2^1 z_1 + K_2^3 \dot{z}_3 + \ldots + K_2^m \dot{z}_{m-1} \]

or in matrix representation

\[
\begin{bmatrix}
  \dot{z}_1 \\
  \dot{z}_2 \\
  \vdots \\
  \dot{z}_m \\
\end{bmatrix} =
\begin{bmatrix}
  0 & 1 & 0 & 0 \\
  -N_2^2 & -D_2 & K_2^1 & 0 \\
  0 & 0 & 0 & 1 \\
  K_2^1 & 0 & -N_2^2 & -D_2 & K_2^3 & 0 \\
  0 & 0 & 0 & 0 & 0 & 1 & \ldots & 0 & 0 \\
  K_2^1 & 0 & K_2^3 & 0 & -N_2^3 & -D_3 & \ldots & K_2^n & 0 \\
  \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
  0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 1 \\
  K_1^n & 0 & K_2^n & 0 & & & & & & \ldots & -N_n^2 & -D_n \\
\end{bmatrix}
\]

\[ \frac{d}{dt} \mathbf{z} = \mathbf{A} \mathbf{z} \]

whence the determinant is formed from the identity matrix, \( I \), with eigenvalues \( \lambda \)

\[
| \lambda I - \mathbf{A} | =
\begin{bmatrix}
  \lambda - 1 & 0 & 0 & \ldots & 0 & 0 \\
  N_2^2 & \lambda + D_1 & -K_2^1 & 0 & -K_2^3 & -K_2^n & 0 \\
  0 & 0 & \lambda - 1 & 0 & 0 & 0 & 0 \\
  -K_2^1 & 0 & N_2^2 & \lambda + D_2 & -K_2^3 & 0 & \lambda \\
  0 & 0 & \lambda + D_3 & \lambda - 1 & \lambda & 0 & 0 \\
  -K_1^n & 0 & -K_2^n & 0 & -K_3^3 & \lambda + D_n & \lambda \\
\end{bmatrix}
= 0.
\]

3.2.3 Properties of the Characteristic Equation

Expansion of this determinant for the \( m \)th order case yields the characteristic equation

\[ a_m \lambda^m + a_{m-1} \lambda^{m-1} + \ldots + a_0 = 0. \]

Since terms in the expansion are products in which one and only one element comes from any row and one and only one element from any column, the following may be observed.

3.2.3.1 The highest term in \( \lambda \) which can be formed is \( \lambda^m \) obtained from the expansion product

\[ \lambda^m (\lambda + D_1) (\lambda + D_2) \ldots (\lambda + D_n) \]

therefore \( a_m = 1 \).

3.2.3.2 In the determinant \( | \lambda I - \mathbf{A} | \) all \( D_i(t) \) appear in elements \( (\lambda + D_i) \). Any \( (\lambda + D_i) \) is in a column with a unity element as the only other non-zero element, and the same unity element is also in a row with \( \lambda \) as the only other non-zero element. Therefore non-zero products in the determinant expansion containing any \( D_i \) contain it in a factor \( (\lambda^2 + \lambda D_i) \) so no terms in \( D_i \) can appear in \( a_0 \).

3.2.3.3 Other coefficients \( a_i \) are formed from certain sums and products of \( D_i(t) \), \( N_i^2(t) \), \( K_j(t) \).

When \( n \) is a large number, these are terms in many parameters. Likewise \( a_0 \) is formed from numerous sums and products in \( N_i^2 \), \( K_j \).
3.3 Consequences of the Characteristic Equation Properties

3.3.1 System Linearity in Time-Invariant Conditions

For coefficients $a_i$ (excluding $a_0$), and under the above listed assumptions regarding the parameters, the Central Limit Theorem of Cramer applies to the value of the coefficients as $n$ tends to a very large number, All $N_i(t_i), K_i(t_i), D_i(t_i)$ may therefore be replaced by $\bar{N}_i, \bar{K}_i, \bar{D}_i$, i.e. the model system is equivalent to a linear, time-invariant system while these mean values remain unchanged.

3.3.2 Constancy of Mode Natural Frequencies with Time-Variant Damping

Since the characteristic equation is of order $m=2n$, it has quadratic factors,

$$(\lambda^2 + \beta_1 \lambda + \gamma_1) (\lambda^2 + \beta_2 \lambda + \gamma_2) \cdots (\lambda^2 + \beta_n \lambda + \gamma_n) = 0. \tag{8}$$

Taking each factor equated to zero yield roots

$$\lambda_i = (-\beta_i \pm \sqrt{\beta_i^2 - 4\gamma_i})/2$$

and when $\beta_i^2 < 4\gamma_i$ the roots are complex conjugate pairs giving the natural frequency of a mode,

$$\bar{\omega_i} = \sqrt{\gamma_i}$$

and damping coefficient of the mode

$$\bar{D}_i = \beta_i$$

where $\bar{\omega}_i > 4\gamma_i$ the roots are real, with a product $\gamma_i$ (Kuo, 1982).

Since $\gamma_1 > \gamma_2 > \cdots \gamma_n = a_0$ which has no parameter in $D_i$ the product of the mode natural frequencies and the real roots is not influenced by time-variation in $D_i$.

This will hold in the most general case where $n$ tends to infinite and $\bar{K}_i, \bar{N}_i, \bar{D}_i$ also become the set of all possible $K_i, N_i, D_i$. Equating to zero the product of any subset of quadratic factors from the characteristic equation of this most general case now also gives the characteristic equation of some sub-system of the most general case, when the sub-system is isolated from the near infinite system. For each sub-system

$$\gamma_1 \gamma_2 \cdots \gamma_k = \bar{a}_0, \tag{9}$$

where $\bar{a}_0$ is likewise a product containing no terms in $D_i(t_i)$. Therefore all $\gamma_i = M_i^2$ must be invariant with changing $D_i(t_i)$ as all possible products of sub-sets of $\gamma_i$ are invariant. A fortiori, this holds for all lower order systems of the same form, i.e. the model system has resonant modes with constant natural frequencies, $M_i$ and time-variant damping $\bar{D}_i$ with time-variation in $D_i(t_i)$.

3.3.3 Clustering of Mode Natural Frequencies About Certain Central Values

If all $\gamma_i$ are significantly clustered about $\bar{\gamma}$ central values, then factors in the coefficients $a_i$ formed from the products of $\gamma_i$ will cluster. Since factors in the coefficients containing $D_i$ do not influence $\gamma_i$ this implies clustering of various products of $N_i$ and $K_i$ in factors of $a_i$ if $\gamma_i$ are clustered.

In the central nervous system, anatomical orderliness and repetition of similar neural circuits leads us to expect that the set of all $K_i(t_i)$ and $N_i(t_i)$ will tend to cluster about a number of central values, thus their products will cluster, i.e. the model system resonant modes will have natural frequencies which are clustered about central values.

3.3.4 Changes in Input Signals Other than Effects on Damping

From Sect. 3.3.1 it will be seen that changes in input signals which do not influence the mean values of $K_i, D_i, N_i$ may greatly perturb their instantaneous values and variance without influence on system linearity. The induction of changes in $\bar{\sigma}_{K_i}, \bar{\sigma}_{D_i}, \bar{\sigma}_{N_i}$ will, on the other hand, appear as a change in the strength of driving signals within the system, i.e. active and non-linear interactions among neuronal components, and input signals contributing to these interactions, can be treated as noise-like driving signals, while the gross waves appear as a passive linear system.

3.4 Summary of System Properties

Assumptions in Sect. 3.2.1 make the system described in (3) equivalent to a system of linear oscillators with simple additive linkage, yielding linear modes, with mode natural frequencies clustered about center frequencies. This lends the system to approximation by single modes at the center frequencies of the dominant rhythms. Assumption Sect. 3.2.1.2 equates the lateral hypothalamus with regulation of the damping coefficients of the modes. Assumption of comparatively constant mean linkage strengths and cycle times in the telencephalon allows that the lateral hypothalamus may also influence the power of noise-like driving signals arising within and imposed upon the system.

For present purposes fibre systems ascending through the lateral hypothalamus are treated as though they acted en mass in the above regards. At best, this minimises the selective roles that might be exhibited by different components. In the experiments to test the theory, which follow later, electrolytic lesions of the lateral hypothalamus are used. If the model is valid for such total interference with the pathway, it must describe the role of either a sub-set of
components, or all in concert. We have further considered the role of the transmitter-specific dopaminergic pathways (see Appendix).

4 Means of Initial Testing of the Model

We wish to test the proposition that the telencephalon is a system with modes of constant natural frequency, damped by input from the lateral hypothalamus and driven by "noise". This can be achieved by unilateral manipulation of the lateral hypothalamus in a between hemispheres, repeated measures design. The undefined transfer function \( A(\omega) \) is eliminated in this treatment.

4.1 Application to Changes in the Power Spectrum

Differential equations for the system are now

\[
\begin{align*}
\ddot{x}_1 + \mathcal{D}_1 \dot{x}_1 + M_1^2 x_1 &= x_0 \quad \text{to} \\
\ddot{x}_n + \mathcal{D}_n \dot{x}_n + M_n^2 x_n &= x_0,
\end{align*}
\]

(10)

where \( x_0(t) \) now represents all processes driving the surface activity, and the surface signal is again \( \sum x_i(t) \).

\[
G^2(\omega) = K \left[ \frac{\sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + D_{IL,A}^2 \omega^2}}{\sum_{i=1}^{n} \frac{-\mathcal{D}_i \omega^2}{(M_i^2 - \omega^2)^2 + D_{IL,A}^2 \omega^2}} \right]^2 + \left[ \frac{\sum_{i=1}^{n} \frac{-\mathcal{D}_i \omega^2}{(M_i^2 - \omega^2)^2 + D_{IL,A}^2 \omega^2}}{\sum_{i=1}^{n} \frac{-\mathcal{D}_i \omega^2}{(M_i^2 - \omega^2)^2 + D_{IL,A}^2 \omega^2}} \right]^2.
\]

Laplace transformation of (10) and addition gives

\[
\sum x_i(s) = x_0(s) \cdot \left\{ \frac{1}{s^2 + \mathcal{D}_s + M_1^2} \right\}.
\]

Multiplying both sides by \( A(s) \) (the surface recording transform in complex variable form), writing \( s = j\omega \), \( j^2 = -1 \), summing real and imaginary parts and obtaining the modulus, then substituting in (1) gives

\[
V^2(\omega) = |x_0|^2(\omega) \cdot |A(\omega)|^2 \cdot \left\{ \left[ \frac{\sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + D_{IL,A}^2 \omega^2}}{\sum_{i=1}^{n} \frac{-\mathcal{D}_i \omega^2}{(M_i^2 - \omega^2)^2 + D_{IL,A}^2 \omega^2}} \right]^2 + \left[ \frac{\sum_{i=1}^{n} \frac{-\mathcal{D}_i \omega^2}{(M_i^2 - \omega^2)^2 + D_{IL,A}^2 \omega^2}}{\sum_{i=1}^{n} \frac{-\mathcal{D}_i \omega^2}{(M_i^2 - \omega^2)^2 + D_{IL,A}^2 \omega^2}} \right]^2 \right\}
\]

(12)

which is the theoretical expression for the EEG power spectrum.

The left and right power spectral densities from symmetrical sites in the same individual are broadly similar and covariant over time although commonly differences in total power left and right are manifest, reflecting in part, different electrode impedances (e.g. Walter et al., 1967; Wright and Craggs, 1976). If (12) is a true description, this implies at least a constant ratio of the power at each frequency in the driving signals left and right, (without implying phase relationship and also parity of \( |A(\omega)| \), \( M_i \), and \( \mathcal{D}_i \) for each side. Following unilateral hypothalamic lesion, this model requires that \( \mathcal{D}_b \) left and right become asymmetrical, the \( M_i \) do not change, and \( |x_0(\omega)| \), the power of the driving signal, may also change relative to the intact side. \( A(\omega) \) are assumed to remain constant at each site. Phase relations of the left and right EEG signals are not considered since the driving signals need not show coherence.

Consider, then, left and right average power spectra obtained before and after unilateral hypothalamic lesion in as close as possible to steady-state conditions. Let the subscripts \( LA, LB, CA, CB \) indicate the lesion and control sides, after and before lesion. Then \( G^2(\omega) \) the relative squared gain attributable to lesion is obtained experimentally as

\[
G^2(\omega) = \left| \frac{V^2_{LA}}{V^2_{LB}} \right|^2 \frac{V^2_{CA}}{V^2_{CB}} (\omega)
\]

(13)

and the theoretical equivalent is obtained by appropriate substitution of (12) in (13) giving

\[
G^2(\omega) = \left| \frac{V^2_{LA}}{V^2_{LB}} \right|^2 \frac{V^2_{CA}}{V^2_{CB}} (\omega)
\]

(14)

Symmetrical factors have cancelled, leaving a function with \( M_i, \mathcal{D}_{IL,A}, \mathcal{D}_{IC,A}, \) and \( K \) as parameters. \( K \) is the relative change in ratio power of the driving noise on each side, consequent to lesion.

The first testable proposition arising from the model is that (14) can be fitted to a variety of experimental data obtained in accord with (13) in such a way that when all parameters are free in the curve fitting, a set of comparable \( M_i \) are obtained from each of a series of similar individuals. This is a direct consequence of the hypothesis that the \( M_i \) of the telencephalic modes of resonance are constant. The \( \mathcal{D}_i \) (\( LA \) and \( CA \)) will of course be expected to be unequal.

An important reservation regarding direct test of this proposition is that \( n \), the number of modes is not known, and cannot be known at this stage without making further assumptions concerning the EEG. This is considered in the following paper describing an experimental test for constancy of the mode natural frequencies.

4.2 Further Test for the Internal Validity of the Model

The above test is a necessary, but not a sufficient test of the model, i.e. while linear systems must exhibit constant natural frequencies this may also be the case
for certain non-linearities. If it can be shown that both the natural frequencies and damping coefficients obtained from curve-fitting the relative squared gain are also consistent with an invariant dispersion relation for electrocortical activity at the site of recording, then a non-linear interpretation will be much less probable. This forms the subject of the second following paper.

5 Discussion

The outlined model rests on drastic simplifications. It bypasses almost all issues of cell-to-cell coupling, details of anatomy etc. Indeed, it involves little more than the treatment of electrocortical activity as arising from a mass of linked oscillators, driven by noise and treated in the second order approximation. No effort is made to formulate the partial differential equations which would be needed to describe the waves in space and time, and to derive a wave equation. This treatment does not, however, reduce to either an empirical mathematical description of the EEG, nor a tautology, for the following reasons:

It gives a mathematically defined role to a specific pathway known to exert influences on the EEG and behaviour, by as of yet undefined anatomical connections to the telencephalon.

It describes, as a linear approximation, regulation of the spatial averaged dendritic signal in the region of the recording electrodes. A signal, input to a local region of the cortical dendritic field, which is represented in generalised Fourier form as

\[ x_g(t) = \int_0^\infty g(\omega) \cos(\omega t) d\omega \]  

becomes converted as a consequence of the effect of each group resonant mode, into

\[ y_g(t) = \int_0^\infty G(\omega) g(\omega) \cos(\omega t + \Phi(\omega)) d\omega , \]  

where

\[ G(\omega) = \left[ (M^2 - \omega^2)^2 + \Omega^2 \omega^2 \right]^{-1/2} \]  

and

\[ \Phi(\omega) = \tan^{-1}(-\Omega \omega/M^2 - \omega^2) . \]

This parallels MacKay's (1960) conception of the regulation of transition probability for neurons in the local network, with \( \Omega \), the damping coefficients, acting as parametric controls i.e. logical properties of the microscopic network are being biased by this central system in a particular way. When all separate modes are considered, the control is multilinear.

The model carries the implicit notion of control of both the spatial and temporal character of electrocortical activity. A linear system with fixed natural frequencies and a specifiable dispersion relation implies fixed boundary conditions, and the existence of spatially determined modes of resonance (Feynman, 1963; Nunez, 1981). A specifiable global, as well as a local, biasing of non-linear elements in the cortical networks could thus be fairly simply achieved by diffuse inhibitory ascending fibres from the brain stem.

Appendix

The following papers concerning tests for the above theory, rely on unilateral electrolytic lesions of the lateral hypothalamus. It might be considered that these results reflect only a non-specific effect, characterising generalised damage. This is not the case, as both earlier control experiments (Wright, 1981) and further selective lesion experiments show. We have repeated the experiments described in the following papers, using unilateral lesion of the dopaminergic neurones of the substantia nigra pars compacta and adjacent dopaminergic ventral tegmental area. Selective cell destruction was obtained with 6 hydroxydopamine local injection, with protection of noradrenergic neurones by desipramine cover. Essentially similar results were obtained to those using electrolytic lesion. Estimated natural frequencies for the telencephalic resonances were demonstrated to not significantly differ for the two types of lesion, by Kolmogorov-Smirnov two sample tests. The later estimate of the dispersion relation gave comparable results also. Control experiments involving injection of non-toxic vehicle alone failed to produce significant asymmetry of the left and right coefficients of damping.

We conclude that this cell group is contributory to the parametric control-by-damping described, as would be expected from their anatomical diffusion of inhibitory fibres to caudate and limbic sites of termination, as well as the behavioural correlates mentioned above.

Since these findings require detailed description, and are of greater relevance to other readings, they will be reported fully elsewhere, and are not further considered in the following papers.

Acknowledgements. This work was supported by the New Zealand Medical Research Council, through a grant to J. J. Wright and a research Fellowship to R. R. Kydd. We gratefully acknowledge the technical assistance of J. A. West.

References


Received: November 25, 1983
Prof. J. J. Wright
Department of Psychiatry
School of Medicine
The University of Auckland
Private Bag Auckland
New Zealand
A Test for Constant Natural Frequencies in Electrocortical Activity Under Lateral Hypothalamic Control

J. J. Wright and R. R. Kydd

Department of Psychiatry, University of Auckland School of Medicine, Auckland, New Zealand

Abstract. An initial test for a theory of lateral hypothalamic regulation of electrocortical activity is undertaken. The theory supposes lateral hypothalamic input directly or indirectly damps telencephalic resonances involving linear wave phenomena, enabling this pathway to act as parametric control of information processing in cortical neural networks. Relative changes in left and right electrocortical power spectra are used to test for the presence of resonant modes with constant natural frequencies in conditions of asymmetrical damping, following unilateral lesion of the lateral hypothalamus. Natural frequency values for the modes clustered about center frequencies in the EEG band are obtained. This method has the advantage of minimizing the effects of time-variation and the recorded signal's distortion from the electrocortical local spatial average, but limits consideration to five dominant modes of resonance. The uncertainty of true model order, and errors in curve-fitting impose limitations on the test.

1 Introduction

A preceding theoretical paper (Wright and Kydd, 1983a) proposes that the lateral hypothalamus exerts an indirect multilinear filtering effect upon electrocortical activity, and that this is done by the inhibitory, diffuse pathways ascending in the lateral hypothalamus damping resonant activity in the rostral systems. The model indicates that changes in the ipsilateral power spectrum following unilateral hypothalamic lesion may be compared to those in the control hemisphere spectrum, so that \( G^2(\omega) \), the relative squared gain consequent to lesion is given by

\[
G^2(\omega) = \frac{V^2_{L4}}{V^2_{L3}} \left| \frac{V^2_{L4}}{V^2_{L3}} \right| = K \left[ \sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \xi_{il}^2} \right]^2 + \left[ \sum_{i=1}^{n} \frac{-\xi_{il} \omega}{(M_i^2 - \omega^2)^2 + \xi_{il}^2} \right]^2
\]

\[
+ \left[ \sum_{i=1}^{n} \frac{-\xi_{il} \omega}{(M_i^2 - \omega^2)^2 + \xi_{il}^2} \right]^2
\]

where \( V^2(\omega) \) is the power spectrum. \( LA, LB, CA, CB \) are subscripts for the lesion and control hemispheres, after and before unilateral lesion, \( K \) is a scaling factor related to the strength of driving signals, \( M_i \) are the natural frequencies, \( \xi \), the damping coefficients, and \( n \) the number of modes, where \( i = 1, \ldots, n \).

A first test of the model, which supposes the \( M_i \) to be constant, is to fit (1) to real data, to determine whether a set of constant \( M_i \) are obtained from data of different individual cases.

A complicating factor is that \( n \), the number of modes, is not known a priori. This might be obtained, for instance, by determining the order of linear model needed to account for the EEG in the time domain, using an optimizing criterion such as that of Akaike (1972). Even so, problems would be encountered, since some of the parameters will be time-variant, and the real EEG might involve activity at groups of modes with closely similar \( M_i \). These problems would interact, to leave \( n \) only approximatively estimated.

For immediate purposes, we have preferred a rough, but computationally much simpler method. Scrutiny of electrocortical power spectra of human, cat, rhesus monkey, and rat shows all of these to exhibit a number of peaks (Walter et al., 1967; Wright and Craggs, 1976, 1979; Wright, 1981 and present data). In the terms of the model to be tested, these must correspond to resonance peaks of modes or groups of modes, as reflected in the recorded signal. Setting \( n \) to five allows for activity at five major peaks. It is accepted that five may be a gross underestimate, or sometimes may be an overestimate of the discriminable peaks. However, if \( n \)
is arbitrarily set at 5 in (1), it leads to the revised prediction that the fitting of (1) to real power spectral data from a number of animals will yield $M_t$ clustered about the $M_t$ of several dominant modes, or groups of modes. This revised prediction is the one tested in the following experiment.

2 Methods

Recordings were obtained from experimental animals before and after unilateral lesion of the lateral hypothalamus, and further processed, as follows:

2.1 Animals

Those to be fully described were eight male Wistar rats with weights in the 300–400 g range. These were selected from a total of eleven animals, as those with well placed lesions in the lateral hypothalamus. Other animals, with off-centre lesions showed minimal or no change in behavioural or EEG characteristics, and are not further reported. Comparable control data has been given elsewhere (Wright, 1981).

2.2 Recording Electrode Implantations

Under general anaesthesia, bipolar extra-dural platinum-ball recording electrodes 0.5 mm in diameter were implanted over the posterior cortex at fixed positions in the stereotaxic plane. These were:

Anterior electrode: P 7.5, L 5.1, Posterior electrode: P 9.5, L 5.1, zeroed on Bregma (Atlas of König and Klippel). Skull screws provided ground electrodes, and the cranial assemblies were embedded in dental acrylic, leaving a connector exposed.

Ten to 14 days of post-operative recovery were allowed before experiments began.

2.3 EEG Recordings

These were obtained in conditions of wakeful rest, after the animals had habituated themselves to the experimental conditions. The rats rested gently restrained in a soft cloth. EEG signals were sensed via a plug to the connector containing unity-gain buffering field-effect transistors. After differential amplification at 5 K gain, on Tektronix AM502 amplifiers and analog filtering by Butterworth filters (low pass 48dB/octave), giving a 1–35 Hz bandpass at the 3dB points, signals were fed to analog-to-digital converters, where they were digitised at 8 ms intervals, before being stored on disk in 34 s continuous epochs, via the core of a PDP 8e computer. The converted signals were visually monitored on-line.

Twenty epochs of recording were obtained 24 h before hypothalamic lesion, and 20 more 48 h following lesion. Further editing, to select epochs reflecting the maintainance of a steady state, reduced the number of epochs used to an average of six prelesion and seven postlesion (range three to eleven). This editing used review of the acquired time-series rejecting those epochs in which any sudden transition of the overall EEG occurred (e.g. from low voltage fast activity to large amplitude slow waves). The later calculated power spectrum of each epoch further helped select epochs with similar spectra in the control hemisphere pre- and post-lesion. This editing was performed by eye, without knowledge of results of later spectral calculations, so that selection was purely with regard to obtaining uniform overall conditions in each animals case.

2.4 Lesion Techniques

After preliminary recordings were obtained, single unilateral lesions aimed at the lateral hypothalamus were made stereotactically (under general anaesthesia with Ketamine and pentobarbitone) using randomised choice of side in each animal. Steel stereotactic needles (0.5 mm tip exposure) were introduced with a 50° obliquity from horizontal, aimed at A6L2H-2.7 from stereotactic zero. Lesions were electrolytic, depth electrode anodal, using 1 mA for 15 s.

2.5 Sensorimotor Neglect

This behavioural procedure was performed to assess the degree of functional lateral hypothalamic damage reflected in impaired somatosensory orientation on the side opposite lesion (Marshall et al., 1971).

The rats were tested pre-lesion, then 2 d, and again 6 d after lesion. Any tendency to turn preferentially in one direction in an open field situation was first noted. Firm cotton buds were simultaneously and repeatedly applied to opposite sides of the body to assess the preferred direction of turning to double stimulation. Sensory stimuli were then applied to each side of the body with random alternation of side. Visual orientation was tested by moving a 2 cm × 2 cm square of white cardboard into the animal's visual field, and olfaction by moving an xylene soaked cotton bud close to each nostril. Touch orientation testing was performed using a cotton bud, and a Von Frey hair of 6 g pressure applied to head, shoulder and mid-trunk. The vibrissae were stroked from behind forwards, and a stick was touched to the head just behind the mouth to
assess the biting response. A 23 g needle was used to assess the response to pinprick. The precision with which the animal could localise and respond to the applied stimulus, was used, together with the turning behaviour, to estimate the degree of neglect (see Marshall et al., 1971; Marshal, 1979; and Table 1).

2.6 Histology
At the conclusion of the experiment, 7–10 d after lesion, all animals underwent transcardiac cerebral perfusion under general anaesthesia. After immediate cerebral flushing with normal saline, perfusion continued with a formaldehyde solution. The brain was removed, and selected sections examined from rostral to the extent of the lesion to caudal to the posterior limit of the substantia nigra. This was performed using Nissl staining.

2.7 Further Analysis of EEG Signals
Data from each recording epoch was divided into segments containing 512 data points and each segment separately underwent 10% cosine-bell windowing then Fast Fourier analysis to obtain the power spectrum of each segment, at 128 point resolution from 0.25 Hz to 32 Hz. The segment power spectra were then averaged to provide an average power spectrum for each channel, before and after lesion respectively. Division of these averaged power spectra yielded the relative squared gain attributable to lesion, as required by (1). Each plot of $G^2(\omega)$ was then slightly smoothed, using low-pass Fourier smoothing to 32 harmonic, of the $G^2(\omega)$ plot's own transform. This minimised cumulative error introduced in divisions, thus improving precision of convergence in subsequent curve fitting. Finally, the relative power estimates at the lowest 1.5 and highest 0.5 Hz were eliminated, to further minimize effects of any residual low or high frequency artifact present in the original recordings. The resulting 120 point plots of relative squared gain from each animal were then transferred to an IBM 4341 computer.

2.8 Curve Fitting
Library routines were used to find the minimum least squares fit of (1) to each of the experimental estimates of $G^2(\omega)$. The procedures utilized machine estimation of the vector to best fit, and its first and second partial derivatives, to converge to local minima from given starting values. Local minima were numerous, and searches for a global minimum were undertaken using starting values systematically distributed to yield diverse initial plots of residuals, until an apparent global minimum with best goodness of fit was found. Values of all sixteen parameters found at best fit were then tabulated for each case.

2.9 Estimation of Optimum Natural Frequencies
The 40 estimates of $M$, the natural frequencies, obtained from all eight animals, were analysed with regard to their tendency to cluster about specific frequencies. This was done using the PDP6/e computer, to find those five values, $M_i$ (optimum) such that the sum of the squares of the deviations of each experimentally obtained $M$, from the nearest value of $M_i$ (optimum) was minimized. This yielded a best estimate of the five dominant, or central, natural frequencies, while the mean square deviation of nearby values about these points gave an estimate of the variation for these values.

2.10 Test of Hypothesis
The principle hypothesis to be tested is that all 40 estimates of $M_i$ obtained by curve fitting exhibit a significant clustering about certain values, the $M_i$ (optimum) obtained above. The null hypothesis may be stated as required randomness in the $M_i$ distribution in the 0–30 Hz EEG band. The test of hypothesis is conservative, since there is no reason why the $M_i$ thus obtained should be limited to the bandwidth of analysis, if (1) has no physical validity.

The test used was the non-parametric Kolomogorov-Smirnov one sample test for significance of clustering of the 40 $M_i$ values within the band, independent of where the cluster centers were distributed.

3 Results

3.1 Behavioural Changes
Following lesion, all animals exhibited contralateral sensorimotor neglect on both occasions of testing, as shown in Table 1. Prior to lesion all had shown unimpaired sensorimotor responses.

3.2 Histological Results
The areas lesioned were marked by a zone of architectural disorganisation and glial cell infiltration extending approximately 3 mm in antero-posterior direction and 2–2.5 mm across. The lesions were pear-shaped tapering dorsolaterally towards the site of introduction of the electrode. Needle tract damage was minimal.

All eight animals had lesions in the lateral hypothalamus centred at the level of the dorsomedial
WORK TOWARD A THEORY OF BRAIN FUNCTION

Table 1. Ratings of contralateral sensorimotor neglect 2d and 6d after unilateral lesion of the lateral hypothalamus

<table>
<thead>
<tr>
<th>Rat name</th>
<th>Argos</th>
<th>Minos</th>
<th>Muldoon</th>
<th>Oedipus</th>
<th>Omega</th>
<th>Paris</th>
<th>Theseus</th>
<th>Zeus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensorimotor neglect*</td>
<td>2d Postlesion</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>6d Postlesion</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+++</td>
<td>++</td>
</tr>
</tbody>
</table>

* Neglect was rated as follows: +++ = Full neglect: marked circling to side of lesion, little or no response in any sensory modality contralateral to lesion. ++ = partial neglect: circling to side of lesion, awareness of stimuli to contralateral side evidenced by squeaking or movement but no head orientation. + = mild neglect: preferential turning to side of lesion, contralateral head orientation in direction of applied stimulus but no localisation. 0 = no neglect: equal turning to both sides, precise head orientation to site of stimulus plus sniffing, biting or scratching

Table 2. Parameters and goodness of fit to relative squared gain for all eight animals

<table>
<thead>
<tr>
<th>Animal</th>
<th>Argos</th>
<th>Minos</th>
<th>Muldoon</th>
<th>Oedipus</th>
<th>Omega</th>
<th>Paris</th>
<th>Theseus</th>
<th>Zeus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural Frequencies (Hz)</td>
<td>4.12</td>
<td>3.30</td>
<td>7.03</td>
<td>4.90</td>
<td>1.86</td>
<td>6.74</td>
<td>3.89</td>
<td>3.19</td>
</tr>
<tr>
<td>6.65</td>
<td>4.30</td>
<td>7.04</td>
<td>6.08</td>
<td>2.37</td>
<td>6.82</td>
<td>4.18</td>
<td>7.44</td>
<td></td>
</tr>
<tr>
<td>8.94</td>
<td>5.71</td>
<td>11.43</td>
<td>10.32</td>
<td>9.00</td>
<td>13.10</td>
<td>8.55</td>
<td>7.63</td>
<td></td>
</tr>
<tr>
<td>11.28</td>
<td>10.64</td>
<td>24.57</td>
<td>10.87</td>
<td>10.02</td>
<td>19.59</td>
<td>10.73</td>
<td>13.09</td>
<td></td>
</tr>
<tr>
<td>18.60</td>
<td>24.09</td>
<td>27.39</td>
<td>26.15</td>
<td>17.23</td>
<td>20.87</td>
<td>16.84</td>
<td>18.89</td>
<td></td>
</tr>
<tr>
<td>Damping Coefficients</td>
<td>0.608</td>
<td>0.943</td>
<td>1.225</td>
<td>2.774</td>
<td>0.730</td>
<td>1.185</td>
<td>0.860</td>
<td>1.456</td>
</tr>
<tr>
<td>Lesion: control</td>
<td>55.670</td>
<td>11.298</td>
<td>3.361</td>
<td>5.486</td>
<td>1.948</td>
<td>1.532</td>
<td>14.105</td>
<td>3.094</td>
</tr>
<tr>
<td>3.761</td>
<td>2.667</td>
<td>2.951</td>
<td>2.739</td>
<td>6.886</td>
<td>4.555</td>
<td>0.462</td>
<td>0.577</td>
<td>4.802</td>
</tr>
<tr>
<td>$K$</td>
<td>1.160</td>
<td>1.037</td>
<td>1.414</td>
<td>0.943</td>
<td>0.477</td>
<td>0.571</td>
<td>0.550</td>
<td>1.113</td>
</tr>
<tr>
<td>Goodness of Fit, $r^2$</td>
<td>0.707</td>
<td>0.939</td>
<td>0.956</td>
<td>0.748</td>
<td>0.559</td>
<td>0.507</td>
<td>0.920</td>
<td>0.633</td>
</tr>
</tbody>
</table>

Table 3. Optimum five natural frequencies and associated cluster variation derived from estimates on all animals

<table>
<thead>
<tr>
<th>Optimum Natural frequency (Hz)</th>
<th>Cluster variation*</th>
<th>Number of estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.89</td>
<td>0.95</td>
<td>9</td>
</tr>
<tr>
<td>7.03</td>
<td>0.85</td>
<td>11</td>
</tr>
<tr>
<td>10.87</td>
<td>1.50</td>
<td>10</td>
</tr>
<tr>
<td>18.6</td>
<td>1.87</td>
<td>6</td>
</tr>
<tr>
<td>26.15</td>
<td>2.07</td>
<td>4</td>
</tr>
</tbody>
</table>

* The average of the squared deviations from the optimum natural frequency for that cluster

3.3 Curve Fitted Estimates of Relative Squared Gain
Values of the parameters obtained for best fit in each animal's case, and the goodness of fit, are given in Table 2. A selection of graphs displaying these fits are shown in Fig. 1.

3.4 Assessed Clustering of Values for Natural Frequency
Table 3 gives the values of $M_i$ (optimum) obtained from all estimates of natural frequency, and the

nucleus (approximately A4.3 atlas of Konig and Klippel, 1963). In the largest lesions, gial cell infiltration extended posteriorly to just beyond the anterior aspect of the substantia nigra compacta (approximately A2.5) and anteriorly to the supraoptic nucleus (approximate-
estimated variation of nearby values. Notably, these are values roughly corresponding to center frequencies of the major cerebral rhythms.

3.5 Test of Hypothesis

The Kolmogorov-Smirnov one sample test showed the tendency to cluster about the center frequencies was significantly removed from randomness. The estimated departure from random clustering gave a value of Kolmogorov-Smirnov $D > 0.28$, with $p < 0.01$ for null hypothesis.

4 Discussion

The above results show the predicted clustering of estimates of $M_i$ about five central frequencies, with the expected left/right inequalities of post-lesion damping coefficients. Although the clustering is statistically significant, considerable variation in each center frequency estimate is present. Since fitting an equation in sixteen parameters to averaged spectral data involves inevitable error in the estimate of each parameter, this might be expected. Arbitrary imposition of a five mode model also must add distortion to the results. Although $M_i$ estimates are scattered about the group optimum frequencies, sometimes several similar values of $M_i$ are obtained in the same animal's case. This might result either from activity at separate modes of similar frequency, or from accumulated errors in the calculation of $G^2(\omega)$ by repeated division, producing distortion in the least squares fitting.

The hypothesis is thus supported, but not exhaustively. However, two logical alternative models for lateral hypothalamic effects on electrocortical activity are quite strongly rejected by this data.

Firstly, in a linear model, if the lateral hypothalamus were the source only of noise-like signals input to the telencephalon, with no other factors being affected by lesion, then the total power, but not the spectral density of the EEG would be altered and the experimental estimates of $G^2(\omega)$ would be best fitted by a straight line, which is not the case.

Secondly, were the lateral hypothalamus the source of rhythmic driving signals which were then imposed on the cortical mantle, then destruction of the pathway should eliminate or reduce these oscillations. Instead as the above plots of $G^2(\omega)$ for different animals show, an increase in the amplitude of the cerebral rhythms often occurs.

The results of this initial test of hypothesis also enable a further test for internal consistency of results and theory to be undertaken. This is the subject of the following paper.

Acknowledgements. This work was supported by the New Zealand Medical Research Council, through a grant to J. J. Wright and a research Fellowship to R. R. Kydd. We gratefully acknowledge the technical assistance of J. A. West.

References


Received: November 25, 1983

Prof. J. J. Wright
Department of Psychiatry
School of Medicine
The University of Auckland
Private Bag
Auckland
New Zealand
Inference of a Stable Dispersion Relation for Electrocortical Activity Controlled by the Lateral Hypothalamus

J. J. Wright and R. R. Kydd
Department of Psychiatry, University of Auckland School of Medicine, Auckland, New Zealand

Abstract. A second test is undertaken for a theory of linear wave motion in electrocortical waves, under lateral hypothalamic control via regulation of damping. This test invokes a general property of linear systems, namely that wave motion with characteristic natural frequencies implies fixed phase velocities associated with each wavelength, independent of the changes in hypothalamic input. A means of testing the invariance of this dispersion relation at the point of recording is derived from a simplified biophysical model for waves in a dipole layer. The method avoids some problems implicit in direct spatio-temporal wave analysis. Results confirm that the model under test is internally consistent, and is also consistent with other findings concerning the origin and spatial nature of the EEG.

1 Introduction

We have proposed that ascending fibres in the lateral hypothalamus exert damping effects upon modes of resonance in more rostral telencephalic structures, thus exerting a multilinear parametric control over total cortical signal processing. (Wright and Kydd, 1984a). As an initial test of this theory we determined that lateral hypothalamic lesion produced power spectral changes consistent with changed damping of activity with constant natural frequencies (Wright and Kydd, 1984b). Using a limited five mode version of the theory, estimates of the natural frequencies were significantly clustered about five central frequencies, corresponding to the center frequencies of the conventional alpha, beta, gamma, delta and theta bands. However, experimental limitations left the degree of constancy of these natural frequencies in some doubt, and a non-linear interpretation was not excluded.

This paper concerns a second and complimentary test of the theory, using a further analysis of the same data.

The theory under consideration depends on the assumption that the rostral telencephalic systems, no matter how non-linear the elements composing them, act as a linear system regarding resonances generated in the dendritic fields. This is identical to asserting that electrocortical activity is comprised of waves in a medium with a specific dispersion relation, and fixed boundary conditions (e.g. Feynman et al., 1963). The phase velocity for each wavelength may be a function of position in space within the system, but must be invariant with time, or constant resonant modes cannot be sustained. The second assumption is that certain lateral hypothalamic fibres contribute to the damping of the resonant modes, without altering either the dispersion relation or boundary conditions, thus leaving the natural frequencies of the modes unchanged.

This restatement of the theory now emphasises the notion of dispersion relation. Direct studies relevant to this, conducted on multichannel EEG (Petsche and Rappelsberger, 1970; Lehman, 1971; Childers, 1977; Nunez, 1981) reveal wave patterns of long wavelength compared to brain dimensions. Estimating a dispersion relation for these waves would encounter difficulties partly because of the limitations imposed by time-variation, poor spatial resolution and the complicated shape of the cortical surface, and would be made even more difficult if the dispersion relation varied with position on the cortex. Instead of attempting a difficult direct test on multichannel EEG, we now infer an indirect test for time-stability of the dispersion relation despite changes in hypothalamic input. This test shows that the parameters obtained by fitting the linear model to relative gain changes are internally consistent with expectation regarding the dispersion relation.
2 Electrode Transfer Characteristics and the Linear Wave Model

Our model gave the relative change in the left/right power spectra consequent to unilateral lateral hypothalamic lesion as \( G^2(\omega) \) the relative squared gain, defined by

\[
G^2(\omega) = \frac{V_{LA}^2}{V_{LB}^2} \frac{V_{CA}^2}{V_{CB}^2}(\omega) = K \left\{ \sum_{n} \left( \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \varpi_{i,LA}^2 \omega^2} \right)^2 + \sum_{n} \left( \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \varpi_{i,CA}^2 \omega^2} \right)^2 \right\}^{1/2}
\]

where \( \omega \) is frequency, \( V^2(\omega) \) the average power spectrum, the subscripts \( LA, LB, CA, CB \) indicate side of lesion and control, after and before lesion, \( K \) is a factor relating relative power left and right of noise-like driving signals, \( M_i \) are natural frequencies, \( \varpi_i \) damping coefficients and \( n \) the number of modes of resonance.

The expression for the individual average power spectra obtained from a single channel was

\[
V^2(\omega) = |X_\theta(\omega)|^2 |A(\omega)|^2 \left\{ \sum_{n} \left( \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \varpi_{i,LA}^2 \omega^2} \right)^2 + \sum_{n} \left( \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \varpi_{i,CA}^2 \omega^2} \right)^2 \right\}^{1/2},
\]

where \( X_\theta(t) \) was the noise-like driving signal, and \( A(\omega) \) was a linear transfer function incorporating effects of the surface electrodes, surface geometry and conductivity of the media in transforming the recorded signal from the spatial average. Whence

\[
|A(\omega)| = \left\{ \frac{V^2(\omega)}{|X_\theta(\omega)|^2} \left\{ \sum_{n} \left( \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \varpi_{i,LA}^2 \omega^2} \right)^2 + \sum_{n} \left( \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \varpi_{i,CA}^2 \omega^2} \right)^2 \right\}^{1/2} \right\}^{1/2}.
\]

The curve-fitting of (1) to experimental data gives values for all parameters needed to calculate \( |A(\omega)| \) in (3) for the lesion and control hemispheres respectively, by using also the appropriate post-lesion power spectrum, and normalising to remove the scaling effect of \( |X_\theta(\omega)|^2 \).

For the approximate case where \( n \), the number of modes is limited to five, the best available estimates for \( M_i (i = 1 \ldots 5) \) are the central values of the five clusters of \( M_i \) obtained in curve fitting all animals data to (1) without constraints on any parameter (previously reported in Wright and Kydd, 1984b). Applying these as constants in (1) and repeating the curve-fitting will then give revised estimates of \( \varpi_{i,LA} (i = 1 \ldots 5) \) and \( \varpi_{i,CA} (i = 1 \ldots 5) \) suitable for an estimate of \( |A(\omega)| \) for left and right hemispheres, after unilateral lesion. The assumptions of the linear wave model require that \( A(\omega) \) should be a linear transfer function which is closely comparable left and right for symmetrical electrode positions, and independent of hypothalamic input. Calculating \( |A(\omega)| \) in the above way therefore tests the values of estimated parameters for their internal consistency.

\( |A(\omega)| \) will also have certain predictable characteristics depending on its relation to the electrocortical dispersion relation, as we now show.

3 Bipolar Recording of Linear Waves

Nunez (1977, 1981) gives a simplified physical analogy for the point electrocortical potential, considered for a spatially distributed wave-form, composed of current sources in a large, plane dipole layer of arbitrary conductivity, overlain by a low conductivity medium.

In the one dimensional case he obtains the point potential \( \Phi_j(t) \) due to the \( 1/2^\text{th} \) spatial Fourier component (of wavelength \( \lambda \)) as

\[
\Phi_j(t) = CJ_0(t) \int_{-\infty}^{\infty} \frac{e^{j\lambda x}}{(D + \xi^2)^{3/2}} \cdot d\xi,
\]

where \( j^2 = -1 \), \( C \) and \( D \) are composite constants depending on electrode separation from cortex and medium conductivity. \( J_0(t) \) is the current density of the \( 1/2^\text{th} \) component immediately below the electrode, \( \xi \) is the distance from the electrode across the surface.

More generally for the activity over the closed surface of a cerebral hemisphere we may write

\[
\Phi_j(t) = CJ_0(t) \int_{-\infty}^{\infty} f(D, \lambda, \xi) \cdot dA
\]

where \( \xi \) is now a three dimensional vector and \( C \) and \( D \) are functions of \( \xi \) allowing for inhomogeneity in conductivity and the cortical geometry.

For two points \((+1)\) and \((-1)\) a small distance, \( a \), apart on the surface, little difference in the space variables in (5) will be apparent. Differential voltage between the two points will therefore be given by

\[
\Delta \Phi_j(t) = R(J_{(+)}(t) - J_{(-)}(t)),
\]
where \( R \) is the expression covering the space variables in (5) and \( J_{+,i}(t) \) and \( J_{-,i}(t) \) are the current densities at each point. The vector differences in (6) weight most highly the contribution to differential voltage of current sources in the line of the electrode array, reduce to zero the contribution of points equidistant from each electrode and give intermediate weights to other points. \( \Delta \Phi(t) \) is therefore approxi-
mate to the difference of voltages obtained equival-
ently in the one dimensional case. Hence the underly-
ing waves can be simply considered as having a
dispersion relation (in the vicinity of the recording electrodes) given as

\[
c_i(\omega) = \omega \lambda,
\]

(7)

where \( c_i \) is the phase velocity of the wave components in the line of the electrodes, at wavelength \( \lambda \) with a temporal frequency \( \omega \).

The signal recorded by differential voltage may now be regarded as the superposition of two waves which are each arising from the same underlying space component but have a phase difference \( \pi / \lambda \) induced by the separation, with one point additionally phase reversed by a half cycle, because of the differential recording convention.

The amplitude of this recorded signal at the frequency \( \omega \) is then \( 2RJ_{\text{max}} \cos \left( \frac{1}{2} \left( 1 - \frac{a \omega}{c_i(\omega)} \right) \right) \) while the underlying wave has an amplitude \( RJ_{\text{max}} \), so \( |A'(\omega)| \) the amplifying effect evident in the recorded time series when this is compared with the spatial average should be

\[
|A'(\omega)| = 2 \cos \frac{1}{4} \left( 1 - \frac{2a \omega}{c_i(\omega)} \right).
\]

(8)

This function will have a first maximum at that \( \omega \) (for any monotonic relation of \( c_i \) and \( \omega \)) where \( a \), the electrode separation is equal to \( \lambda / 2 \).

\( |A(\omega)| \) the total electrode transfer characteristic, has \( |A'(\omega)| \) as a factor. Other factors are any attenuations arising at the site of the electrode, (which we here treat as frequency independent) and the band-pass characteristics of the recording system.

4 Predictions

Similarity of \( |A(\omega)| \) calculated for the lesion and control hemispheres [via (1) and (3)] is therefore a test of the stability of the dispersion relation in the face of marked changes in the lateral hypothalamic input. Where all wavelengths are longer than twice the interelectrode separation, then \( |A(\omega)| \) will be a rising monotonic function across \( \omega \), if other factors influencing \( |A(\omega)| \) are corrected for.

Two Further Aspects Require Consideration

Firstly, to be consistent with the model’s origin from Elul’s (1972) treatment of EEG origin from neuronal waves, \( |A(\omega)| \) ought not to vary too greatly with \( \omega \) or the spectral density of neuronal waves would vary markedly from that of the electrocorticogram, which is not the case as reported by Elul (1972).

Secondly, the distribution of errors in the calculation of \( |A(\omega)| \) can be anticipated from the limitations implicit in the model’s idealisations. The experimentally obtained power spectrum to be used in (3) is an averaged power spectrum, in which time-variation of damping parameters, and departures from the idealised one-dimensional case represented in (6) will “smear” the spectrum i.e. reduce frequency resolution. The expression calculated in the denominator of (3) is inevitably clearly resolved, with peaks near the natural frequencies imposed in the curve-fitting of (1). We can anticipate that plots of \( |A(\omega)| \) will contain autocorrelated errors, imposing minima near the center natural frequencies used in (1). Since this technique renders the pattern of errors predictable, we can avoid confusion with true curvilinear trends in \( |A(\omega)| \). The linear trends in \( |A(\omega)| \), control and lesion, will therefore be relied on in comparing these estimates, in the absence of other marked curvilinear trends.

5 Methods

This experiment, and the further processing of data, begin from the results reported in the preceding paper (Wright and Kydd, 1984b) which concluded with estimates for the center frequencies of the five dominant groups of frequencies.

5.1 Curve Fitting to Obtain Model Parameters with Imposed Natural Frequencies

The above obtained five center frequencies (3.89, 7.03, 10.87, 18.6, and 26.15 Hz) were imposed as values of \( M_i(i = 1 \ldots 5) \) in (1). Curve fitting to \( G^2(\omega) \), the relative squared gain for each rat’s spectral data was then repeated with the fixed \( M_i \) to obtain revised estimates of \( \mathcal{R}_{IL,i}(i = 1 \ldots 5) \), \( \mathcal{R}_{II,i}(i = 1 \ldots 5) \), and \( K \), the scaling parameter. These new parameters now approximate \( G^2(\omega) \) as though it arose only from activity with a single natural frequency at each “cluster”.

Optimisation was again by local search for minimised least square errors, using the IBM 4341. Repeated searches for a global minimum in each case were again undertaken using five sets of starting values yielding diverse patterns of initial residuals. That minimum yielding the highest goodness of fit was taken as the global minimum.
Table 1. Damping coefficients, scaling factor and goodness of fit obtained for each animal, when relative squared gain in the EEG power spectrum (attributable to the effect of unilateral lesion) is fitted to (1). The natural frequencies, $M_i (i = 1 \ldots 5)$ were imposed on the basis of earlier estimates, at 3.9, 7.0, 10.9, 18.6, and 26.1 Hz

<table>
<thead>
<tr>
<th>Animal</th>
<th>Argos</th>
<th>Minos</th>
<th>Muldoon</th>
<th>Oedipus</th>
<th>Omega</th>
<th>Paris</th>
<th>Theseus</th>
<th>Zeus</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{1\text{LA}}:D_{1\text{CA}}$</td>
<td>16.73:1000</td>
<td>4.71:12.80</td>
<td>57.72:15.11</td>
<td>26.50:20.07</td>
<td>75.96:1000</td>
<td>0.36:0.31</td>
<td>1.08:2.17</td>
<td>3.11:3.68</td>
</tr>
<tr>
<td>$D_{2\text{LA}}:D_{2\text{CA}}$</td>
<td>3.39:6.05</td>
<td>5.42:11.91</td>
<td>0.97:2.06</td>
<td>12.03:444.39</td>
<td>131.45:651.42</td>
<td>3.85:3.77</td>
<td>402.77:999.88</td>
<td>1.37:2.42</td>
</tr>
<tr>
<td>$D_{3\text{LA}}:D_{3\text{CA}}$</td>
<td>1.27:1.65</td>
<td>1.21:1.58</td>
<td>999.88:999.88</td>
<td>285.65:16.73</td>
<td>8.14:9.33</td>
<td>7.62:5.61</td>
<td>5.82:4.82</td>
<td>0.16:0.25</td>
</tr>
<tr>
<td>$D_{4\text{LA}}:D_{4\text{CA}}$</td>
<td>6.79:6.54</td>
<td>7.30:7.17</td>
<td>397.64:232.95</td>
<td>74.56:1000</td>
<td>1.98:2.41</td>
<td>3.68:3.86</td>
<td>101.33:18.68</td>
<td>49.25:26.06</td>
</tr>
<tr>
<td>$D_{5\text{LA}}:D_{5\text{CA}}$</td>
<td>49.24:1000</td>
<td>0.41:0.49</td>
<td>69.72:30.65</td>
<td>0.71:0.98</td>
<td>2.54:2.56</td>
<td>2.56:2.76</td>
<td>51.87:74.62</td>
<td>3.12:2.32</td>
</tr>
<tr>
<td>$K$ (Scaling factor)</td>
<td>0.57</td>
<td>1.02</td>
<td>2.78</td>
<td>0.80</td>
<td>0.66</td>
<td>0.83</td>
<td>1.00</td>
<td>1.37</td>
</tr>
<tr>
<td>Goodness of fit $r^2$</td>
<td>0.615</td>
<td>0.885</td>
<td>0.949</td>
<td>0.606</td>
<td>0.414</td>
<td>0.271</td>
<td>0.909</td>
<td>0.521</td>
</tr>
</tbody>
</table>

5.2 Calculation of the Denominator in (3)

Each set of $X_i$ (LA versus CA) and the imposed $M_i$ yield the terms for the calculation of the expression in curled brackets in (3). This was computed for all frequencies of interest for the LA and CA cases respectively. The functions were then normalized and averaged over the animals to yield a mean, normalized expression for lesion and control respectively.

5.3 Correction of the Experimentally Obtained Post-lesion Power Spectra for Characteristics of the Electronic Recording Apparatus

It has been noted above that the sought expression, $|A'(\omega)|$ is equal to $|A(\omega)|$ only if other transforming processes influencing the recorded spectral power are minimised.

For convenience the original EEG signals had been recorded through an amplification system imposing analog filtering on the signal. [See earlier report - Wright and Kydd (1984b).]

Therefore system band-pass characteristics were calibrated in their range with less than 1 mV sine waves over the EEG frequency band. From these system filter gain characteristics, the post-lesion power spectra from each animal were corrected to give their equivalent unfiltered spectral density.

5.4 Calculation of $|A(\omega)|$

The corrected post-lesion spectra were each normalised, and averaged over animals for the lesion and non-lesion data respectively. These plots provide the numerators for (3). All values to compute $|A(\omega)|$ in (3) were then available, except for $|X_0(\omega)|^2$ which being constant at each $\omega$, is scaled to unity in normalized data. The resulting plots of $|A(\omega)|$ (lesion and control) therefore contain all frequency-dependent gains or attenuations attributable to $|A'(\omega)|$ with extraneous factors minimised as much as possible.

The two plots of $|A(\omega)|$ were then fitted by least squares with a straight line, permitting comparison of their linear trends.

6 Results

The parameters found from curve-fitting $G^2(\omega)$ in each case with imposed natural frequencies are given in Table 1. Plots of $|A(\omega)|$ for the lesion and non-lesion hemispheres, and their linear trends, are shown in Fig. 1.

![Fig. 1. Dotted lines: Plots of $|A(\omega)|$; calculated estimates of the modulus of the transfer function transforming the electrocorticogram from the spatial average. Each graph is based on the average of eight cases. Top graph: Average from recordings on the side of unilateral lateral hypothalamic lesion. Bottom graph: Average obtained from the side opposite lesion. Solid lines: Straight lines of least-squares best fit, to linearize the trend in $|A(\omega)|$ against autocorrelated errors. Coefficients of slope and displacement are given above each graph](image-url)
7 Discussion

The mean linear trend for the control and lesion data on $|A(\omega)|$ is 0.974 dB/octave for the side of lesion, and 1.124 dB/octave for the normal side; a difference in gain of about 2% per octave.

Non-linearities in the plots of $|A(\omega)|$ are distributed in accord with expected autocorrelated error, showing minima close to the imposed natural frequencies. It will be seen that these errors are considerable, as is to be expected with a model of low and arbitrary order.

Linear trends in $|A(\omega)|$ are rather flat, consistent with Elul's (1972) comparison of neuronal waves with the electrocorticogram.

The slope of the linear trends in $|A(\omega)|$ is slightly upward with increasing frequency, consistent with that expected for electrocortical waves all of long wavelength compared to electrode separation.

It remains wholly possible that $|A(\omega)|$ is influenced by other factors than $|A'(\omega)|$ and the recording attenuations and changed driving signals allowed for by normalisation of the spectra. Whatever the case these factors also must have remained invariant following lesion to account for the comparability of left and right $|A(\omega)|$. If electrocortical activity were a markedly non-linear phenomenon, or one in which the influence of the lateral hypothalamus is not associated with damping parameters, it would appear unlikely that the back-calculation of $|A(\omega)|$ would be similar on each side. Instead dissimilar results would be expected, revealing the arbitrary nature of the relative gain function in fitting the experimental results.

One special case requires consideration: electrocortical activity which was totally spatially incoherent, but still retained the local temporal characteristics of damped oscillation with fixed natural frequencies could have given rise to these results. This is not compatable with the results of other work in spatial analysis cited above. Even if this interpretation were correct, it has still been shown that parameters obtained by fitting the relative changes in spectral power "unfold" into appropriate and similar estimates of the electrode transfer characteristics on each side. The initial theoretical assumption of linear transfer characteristics for the transformation of the electrocortical average into the recorded signal is thus also vindicated.

The collective effect of these findings is to show that the linear control model on which they are based is internally consistent. Therefore an appropriate linear description fitted to the EEG in the time-domain may describe the movement and regulation of real waves subject to a superposition principle, rather than arbitrarily representing the data. The limits of this linearity are not yet defined.

Given these grounds to believe a linear model appropriate for initial studies in this field, several further applications and developments seem warranted, principally in those areas we have circumvented for given reasons.

Left/right EEG comparisons before and after lateral hypothalamic lesion should be undertaken using linear analysis techniques in the time-domain, aiming both to optimise estimates of the correct number and frequencies of natural modes, and to obtain estimates of $\mathcal{S}_i$ as time-varying parameters.

Further analysis of the relevant anatomical pathways is needed. A beginning has been made by assessing the role of dopaminergic neurons (see Wright and Kydd, 1984a). A large literature exists which describes centripetal and corticofugal pathways involved in "cortical activation" (e.g. Lindsley et al., 1950; French et al., 1955; Scheibel and Scheibel, 1967; Zanchetti, 1967). If application of a linear model continues to prove practicable, then repetition of these anatomical analyses in association with time-domain linear modelling would cast light on which structures are to be considered mediators of control by damping, and which are involved in the generation of resonances. These need not be exclusive attributes for most structures. It might, for instance, prove the case that a number of resonant systems interact by multiplicative coupling (Mohler, 1973) as well as being subject to overall supervision via the brain-stem.

It will be seen that (8) and the experimental estimates of $|A(\omega)|$ lead to a first approximation of the relation of $c_k$ to frequency, at least for waves moving in one direction at the local cortical region. Multichannel spatial EEG analysis might be considered in association with development of the control model in the time-domain, to see if a dispersion relation consistent with this first approximation can be demonstrated (cf. Nunez, 1981).

Finally, we urge that the implications for control of wave movement of this type be explored in network theory, to aid convergence of EEG modelling with simulated microscopic networks. In MacKay's (1960) concept the EEG waves were thought of as reflecting the shift toward depolarisation of whole populations of cells, and the significance of this as a change in transition probability was pointed out. A realistic network model would need to show lawful changes in logical properties as a consequence of the overall control envisaged in these studies and might be expected to parallel the psychological changes in attention, arousal and motivation associated with manipulation of the lateral hypothalamus.

Acknowledgements. This work was supported by the New Zealand Medical Research Council, through a grant to J. J.
Wright and a research Fellowship to R. R. Kydd. We gratefully acknowledge the technical assistance of J. A. West.

References
Scheibel, M.E., Scheibel, A.B.: Structural organisation of non-specific thalamic nuclei and their projection toward cortex. Brain Res. 6, 60–94 (1967)

Received: November 25, 1983

Prof. J. J. Wright
Department of Psychiatry
School of Medicine
The University of Auckland
Private Bag
Auckland
New Zealand
Amplitude and Phase Relations of Electro cortical Waves Regulated by Transhypothalamic Dopaminergic Neurones: A Test for a Linear Theory

J. J. Wright, R. R. Kydd, and G. J. Lees

Department of Psychiatry, University of Auckland School of Medicine, Auckland, New Zealand

Abstract. We have previously proposed that electrocortical activity (EEG) arises as a manifestation of linear waves generated by resonance among telencephalic neurones, and that this activity is controlled in part by ascending neurones from the brain-stem, which regulate the damping of each resonance. The present experiments focus on a specific class of ascending neurones, the mesotelencephalic dopaminergic cells, because these cells are thought to mediate important psychological effects, and are conveniently subject to selective lesion. A critical test of the theory is undertaken, by performing selective unilateral lesion, assessing the changes in the power spectrum of the EEG attributable to lesion, and determining whether the changes in phase of the EEG correspond to that predicted from the changes in power. Results support the theory, although the model order applicable in these experiments is inadequate. The consequences of these findings for automata theory, linear network theory and their application to mammalian brains are briefly discussed.

1 Introduction

This paper concerns a critical test for a linear theory of the origin of electrocortical waves and their internal regulation. In this experimental test, emphasis is placed upon the actions of a particular class of neurones - the mesotelencephalic dopaminergic neurones.

This class of neurones has cell bodies located in the substantia nigra pars compacta and related nuclei, and axonal fibres project via the lateral hypothalamus to terminate ipsilaterally in the caudate nuclei, parts of the limbic system, and frontal cortex (Jacobowitz and Palkovits, 1974).

These neurones appear involved in the regulation of attention and motivation (e.g. Marshall et al., 1974; Ljungberg and Ungerstedt, 1976; Wise, 1978; Wise et al., 1981). Since electrocortical activity is an observable physiological variable closely reflecting cognition, attention and motivational set (Walter et al., 1967; Berkhout et al., 1969; Gevins et al., 1979) it may be expected that activity of dopaminergic neurones will partially govern electrocortical function, albeit by indirect mechanisms. The type of control exerted upon overall brain processes by central dopaminergic neurones might therefore be determined using the EEG. Partly toward this end, we have previously advanced a mathematical theory and supporting evidence relating to the role of transhypothalamic fibres in controlling electrocortical activity (Wright and Kydd, 1984a–c). These findings concerned EEG changes produced by electrolytic lesion of the entire lateral hypothalamic fibre system, but we indicated that further experiments we had conducted showed that lesion of the dopaminergic neurones alone gave very similar results. A detailed account of these further findings will be published elsewhere.

The present findings, also obtained using selective dopaminergic lesions, lead us to believe that further development of the model under consideration might embrace certain findings regarding electrocortical activity, psychophysiology, linear network theory and the theory of automata within a single general framework.

2 Review of the Theoretical Model

A detailed account of this theory and experimental evidence in support has been given previously (Wright and Ihaka, 1981; Wright and Kydd, 1984a–c). Essential theoretical features are as follows:

2.1

Electrocortical recordings reflect the transformed spatial average of cortical dendritic potentials.
2.2
The fixed circuitry of the telencephalon and stochastic considerations of the linkages between neural elements render the telencephalon a linear wave medium, with regard to the gross wave potentials, although the underlying microscopic interactions may be extremely non-linear.

2.3
Closed and constant boundary conditions lead the linear waves to generate activity at a large number of resonant modes, each associated with a constant natural frequency (and presumably a specific spatial configuration).

2.4
Consideration of the modular and architectonic orderliness of the telencephalon requires that the values for the natural frequencies of the resonant modes be clustered about certain central values, each cluster within the frequency band width of a major cerebral rhythm.

2.5
Ascending inhibitory systems (including the mesotelencephalic dopaminergic neurones) act partly to damp resonant activity and partly as a source of noise-like driving signals, by their input to the telencephalon at their fields of termination.

To be more mathematically explicit; starting from the general description of the voltage variation for each segment of the dendritic field of cortical neurones by the equation

\[ x(t) + D(t)x(t) + N(t)x(t) = 0, \]

where \( x(t) \) is the segment contribution to the surface potential, and \( D(t), N(t) \) time-varying parameters, then (1) is assumed to describe the “isolated” activity arising from activity in some dominant circuit of which this segment is a part. Time variation in the parameters is included to descriptively account for non-linearities within the segment activities.

The telencephalon is then considered as composed of millions of such units, linked in a complicated, but non-random fashion, the whole being perturbed by input signals and active non-linear, inhibitory and excitatory interactions. Stochastic considerations show that the whole is equivalent to a linear network, and the input of diffuse inhibitory neurones arising outside the telencephalon is shown to regulate the damping of activity at each of the numerous resonant modes, and to simultaneously regulate the strength of the background driving “noise”.

Thus the point potential attributable to each mode of resonance will have a power spectrum, \( V_i^2(\omega) \) given by

\[ V_i^2(\omega) = (M_i^2 - \omega^2) + D_i^2 \omega^2, \]

where \( M_i \) is the natural frequency of the mode, \( D_i \) is the damping coefficient and \( \omega \) is frequency.

The phase of the mode potential \( \Phi(\omega) \) (referred to the average phase of the driving noise) will be given by

\[ \Phi(\omega) = \tan^{-1} \left( \frac{-D_i \omega}{M_i^2 - \omega^2} \right). \]

The total signal is the sum of the mode activities, transformed by other effects attributable to volume conduction, media conductivities, electrode position and point phase velocities for wave components.

3 The Proposed Test of Theory

From (2) and (3) it will be seen that \( D_i \), the damping coefficients, act as parametric controls of amplitude and phase, since all \( M_i \) are theoretically constant, and were provisionally shown to be so, in earlier experiments (Wright and Kydd, 1984b, c). The \( D_i \) are theoretically supposed to depend on trans-hypothalamic fibres, including the dopaminergic group.

Advantage may now be taken of the degree of left-right phase coherence exhibited by electrocortical signals from opposite recording sites (Walter et al., 1967). Their left-right coherence has been shown in other experiments to be attributable in part to signals transferred via the interhemisphere commissures (Nunez, 1981). We assume this gives rise to a degree of left-right synchrony of the noise-like driving signals, and possibly (but not necessarily) of the mode activities also.

As a highly idealised abstraction we may consider the left and right driving signals as perfectly coherent, and the time-varying damping coefficients as constant in steady-state conditions. These coefficients will then be equal left and right in the intact animal, and unequal following unilateral lesion of the dopaminergic neurones. Likewise, the power of the driving signals will be equal left and right prior to lesion, but dissimilar following lesion. If other factors transforming the recording signal at each site also remain invariant, then the relative change in power of left and right electrocortical signals following lesion is obtained experimentally as

\[ G^2(\omega) = \frac{V_{LA}^2}{V_{LB}^2}, \]

where

\[ G^2(\omega) = \frac{V_{LA}^2}{V_{LB}^2}, \]

\[ V_{LA}^2, V_{LB}^2, V_{LA}^2, V_{LB}^2. \]
where the subscripts LA, LB, CA, CB, refer to epochs of recording from the lesion and control sides after and before lesion. Theoretically, this relative power function is given by

\[
G^2(\omega) = K \left[ \frac{\sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + D_{LA}^2 \omega^2}}{\sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + D_{CA}^2 \omega^2}} \right]^2 + \frac{\sum_{i=1}^{n} \frac{-D_{LA} \omega}{(M_i^2 - \omega^2)^2 + D_{LA}^2 \omega^2}}{\sum_{i=1}^{n} \frac{-D_{CA} \omega}{(M_i^2 - \omega^2)^2 + D_{CA}^2 \omega^2}}^2.
\]

where K indicates the ratio change in power of the driving signal, and \( n \) is the number of modes.

Similarly the relative change in phase of the signals attributable to lesion, \( \Delta \Phi(\omega) \) can be found experimentally as

\[
\Delta \Phi(\omega) = \Delta \Phi_{LA,CA} - \Delta \Phi_{LB,CB},
\]

where \( \Delta \Phi_{LA,CA} \) and \( \Delta \Phi_{LB,CB} \) refer to the mean difference in phase angle over numerous short epochs.

Since, in steady-state conditions for the control hemisphere, \( D_{CA} = D_{CB} = D_{LB} \), this difference in mean phase angle is given theoretically by

\[
\Delta \Phi(\omega) = \tan^{-1} \left[ \frac{\sum_{i=1}^{n} \frac{-D_{LA} \omega}{(M_i^2 - \omega^2)^2 + D_{LA}^2 \omega^2}}{\sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + D_{LA}^2 \omega^2}} \right] - \tan^{-1} \left[ \frac{\sum_{i=1}^{n} \frac{-D_{CA} \omega}{(M_i^2 - \omega^2)^2 + D_{CA}^2 \omega^2}}{\sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + D_{CA}^2 \omega^2}} \right].
\]

It follows that, given ideal conditions, knowing \( n \), the number of modes, and with error free recordings, (5) could be curve-fitted to \( G^2(\omega) \) found experimentally via (4). The values of the parameters for \( M_i \) and \( D_i \) (LA and CA) would then be known. \( \Delta \Phi(\omega) \) found experimentally in (6) should then be perfectly predicted via (7) using these parameter values.

The actual test used involves an approximation to this idealized critical test.

4 Experimental Limitations, and Error Analysis

In practice, this test confronts severe limitations, arising because of departures from idealized conditions.

Firstly, left-right EEG phase coherence (before lesion) in these experiments, averaged over all frequencies of interest is found to be only 0.5. Time variation in behavioural state is bound to be considerable over the duration of EEG epochs long enough for adequate analysis, and such behavioural variation is equivalent to time-variation of the damping coefficients, as conceived in this model. Such variations can be both systematic and random. These factors induce a great degree of error into estimates of \( \Delta \Phi(\omega) \).

Secondly, the model order, \( 2n \) is not known, and is almost certainly very high. Only a very rough approximation of the true relative power changes can be expected from a low order model fitted to averaged power spectral data. Conversely, a practical upper limit to the order of model which can be fitted is soon reached, as with rising order, less and less accurate assessments of individual parameter values are obtained. Both these limits to accurate approximation, and the possible failure to find a true global minimum during curve-fitting can introduce major errors in the pattern of residuals when theoretical and experimental relative phase estimates are compared.

For these reasons we have not attempted a fit to power changes with more than a tenth order linear model, allowing for five dominant modes of resonance, and instead of measuring goodness of fit of theoretical and experimental phase, have measured their non-parametric correlation. This technique enables a test of significance to be applied, despite the effects of distortion of residuals described above.

If, as expected, the limitations introduced by inaccurate model order, left/right signal incoherence and time-variation are considerable, then only a low correlation can be found between the phase predicted and that found experimentally. By itself, this would leave the physical validity of the model somewhat undecided. That is, a low correlation might mean that the model is only marginally physically valid, or it might mean that the experimental limitations prohibited finding a more exact fit for a model of high physical validity if correctly optimised. To distinguish between these outcomes as much as possible, we performed curve-fitting to the relative power changes for models of the second, fourth, sixth and eighth orders in addition to the tenth order, and examined the correlations of predicted and experimental phase at all
orders, along with explicit consideration of the pattern of residuals. Details of this analysis are dealt with in Sect. 6.

5 Methods

5.1 Animals

These were seventeen young male Wistar rats with body weights of 300–400 g.

5.2 Implantation of Recording Electrodes

Under general anaesthesia, bipolar extra-dural platinum-ball recording electrodes 0.5 mm in diameter were implanted over the posterior cortex at fixed positions in the stereotaxic plane.

These were:

Anterior electrode: P 7.5, L 5.1. Posterior electrode: P 9.5, L 5.1, zeroed on Bregma. Skull screws provided ground electrodes, and the cranial assemblies were embedded in dental acrylic leaving a connector exposed.

Ten to fourteen days of post-operative recovery were allowed before experiments began.

5.3 EEG Recordings

These were obtained in conditions of wakeful rest, after the animals had habituated themselves to the experimental conditions. The rats rested gently restrained in a soft cloth. EEG signals were sensed via a plug to the connector containing unity-gain buffering field-effect transistors. After differential amplification at 5 K gain, on Tektronix AM502 amplifiers and analog filtering by Butterworth filters (low pass 48 dB/octave), giving a 1–35 Hz bandpass at the 3 dB points, signals were fed to analog-to-digital convertors, where they were digitised at 8 ms intervals, before being stored on disk in 34 8 continuous epochs, via the core of a PDP 8e computer. The converted signals were visually monitored on-line.

Twenty epochs of recording were obtained 24 h before mid-brain dopaminergic lesion, and twenty more 4 days following lesion. Further editing, to select epochs reflecting the maintenance of a steady state, reduced the number of epochs used to a median of seven prelesion and seven postlesion (range five to thirteen). This editing used review of the acquired time- series, rejecting those epochs in which any sudden transition of the overall EEG occurred (e.g. from low voltage fast activity to large amplitude slow waves). The later calculated power spectrum of each epoch further helped select epochs with similar spectra in the control hemisphere pre-and post-lesion. This editing was performed by eye, without knowledge of results of later spectral calculations, so that selection was purely with regard to obtaining uniform overall conditions in each animals case.

5.4 Lesion Techniques

After the initial EEG data had been obtained, unilateral lesion of the dopaminergic cell group was produced by intracerebral injection of the selective toxin, 6 hydroxy-dopamine hydrogen bromide. (6-OHDA) (Hökfelt and Ungerstedt, 1973). This was performed under ketamine/barbiturate anaesthesia, with a stereotactically mounted Hamilton syringe, introduced at a 50 deg. obliquity from vertical. This infused 0.5 μL/min into the rostral extremity of the substantia nigra [coordinates A3L1.8V-2 from stereotaxic zero – Atlas of Konig and Klippel (1963)] of a 6-OHDA solution (2 μgm in 2 μL of phosphate buffer with 0.2% ascorbic acid). Noradrenergic neurones were protected from damage with desipramine HCl 25 mg/kg given i.p. Control animals underwent injection of buffer alone.

5.5 Sensorimotor Neglect

This behavioural procedure was performed to assess the degree of functional transhypothalamic dopamine cell damage, reflected in impaired somatosensory orientation on the side opposite lesion (Marshel et al., 1971).

The rats were tested pre-lesion, then four days, and again ten days after lesion. Any tendency to turn preferentially in one direction in an open field situation was first noted. Firm cotton buds were simultaneously and repeatedly applied to opposite sides of the body to assess the preferred direction of turning to double stimulation. Sensory stimuli were then applied to each side of the body with random alternation of side. Visual orientation was tested by moving a 2 cm × 2 cm square of white cardboard into the animal's visual field, and olfaction by moving a xylene soaked cotton bud close to each nostril. Touch orientation testing was performed using a cotton bud, and a Von Frey hair of 6 gm pressure applied to head, shoulder and mid-trunk. The vibrissae were stroked from behind forwards, and a stick was touched to the head just behind the mouth to assess the biting response. A 23 g needle was used to assess the response to pinprick. The precision with which the animal could localise and respond to the applied stimulus was used, together with the turning behaviour, to estimate the degree of neglect (see Marshall et al., 1971; Marshall, 1979). Neglect was rated as follows: + + + = Full neglect: marked circling to side of lesion, little or no response in any sensory modality contralateral to lesion. + + = partial neglect: circling to side of lesion, awareness of stimuli to contralateral side evidenced by
squeaking or movement but no head orientation. + = mild neglect: preferential turning to side of lesion, contralateral head orientation in direction of applied stimulus but no localisation. 0 = no neglect: equal turning to both sides, precise head orientation to site of stimulus plus sniffing, biting or scratching.

5.6 Histology

Ten days following intracerebral injection the animals were sacrificed and subjected to histochemical evaluation in order to determine the extent of the 6-OHDA-induced lesion. Catecholamine-containing cells were visualized by the FAGLU method of Furness et al. (1978). The animals were perfused via the thoracic aorta with approx. 100 ml 0.9% saline followed by approx. 200 ml of FAGLU (a mixture of 4% formaldehyde and 0.5% glutaraldehyde in 0.1 M phosphate buffer, pH 7). The brain was removed and allowed to soak overnight in the FAGLU mixture before being frozen-sectioned on a Jung Hn 40 sliding microtome at a 30 μm spacing. Sections were mounted on microscope slides, dried over phosphorous pentoxide for at least 1 h and mounted in DPX-mountant. Catecholamine-induced fluorescence was visualized under a Wild M 20 transmitted light fluorescent microscope fitted with a 200 W mercury lamp, a BG 23 red absorbing filter, a FITC exciting filter (emission 400–500 nm) and a FITC barrier filter (light transmission above 510 nm). Alternate sections were also stained with thionin (Rucker-Koithan procedure) and used to confirm the extent of damage to the dopaminergic cells, the stereotaxic needle placement and the specificity of the lesion (particularly with respect to the extent of microglial infiltration).

The presence of yellow fluorescing cells, or the absence of all fluorescing cells in areas normally occupied by dopaminergic cell bodies was taken as an indication of cellular death. Histological patterns of cellular degeneration were mapped over the whole rostro-caudal extent of the midbrain dopaminergic system (cell groups A 8, A 9, and A 10). A semiquantitative assessment of cellular damage was made from these maps at 4 planes; A 1.3, A 1.8, A 2.2, and A 2.6. At the appropriate planes, damage to the lateral and medial substantia nigra and the ventral tegmental area were each assessed on a scale of 0 to 3 (0, no damage; 1, 1–10% cell loss; 2, 11–90% cell loss; 3, 91–100% cell loss). The results from each area were summed for each animal to give an estimate of total unilateral damage to the midbrain dopaminergic system. Complete ipsilateral loss of the mesencephalic dopamine system results in a damage score of 33.

The maximal lateral, vertical and rostro-caudal extents of the zone of architectural disorganisation and glial cell infiltration evident on thionin stained sections were used as a measure of nonspecific damage.

5.7 Further Analysis of EEG Signals

Data from each recording epoch was divided into segments containing 512 data points and each segment separately underwent 10% cosine-bell windowing then Fast Fourier analysis.

The Fourier coefficients obtained for each frequency ω, from each segment, V(t) of duration T, that is

\[ a(ω) = \frac{1}{T} \int_{-T/2}^{+T/2} V(t) \cos ωt \cdot dt \]  

and

\[ b(ω) = \frac{1}{T} \int_{-T/2}^{+T/2} V(t) \sin ωt \cdot dt \]  

gave the power spectrum for each segment

\[ V^2(ω) = a^2(ω) + b^2(ω) \]  

and the phase for the segment

\[ ϕ_{seg}(ω) = \tan^{-1} \left( \frac{b(ω)}{a(ω)} \right) \]  

Both were obtained at 128 point resolution from 0.25 Hz to 32 Hz. The Fourier coefficients also enabled the calculation of left/right coherence, over a sequence of segments. The segment power spectra were then averaged to provide an average power spectrum for each channel, before and after lesion respectively. Division of these averaged power spectra yielded the relative squared gain attributable to lesion, as required by (4). Subtraction of the corresponding phase estimates gave the relative phase required by (6). Each plot of \( G^2(ω) \) was then slightly smoothed, using low-pass Fourier smoothing to remove all components above the 32 harmonic, from the \( G^2(ω) \) plot's own transform. This minimised cumulative error introduced in divisions thus improving precision of convergence in subsequent curve fitting. The \( Δϕ(ω) \) plot was similarly treated. Finally, the relative power estimates at the lowest 1.5 and highest 0.5 Hz were eliminated, to further minimize effects of any residual low or high frequency artifact present in the original recordings. Estimates of \( G^2(ω) \) and \( Δϕ(ω) \) thus obtained for each animal were then respectively pooled into three animal groups (see results). These groups were constituted of controls, animals with marked dopaminergic unilateral damage and neglect, and those with lesser damage and no neglect.
5.8 Curve Fitting

Library routines were used to find the minimum least squares fit of (5) to each of the pooled estimates of $G^2(\omega)$. The procedures utilized machine estimation of the vector to best fit, and its first and second partial derivative, to converge to local minima from given starting values. Local minima were numerous, and searches for a global minimum were undertaken using starting values systematically distributed to yield diverse initial plots of residuals, until an apparent global minimum with best goodness of fit was found. Values of all parameters found at best fit were then tabulated for each case. This procedure was repeated with (5) set at $n = 1, n = 2, n = 3, n = 4$, and $n = 5$.

5.9 Test of Hypothesis

The parameters obtained for best fit to the relative squared gain, for each animal group, and each order of model, were then used to generate a corresponding predicted relative phase using the PDP 8e once again. The theoretical phase for each order was then correlated non-parametrically using Spearman's rho, with the experimental phase. The value of rho was converted to the equivalent student's $t$, and this tested for significance by reference to tables. Degrees of freedom in the test of significance were considered halved from that appropriate to the number of data points correlated, since the raw relative phase plot had been Fourier smoothed to the corresponding extent. Correlation was performed over the entire band width of interest, and in selected cases over a narrower band width to delineate special effects (see results).

6 Results

6.1 Sensorimotor Neglect

The four animals injected with control vehicle showed no evidence of sensorimotor neglect. Ten of the thirteen 6-OHDA animals developed contralateral sensorimotor neglect post-operatively, and this was manifest on each day of testing. Neglect scores were between $+++$ and $+$ on the day of post-operative EEG recording, with a median score of $+++$ for this group. The remaining three 6-OHDA animals did not manifest sensorimotor neglect as measured by these tests.

6.2 Histological Findings

No damage to the mesotelencephalic dopaminergic system was discerned in the control animals. The ten animals with contralateral neglect showed widespread unilateral dopamine-cell damage, scattered throughout the A8, A9, and A10 cell groups. The mean damage score for this group was 23.4, with a damage score ranging from 13 to 29 for the individuals. The three 6-OHDA animals without evidence of sensorimotor neglect showed a lesser degree of dopaminergic cell damage ipsilateral to the injection. In these cases the group mean damage score was 19 with an individual range of 16 to 21.5. The animals showing no neglect after lesion had significantly lower damage ($p < 0.02$, Mann-Whitney $U$ test, two tailed) in the A8 and A10 mesolimbic area (average 3.8) than those animals showing neglect (average 8.0). Loss of cells in the A9 dopamine cells was equal for the two groups (15.2 and 15.5, respectively). An associated study with a larger animal group (Lees et al., submitted for publication) shows that A8, and A10 damage is additive to A9 damage in the severity of sensorimotor neglect.

Conventional histology revealed no significant difference in needle track damage, local necrosis at the site of injection, or gliosis, all of which were minimal. The dorsal noradrenergic bundle was intact in all cases.

6.3 Grouping of Results

For the purposes of pooling relative power and relative phase estimates, the animals were subsequently treated as comprising three groups.

Data set one: the ten animals with contralateral sensorimotor neglect and marked unilateral mesencephalic dopaminergic damage.

Data set two: the three animals with limited dopaminergic damage and no apparent sensorimotor neglect.

Data set three: controls.

6.4 EEG Changes in Animal Groups

Pooled estimates of relative power, $G^2(\omega)$, for data sets one and two showed that a manifest asymmetry of spectral power had developed in both groups. Data set three showed no such significant changed power ratio following injection and $G^2(\omega)$ appeared to be a near random plot, scattered about a mean of 0.8 at all frequencies. The only effect therefore, was a slight drop of total power on the side of the lesion, presumably a consequence of non-specific damage. It was taken that control injections had been without consequences of interest, and subsequent analysis was directed to data sets one and two.

6.5 Outcomes of Model Fitting

Table 1 shows the parameters, goodness of fit to relative power and the tests of correlation between
Table 1. Parameter values for best goodness of fit to relative power, and consequent correlations of predicted and experimentally determined relative phase. Model parameters limited to order ten (five resonant modes)

<table>
<thead>
<tr>
<th></th>
<th>Data set 1 (Fig. 1a and b)</th>
<th>Data set 2 (Fig. 1c and d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scaling factor, ( K )</td>
<td>0.354</td>
<td>0.749</td>
</tr>
<tr>
<td>Natural frequencies</td>
<td>1.775, 2.435, 3.302, 12.415, 24.879</td>
<td>3.801, 4.093, 6.103, 19.852, 27.594</td>
</tr>
<tr>
<td>Damping coefficients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(cycles/s) on lesion sides</td>
<td>0.868, 0.813, 2.465, 6.970, 22.073</td>
<td>2.248, 0.701, 2.466, 15.566, 13.110</td>
</tr>
<tr>
<td>Damping coefficients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(cycles/s) on control sides</td>
<td>1.693, 1.334, 1000, 8.528, 23.342</td>
<td>6.886, 0.859, 2.857, 11.133, 9.599</td>
</tr>
<tr>
<td>Goodness of fit to relative squared gain</td>
<td>0.847</td>
<td>0.911</td>
</tr>
<tr>
<td>Correlation of theoretical and experimental phase (Spearman’s rho)</td>
<td>Whole bandwidth 0.186</td>
<td>Above 6 Hz 0.127</td>
</tr>
<tr>
<td>Student’s ( t ) (60 df)</td>
<td>2.049</td>
<td>1.283</td>
</tr>
<tr>
<td>Significance, ( p ) (one tailed)</td>
<td>&lt;0.025</td>
<td>&lt;0.15</td>
</tr>
</tbody>
</table>

**Fig. 1a and c.** Experimental relative power consequent to lesion (dots) and best fit for tenth order theoretical gain function (solid lines). **Fig. 1b and d.** Experimental relative phase shift consequent to lesion (dots and connecting lines) and superimposed theoretical phase shift (solid lines) predicted from limited order model parameters (corresponding phase plots centered to their mean values across frequency).
Table 2. Parameter values, goodness of fit and phase correlations (as in Table 1) determined for model orders two to eight (one to four resonant modes)

<table>
<thead>
<tr>
<th>Order of model</th>
<th>Data set 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Best fit 1</td>
<td>Best fit 2</td>
<td>Best fit 2</td>
<td>Best fit 2</td>
<td>Best fit 2</td>
</tr>
<tr>
<td>Natural frequencies (Hz)</td>
<td>26.689</td>
<td>3.999</td>
<td>0.00</td>
<td>1.879</td>
<td>0.0001</td>
</tr>
<tr>
<td>Corresponding damping coefficients</td>
<td>46.449</td>
<td>36.579</td>
<td>13.045</td>
<td>17.979</td>
<td>2.709</td>
</tr>
<tr>
<td>lesion: control (cycles²/s)</td>
<td>6.468</td>
<td>11.346</td>
<td>2.999</td>
<td>2.968</td>
<td>0.690</td>
</tr>
<tr>
<td>K (scaling factor)</td>
<td>0.91511</td>
<td>0.479133</td>
<td>0.531438</td>
<td>0.459254</td>
<td>0.81940</td>
</tr>
<tr>
<td>Goodness of fit, r², to power</td>
<td>0.739948</td>
<td>0.73988</td>
<td>0.815962</td>
<td>0.826154</td>
<td>0.836285</td>
</tr>
<tr>
<td>Spearman’s rho</td>
<td>0.0706855</td>
<td>0.215203</td>
<td>0.152531</td>
<td>0.142871</td>
<td>0.1203</td>
</tr>
<tr>
<td>Students t</td>
<td>0.769766</td>
<td>2.39378</td>
<td>1.67653</td>
<td>1.56806</td>
<td>1.31635</td>
</tr>
<tr>
<td>Significance p (one tailed)</td>
<td>NS</td>
<td>&lt;0.01</td>
<td>&lt;0.05</td>
<td>&lt;0.10</td>
<td>&lt;0.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Order of model</th>
<th>Data set 2</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Natural frequencies (Hz)</td>
<td>2.290</td>
<td>4.630</td>
<td>3.077</td>
<td>0.00</td>
<td>2.427</td>
</tr>
<tr>
<td>Corresponding damping coefficients</td>
<td>11.513</td>
<td>19.332</td>
<td>3.830</td>
<td>31.088</td>
<td>3.627</td>
</tr>
<tr>
<td>lesion: control (cycles²/s)</td>
<td>2.962</td>
<td>3.530</td>
<td>8.927</td>
<td>7.075</td>
<td>2.951</td>
</tr>
<tr>
<td>K (scaling factor)</td>
<td>0.456003</td>
<td>0.343614</td>
<td>0.84207</td>
<td>0.8289</td>
<td></td>
</tr>
<tr>
<td>Goodness of fit, r², to power</td>
<td>0.853</td>
<td>0.887207</td>
<td>0.905875</td>
<td>0.909345</td>
<td></td>
</tr>
<tr>
<td>Spearman’s rho</td>
<td>0.0516147</td>
<td>-0.102111</td>
<td>0.0766859</td>
<td>0.255539</td>
<td></td>
</tr>
<tr>
<td>Students t</td>
<td>0.561427</td>
<td>-1.11504</td>
<td>0.835483</td>
<td>2.87119</td>
<td></td>
</tr>
<tr>
<td>Significance p (one tailed)</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>&lt;0.005</td>
<td></td>
</tr>
</tbody>
</table>

predicted and experimental relative phase, for the best fits of the tenth order model, to data sets one and two.

Table 2 shows corresponding values for all orders of model from two to eight, for each of these data sets. The second order fit to data set one yielded two fits of almost identical goodness of fit so both are given.

Figure 1 shows the experimental and theoretically fitted data of Table 1 (the highest order cases) in graphical form.

Coherence ranged between 0.4 and 0.7 over various frequencies and samples, and is not further reported here.
6.6 Consideration of Results: Phase Residuals and Model Order

Goodness of fit of the equation for relative squared gain (5) to the experimentally found relative power shifts following lesion, rises as is to be expected with increasing model order, for best fit at each order.

The two cases of near identical goodness of fit shown in Table 2 (second order, data set one) reveal the potential ambiguities in curve-fitting in these circumstances. One case describes the relative power as if it were attributable to a relative release of damping at a high frequency mode on the side opposite lesion. This case is almost certainly physiologically spurious, as control hemisphere spectra were approximately matched before and after lesion (see Sect. 5.3). The other case is consistent with that found by higher order best fits, and this is appropriate to the increase in low frequency power that was apparent in the raw spectra on the side of lesion.

The relative power shifts in data sets one and two are disparate in form, and it is not at present possible to say whether this signifies a systematic difference between animals with localised versus widespread dopaminergic lesion.

Of the relevant correlations of theoretical and experimental relative phase estimates, nine are positive correlations, some very significantly so, and the tenth is not significantly negative.

In analysing the adequacy of phase description in detail the following must be kept in mind: (5) and (7) respectively exhibit maxima for power, and a tendency to rapid positive to negative transitions for phase, about the natural frequencies. When (5) is fitted to real data at inadequate order, parameters approximating activity at certain dominant modes are obtained, and the predicted phase will be best correlated with real relative phase near these dominant modes. For data set one, most significant phase correlations are found at the tenth, fourth and second orders, and less so at other orders. Each of these best-fit approximations to the relative power giving $p < 0.05$ correlations for phase, find the dominant natural frequencies, with marked left/right disparities of damping coefficients, at below 6 Hz. Partial range correlation of phase above 6 Hz, for the tenth order fit (Table 1) shows that for this data set the phase is best described in the low frequencies. For data set two, the eighth order model gives a highly significant correlation to phase, and the tenth order model a weaker positive correlation. In the tenth order instance left/right disparities of the damping coefficients are apparent at natural frequencies from approximately 4 Hz upwards, with more scatter of natural frequencies across the frequency spectrum. Partial range correlation (performed from 6 Hz upwards to be consistent with that of data set one) in this tenth order case shows that the phase is well described at the higher frequencies, and is negatively correlated overall below 6 Hz, although still reasonably described near the lower estimated natural frequencies. This low frequency part of the spectrum appears generally better accounted for in the eighth order description (not shown graphically here), at which very low frequency modes were estimated.

In both data sets a successful prediction of phase was achieved by lower order models which estimated very low frequency modes (below 1 Hz). This might suggest that such low frequency activity has been "missed out" in the tenth order fit. Similar considerations, or spurious global minima might also account for the indeterminate outcome of certain other cases. In summary, the correlations found support hypothesis. In the highest order fits attempted, it can be seen that the fine-structure of phase is best correlated at frequencies near those at which the dominant natural frequencies have been estimated on the relative power plot. These features are as expected for limited model order estimations of relative power changes.

7 Discussion

These results indicate that mesotelopechalic dopaminergic neurons exert controlling influences upon electrocortical activity which are appropriate to the regulated damping of linear resonant phenomena envisaged in the outlined model. Relative power changes are significantly in accord with shifts of phase, showing the description of relative power has physical, as opposed to arbitrary, validity. The above considerations of experimental error and the types of residual distribution show that time-variation, low model order and left/right signal incoherence set an upper limit to the degree of correlation found, rather than theoretical considerations. The approximation to physical reality should be much better if adequate model orders can be utilised, and allowance made for time-variation of the damping parameters—this could only be achieved in the time-domain, and is a formidable (but not insurmountable) task of computation.

The earlier published tests of this model (Wright and Kydd, 1984b, c) and their later replication using selective dopaminergic lesions of dopaminergic neurons appear to offer parallel support for the same theory, and indicate that a general linear analysis of electrocortical activity is possible. Such an analysis could have fairly considerable weight in interpreting electrocortical correlates of behavioural and cognitive states. This seems implied for the following reasons. The initial model (Wright and Ihaka, 1981; Wright and
Kydd, 1984a) dealt with all neurones ascending through the lateral hypothalamus as a single class, and accordingly the first experiments to test the model (Wright and Ihaka, 1981; Wright and Kydd, 1984b, c) used gross electrolytic destruction, or electrical stimulation (Wright and Craggs, 1979; Wright and Ihaka, 1981) of this pathway. Duplication of these results with selective 6-OHDA lesions, and the present data indicate that the dopaminergic cell class alone can account for those findings. The other classes of cells passing through the lateral hypothalamus must therefore either produce effects of the same general type, or alternatively no marked effect upon EEG activity. The theory itself suggests that any diffuse inhibitory cell class should exert similar effects, and this is subject to test by further selective lesion experiments, concerning the noradrenergic and serotonergic cell groups, for example. If this is generally the case, then there is no need to confine application of these concepts to the supervisory and diffuse cell groups arising from the brain-stem. Similar reasoning would apply to interactions among all cell elements in the telencephalon, which would not only be stochastically additively cross-coupled to each other (as described in the present theory) but multiplicatively coupled; that is, exerting cross-damping and noise-like driving effects upon each other. In view of the known anatomy (which is too complicated to be reviewed here) it is a very reasonable assumption that certain limbic and cortical structures may exert control from the telencephalon upon the ascending brain-stem systems. This supposition renders a means for self-control of the entire telencephalon and brain-stem system apparent, and means that the present theory of electrocortical resonances and their control can be developed into a minimal, "skeletal" theory of overall brain information processing. That is, within this theory, the marked non-linearity of real neural elements and their interactions reduces to that of an equivalent linear network, driven by both input signals and a large amount of internally generated "noise", arising from the non-linearity of the interactions themselves. State-changes in this network would reflect both the input signals, and internally generated shifts of state, arising from changes in the damping coefficients, which in turn reflect the telencephalic control of brain-stem supervisory systems with diffuse effects, and multiplicative interactions within the telencephalon itself. The entire system would thus have the properties of a finite-state machine, which is perhaps the most fundamental attribute needed by any brain model with aspirations to theoretical adequacy (Arbib, 1964). Within this view, a large subsystem (the telencephalic surface) is both linear and observable via the EEG.

Exactly what is implied by this model for our conceptions of whole brain function depends heavily upon questions related to the true model order. As formulated in the initial paper (Wright and Kydd, 1984a), the true model order might be many millions, since for coupled oscillators there are generally as many resonant modes as there are coupled units. We considered these resonant modes to have natural frequencies clustered tightly about certain center frequencies, as a consequence of the architectural orderliness and modular structure of the telencephalon, and this feature permitted approximation of electrocortical activity by only a few modes of resonance. Present results support this, even at a very low model order. The degree to which such modes are functionally distinct is crucial to their significance in cerebral information processing. If modes of closely similar natural frequency are not functionally distinct, then the state-space of the envisaged finite-state machine would be limited, and this model would describe a broad form of control of cerebral information processing, in which other features of cortical neurone interaction performed the detailed information processing under supervision via control of the overall linear waves. On the other hand, if the millions of modes were functionally distinct, then the possible state space could be of very great complexity, and of itself could represent much of the brain's total information processing, without reliance upon the exact local interactions among neurones. Arbib (1964) in discussing Greene's (1962) treatment of resonant phenomena in neural nets, highlights the potential for information encoding implicit in resonance, as opposed to step-by-step state-changes in completely stable networks, and points out the contrast to classic network theories. From a more biological perspective, Pribram has dealt extensively with the virtues of a shift from preoccupation with individual neurones to consideration of network properties in the frequency domain, if the redundancy and regional non-specificity of cortical information storage and processing is to be explained (Pribram, 1972). The present findings seem supportive of these views, which have been greatly limited in applicability by the absence of a sufficient experimental framework within which to test them. A practical procedural advantage may arise from the present results, relating a linear analysis of real EEG events to network properties. Hypothetical networks might then be studied with real EEG events and their behavioural correlates as two of the criteria to be met by a realistic simulation.

This approach would seem a considerable simplification from conventional expectation concerning the brain, in which the complexity and
non-linearity of neural elements is often thought of as miliating against any practicable analysis of global information processing. Even so, non-linearities in interactions between telencephalon and brain-stem and within the multiplicative couplings, as well as systematic properties of the microscopic neural interactions (treated as simply noise-like in this model) should be expected. If there are many functionally distinct modes, then these could be only roughly approximated by existing recording and analytic techniques. These problems would greatly complicate any attempt to reduce whole-brain properties to those of a multiplicatively coupled resonant linear network, as suggested above. Nonetheless, in strictly circumscribed behavioural, cognitive and motivational conditions, it might be expected that analysis of the time-course and spatial conformation of activity at each major mode of resonance should go some way to revealing the system's changes of state, and hence a part of its next-state function.

We consider that a unifying framework within the domain of linear network and automata theory might be found for both motivation and attention related spectral changes and sensory evoked potentials – those relating to experimental manipulation of lateral hypothalamic fibres in particular, and possibly to a much larger class of phenomena.

Acknowledgements. This work was supported by the New Zealand Medical Research Council, through a grant to J. J. Wright and a research Fellowship to R. K. Kydd. We gratefully acknowledge the technical assistance of J. A. West, P. A. Briscoe, and A. R. Williamson.

References


Received: February 2, 1984

Prof. J. J. Wright
Department of Psychiatry
University of Auckland
School of Medicine
Private Bag
Auckland
New Zealand
Relationship between sensorimotor neglect and the specificity, degree and locus of mesotelencephalic dopaminergic cell loss following 6-hydroxydopamine

G. J. Lees, R. R. Kydd, and J. J. Wright

Department of Psychiatry, School of Medicine, University of Auckland, Auckland, New Zealand

Abstract. The involvement of cell groups within the dopaminergic mesotelencephalic system in the development of the sensorimotor neglect syndrome was re-evaluated in two ways. Firstly, dopaminergic specificity of the neglect was further established by studying the relationship between nomifensine protection of dopamine cells against 6-hydroxydopamine damage and the degree of neglect which resulted. The sensorimotor neglect syndrome which developed following injection of 6-hydroxydopamine was diminished by concomitant treatment with nomifensine in parallel with the degree of protection afforded the dopaminergic cells. Non-specific damage produced by 6-hydroxydopamine was unaltered by nomifensine. Secondly, the role in sensorimotor neglect of both total cell damage, and damage to regional sub-classes of dopaminergic cells was considered. It was found that the extent of the resulting neglect was correlated with the overall damage to the substantia nigra and ventral tegmental area, rather than to any individual region within this dopaminergic system. There was a threshold, involving destruction of approximately one third of the system, below which no neglect syndrome developed. Certain regions, including the ventral tegmental area (VTA), showed a higher partial correlation with the extent of neglect than other regions. While specific lesioning of the A8 or A10 dopaminergic neurons is probably insufficient to produce a neglect syndrome, damage to these areas potentiates the severity of the neglect produced by nigrostriatal lesions. It appears that the involvement of the individual subclasses of the mesotelencephalic dopaminergic neurons in the neglect syndrome is more widespread than previously thought.

Key words: Sensorimotor neglect — 6-Hydroxydopamine — Nomifensine — Nigrostriatal dopamine neurons — Mesolimbic dopamine neurons — Rats

Previous investigators have examined the role of dopamine neurons in the development of sensorimotor neglect by inducing unilateral lesions of the dopaminergic nigrostriatal pathway with 6-hydroxydopamine (Ungerstedt 1973; Ljungberg and Ungerstedt 1976; Siegfried and Bureš 1978; Marshall 1979; Marshall and Gotthelf 1979; Dunnett and Iversen 1982; Altar et al. 1984). Intracerebral injection of 6-hydroxydopamine, however, induces destruction of neurons containing other catecholamines, in addition to those containing dopamine (Kostrzewa and Jacobowitz 1974; Jonson 1980). Non-specific damage resulting from 6-hydroxydopamine injections also occurs to a variable extent (Poirier et al. 1972; Hokfelt and Ungerstedt 1973; Butcher et al. 1974; Javoy 1975; Javoy et al. 1976; Willis et al. 1976).

While various techniques have been employed to overcome these objections (Ljungberg and Ungerstedt 1973; Marshall 1979; Marshall and Gotthelf 1979), it is still not certain what part is played by non-specific cell damage in the immediate vicinity of the dopaminergic cell bodies or terminals. We first report a further selective method to control for the non-specific effects of 6-hydroxydopamine, by using nomifensine to prevent its selective uptake into dopamine cells. Nomifensine is a potent inhibitor of the uptake of dopamine and noradrenaline by nerve terminals and synaptosomes (Hunt et al. 1974; Kruse et al. 1977; Randrup and Braestrup 1977; Schacht et al. 1977; Tuomisto 1977) and has been used in several biochemical studies to protect catecholamine systems from the toxic effects of 6-hydroxydopamine (Samanin 1975; Broch 1979; Waddington 1980).

Earlier research has also endeavoured to define the role of sub-classes of dopaminergic cells in sensorimotor neglect (Ungerstedt 1973; Ljungberg and Ungerstedt 1976; Marshall 1979; Robbins and Everitt 1982). These used selective lesions, and results indicate that the laterally placed cells in the substantia nigra (with mainly caudate terminations) are more significant, while the medial cells of the ventral tegmental area, with their fronto-cortical and limbic terminations, appear unimportant. While association of the laterally placed fibres with sensorimotor neglect may accord well with an interpretation placing some emphasis on impaired motor responses (Robbins and Everitt 1982) it is harder to reconcile with our understanding of limbic and frontal neglect syndromes (Reeves and Hageman 1971; Heilman et al. 1973). It might be expected that destruction of cells afferent to these areas would also give rise to neglect consequent to functional disruption in the area of projection. The dopaminergic innervation of the nucleus accumbens, amygdala and cortex is mainly from neurons of the ventral tegmental area (Lindvall and Björklund 1982). Our second goal was to assess the possible contribution of cells in the ventral tegmental area to the severity of neglect.
Methods

Stereotaxic surgery. Female Wistar rats (190–210 g), bred within the animal laboratories of the Medical School, were housed in groups of up to five animals under a 12 h : 12 h diurnal light cycle. Animals had access to standard lab chow and water at all times. Groups of animals were treated with 2 μl 6-hydroxydopamine hydrogen bromide (2–8 μg free base), infused intracerebrally via a Hamilton syringe at a rate of 0.4 μl/min into the rostral extent of the substantia nigra [co-ordinates A3, L1.8, V-2, atlas of König and Klippel (1963)]. The 6-hydroxydopamine was dissolved in phosphate-buffered saline (pH 7) containing 0.2% ascorbic acid. The effect of nomifensine on this treatment was examined by pretreating groups of rats intraperitoneally with nomifensine maleate (6.7 or 20 mg/kg, calculated as free base) 30–45 min prior to injection of 6-hydroxydopamine. The nomifensine was dissolved in dimethyl sulphoxide (0.2 ml injected). Control animals were injected intracerebrally with vehicle. Each experimental group contained between three and ten animals (Table 1).

Histochemistry. Two or ten days following treatment the animals were subjected to histochemical evaluation in order to determine the extent of the 6-hydroxydopamine-induced lesion. Catecholamine-containing cells were visualized by the FAGLU method of Furness et al. (1978). In brief, the animals were perfused via the thoracic aorta with approximately 100 ml 0.9% saline followed by approximately 200 ml of a mixture of 4% formaldehyde and 0.5% glutaraldehyde in 0.1 M phosphate buffer, pH 7. The brain was removed and allowed to soak overnight in the FAGLU mixture before being frozen-sectioned on a Jung H410 sliding microtome at a 30 μm spacing. Sections were mounted on microscope slides, dried over phosphorus pentoxide for at least 1 h and mounted in DPX-mountant. Catechol-amine-induced fluorescence was visualized under a Wild M20 transmitted light fluorescent microscope fitted with a 200 watt mercury lamp, a BG23 red absorbing filter, a FITC exciting filter (emission 400–500 nm) and a FITC barrier filter (light transmission above 510 nm). The distribution of catecholamine containing fluorescent cells in the midbrain (Fig. 1) shows a good correlation with the distribution of dopamine cells as detailed by other investigators (Jacobowitz and Palkovits 1974; Palkovits and Jacobowitz 1974; Fallon and Moore 1978; Guyenet and Crane 1981). The extent of the ventral tegmental area was taken from the data of Palkovits and Jacobowitz (1974), and also as the difference between the total catecholamine cell population and those cells labelled following injections of horse radish peroxidase into the striatum (Fallon and Mehler 1978). Alternate sections were stained with thionin (Rucker-Kothan procedure, Skinner 1971) and used to confirm the extent of damage to the dopaminergic cells, the electrode placement and the specificity of the lesion (particularly with respect to the extent of microglial infiltration).

Estimation of cellular degeneration. The presence of yellow fluorescing cells, or the absence of all fluorescing cells in areas normally innervated by dopaminergic cells, was taken as an indication of cellular death. Histological patterns of cellular degeneration were mapped over the whole rostro-caudal extent of the midbrain dopaminergic system (cell groups A8, A9 and A10). A semiquantitative assessment of cellular damage was made from these maps at four planes, A1.3, A1.8, A2.2, and A2.6. At these planes, damage to the lateral and medial substantia nigra and (where present in these planes) to the ventral tegmental area was assessed on a scale of 0–3 (0, no damage; 1, 10–100% cell loss; 2, 10–90% cell loss; 3, 90–100% cell loss). The 11 areas were arithmetically summed in nine different combinations (Table 2) to give an estimate of damage to individual areas of the midbrain dopaminergic system.

The maximal lateral, vertical and rostro-caudal extent of the zone of architectural disorganisation and glial cell infiltration evident on thionin stained sections was used as a measure of non-specific damage.

Sensorimotor neglect. Sensorimotor neglect was assessed by an experimenter blind to the site (and nature) of the lesions and took place 3–4 days post-lesion. Testing was per-

Table 1. Experimental groups

<table>
<thead>
<tr>
<th>Pretreatment with nomifensine (mg/kg)</th>
<th>Number of animals injected intracerebrally with</th>
<th>2 μg 6-OHDA</th>
<th>8 μg 6-OHDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle</td>
<td>3 (0)</td>
<td>10 (10)</td>
<td>88 (6)</td>
</tr>
<tr>
<td></td>
<td>6.7</td>
<td>9 (0)</td>
<td>9 (2)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>5 (0)</td>
<td>9 (0)</td>
</tr>
</tbody>
</table>

Animals sacrificed at 2 days. These were not tested for neglect although brain sections were examined histochemically. All other animals were tested for neglect 4 days following treatment, and sacrificed at 10 days for brain histochemical evaluation. Animals were pretreated with nomifensine at the doses indicated 30–45 min prior to intracerebral injection. Values in parentheses are the number of animals in which lesions of the dorsal noradrenergic bundle were apparent.

Table 2. Division of the substantia nigra and ventral tegmental area into midbrain regions

<table>
<thead>
<tr>
<th>Rostro-caudal co-ordinate</th>
<th>Brain area</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2.6</td>
<td>A</td>
</tr>
<tr>
<td>A2.2</td>
<td>C</td>
</tr>
<tr>
<td>A1.8</td>
<td>F</td>
</tr>
<tr>
<td>A1.3</td>
<td>I</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approx. lateral co-ordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.0–1.7 mm</td>
</tr>
<tr>
<td>1.7–0.8 mm</td>
</tr>
</tbody>
</table>

Division of the substantia nigra and ventral tegmental area into regions was based on the co-ordinates of König and Klippel (1963). Fluorescent histochemical estimates of dopaminergic cellular degeneration (see Methods) were made at A2.6, A2.2, A1.8, and A1.3. The above areas were combined into regions as follows.

Restral = A + B + C + D, caudal = F + G + I + J, lateral = A + C + F + I, medial = B + D + G + J, VTA = E + H + K.

On this basis, the groups of dopamine perikarya as defined by Dahlström and Fuxe (1965) are: A8 = 1 + 3; A9 = A + B + C + D + F + G; A10 = E + H + K.
Fig. 1a–d. Schematic representation of the protection afforded by nomifensine against 6-hydroxydopamine-induced loss of dopamine cells in the substantia nigra. 6-Hydroxydopamine was injected intracerebrally into the most rostro-lateral extent of the substantia nigra (A3, L1.8, V-2) of rats pretreated with nomifensine (IP). Ten days subsequently, the animals were sacrificed and the dopamine cells visualized by aldehyde-induced fluorescence as described in Methods. Representative sections from nine areas of the substantia nigra are shown. ● normal cells; ○ former dopamine cells showing yellow fluorescence; // non-specific damage with no dopamine cells visible. Areas have been left blank when histofluorescent evidence of dopamine cells could not be detected, in spite of normal innervation on the contralateral side. a normal innervation, no drug. b 8 µg 6-hydroxydopamine. c 8 µg 6-hydroxydopamine, pretreatment with nomifensine (6.7 mg/kg IP). d 8 µg 6-hydroxydopamine, pretreatment with nomifensine (20 mg/kg IP)

formed in a 100 × 100 cm open square cardboard box with 20 cm sides. The technique used was a modification of that described by Ljungberg and Ungerstedt (1976) for assessing somatosensory orientation and a total of seven separate tests were performed. Any tendency to turn preferentially in one direction was first noted. Sensory stimuli were then applied to each side of the body with random alternation of side. Visual orientation was tested by moving a 2 × 2 cm square of white cardboard into the animal’s visual field and olfaction by moving an xylene soaked cotton bud close to each nostril. Touch orientation testing utilised a cotton bud, and a von Frey hair of 6 g pressure applied to various parts of the body. A stick was touched to the head just behind the mouth to assess the biting response and the sensitivity of the vibrissae was assessed by stroking them from behind forwards. The precision with which the animal could localize and respond to the applied stimulus was used, together with the turning behaviour, to estimate the degree of neglect as follows: 3 = full neglect; marked circling to side of lesion, little or no response in any sensory modality contralateral to lesion; 2 = partial neglect; circling to side of lesion, awareness of stimuli to contralateral side as evidenced by squeaking or movement but no head orientation; 1 = mild neglect; preferential turning to side of lesion, contralateral head orientation in direction of applied stimulus but no localization; 0 = no neglect; equal turning to both sides, precise head orientation to site of stimulus plus sniffing, biting or scratching.
**Statistical analysis of data.** Non-parametric correlation (Kendall's tau) was used to establish the relations between
(a) Regional cell damage to dopaminergic cells, and neglect grading.
(b) Regional cell damage and total cell damage, for dopaminergic cells.
(c) Cannular co-ordinates and total cell damage to dopaminergic cells.
(d) Cannular co-ordinates and degree of neglect.
(e) Partial correlation (partial tau) was then used to isolate the correlations of cell damage in specific regions with

sensorimotor neglect grading, with total damage as the controlled variable.

**Results**

**Histochecmistry.** Extensive loss of the mesotelencephalic dopamine neurons occurred after the injection of either 2 or 8 µg 6-hydroxydopamine intracerebrally into the substantia nigra (Figs. 1 and 2). Three categories of damage to the nigral cells could be detected. Normal cells showed a bright green fluorescence, and dark staining of cell bodies (Nissl staining) with thionin. Within 2 days after a high dose (8 µg) of 6-hydroxydopamine, the degenerating catecholaminergic cells developed a bright yellow fluorescence, as previously described by Hokfelt and Ungerstedt (1973). A proportion of yellow cells in this category still showed dark Nissl staining with thionin. In the second category (as exemplified by treatment with 2 µg 6-hydroxydopamine) the thionin staining of cells was lost within 10 days, although the cells retained their bright yellow fluorescence. In the third category, dopamine cells treated with 8 µg 6-hydroxydopamine lost all fluorescence within 10 days and failed to stain with thionin. This last stage presumably corresponded to phagocytosis of the dead cells.

Lesioning of the dorsal noradrenergic bundle, visualized by the presence of swollen bright green fluorescence axons, was a common feature in animals treated with 6-hydroxydopamine alone (22 out of 24 animals, Table 1). The noradrenergic cell bodies in the locus coeruleus did not show any corresponding loss of catecholamine fluorescence when examined in three animals showing lesions of the noradrenergic bundle. These results are in agreement with

---

**Fig. 3.** Correlation between total dopamine cell degeneration, as estimated by histofluorescence, and the extent of neglect shown by the animals. A variable degree of dopamine cell loss was obtained by the use of two doses of 6-hydroxydopamine in association with a varying degree of protection by nomifensine. 6-Hydroxydopamine (2 µg) injected intracerebrally into animals without pretreatment (●), or 30–45 min following nomifensine (IP) at doses of 6.7 mg/kg (△), or 20 mg/kg (■). 6-Hydroxydopamine (8 µg) injected intracerebrally into animals without pretreatment (◇), or 30–45 min following nomifensine (IP) at doses of 6.7 mg/kg (◇), or 20 mg/kg (□). Neglect testing was carried out on day 3 or 4, histofluorescence on day 10 or 11 following 6-hydroxydopamine. Values given are the mean ± SEM for groups of five to ten animals.
those of Ungerstedt (1973) who demonstrated that noradrenergic cells in the locus coeruleus do not undergo retrograde degeneration within this time period following 6-hydroxydopamine-induced lesions of the noradrenergic bundle. There was no evidence for a loss of neurons in the A7 or A11 catecholamine cell groups.

6-Hydroxydopamine (8 μg) produced an area of non-specific damage measuring on average 1.12 mm vertically, 0.84 mm laterally and 1.17 mm rostro-caudally round the cannula tract. Pretreatment with nomifensine (20 mg/kg) did not significantly alter this area (0.96 mm vertically, 0.91 mm laterally and 1.12 mm rostro-caudally). Non-specific damage after 2 μg 6-hydroxydopamine was generally confined to the cannula tract.

The degree to which specific loss of the dopamine cells was attenuated by pretreatment with nomifensine (6.7 or 20 mg/kg) depended on the relative concentrations of nomifensine and 6-hydroxydopamine (Fig. 3). Nomifensine protection was almost complete at a dose of 20 mg/kg when 2 μg 6-hydroxydopamine was injected intracerebrally. However, this dose of nomifensine failed to prevent a substantial loss of dopamine neurons when the amount of 6-hydroxydopamine was increased to 8 μg. On the other hand, nomifensine at either 6.7 mg/kg or 20 mg/kg completely protected the noradrenergic axons from 6-hydroxydopamine-induced damage (40 out of 43 animals).

No damage to the noradrenergic fibres or to any group of dopamine cells occurred in animals pretreated intraperitoneally with saline or nomifensine (20 mg/kg) and injected intracerebrally with vehicle.

Relationship between the extent of degeneration of the dopaminergic system and the degree of sensorimotor neglect. Of the 16 rats injected with either 2 μg or 8 μg 6-hydroxydopamine, 14 developed neglect syndromes within 3–4 days. In 18 and 12 animalspretreated with nomifensine (6.7 mg/kg and 20 mg/kg respectively) before injection of 6-hydroxydopamine, only 10 and 4 animals respectively developed neglect. 14 animals which showed neglect on days 3–4 were retested at 10 days. The average neglect score decreased by 2.39 ± 0.25 (SEM) to 1.70 ± 0.35 (SEM). The decrease was not significant (Mann-Whitney U).

The relationship between the degree of neglect exhibited by the animals and the extent of dopaminergic cell degeneration as estimated semiquantitatively was then examined. Animals with extensive damage to the nigra developed neglect syndromes similar to that described by Ljungberg and Ungerstedt (1976), while those with slight or no damage did not (Fig. 3). There was a damage threshold (a damage score of less than 9, of a possible total of 33) below which no neglect could be detected. This is equivalent to damage of approximately one third of the dopaminergic neurons on the side of the lesion. Above this threshold, the degree of global neglect was directly proportional to the extent of damage to the dopaminergic system. Groups of animals pretreated with nomifensine (6.7 or 20 mg/kg) showed a reduced level of neglect following intracerebral injection of 6-hydroxydopamine in direct proportion to the amount of protection afforded the dopaminergic cells.

The statistical analysis showed a high correlation between regional cell damage in any area, and sensori motor neglect (Table 3). No region showed a manifestly higher weighting by this direct correlation, and indeed, all regions showed a higher correlation of damage score with total damage, than with neglect. It appears then that total damage is the dominant variable in determining degree of neglect. In accord with this, partial tau statistics show lower values when regional damage is correlated with neglect, after the effect of total damage is removed. In spite of this, the partial tau values are heterogeneous, and indicate that a higher value for neglect correlation is found for cell damage in the VTA, for the region comprising the A8 group of cells, and the most rostral extension of the substantia nigra, than for other cells more laterally placed (Table 3).

There was much less correlation between the co-ordinates of the placement of the cannula in either anterior–posterior or lateral directions and the resulting total cellular damage or neglect (Table 3). There was no significant correlation between cell damage in any individual area or

<table>
<thead>
<tr>
<th>Anatomical variable</th>
<th>( r_{xy} )</th>
<th>( r_{xz} )</th>
<th>( r_{yz} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regional cell damage to midbrain</td>
<td>0.673*</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>A9 cell group (A + B + C + D + F + G)</td>
<td>0.623*</td>
<td>0.885*</td>
<td>0.080</td>
</tr>
<tr>
<td>A10 cell group (E + H + K)</td>
<td>0.724*</td>
<td>0.809*</td>
<td>0.414</td>
</tr>
<tr>
<td>A8 cell group (I + J)</td>
<td>0.692*</td>
<td>0.781*</td>
<td>0.359</td>
</tr>
<tr>
<td>Mesolimbic cell group (E + H + K + I + J)</td>
<td>0.714*</td>
<td>0.869*</td>
<td>0.352</td>
</tr>
<tr>
<td>Rostral (A + B + C + D)</td>
<td>0.651*</td>
<td>0.835*</td>
<td>0.218</td>
</tr>
<tr>
<td>Caudal (F + G + I + J)</td>
<td>0.642*</td>
<td>0.851*</td>
<td>0.179</td>
</tr>
<tr>
<td>Medial (B + D + G + J)</td>
<td>0.652*</td>
<td>0.895*</td>
<td>0.150</td>
</tr>
<tr>
<td>Lateral (A + C + F + I)</td>
<td>0.630*</td>
<td>0.810*</td>
<td>0.195</td>
</tr>
<tr>
<td>Rostralmedial (B + D)</td>
<td>0.659*</td>
<td>0.855*</td>
<td>0.238</td>
</tr>
<tr>
<td>Rostrolateral (A + C)</td>
<td>0.454*</td>
<td>0.660*</td>
<td>0.018</td>
</tr>
<tr>
<td>Caudomedial (G + J)</td>
<td>0.635*</td>
<td>0.816*</td>
<td>0.185</td>
</tr>
<tr>
<td>Caudolateral (F + I)</td>
<td>0.596*</td>
<td>0.748*</td>
<td>0.196</td>
</tr>
<tr>
<td>Rostral (1) (A + B)</td>
<td>0.699*</td>
<td>0.779*</td>
<td>0.377</td>
</tr>
<tr>
<td>Rostral (2) (C + D)</td>
<td>0.579*</td>
<td>0.804*</td>
<td>0.087</td>
</tr>
<tr>
<td>Caudal (1) (F + G)</td>
<td>0.561*</td>
<td>0.802*</td>
<td>0.049</td>
</tr>
<tr>
<td>A</td>
<td>0.524*</td>
<td>0.634*</td>
<td>0.171</td>
</tr>
<tr>
<td>B</td>
<td>0.705*</td>
<td>0.786*</td>
<td>0.385</td>
</tr>
<tr>
<td>C</td>
<td>0.457*</td>
<td>0.651*</td>
<td>0.033</td>
</tr>
<tr>
<td>D</td>
<td>0.535*</td>
<td>0.755*</td>
<td>0.055</td>
</tr>
<tr>
<td>E</td>
<td>0.711*</td>
<td>0.789*</td>
<td>0.397</td>
</tr>
<tr>
<td>F</td>
<td>0.403*</td>
<td>0.617*</td>
<td>0.022</td>
</tr>
<tr>
<td>G</td>
<td>0.560*</td>
<td>0.810*</td>
<td>0.054</td>
</tr>
<tr>
<td>H</td>
<td>0.602*</td>
<td>0.638*</td>
<td>0.303</td>
</tr>
<tr>
<td>I</td>
<td>0.655*</td>
<td>0.727*</td>
<td>0.326</td>
</tr>
<tr>
<td>J</td>
<td>0.710*</td>
<td>0.727*</td>
<td>0.434</td>
</tr>
<tr>
<td>K</td>
<td>0.417*</td>
<td>0.392*</td>
<td>0.225</td>
</tr>
</tbody>
</table>

* \( P < 0.001 \)
† \( P < 0.03 \)

\( r_{xy} \) = correlation of neglect and anatomical variable
\( r_{xz} \) = correlation of anatomical variable with total damage
\( r_{yz} \) = partial correlation neglect and anatomical variable, independent of total damage

For definition of areas A to K see Table 2.
region and the cannula co-ordinates (results not shown). Actual cannula placement for 46 animals was at A2.5 ± 0.4 mm (SD) and 1.8 ± 0.2 mm (SD) laterally.

Control animals (n = 6) given vehicle alone or vehicle subsequent to nomifensine pretreatment did not develop neglect.

**Discussion**

**Specificity of the dopaminergic system in relation to neglect.** The current results support earlier evidence that the dopaminergic cells of the nigrostriatal tract must be extensively lesioned before a sensorimotor neglect syndrome becomes apparent (Ljungberg and Ungerstedt 1973; Marshall 1979). Marshall (1979) showed that animals with a 50% depletion of striatal dopamine had only minimal or no sensorimotor deficits. Our results using histochemical evaluation indicate that at least one third of the dopaminergic system must degenerate before a significant neglect syndrome develops (Fig. 3). The correlation obtained demonstrates that semiquantitative histochemical analysis can adequately predict behavioural deficits. The major advantage obtained from using histochemical rather than biochemical estimates of cellular degeneration is that a more precise anatomical mapping of the damaged areas, including lesions of the noradrenergic bundles, can be obtained. In addition, the extent of non-specific damage can be estimated. Our results show that non-specific damage can be extensive after 8 μg 6-hydroxydopamine is injected intracerebrally. Non-specific damage was minimal after 2 μg 6-hydroxydopamine. Since the degree of degeneration of dopaminergic cells is equally extensive with either dose of 6-hydroxydopamine (Figs. 1 and 2), the use of 8 μg or higher amounts in rat behavioural studies is not justified, provided the 6-hydroxydopamine is adequately protected against chemical oxidation. The non-specific damage was not significantly altered by nomifensine. Hence, the results demonstrate that non-specific damage produced by 6-hydroxydopamine is not a necessary factor in producing sensorimotor neglect.

The ascending noradrenergic system arising from the locus coeruleus and passing through the midbrain as the dorsal noradrenergic bundle also does not appear to play an essential role in the development of sensorimotor neglect. Nomifensine at either 6.7 mg/kg or 20 mg/kg protected the dorsal noradrenergic bundle in 29 of 30 animals examined for neglect. The same correlation between degree of neglect and loss of dopamine cells was obtained as in untreated animals (Fig. 3). These results thus confirm those of Ljungberg and Ungerstedt (1976) and Marshall (1979). It is still possible that noradrenergic innervation might be involved, although to a minor extent.

**Relationship between regional damage to the dopaminergic system and neglect.** This study used differing concentrations of nomifensine and 6-hydroxydopamine to induce a variable degree of specific lesion to the dopamine neurons. As a result, it was found that the best predictor of the degree of neglect produced by 6-hydroxydopamine was the global extent of damage to the dopaminergic system. After minimizing the effect of global damage by using Kendall partial tau correlations, dopaminergic cell degeneration in any region of this system shows a poorer association with neglect. Moreover, these studies indicate that cellular degeneration in the A8 and A10 cell groups contribute to the neglect, in addition to the A9 damage. The higher partial correlations with neglect shown by A8 and A10 groups confirm their importance, but do not necessarily indicate a greater functional weight, since A9 damage was ubiquitous, and A8 and A10 damage was superordinate to A9 damage. It is of interest to note in this regard that the A8 cell group can be regarded anatomically as being mainly a part of the mesolimbic/mesocortical dopaminergic pathway in the rat (Nauta et al. 1978; Lenard and Nauta 1979).

Further studies, in which dopaminergic damage is induced over more variable fields of cells would enable true functional weighting of the various cell groupings to be deduced by partial correlation. The prior emphasis on the nigral dopamine cells (A9) (Ljungberg and Ungerstedt 1973; Marshall 1979; Altar et al. 1984), appears due to the fact that there is an insufficient mass of dopamine cells in the VTA alone to produce sensorimotor neglect on lesion, without an additional involvement also of part of the nigrostriatal system. Thus injections of 6-hydroxydopamine directly into the limbic or cortical dopamine terminal regions alone fail to cause sensorimotor neglect (Marshall et al. 1980). In addition, mesolimbic dopaminergic cells are more resistant to the effects of 6-hydroxydopamine (Ungerstedt 1973) and hence even injections of 6-hydroxydopamine directly into the VTA result in a greater loss of dopamine in the striatum (both absolutely and as a percentage of normal) than in the nucleus accumbens (Marshall 1979).

On the other hand, the argument could be made that damage to the VTA is likely to be associated with lesioning of the nigro striatal dopaminergic axons which ascend in this area. Against this is the lack of correlation between the lateral placement of the cannula (as determined by histology) and the degree of neglect. In addition, the high degree of correlation between total dopaminergic cellular damage and the neglect score (Fig. 3) suggests that any lesioning of the axons close to the cell bodies (as in these experiments) results in retrograde degeneration of the perikarya. In support of this interpretation it has been shown that degeneration occurs later in caudal than in rostral regions of the nigra (Lees et al. 1984). Similarly, damage to the A8 cell bodies is contributory to the development of the neglect, and likewise this would not be apparent if axonal damage alone had been inflicted on these cells.

A different approach has been used by Dunnett and Iversen (1982). These workers have shown that localized 6-hydroxydopamine-induced lesions of the dopaminergic terminal fields in the striatum induce varying degrees of intensity and permanency of neglect. The advantage of this approach is in defining which of the subsequent connections between the striatum and other parts of the brain may be preferentially involved in the development of neglect. Their results indicate that the relationship of dopaminergic damage to neglect is not wholly described by “mass action” concept of their function. On the other hand, this approach is less effective in determining which dopaminergic cells are involved in the syndrome, due to the overlapping projections of the A9 and A10 groups of dopamine neurons (Lindvall and Björklund 1982). Dunnett and Iversen (1982) have identified a ventrolateral region in the striatum which is particularly associated with the development of neglect.
This region is sparsely innervated by A10 dopamine neurons (Lindwall and Björklund 1982). However, our results indicate that lesions of the A8 and A10 dopamine cells contribute to the development of the syndrome. This suggests that the areas to which these cells project preferentially should, on lesion, prove to potentiate the severity of the neglect syndrome produced by lesions of the A9 cells or caudate nucleus. As the neglect detected by Dunnett and Iversen (1982) was relatively mild (less than half the maximum score possible) such potentiation should be readily observable.

Neglect arising from critical area damage as described by Dunnett and Iversen (1982) might involve different mechanisms to the contributions of other sites. These effects might sum in a variety of ways to give the increasing manifest neglect we have shown to occur with increasing damage.

Acknowledgements. This research was supported by a grant to JIW from the Medical Research Council of New Zealand, and a Fellowship to RRK. The technical help of Mr R. Horsham, Ms P. Briscoe and Ms R. Williamson is gratefully acknowledged.

References


Dahlstrom A, Fuxe K (1965) Evidence for the existence of monoamine neurons in the central nervous system. II. Experimentally induced changes in the intraneuronal amine levels of bulbopinal neuron systems. Acta Physiol Scand (Suppl 64) 247: 1–80


Received April 19, 1983; Final version August 21, 1984
Quantitation of a Mass Action of Dopaminergic Neurones Regulating Temporal Damping of Linear Electro cortical Waves

J. J. Wright, R. R. Kydd, and G. J. Lees
Department of Psychiatry, University of Auckland School of Medicine, Auckland, New Zealand

Abstract. We have previously proposed that electro cortical waves are linear waves, subject to regulation by mesotelencephalic dopaminergic neurones. As a further means to test this theory, selective unilateral lesions of varying extent were made in the nuclei of origin of the dopaminergic mesotelencephalic tract. Changes in the electro cortical power spectrum were assessed by a repeated measure, between hemispheres comparison of ratio changes in power. With increasing unilateral dopamine cell damage, the animals showed increasing contralateral sensorimotor neglect. Curve fitting the ratio changes in power attributable to lesion, showed that estimates of the power of driving signals and the temporal damping moved in reverse directions with increasing extent of lesion, as expected from the theory. A further test was undertaken, to determine whether equal estimates for a transformation of surface signals were obtained from each side. Equality would not be expected if the equation for relative power were invalid. Left and right equality was found for three grades of unilateral lesion.

1 Introduction

The electrocorticogram (EEG) reflects global activity of the brain and can be related to mental state (Walter et al. 1967; Berkhout et al. 1969; Gevins et al. 1979). This observation raises a series of questions concerning the cell systems and mechanism governing this relationship. A cell system of considerable interest to the control of overall brain function is the mesotelencephalic tract. Dopaminergic neurones in the mesotelencephalic tract project from the brain-stem to limbic, striate and frontal cortical regions of the forebrain (Lindvall and Bjorklund 1974). Manipulation of these neurones influences cognition, motivation and attention. For example, lesion of the dopaminergic mesotelencephalic tracts gives rise to an inattentive syndrome (sensorimotor neglect) (Ljungberg and Ungerstedt 1976; Marshall et al. 1971; Marshall 1979) while stimulation of these pathways causes non-specific motivation (intracranial self-stimulation) (Valenstein et al. 1970; Wise and Bozarth 1981). Findings to date relating dopaminergic activity to the EEG have been equivocal, partly because of inadequate control for nonspecific effects and a failure to manipulate dopaminergic pathways alone. These studies appear in contradiction about whether power increases, decreases, or moves to lower frequencies with dopaminergic manipulation. (Neale and Keane 1980; Wright and Craggs 1979; Hansen and Whishaw 1973; DeRyck and Teitelbaum 1978; Jones et al. 1973; Whishaw et al. 1978; Krieglstein et al. 1979; Monti 1982; Rouqel 1982). The absence of an adequate theoretical description with which to interpret findings may also have increased the difficulties. Our present intention is to explain the role of dopaminergic neurones in regulating the EEG as one case of a more general principle of EEG control.

Earlier work from our group has aimed at determining how brain waves and their control can be correctly paramatised (Wright and Kydd 1984a). We have advanced a model in which electrocortical waves obey simple linear laws akin to waves in an elastic medium subject to closed boundary conditions. Driving signals and temporal damping are the controlling parameters influenced by certain centrifugal pathways. While this treatment is very simple in principle, the system will exhibit very many resonant modes, making practical analysis still a very difficult proposition. A summary of this preceding model and prior tests for its validity, is now given, before a further test is proposed and results reported. The further test involves quantitation of the effect of dopaminergic neurones upon the EEG, and enables a reconciliation of previous disparate results in this field. The following review is also
relevant to the first following paper (Wright et al. 1985b) and the model itself is further developed in the final paper of this series (Wright et al. 1985c).

1.1 A Linear Wave Model

We consider the telencephalon as a mass of linked nonlinear oscillatory circuits. These circuits are considered to each have cycle times and coupling coefficients which vary continuously about certain mean values specific to each circuit, but are perturbed by such complicated interactions that the instantaneous cycle times and coupling coefficients are effectively stochastically independent. The electrocorticogram is considered as the transformed local spatial average of dendritic potentials generated by the group actions of the system (Elul 1972). A state-transition matrix is developed, and the system characteristic equation for steady-state is obtained from this. The Central Limit theorem applies to the values of the coefficients in this characteristic equation, which tend near constant values of low variance, since they are sums of very many stochastically independent instantaneous values. The group properties of electrocortical waves will thus exhibit linear characteristics -- i.e., the waves are subject to the superposition rule, and must have definite phase velocities for waves of each temporal frequency, at each site on the cortex. Within the closed telencephalic system, resonant wave patterns will then develop at specific frequencies. While millions of specific resonant patterns are possible, consideration of the anatomical orderlines of the system requires that these resonant modes will have natural frequencies clustered about certain principle values.

The ascending lateral hypothalamic pathways (including the dopaminergic pathways) are ascribed two roles -- perturbation of the forebrain resonant systems, akin to a noise-like driving signal, and regulation of temporal damping of the resonant modes.

1.2 Methods of Testing the Models Validity

The very high model order, the time-varying nature of damping coefficients, the complicated surface in signal transformation, and difficulties in determining exact input and output relations make testing of this model problematic. Our approach has been to eliminate some of the difficulties by working with time-averaged EEG spectra, obtained from symmetrical sites, before and after unilateral lesion of supposed controlling pathways ascending through the lateral hypothalamus. The model is then tested via a series of hypothetico-deductive inferences, as follows.

1.2.1 Consistency of Natural Frequencies with Changing Damping. By experiment we can obtain \( G^2(\omega) \), the relative power, defined as

\[
G^2(\omega) = \frac{V_{LA}^2}{V_{LB}^2} \frac{V_{CA}^2}{V_{CB}^2} = \frac{V_{LA}^2}{V_{LB}^2} \left( \frac{V_{CA}^2}{V_{CB}^2} \right)
\]

where \( V^2(\omega) \) is an EEG power spectrum, and \( LA, LB, CA, CB \) represent the epochs before and after lesion, on the side of lesion and the control side. In theory, \( G^2(\omega) \) is given by

\[
G^3(\omega) = K \left[ \sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \Delta_i^2(\omega)^2} \right]^2 + \left[ \sum_{i=1}^{n} \frac{-\Delta_i \omega}{(M_i^2 - \omega^2)^2 + \Delta_i^2(\omega)^2} \right]^2 \]

(2)

where \( K \) is a scaling factor reflecting the changed power of noise-like driving signals, \( M_i \) are constant natural frequencies, and \( \Delta_i \) are corresponding damping coefficients for the \( i \)-th resonant mode on each side, following lesion.

The first test is to determine whether curve-fitting (2) to the experimentally derived (1), yields from the ratio power change, a set of \( M_i \) which are comparatively constant, and of values typical of the frequencies of electrocortical rhythms.

1.2.2 Equality of Electrode Transfer Characteristics. Equation (2) is theoretically derived from expressions of the form

\[
V_{LA}(\omega) = K_1 |A(\omega)|_{LA}^2 \left[ \sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \Delta_i^2(\omega)^2} \right]^2 + \left[ \sum_{i=1}^{n} \frac{-\Delta_i \omega}{(M_i^2 - \omega^2)^2 + \Delta_i^2(\omega)^2} \right]^2,
\]

(3)

\[
V_{CA}(\omega) = K_2 |A(\omega)|_{CA}^2 \left[ \sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \Delta_i^2(\omega)^2} \right]^2 + \left[ \sum_{i=1}^{n} \frac{-\Delta_i \omega}{(M_i^2 - \omega^2)^2 + \Delta_i^2(\omega)^2} \right]^2.
\]

(4)

The equality of \( |A(\omega)|_{CA} \); \( |A(\omega)|_{LA} \) was assumed and thus these terms cancel to give (2).

But terms for the calculation of \( |A(\omega)|_{LA} \) and \( |A(\omega)|_{CA} \) are available after curve fitting as performed in 1.2.1, and from the post-lesion power spectra (after correction for simple effects of the recording system).
Thus, the second test is to determine if equal values for $|A(\omega)|$ left and right are obtained by calculation—i.e., to show that the model parameters for relative power ratios, and the absolute power spectra, are consistent.

1.2.3 Correspondece of Changes in Relative Amplitude and Relative Phase of Electrocoical Waves. The theoretical expression for $G^2(\omega)$ in (2) is composed of real and imaginary components from which the phase of signals can also be obtained. Thus, in steady-state conditions before and after lesion, where $G^2(\omega)$ is parametrised by (2) the relative changes in the phase of electrocortical signals, $\Delta \Phi(\omega)$ is predicted to be

$$\Delta \Phi(\omega) = \tan^{-1} \left[ \frac{\sum_{i=1}^{n} \frac{-\mathcal{D}_{1LA} \omega}{(M_i^2 - \omega^2)^2 + \mathcal{D}_{1LA} \omega^2}}{\sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \mathcal{D}_{1LA} \omega^2}} \right] - \tan^{-1} \left[ \frac{\sum_{i=1}^{n} \frac{-\mathcal{D}_{1CA} \omega}{(M_i^2 - \omega^2)^2 + \mathcal{D}_{1CA} \omega^2}}{\sum_{i=1}^{n} \frac{M_i^2 - \omega^2}{(M_i^2 - \omega^2)^2 + \mathcal{D}_{1CA} \omega^2}} \right].$$

(5)

As $\Delta \Phi(\omega)$ can be independently obtained from experiment, comparison of theoretical and experimental relative phase also tests the validity of the theory.

1.2.4 Means of Applying These Tests with a Model of Order Much Lower than Optimal. The number of resonant modes appearing in the summations within (2) would be very high, probably millions. Since $M_i$, the mode frequencies are expected to be clustered about a much smaller number of central frequencies, this lends the system to approximation at lower order. We have obtained this lower order approximation in two steps, where necessary. First, curve fitting of (2) to (1) has been performed with data from a number of different animals, and all parameters were allowed free during curve-fitting. By obtaining estimates of the natural frequency values of cluster centers, curve fitting can then be repeated with $M_i$ constant at the center values. Both stages of approximation yield results which are distorted in various ways.

Further practical problems include the difficulty of obtaining recordings in steady state conditions, low coherence of left and right EEG signals, relative insensitivity of the minimisation function to errors in damping coefficients, and lack of a clear criterion for global minimum during the curve-fitting.

Findings on animals which have undergone unilateral lateral hypothalamic lesion by electrolysis have confirmed the predictions of Sects. 1.2.1 and 1.2.2. Test of prediction 1.2.3 was not attempted in this group. (Wright and Kydd 1984a, b, c).

Utilising animals which had undergone selective unilateral lesion of dopaminergic neurones of the mesotelencephalic group, all three predictions were fulfilled, and where comparable, results were similar to the electrolytic lesion group. The phase/amplitude relation was confirmed in a split-group replication and limiting effects of model order were explored (Wright et al. 1984).

1.3 A Further Test, Quantitating the Massed Action of Dopaminergic Neurones in this Control Function

If the model under consideration is correct, then varying extents of selective unilateral lesion of mesotelencephalic dopaminergic neurones should cause varying degrees of release of resonant activity from temporal damping on the side of lesion, and variation in the strength of driving signals on the side of lesion. Specifically, with increasing unilateral dopaminergic damage, the mean ratio of $\mathcal{D}_{1LA}/\mathcal{D}_{1CA}$ across the ensemble of modes should decrease, while $K$, the factor reflecting relative strength of the driving signals on each side, should also decrease. These separate actions will have opposite effects upon the amplitude of the total EEG signal—a point which may explain the inconsistent results of earlier studies concerning dopaminergic effects upon the EEG.

The present paper reports quantification of this mass action effect of mesotelencephalic dopamine neurones. In addition, we have repeated the tests described above in 1.2.2 to show whether similar results are obtained with varying degrees of dopaminergic lesion.

2 Methods

2.1 Animal Recordings

The methods of animal surgery, recording, and data analysis were identical to those reported in detail in earlier papers, unless stated (Wright and Kydd 1984a, b, c; Wright et al. 1984).

Animals were seventeen young adult male Wistar rats with implanted bipolar extradural platinum ball electrodes situated bilaterally at P7.5, L4.5 and P9.5, L4.5 relative to bregma. Electrococticograms were obtained from animals in a condition of habituated wakefulness, using low noise amplification systems, 48 dB/octave filters (1-30 Hz band-pass), and analog-to-digital conversion of about 300 s of EEG at 8 ms interrupts, on two occasions of recording. The first occasion was 24 h before unilateral injection near the cells of origin of the mesotelencephalic dopaminergic pathway, the second 3–4 days later.
2.2 Selective Unilateral Dopaminergic Lesion

This injection was performed upon 13 animals, using the selective neurotoxin 6-hydroxydopamine (6-OHDA) (2 µgm in 2 µl of phosphate buffer with 0.2% ascorbic acid), at stereotaxic coordinates A3 L1.8 V−2.3 (from stereotaxic zero, Atlas of König and Klippel, side alternated between animals).

Four control animals underwent injection of carrier solution alone. Noradrenergic neurons were protected from 6-OHDA damage by desipramine HCl 25 mg/Kg given intraperitoneally.

2.3 Behavioural Assessment

On the day of the second EEG recording animals underwent assessment for contralateral sensorimotor neglect. This provided a behavioural measure of the amount of damage to dopamine neurons (Ljungberg and Ungerstedt 1976; Marshall 1979; Lees et al. 1985). The technique used for the assessment was a modification of the methods of Marshall (Marshall et al. 1971; Marshall 1979) which permits grading of the severity of neglect from + to ++++; Neglect was graded: ++++; marked circling to side of lesion, little or no response in any sensory modality contralateral to lesion. ++ = circling to side of lesion, awareness of stimuli to contralateral side evidenced by squeaking or movement but no head orientation. + = preferential turning to side of lesion, contralateral head orientation in direction of stimulus but no localised response. 0 = equal turning to both sides, precise orientation to site of stimulus, plus sniffing, biting or scratching.

2.4 Histological Assessment

Ten days after intracerebral injection animals were killed and underwent analysis by fluorescent histochemistry (Furness et al. 1978) and conventional histology. A semiquantitave assessment of dopamine cell damage was made at 4 stereotaxic planes, A1.3, A1.8, A2.2, and A2.6. Damage to equal areas in the substantia nigra (A9) and the mesolimbic and mesocortical ventral tegmental areas (A8 and A10) were each assessed on a scale of 0 to 3 (0 = no damage, 1 = 1–10% cell loss, 2 = 11–90% cell loss, 3 = 91%–100% cell loss), and the results for each animal were summed in total, and over the cell sub-groups to give a maximum damage score of 33. This technique was preferred to enzymic analysis because of the ability to precisely localize the damage to dopamine neurons, and to distinguish damage to other important pathways, most especially ascending noradrenergic fibres.

2.5 Spectral Analysis

Meanwhile, the electrocortical signals underwent Fast Fourier transformation in 512 point segments, and the segment power spectra were averaged over the relevant channel and epoch. These spectra were used to calculate $G^2(\omega)$ for each animal, in accord with Eq. (1).

2.6 Curve Fitting, and Subsequent Calculations

2.6.1 Curve Fitting. The estimate of $G^2(\omega)$ from each animal which had been injected with 6-OHDA was fitted to Eq. (2) with $n$ set at five, and optimisation performed under least squares criterion. All parameters were free. A space search from a wide variety of starting parameters was performed, and the best fit taken as a global minimum.

We then found the 5 values which represented cluster medians, about which the mean absolute deviations of all values were minimised (this is a slight change from earlier procedure, in which a less formal minimisation by least squares was used, but results were in fact closely similar).

Having obtained estimates of the cluster centre values for $M_i$, curve-fitting was repeated with these values of $M_i$ imposed, in every subjects case.

2.6.2 Quantifying the Mass Effect. Individual estimates of $G^2(\omega)$ obtained from Eq. (1) were pooled into four groups according to grade of sensorimotor neglect. These groups were ++++; ++; 0/++; and the control injection group.

The pooled group estimates of $G^2(\omega)$ were then curve fitted by the first (free parameter) method, to obtain model parameters for the group.

2.6.3 Calculation of Electrode Transfer Characteristics. Following earlier procedure, and utilising Eqs. (3) and (4), $|A(\omega)|_{LA}$ and $|A(\omega)|_{CA}$ were obtained using the second set of curve-fitted estimates. Appropriate normalisation procedures and correction for band-pass characteristics of the recording apparatus were used to correct post-lesion power spectra. Use of the curve fits with imposed centre $M_i$ enables the pattern of error in $|A(\omega)|$ to be more readily interpreted.

Average lesion and non-lesion $|A(\omega)|$ was found respectively for the ++++; ++; and 0/+ neglect groups.

3 Results

3.1 Results of Curve-Fitting

Parameters obtained from curve-fitting with free parameters are shown in Table 1.

The five median values for $M_i$ obtained in this manner were 2.33, 7.1, 10.77, 19.65, and 25.01 Hz.

Parameters obtained by repeating the curve-fitting with these values of $M_i$ imposed are given in Table 2.
Table 1. Neglect grade, Model parameters and goodness of fit to relative squared gain for each experimental animal (excluding animals without dopaminergic damage). Curve fitting accounts for the five principle resonant modes with all parameters free during minimisation.

<table>
<thead>
<tr>
<th>Animal</th>
<th>Natural Frequencies (Hz)</th>
<th>Corresponding Damping Coefficients (cycles/sec)</th>
<th>( K )</th>
<th>Goodness of fit, ( r^2 )</th>
<th>Neglect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lesion (Left) : Control (Right)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anton</td>
<td>7.09</td>
<td>1.736 : 0.408</td>
<td>0.396</td>
<td>0.655</td>
<td>+ + +</td>
</tr>
<tr>
<td></td>
<td>7.15</td>
<td>0.257 : 10.398</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.19</td>
<td>7.130 : 45.132</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.90</td>
<td>4.656 : 9.056</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.06</td>
<td>0.970 : 1.689</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joseph</td>
<td>6.53</td>
<td>1.494 : 48.830</td>
<td>0.16</td>
<td>0.889</td>
<td>+ + +</td>
</tr>
<tr>
<td></td>
<td>7.76</td>
<td>0.764 : 1.526</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.24</td>
<td>4.539 : 4.034</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.65</td>
<td>10.361 : 14.307</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26.05</td>
<td>4.371 : 5.607</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likka</td>
<td>3.54</td>
<td>1.33 : 0.911</td>
<td>0.880</td>
<td>0.553</td>
<td>+ + +</td>
</tr>
<tr>
<td></td>
<td>4.79</td>
<td>0.18 : 0.044</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.77</td>
<td>5.79 : 4.159</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.13</td>
<td>192.099 : 32.021</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23.08</td>
<td>0.564 : 0.385</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coney</td>
<td>2.58</td>
<td>0.163 : 0.265</td>
<td>0.777</td>
<td>0.273</td>
<td>+ + +</td>
</tr>
<tr>
<td></td>
<td>8.16</td>
<td>3.286 : 3.742</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.13</td>
<td>0.694 : 0.459</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23.22</td>
<td>1.775 : 2.136</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25.35</td>
<td>0.619 : 0.848</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merv</td>
<td>1.98</td>
<td>3.986 : 5.542</td>
<td>0.659</td>
<td>0.622</td>
<td>+ + +</td>
</tr>
<tr>
<td></td>
<td>6.47</td>
<td>4.290 : 1.394</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.56</td>
<td>0.990 : 18.363</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.17</td>
<td>0.724 : 1.168</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23.33</td>
<td>2.352 : 2.449</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aron</td>
<td>2.33</td>
<td>8.056 : 1.039</td>
<td>0.436</td>
<td>0.619</td>
<td>+ +</td>
</tr>
<tr>
<td></td>
<td>2.52</td>
<td>0.764 : 37.473</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.96</td>
<td>2.018 : 2.167</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.78</td>
<td>10.648 : 20.697</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25.01</td>
<td>4.256 : 4.992</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leon</td>
<td>7.34</td>
<td>22.328 : 13.370</td>
<td>0.517</td>
<td>0.766</td>
<td>+ +</td>
</tr>
<tr>
<td></td>
<td>9.08</td>
<td>4.316 : 170.216</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.14</td>
<td>0.761 : 1.362</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26.45</td>
<td>0.225 : 0.388</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27.92</td>
<td>12.030 : 36.306</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Richard</td>
<td>1.58</td>
<td>0.700 : 1.106</td>
<td>1.102</td>
<td>0.806</td>
<td>+ +</td>
</tr>
<tr>
<td></td>
<td>8.78</td>
<td>2.574 : 9.166</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.22</td>
<td>13.500 : 2.746</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.28</td>
<td>50.535 : 14.528</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25.07</td>
<td>150.797 : 21.323</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edgar</td>
<td>6.81</td>
<td>7.235 : 13.783</td>
<td>0.490</td>
<td>0.913</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>13.91</td>
<td>1.344 : 0.801</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.98</td>
<td>1.161 : 2.777</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24.87</td>
<td>1.168 : 0.954</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26.28</td>
<td>2.524 : 1.815</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duncan</td>
<td>2.37</td>
<td>2.164 : 5.041</td>
<td>0.388</td>
<td>0.884</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>4.75</td>
<td>2.704 : 8.735</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.97</td>
<td>2.364 : 6.607</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.38</td>
<td>2.105 : 3.532</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24.70</td>
<td>8.57 : 10.466</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1 (continued)

<table>
<thead>
<tr>
<th>Odysseus</th>
<th>Hector</th>
<th>Frank</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.474</td>
<td>4.965</td>
<td>3.051</td>
</tr>
<tr>
<td>3.560</td>
<td>8.132</td>
<td>4.414</td>
</tr>
<tr>
<td>9.272</td>
<td>10.522</td>
<td>6.253</td>
</tr>
<tr>
<td>13.294</td>
<td>22.207</td>
<td>17.079</td>
</tr>
<tr>
<td>13.483</td>
<td>25.710</td>
<td>25.321</td>
</tr>
<tr>
<td>1.571</td>
<td>5.110</td>
<td>0.837</td>
</tr>
<tr>
<td>0.769</td>
<td>4.644</td>
<td>1.766</td>
</tr>
<tr>
<td>8.518</td>
<td>0.807</td>
<td>2.205</td>
</tr>
<tr>
<td>2.541</td>
<td>0.593</td>
<td>19.447</td>
</tr>
<tr>
<td>14.606</td>
<td>0.370</td>
<td>11.566</td>
</tr>
<tr>
<td>3.288</td>
<td>10.073</td>
<td>12.474</td>
</tr>
<tr>
<td>0.9302</td>
<td>0.8617</td>
<td>0.5288</td>
</tr>
<tr>
<td>0.826</td>
<td>0.708</td>
<td>0.901</td>
</tr>
</tbody>
</table>

3.2 Unilateral Dopaminergic Damage and Sensorimotor Neglect

Table 3 gives the estimates of unilateral cell damage characterising each category, and sub-divided over specific cell groupings within mesotelencephalic dopaminergic cells. It can be seen that increasing degree of neglect reflects cell damage across all cell groups.

Fluorescent histochemistry revealed no damage to the noradrenergic bundles, and conventional histology showed that non-specific damage was minimal in all groups.

3.3 Difference in Relative Power in the Groups Graded by Severity of Neglect

Figure 1 shows the results of this exercise graphically, and the specific parameters are tabulated in the caption. Notice that both the experimental data points and the fitted curve have been re-scaled by the reciprocal of $K$, so that curvilinear trends are emphasised. The raw, unscaled group estimates do not indicate any strong trend in the data, which is revealed only by the results of the curve-fitting.

The changes with increasing damage appear to be an orderly decrement of two separate processes, having converse effects upon the gross amplitude of the EEG waves. That is, increasing cell damage was accompanied both by (i) a frequency independent fall in total power on the side of lesion (diminishing value of $K$, a measure of the strength of driving signals) and (ii) a frequency dependent change in power consistent with a progressive release from damping of various resonant modes - i.e., diminishing lesion/control damping ratios. With higher damage this release involves modes with natural frequencies (the $M_i$ values) across the EEG spectrum. With lesser damage the effect appears to predominate at low frequency modes.

The associated trends of neglect, $K$ (the relative driving power) and of the mean ratios of lesion/control damping, are highly significant ($p < 0.002$ for the joint probability of the permutations of neglect rating, and the values of $K$ and average ratio damping).

3.4 Electrode Transfer Characteristics for the Graded Neglect Groups

These are shown in Fig. 2. It can be seen that a close parity obtains for lesion and control estimates of this value within a neglect group, and indeed a similar linear trend is apparent between the severe and mild groups. These also are comparable trends to those obtained using electrolytic lesion.

All estimates show autocorrelated errors, introduced by the low order approximation of a much higher order phenomenon. These are less marked in the medium damage group, and this group shows the biggest lesion/control disparity in linear trend. It is also the group with the smallest number of animals. These features lead to the interpretation of the medium damage group as the most error prone estimate.

It will be seen that a minor, but consistent lesion/control disparity is seen across all groups, in that the linear trend has the higher slope on the control side. This is also seen in data from electrolytic lesion, and the later reported noradrenergic lesion group. This might indicate a consistent non-linear quality to the electrocortical waves, in the terms outlined above (1.2.2) for this test of the model. However, since this difference does not show any trend with severity of unilateral lesion (which should increasingly unmask the non-linearity) we conclude that this difference has arisen from some systematic error in the estimates of $D_b$ during the curve fitting to $G^2(\omega)$. 

WORK TOWARD A THEORY OF BRAIN FUNCTION | 109
Table 2. Model parameters and goodness of fit to relative squared gain for each animal when curve fitting is repeated using imposed values of five natural frequencies. The values imposed were the median values of the five principle groups found from free-parameter fitting (Table 1), i.e., $N_1=2.33$, $N_2=7.1$, $N_3=10.77$, $N_4=19.65$, $N_5=25.01$ Hz.

<table>
<thead>
<tr>
<th>Animal</th>
<th>Damping coefficients (cycles$^2$/s) side of lesion, $D_1$ to $D_3$</th>
<th>Damping coefficients (cycles$^2$/s) side opposite lesion $D_1$ and $D_3$</th>
<th>$K$</th>
<th>Goodness of fit, $r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anton</td>
<td>0.583, 38.612, 23.906, 9.258, 16.247</td>
<td>0.349, &gt; 1000, 1.753, &gt; 1000</td>
<td>8.144, 15.685</td>
<td>0.647, 0.550</td>
</tr>
<tr>
<td>Joseph</td>
<td>78.513, 1.208, &gt; 1000, 17.973, 13.965</td>
<td>17.695, 6.044, &gt; 1000,</td>
<td>11.139, 8.341</td>
<td>0.470, 0.841</td>
</tr>
<tr>
<td>Likka</td>
<td>12.966, 7.197, 45.623, 18.867, 11.764</td>
<td>80.216, 14.184, 109.969, 612.376, 20.933</td>
<td></td>
<td>0.187, 0.315</td>
</tr>
<tr>
<td>Coney</td>
<td>39.585, 115.246, 34.478, 52.698, 1.530</td>
<td>&gt; 1000, 18.501, 18.343, 1.722</td>
<td>525.414, 999.878, 2.001</td>
<td>0.807, 0.244</td>
</tr>
<tr>
<td>Merv</td>
<td>11.546, 3.218, 4.437, 226.778, 2.164</td>
<td>16.510, 7.540, 6.026, 999.878, 2.001</td>
<td>4.906, 1.722</td>
<td>0.703, 0.539</td>
</tr>
<tr>
<td>Leon</td>
<td>5.022, 1.647, 105.084, 12.188, 11.667</td>
<td>3.787, 1.955, 25.612, 8.780, 7.107</td>
<td>2.512, 0.633</td>
<td>1.239, 0.633</td>
</tr>
<tr>
<td>Richard</td>
<td>18.105, 6.700, 966.950, 277.335, 812.264</td>
<td>999.878, 6.972, 12.246, 528.635, 26.620</td>
<td>1.119, 0.782</td>
<td>0.379, 0.782</td>
</tr>
<tr>
<td>Edgar</td>
<td>24.084, 7.427, 2.512, &gt; 1000, 9.926</td>
<td>102.176, 18.841, 2.469, 83.319, 9.861</td>
<td>0.367, 0.898</td>
<td>0.367, 0.898</td>
</tr>
<tr>
<td>Duncan</td>
<td>6.019, 6.027, 7.258, 14.150, 5.825</td>
<td>18.079, 22.432, 10.377, 35.405, 7.828</td>
<td>0.273, 0.833</td>
<td>0.273, 0.833</td>
</tr>
<tr>
<td>Odysseus</td>
<td>7.098, 5.126, 7.468, 7.580, 345.517</td>
<td>10.467, 5.109, 6.800, 6.199, 15.752</td>
<td>0.967, 0.777</td>
<td>0.967, 0.777</td>
</tr>
<tr>
<td>Hector</td>
<td>297.694, 15.762, 0.839, 21.651, 0.314</td>
<td>22.869, &gt; 1000, 1.177, 17.119, 0.350</td>
<td>0.825, 0.691</td>
<td>0.825, 0.691</td>
</tr>
<tr>
<td>Frank</td>
<td>415.827, 15.724, 10.628, 9.668, 191.567</td>
<td>29.880, 200.063, 9.868, 8.380, 11.187</td>
<td>0.600, 0.809</td>
<td>0.600, 0.809</td>
</tr>
</tbody>
</table>

Table 3. Degree of unilateral dopaminergic cell damage for each group, sub-classified according to cell groups.

<table>
<thead>
<tr>
<th>Animal group</th>
<th>Neglect rating</th>
<th>Total unilateral dopamine cell damage median and (range) (Maximum score possible = 33)</th>
<th>A8 Damage median and (range) (Maximum score possible = 6)</th>
<th>A9 Damage median and (range) (Maximum score possible = 18)</th>
<th>A10 Damage median and (range) (Maximum score possible = 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$ (n=5)</td>
<td>++ +</td>
<td>27.5 (24–29) 6 (6–6) 18 (15–18) 3 (2–5)</td>
<td>27.5 (24–29) 6 (6–6) 18 (15–18) 3 (2–5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B$ (n=3)</td>
<td>++</td>
<td>22.5 (22.5–23.5) 6 (6–6) 16 (15–17) 0 (1–0)</td>
<td>22.5 (22.5–23.5) 6 (6–6) 16 (15–17) 0 (1–0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C$ (n=5)</td>
<td>+/+</td>
<td>19 (13–21) 5 (2–6) 12 (7–16) 0 (1–0)</td>
<td>19 (13–21) 5 (2–6) 12 (7–16) 0 (1–0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D$ (n=4)</td>
<td>0</td>
<td>0 (0–0) 0 (0–0) 0 (0–0) 0 (0–0)</td>
<td>0 (0–0) 0 (0–0) 0 (0–0) 0 (0–0)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4 Discussion

Both aspects of the results reported above – the trends in relative power accompanying increasing severity of unilateral dopaminergic lesion, and the transfer characteristics of the surface electrodes estimated from the changes in absolute and relative power – further support the linear model of electrocortical activity.

It appears that a quantitative relationship exists between the intensity of dopaminergic input to all the fields of termination (in caudate nucleus, nucleus accumbens and also, to the frontal cortical fields) and the degree of damping and driving of linear resonant waves generated by the collective properties of telencephalic circuitry. It may be argued that only one of these fields of termination may be critical – e.g., the caudate projections. Most experimental data suggest that sensorimotor neglect depends principally upon denervation of the caudate nucleus (Ljungberg and Ungerstedt 1976; Marshall 1979; Dunnet and Iversen 1982). In an allied study we have shown that sensorimotor neglect severity is related to the amount of loss of all dopaminergic cells and that lesion of the allied cell groups as well as nigro-striatal cells, also contributes (Lees et al. 1985). It is parsimonious to interpret sensorimotor neglect as the behavioural

Fig. 1. The ratio change in spectral power attributable to unilateral mesencephalic injection, for each animal group, with fitted theoretical curves. Each graph has been multiplied by 1/K to eliminate the effect of frequency independent changes in total power. The baseline at unity represents no change in the relative power spectrum.

From above down
- Group A (n = 3), maximum unilateral dopamine cell damage (median damage score = 27.5), neglect rating 3 + .
- Dominant resonances found at > 1, 7, 7.8, 12.6, 20.5 Hz.
- Lesion/Control damping coefficients at each resonance = 9.4/64.4, 3.7/19, 1.8/2.6, 6.6/9.6, 17.6/31.4.
- Mean lesion/control ratio of damping for ensemble = 0.46.
- Group B (n = 3), moderate dopamine cell damage (median score = 22.5), neglect 2 + .
- Dominant resonances = 3.6, 5.3, 8.8, 10, 28.8 Hz.
- Mean ratio of damping = 0.51.
- Group C (n = 5), least dopamine cell damage (median score = 19), neglect + /–.
- Dominant resonances = 2.8, 2.9, 10.8, 21.4, 25.3 Hz.
- Lesion/control damping coefficients = 2.5/2.9, 5.7/12.7, 8.1/8.6, 4.6/4, 17.4/10.8.
- Mean ratio of damping = 1.
- Group D (n = 4), no dopamine cell damage (control group), no neglect.
- Dominant resonances = 2.4, 8.5, 10.2, 15.7, 15.8 Hz.
- Lesion/control damping coefficients = 10.4/9.9, 0.6/0.5, 3.1/2.4, 1/0.6, 60.6/2.2.
- Mean ratio of damping = 1.5.
consequence of disrupted cerebral activity and information processing, consequent to lesion of a sufficient mass of dopaminergic neurones.

We may now ask how general the principle of control is. Does it extend across other cell classes, utilising different neurotransmitter substances, and with wholly different fields of termination? The theoretical model suggests that the principle should hold true for any cell class arising from the brain-stem and widely distributed to telencephalic cells. For this reason, the following paper (Wright et al. 1985b) directs tests of the linear wave theory to noradrenergic neurones of the locus coeruleus.

Acknowledgements. This work was supported by the New Zealand Medical Research Council, through a grant to J.J. Wright and a research Fellowship to R.R. Kydd.

We gratefully acknowledge the technical assistance of J.A. West.
References


Lees GJ, Kydd RR, Wright JJ (1985) Relationship between sensorimotor neglect and the specific, degree and locus of mesotolencephalic dopaminergic cell loss following 6-hydroxydopamine. Psychopharmacology (in press)


Wright JJ, Kydd RR, Lees GJ (1985e) State changes in the brain viewed as linear steady-states and non-linear transitions between steady-states. Biol Cybern (in press)

Received: January 10, 1985

Prof. J. J. Wright
The University of Auckland
Psychiatry and Behavioural Science
School of Medicine
Auckland
New Zealand
Contributions of Noradrenergic Neurones of the Locus Coeruleus to the Temporal Damping of Linear Electro cortical Waves

J. J. Wright, R. R. Kydd, and G. J. Lees
Department of Psychiatry, University of Auckland School of Medicine, Auckland, New Zealand

Abstract. The preceding paper (Wright et al. 1985a) gives evidence that mesotelencephalic dopaminergic neurones regulate gross electrocortical waves with linear properties, by influencing the strength of their driving signals and temporal damping. The present study further generalises the findings to ascending noradrenergic neurones, which have different fields of termination to dopaminergic fibres. It is shown that:

(a) Estimates of the major groups of natural frequencies for the telencephalic system obtained from curve-fitting the ratio changes in the power spectrum attributable to unilateral noradrenergic neurone lesion, are again centered about the frequencies of the major cerebral rhythms.

(b) Estimates of electrode transfer characteristics, using parameters obtained from curve fitting ratio changes in power, in conjunction with the raw left and right power spectra, are again found to be equal left and right, as required by the theoretical derivation.

(c) Changes in relative amplitude of electrocortical waves and their relative phase are significantly in accord with the relationship expected from theory.

1 Introduction

1.1

The noradrenergic neurones which ascend from the brain-stem to more rostral structures may be loosely categorised into two systems – a ventral system distributed principally to diencephalic sites, and the dorsal noradrenergic bundle. This latter system has cell bodies in the locus coeruleus, and projects axons very diffusely in the brain, especially to cerebral cortex and hippocampus (Lindvall and Bjorklund 1974; Robbins and Everitt 1982). These projections are largely ipsilateral, and the fibres of passage pass through the lateral hypothalamus. The function of these cells is far from clear, although they have been implicated in neuroendocrine regulation, rapid eye movement sleep, cortical “activation”, intracranial self-stimulation, and even in learning processes (Oids and Fobes 1981; Monti 1982; Robbins and Everitt 1982).

Our present concern is whether or not these cells contribute to the regulation of driving and temporal damping of electrocortical waves. The noradrenergic neurones offer a test-case for the model of electrocortical control we have previously proposed. Experiments to date tested the theory utilising gross lesions of the lateral hypothalamus, and selective lesions of the mesotelencephalic dopaminergic system (Wright and Kydd 1984a–c; Wright et al. 1984, 1985a). While these results support the theory, the findings might be whole description of dopaminergic actions. In the following paper we discuss the theory’s wider applicability to cerebral changes of state (Wright et al. 1985c). In the limited context of ascending brain-stem control of telencephalic activity, the same principle of control should apply to noradrenergic fibres as to dopaminergic ones. This is quite different to stating that the same specific effects are exerted by the two systems. Both should influence the damping and driving of telencephalic resonances, but may produce quite different changes of state in otherwise similar conditions.

The critical tests we have contrived to test the validity of the theory for dopaminergic control, assess only the validity of the control principle. Therefore, repetition of these tests using unilateral noradrenergic lesion will test the wider generality of the theory.

We here report the results of this replication using noradrenergic lesion, and compare these results to those obtained earlier. To avoid repetition we now list the tests used under the same headings used in the immediately preceding paper, and will subsequently refer to equation numbers and sections as they are referred to in (Wright et al. 1985a).
1.2 The Critical Tests

1.2.1 Consistency of Natural Frequencies with Changing Damping. This test requires that curve fitting the proportional, between hemispheres, change in spectral power following unilateral lesion will yield estimates of natural frequencies clustered about frequencies of the major cerebral rhythms.

1.2.2 Equality of Electrode Transfer Characteristics. This test requires that estimates of the transfer characteristics (describing the transformation of surface signals into the recorded signals) be equal on opposite sides. Since the theoretical equations for relative power is based on assumption of this equality, estimation of the transfer characteristics from the initial absolute spectra and parameters from curve-fitting the ratio changes in power, is a test of theory.

1.2.3 Correspondence of Relative Amplitude and Relative Phase of Electrocortical Waves. Since the theory predicts a unique relation between the ratio changes in left and right amplitude, and relative phase, following unilateral lesion, curve fitting the relative power yields parameters which predict the changes in relative phase, and can be checked against actual changes in phase.

1.2.4 Problems of Limited Model Order Approximation of a Very High-Order System. The caveats given in the corresponding section of the preceding paper continue to apply.

2 Methods

In most respects, these have been described in detail in earlier papers, and are only outlined here except where new procedures were used (Wright and Kydd 1984a–c; Wright et al. 1984, 1985a).

2.1 Animals

These were 10 male Wistar rats. Implanted extradural bipolar platinum recording electrodes were installed symmetrically at the stereotaxic sites P7.5, L5.1 (anterior electrode) and P9.5, L5.1 (posterior electrode). Two weeks recovery was allowed before recordings began.

2.2 EEG Recordings

These were obtained with the animals in a state of habituated wakefulness, under gentle restraint in a soft cloth. About 300 s of EEG was obtained 24 h before lesion, and 4–5 days following lesion, these being obtained from larger epochs of recording by an editing process aimed at ensuring that a comparatively steady set of conditions prevailed. Signals were band-pass filtered, digitised at 8 ms intervals, and stored on computer. Off-line Fourier analysis was then used to compute the average power spectrum of each channel, before and after selective lesion, and the average phase relation at each frequency interval, between the left and right channels. From these estimates we then obtained the relative power change attributable to lesion, $G^2(\omega)$ [see Eq. (1), preceding paper] and $\Delta \Phi(\omega)$ the relative change in left/right phase attributable to lesion (see Wright et al. 1984).

2.3 Method of Unilateral Selective Lesion

Under ketamine and barbiturate anaesthesia, unilateral stereotaxic injection was performed, with the cannular tip aimed at $A - 1.0$, $L1.3$, $V - 0.4$ (relative to stereotaxic zero). Two microlitre of 6 hydroxydopamine hydrogen bromide (6-OHDA) ($2 \mu$g free base) in 0.2% ascorbic acid carried in phosphate buffer solution was thus injected at a site rostral to the locus coeruleus and near the dorsal noradrenergic bundle.

2.4 Histological Assessment

Six days after selective lesion, animals were perfused to permit histofluorescent and routine histological examination of their brain-stems. Sections were cut extensively to permit examination of needle tract damage and all related catecholaminergic neurone systems.

2.5 Further Processing of EEG Data

2.5.1 For each animal the plot of $G^2(\omega)$ was curve-fitted to Eq. (2) for the five resonant mode approximation, with all parameters free.

2.5.2 All estimates of $M_\alpha$, the natural frequencies obtained in 2.5.1 from animals which subsequently were shown to have unilateral dorsal bundle lesion (seven animals) were then analysed to find the five median values best representing center values for five clusters of natural frequency, using least absolute deviation minimisation.

2.5.3 Curve fitting of individual estimates of $G^2(\omega)$ was repeated, imposing $M_\alpha$ values equal to the five median values found in 2.5.2 (for the seven lesioned animals).

2.5.4 After appropriate normalisation and correction for equipment band-pass characteristics, the parameters obtained in 2.5.3, were used in conjunction with the post-lesion power spectra, to obtain the mean left and right electrode transfer characteristics, for the dorsal bundle lesioned group. These $|A(\omega)|$ were calculated in accours with Eqs. (3) and (4).

2.5.5 Estimates of $G^2(\omega)$ were pooled for the seven animals with "specific" dorsal bundle lesion, and the three animals with "non-specific" lesion, respectively.
Curve fitting was then repeated on the group estimate from the specific lesion group, as in 2.5.1, and again with the five imposed values of centre $M_i$ obtained in 2.5.2.

2.5.6 The parameters obtained in each type of approximation of $G^2(\omega)$ in 2.5.5 were then used to predict the group average change in relative phase, according to Eq. (5). This was then correlated with the group average shift in relative phase attributable to lesion. Here a minor departure from earlier procedure was introduced (cp. Wright et al. 1984). Theoretical and experimental phase was correlated over the entire bandwidth analysed (i.e. without deletion of the few highest and lowest frequencies). Also, since correlation of this sort may find significance simply because of a simple linear trend across the bandwidth, correlation was repeated after both theoretical and experimental relative phase plots had been fitted with a straight line, and the residuals (without linear trend) then used for correlation.

### 3 Results

#### 3.1 Histological Results

Seven of the ten experimental animals showed histological evidence of unilateral damage to the dorsal noradrenergic bundle (the axons of passage from locus coeruleus). In 6 out of the 7 this was apparent as moderate to gross swelling of the noradrenergic axons, and in the seventh (James) unilateral damage was mild both in degree and in rostro-caudal extent. All these animals had in addition moderate to extensive damage of the ventral noradrenergic bundle, most apparent on the side of injection. Non-specific damage was slight and limited to the needle tract. In no animal were the noradrenergic or dopaminergic cell bodies damaged ($A7$, $A8$, and $A10$ groups). In the case of the noradrenergic system this is expected since locus coeruleus cell bodies are relatively insensitive to $6\text{OHDA}$. Preservation of the dopaminergic cell bodies, in contrast, indicates intactness of the dopaminergic system (Hokfelt and Ungerstedt 1973; Ungerstedt 1973).

Three animals showed no evidence of damage to the dorsal bundle, but had evidence of ventral noradrenergic damage comparable to the “specific” lesion group. General cell damage was also comparable in this group, which were thus treated as controls for unilateral dorsal bundle lesion.

#### 3.2 Results of Curve-Fitting

The results of curve-fitting $G^2(\omega)$ for each animal with all parameters free are listed in Table 1, and Fig. 1 shows examples of this graphically. Results of repeating this curve-fitting with imposed values of $M_i$ obtained from a five centre cluster analysis are listed in Table 2. Table 3 gives the values of the cluster centers and their deviation, with earlier values from other experiments given for comparison.

<table>
<thead>
<tr>
<th>Rat name</th>
<th>Quincy</th>
<th>James</th>
<th>Tom</th>
<th>Herb</th>
<th>Dwight</th>
<th>Ted</th>
<th>Martin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated natural frequencies of the dominant (asymmetric) modes (Hz)</td>
<td>4.58</td>
<td>2.40</td>
<td>3.27</td>
<td>1.79</td>
<td>&lt; 1.00</td>
<td>2.81</td>
<td>2.45</td>
</tr>
<tr>
<td>Corresponding damping coefficients (side of lesion) (cycles²/s)</td>
<td>2.84</td>
<td>1.00</td>
<td>0.14</td>
<td>0.95</td>
<td>2.61</td>
<td>1.13</td>
<td>1.43</td>
</tr>
<tr>
<td>Corresponding damping coefficients (side opposite lesion) (cycles²/s)</td>
<td>1.43</td>
<td>4.16</td>
<td>6.80</td>
<td>0.93</td>
<td>10.81</td>
<td>0.67</td>
<td>10.42</td>
</tr>
<tr>
<td>$K$ (scaling parameter)</td>
<td>5.53</td>
<td>37.99</td>
<td>11.12</td>
<td>67.87</td>
<td>0.72</td>
<td>0.85</td>
<td>5.47</td>
</tr>
<tr>
<td>$r^2$, goodness of fit to relative power, $G^2(\omega)$</td>
<td>0.95</td>
<td>0.58</td>
<td>5.51</td>
<td>1.92</td>
<td>0.40</td>
<td>16.36</td>
<td>100000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rat name</th>
<th>Quincy</th>
<th>James</th>
<th>Tom</th>
<th>Herb</th>
<th>Dwight</th>
<th>Ted</th>
<th>Martin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated natural frequencies of the dominant (asymmetric) modes (Hz)</td>
<td>7.48</td>
<td>0.51</td>
<td>2.24</td>
<td>5.60</td>
<td>18.68</td>
<td>132.89</td>
<td>7.10</td>
</tr>
<tr>
<td>Corresponding damping coefficients (side of lesion) (cycles²/s)</td>
<td>11.76</td>
<td>1.45</td>
<td>0.23</td>
<td>1.84</td>
<td>3.38</td>
<td>1.62</td>
<td>0.79</td>
</tr>
<tr>
<td>Corresponding damping coefficients (side opposite lesion) (cycles²/s)</td>
<td>2.38</td>
<td>33.25</td>
<td>3.64</td>
<td>1.47</td>
<td>7.19</td>
<td>0.97</td>
<td>3.45</td>
</tr>
<tr>
<td>$K$ (scaling parameter)</td>
<td>2.36</td>
<td>3.77</td>
<td>4.23</td>
<td>5.22</td>
<td>0.54</td>
<td>1.26</td>
<td>15.45</td>
</tr>
<tr>
<td>$r^2$, goodness of fit to relative power, $G^2(\omega)$</td>
<td>4.89</td>
<td>0.42</td>
<td>2.90</td>
<td>3.99</td>
<td>11.25</td>
<td>11.11</td>
<td>5.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rat name</th>
<th>Quincy</th>
<th>James</th>
<th>Tom</th>
<th>Herb</th>
<th>Dwight</th>
<th>Ted</th>
<th>Martin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated natural frequencies of the dominant (asymmetric) modes (Hz)</td>
<td>1.153</td>
<td>0.718</td>
<td>0.690</td>
<td>0.577</td>
<td>0.413</td>
<td>0.738</td>
<td>0.795</td>
</tr>
<tr>
<td>Corresponding damping coefficients (side of lesion) (cycles²/s)</td>
<td>0.73</td>
<td>0.69</td>
<td>0.82</td>
<td>0.81</td>
<td>0.75</td>
<td>0.81</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Table 1. Model parameters and goodness of fit to relative squared gain for each animal with unilateral dorsal noradrenergic bundle damage. Curve fitting performed with all parameters free during minimisation.

116 I WORK TOWARD A THEORY OF BRAIN FUNCTION
cantly removed from random and (b) whether each of the estimates obtained for the differing lesion conditions is significantly different from the others.

These show that the scatter of estimates for the present, noradrenergic data are marginally short of formal statistical significance ($D = 0.202, p < 0.075, one
tailed$). In contrast the electrolytic group showed highly non-random results ($D = 0.333, p < 0.01$) as did the dopaminergic lesion group ($D = 0.227, p < 0.01$).

Two sample tests showed no significant differences between any pair of these results (electrolytic versus dopamine $D = 0.155$, electrolytic versus noradrenaline $D = 0.171$, dopamine versus noradrenaline $D = 0.0929$).

Figure 2 shows the raw averaged data for relative power for the dorsal bundle lesion group and the "non-specific" damage group.

### 3.3 Estimation of Lesion and Control Electrode Transfer Characteristic

These are shown in Fig. 3. It is seen that control and lesion estimates are again comparable, with similar linear trends, and pattern of autocorrelated error showing similar distribution on each side. While similar on either side, these estimates are only moderately in accord with the most reliable results found from earlier experiments with the dopaminergic cell group, but are within the error limits found within this group (Wright et al. 1985a).

### 3.4 Correspondence of Changes in Relative Amplitude and Relative Phase Attributable to Unilateral Lesion

Figure 4 shows the curve-fitted relative power changes for the lesion group, and the superimposed experi-

---

**Table 2.** Model parameters and goodness of fit to relative squared gain for each animal when curve-fitting is repeated using imposed natural frequencies (median values of five principle natural frequency clusters, from data in Table 1) $M_1 = 2.81, M_2 = 7.67, M_3 = 12.06, M_4 = 16.31, M_5 = 25.09$

<table>
<thead>
<tr>
<th>Rat name</th>
<th>Quincy</th>
<th>James</th>
<th>Tom</th>
<th>Herb</th>
<th>Dwight</th>
<th>Ted</th>
<th>Martin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damping coefficients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(side of lesion)</td>
<td>$\mathcal{D}_1$</td>
<td>647.89</td>
<td>4.96</td>
<td>0.67</td>
<td>20.249</td>
<td>6.39</td>
<td>292.62</td>
</tr>
<tr>
<td>(cycles/s)</td>
<td></td>
<td>9.34</td>
<td>7.10</td>
<td>8.68</td>
<td>8.84</td>
<td>6.59</td>
<td>36.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.62</td>
<td>2.86</td>
<td>16.77</td>
<td>4.13</td>
<td>14.52</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.95</td>
<td>9.89</td>
<td>6.98</td>
<td>20.92</td>
<td>1.27</td>
<td>10.30</td>
</tr>
<tr>
<td></td>
<td>$\mathcal{D}_3$</td>
<td>4.84</td>
<td>7.52</td>
<td>15.91</td>
<td>$&gt;1000.00$</td>
<td>7.35</td>
<td>41.35</td>
</tr>
<tr>
<td>Damping coefficients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(side opposite lesion)</td>
<td>$\mathcal{D}_1$</td>
<td>&gt;1000.00</td>
<td>11.14</td>
<td>0.72</td>
<td>$&gt;1000.00$</td>
<td>8.27</td>
<td>41.20</td>
</tr>
<tr>
<td>(cycles/s)</td>
<td></td>
<td>66.65</td>
<td>15.97</td>
<td>4.43</td>
<td>816.56</td>
<td>6.92</td>
<td>$&gt;1000.00$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.74</td>
<td>3.61</td>
<td>8.11</td>
<td>7.30</td>
<td>10.21</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.66</td>
<td>64.9</td>
<td>6.49</td>
<td>33.32</td>
<td>0.91</td>
<td>8.59</td>
</tr>
<tr>
<td></td>
<td>$\mathcal{D}_3$</td>
<td>4.93</td>
<td>10.70</td>
<td>38.12</td>
<td>431.55</td>
<td>5.72</td>
<td>18.88</td>
</tr>
<tr>
<td>$K$ (scaling parameter)</td>
<td>0.814</td>
<td>0.345</td>
<td>0.688</td>
<td>0.085</td>
<td>0.396</td>
<td>0.701</td>
<td>0.493</td>
</tr>
<tr>
<td>$r^2$ (goodness of fit)</td>
<td>0.78</td>
<td>0.63</td>
<td>0.74</td>
<td>0.72</td>
<td>0.79</td>
<td>0.67</td>
<td>0.35</td>
</tr>
</tbody>
</table>
Table 3. Summary of the group median natural frequencies and the mean absolute deviations about the cluster centres, under three conditions of unilateral lesion. Kolmogorov-Smirnov two sample tests reveal no significant differences between distributions.

<table>
<thead>
<tr>
<th>Class of unilateral lesion used to induce left/right EEG asymmetry</th>
<th>Cluster medians for natural frequencies estimated from ratio changes in power spectra</th>
<th>Corresponding mean absolute deviations about each cluster median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noradrenergic</td>
<td>2.81, 7.67, 12.06, 16.31, 25.09</td>
<td>0.78, 0.77, 1.07, 0.73, 2.43</td>
</tr>
<tr>
<td>Dopaminergic</td>
<td>2.33, 7.1, 10.77, 19.65, 25.01</td>
<td>0.85, 0.55, 1.96, 1.16, 1.08</td>
</tr>
<tr>
<td>Lateral hypothalamic electrolytic</td>
<td>3.89, 7.03, 10.73, 18.6, 24.57</td>
<td>0.75, 0.69, 0.91, 1.11, 1.22</td>
</tr>
</tbody>
</table>

Fig. 2. Group average relative squared gain. *Top graph*: Group with unilateral lesion of the dorsal noradrenergic bundle. *Bottom graph*: Group without damage to the dorsal noradrenergic bundle.

A goodness of fit $r^2 = 0.907$. Direct non-parametric correlation of theoretical and experimental relative phase gave no evidence of significant relationship ($\rho = -0.07$).

However, when correlation was repeated after removal of linear trends, a highly significant relationship was found ($\rho = 0.279$, equivalent Students $T = 3.265$, $p < 0.0025$, one-tailed, with 64 df).

This anomalous result appeared to be due to gross distortion in the estimation of damping coefficients for the very low frequency mode, which induce a marked upward trend in the predicted phase, while preserving other curvilinear trends. This result caused us to elect further confirmation by a more demanding test, using approximation of gain with imposed center natural mental and predicted relative phase. It must be noted that the fitted curve to gain and the predicted phase are those obtained using the imposed values of $M_i$ from the cluster centers.

Performance of this test using the free parameter estimation yielded a best fit with parameters $M_1 = < 1$ Hz, $M_2 = 2.88$, $M_3 = 5.23$, $M_4 = 11.55$, $M_5 = 15.20$, with $\delta_{1L} = 0.35$, $\delta_{2L} = 3.16$, $\delta_{3L} = 5.56$, $\delta_{4L} = 4.90$, $\delta_{5L} = 3.5$ cycles/s. Values for $\delta_{1C}$ to $\delta_{5C}$ were 6.17, 7.89, 8.98, 7.63, and 3.52, with

Fig. 3. Plots of $|A(\omega)|$ for unilateral dorsal noradrenergic bundle group. *Top graph*: Side of lesion. *Bottom graph*: Side opposite lesion. Percentage difference in gain for the linear trend is 5% per octave.
WORK TOWARD A THEORY OF BRAIN FUNCTION

Fig. 4. Top graph: The group relative squared gain for animals with unilateral dorsal bundle damage (dots) with curve fitted line obtained by imposing the five resonance model with values for the natural frequencies imposed at the group centres (see Table 2).

Parameters were

\[ \delta_{1,} \text{LA; 23.74, 10.22, 3.90, 9.83, 16.07} \]

\[ \delta_{2,} \text{LB; > 1000, 48.39, 5.87, 18.24, 26.15} \]

Corresponding to ascending values of center natural frequency. \( r^2 = 0.886 \). Bottom graph: Superposition of the corresponding group relative phase shift (dots and connecting lines) and predicted phase from curve fitting to relative squared gain (solid line). Note phase estimates are displaced to share a common ordinate at their respective mean value across frequency. Correlation of theoretical and experimental phase without removal of linear trend: Spearman's \( r = 0.1299 \), \( T = 1.470 \), \( p < 0.1 \) for 64 df. Correlation after removal of linear trend: Spearman's \( r = 0.1450 \), \( T = 1.645 \), \( p = 0.05 \)

While problems of the same sort are encountered in this test also, it will be seen that results again reach statistical significance (see caption, Fig. 4).

4 Discussion

The above results show that the influences exerted upon the EEG by ascending dorsal noradrenergic neurones conform to the same principles which apply to dopaminergic neurones.

Both noradrenergic neurones and dopaminergic neurones can influence the temporal damping and the strength of internal noise-like signals driving a rostral resonant system, by virtue of their inhibitory influences on some of the telencephalic cells. The changes each ascending cell-system would impose upon the activity of the telencephalon (if either were equally stimulated, under similar initial conditions) would necessarily be quite different in degree and relative distribution over different modes.

In the following paper (Wright et al. 1985c), the implications of these findings are further explored. It is asked whether a generalisation regarding global regulation of all electrocortical activity is possible, and what may then be expected to follow for concepts of cerebral information processing.

Acknowledgements. This work was supported by the New Zealand Medical Research Council, through a grant to J. J. Wright and a research Fellowship to R. R. Kydd.

We gratefully acknowledge the technical assistance of J. A. West.

References


Wright JJ, Kydd RR, Lees GJ (1985c) State changes in the brain viewed as linear steady-states and non-linear transitions between steady-states. Biol Cybern (in press)

Received: January 10, 1985

Prof. J. J. Wright
The University of Auckland
Psychiatry and Behavioural Science School of Medicine
Auckland
New Zealand
Part II. The electrocorticogram

By this stage we were in personal contact with both Paul Nunez and Walter Freeman, the field luminaries. Totally disparate though their views were, both used systems known as neural mean-field equations, similar to those also introduced by Wilson and Cowan and Amari, in the 1970s. Generalising the form of these equations (and modernising their symbolism) so that any required sophistication of axonal and dendritic delays, synaptic connectivity, etc. can be introduced enables representation of the cortex in two dimensions, as a mesh of excitatory and inhibitory neurons.

$$q_{pq}^r(t) = f_p^r \times Q_p^r (r, t - \frac{|q-r|}{v})$$

1

$$\psi_p^q(t) = M_p^q \times q_p^q(t)$$

2

$$\Psi_p^q (q,t) = \int \psi_p^q(t) \, dr$$

3

$$V_p^q (q,t) = G_e^q \times \Psi_p^q (q,t) + G_i^q \times \Psi_i^q (q,t)$$

4

$$Q_p^q (q,t) = f_s (V_p^q (q,t)) + E_p^q (q,t)$$

5

Subscript $p=e,i$ refers to excitatory or inhibitory neurons; superscript $qr$ refers to synaptic connection from $r$ to $q$ where $q,r$ are cortical positions occupied by single neurons, represented by a complex number.

$q_{pq}^r(t)$ is the flux of pulses reaching pre-synapses at the neuron at $q$, from the neuron at $r$.

$\psi_p^q(t)$ is the synaptic current generated by $q_{pq}^r(t)$.

$\Psi_p^q (q,t)$ is the aggregate synaptic current of type $p$ generated at $q$.

$V_p^q (q,t)$ is the soma membrane potential (relative to the resting potential) generated at $q$.

$Q_p^q (q,t)$ is the pulse emission rate at $q$.

$f_p^q$ is the probability density of occurrence of pre-synapses generated by axons of the neuron at terminating at $qr$.

$V$ is axonal conduction speed.

$M_p^q$ is the steady-state term in a convolution transforming pre-synaptic flux to synaptic current.

$G_p^q$ is the steady-state term in a convolution transforming synaptic current into dendritic potentials.

$f_s (V_p^q (q,t))$ is a sigmoid function describing the local conversion of dendritic potentials into the rate of generation of action potentials.

$E_p^q (q,t)$ is a driving signal noise, arising from intrinsic random cell action potentials, and subcortical inputs.

This very general representation begs the question of exactly what cellular details need to be included to capture essential properties of cortical function and, indeed, what the essential properties might be – and that, of course, poses a set of research questions to be answered.
Before applying such equations, I tried to improve upon the theoretical position with regard to the range of linear analysis, and also to formulate the equations in a way that made the subcortical inputs central to explanation of the electrocortical effects. This led to the paper


in which I was trying to come to grips with one of the objections to the mean-field approach – that it was describing merely the “hum of the engine”, but not anything of importance for the pulse activity widely considered the “real” activity. I hoped one could get round the enormous complexity of neural activity, without ignoring the fact of complexity, using a statistical approach, describing the cortex as composed of linked local excitatory/inhibitory oscillating components, given by:

\[
d_2 V_e(q,t) \frac{dt^2}{dt^2} + D(q,t) \frac{dV_e(q,t)}{dt} + N^2(q,t)V_e(q,t) = K(q,r,t)V_e(r,t)
\]

where \( V_e(q,t) \) are excitatory dendritic potentials and \( \{D,N,K\} \) are time-varying stochastic parameters.

Assuming extreme complexity in the factors governing parameters \( \{D,N,K\} \), so that all are statistically independent, or have dependencies only within small groups, it can be shown that, in sequential short epochs, the natural frequencies of low frequency resonant modes would have low time-variance, with their damping coefficients approximately equal at all frequencies, while high frequency resonant modes would show high time-variance of both frequency and damping. Since the low-frequency resonances approximate the time-courses of electrocortical (dendritic) fields, and the high-frequency resonant modes approximate rapid events such as pulses, apparent dissociation of dendritic waves and individual pulses is an expected property, not a paradox, and the pulses and dendritic voltage waves can be considered aspects of a single linked system – not a motor with a superfluous hum.

Conversely, if sudden departures from steady-state occur, coincident changes in members of \( \{D,N,K\} \) will be reflected in highly nonlinear wave events. Sudden bursts of high speed firing in cortex would go outside the linear, near-equilibrium range, but in doing so would initiate roughly linear travelling waves in the surrounding dendritic fields. For reasonably steady conditions it ought to be possible to experimentally test the conclusions reached about the background, low-firing rate, linear state.

A test of one critical proposition – a more direct test for constancy of natural frequencies and equality of damping for all lower frequencies of electrocortical activity – followed in


As well as confirming the equality of damping over the frequency band, this paper, which was heavily dependent on the technical contributions of Alex Sergejew, reported use of the Akaike Information Criterion, to show that electrocortical waves were apparently a simple phenomenon, not one loaded with complex information.

Further tests followed. In alliance with Hans Stampfer, of UWA, Perth, who provided us with human EEG data, we followed up a lead he had earlier established, and also a suggestion from Paul Nunez, that the
cortical-evoked potentials might be the linear impulse response of the cortex. We tested that hypothesis and showed that the interpretation of the evoked potential as a linear impulse response to a brief sensory input delivered into the ongoing electrocortical activity was internally consistent.


Then, taking our lead from Walter Freeman, using closely spaced small recording arrays on the cortical surface of cats, we sought to further explore the electrocortical waves. To do so, Nick Hawthorn constructed 64 matched band-pass amplifiers for us (otherwise far beyond our budget, and rare worldwide in electrophysiology at the time) and Alex Sergejew implemented coherence analysis programs, as reported in


This work yielded an estimate of electrocortical wave velocity that accorded with the only other measure available. It also indicated that the electrocortical waves were arising from many small, highly local, sources scattered over the cortical surface, and generating waves that travelled in all directions. This result anticipated some features of Walter Freeman’s much more detailed analysis at a later date. With Paul Nunez’s help we were able to apply frequency/wavenumber analysis to the same data. This showed the electrocortical waves to have wavelengths much greater than the size of the arrays, so that the electrocortical voltage was close to synchrony across the cortical surface. The waves themselves were not complex, and could carry little information – but it appeared they were being driven by manifold small foci of activity, as the analysis of equation 6 suggested may be the case

Volume conduction of electrical flow in the conductive media of the head, blurring out all detail, could not be excluded in these experiments. Yet the contemporary discovery by Eckhorn et al., and Singer et al., that synchronous pulse activity of individual neurons, even at considerable distances of separation, characterised responses to specific stimuli, “binding” co-active cells, suggested that the global synchrony of electrocortical signals (wider still than reports of pulse synchrony) might be functionally meaningful, rather than merely a consequence of poor spatial resolution in recordings.

David Liley joined our small team, and about that time I made contact with Daniel Amit, an authority on Attractor Neural Networks – a branch of physics applied in analogy to neurons, yielding properties of pattern recognition, but still short of biological realism. Daniel invited me to write a manifesto of what would have to be included in a realistic model, and this was published as


We now had a clearer goal. It was to develop neural field equations progressively from their simplest form,

---

10 Note added in proof. This paper contained an arbitrary frequency-dependent term in the spatial damping. Had the mechanism of origin for synchronous oscillation been recognised by us at that time, this term could have been derived rather than (somewhat spuriously) assumed.
adding accurate parameter values for cellular details, until a sufficient explanation for electrocortical waves was apparent, with the hope that as we did so, an account of cortical information processing would emerge. David began work on calculations of synaptic connectivity that would provide the first estimates of parameter values needed for our own version of neural field equations and a simulation of electrocortical waves based on these – published in due course as


Meantime, storm clouds were gathering in the little world of Auckland psychiatry, which was profoundly disorganised, fraught with controversy, and developing at a high rate. During the turbulent latter years of the Fourth Labour Government a crisis of services arose. I had been Head of Department for seven years, and had combined my beliefs about the best path of psychiatric service development and research with a degree of political agitation and a cavalier disregard for administrative detail. When the Departmental Professorial Chair became vacant I was not, needless to say, appointed to the Chair. Alex Sergejew and David Liley had decamped to Australia, and from there they helped influence an offer made to me of a Professorial Fellowship at the Mental Health Research Institute (MHRI) in Melbourne. The shift was difficult for Adrienne, who had established herself as a significant figure in community service, for our elderly parents, and for our children, then becoming newly adult. With a sigh, Adrienne packed up house. In contrast to the joy with which we had departed for Caltech, it was as if we had taken to the lifeboats.

In Melbourne, things looked up for me. Adrienne established us, and tried to keep the family afloat at long range. The Australian magnate Richard Pratt made available a commercially linked grant of 4.3 million Australian dollars, and a well-equipped Brain Dynamics Laboratory grew into existence at MHRI. I began to collaborate with Evian Gordon in Sydney, whose interests closely complemented my own and who became an inspiring friend. Through Evian, the theoretical physicists Peter Robinson and Christopher Rennie, and the mathematician Clare Chapman, joined forces with us. Nick Hawthorn joined David and Alex with me, that side of the Tasman.

David Liley and I completed a defence of arguments for the essential linearity of electrocortical waves and their relation to highly nonlinear pulse summations, in


This Open Commentary article brought largely critical responses, but no substantive rejection. There was one important consequence for our work. A critic stated that if our work was able to reproduce long-range synchrony that would be a finding of substance. A simple test revealed that synchrony appeared between separated points in the simulation, even when the two points were driven by separate, independent white noise inputs. This was described in our reply to the critic.

Further numerical work showed the circumstances in which synchronous oscillation developed in the cerebral cortex, and began definition of the factors controlling spectral properties of the simulated electrocortical waves, in


From this work, Peter and Chris began a more formal physical approach, introducing stability analysis to the state equations. Peter formulated a partial differential wave equation, summarising the medium’s wave transmission properties.


and further applied the equations to explain analytically the occurrence of synchronous firing in the simulations, in


I now became a little concerned at how manipulation of parameter values, and arrangement of the state equations in different ways, could produce equally good approximations to physiological spectra and reported something of the range of these sensitivities in


Trying to find the most physiologically realistic means of the equations’ application, we applied the new mathematical techniques to gamma rhythm


and Peter and Chris then found introduction of lagged interaction of cortex with thalamic neurons accurately reproduced the characteristic spectral sequence of the EEG’s theta, alpha, beta and gamma background peaks, without appeal to intracortical resonances in the way Paul Nunez had asserted, and I had previously accepted.


Concurrently, Clare Chapman’s work using eigenmode analysis of synchronous firing in the simulated field provided a clear visualisation and explanation of the origin of synchrony, which was belatedly published as

This showed synchrony arises as a universal property of networks with summing junctions. It is generated by the constructive and destructive interference of travelling waves in this medium. As pulse trains are summed in dendrites, in-phase components sum, and out-of-phase components are eliminated, and the field rapidly attains maximum correlation at zero lag.

I had by this time met Wolf Singer, one of the co-discoverers of synchronous firing in the brain, who asked if this explanation could reproduce the classic experimental data he had used to indicate the relationship between synchrony and specific stimulus features. This we were able to do in


Chris linked the PDE wave model back to the earlier work in our group on the evoked cortical potential, in


Meantime and subsequently we all collaborated in papers that were fusions of these sets of results.


Leading a large group of his students, Peter Robinson has since developed applications of these techniques elsewhere in the brain.

I remained unhappy that we were not getting closer to an explanation of how the properties of the electrocortical field were relevant to cortical information processing. I would come as close as I could, in two papers published considerably later, the first of which owed much to the advice of Gordon Lees, and the help of Nick Hawthorn.

This paper introduced an effect I thought very important, that had been recently discovered experimentally\textsuperscript{11} – the retrograde propagation of action potentials into the dendritic tree of neurons – an effect that must drastically, but transiently, change the weight of synapses in the distal trees in determining subsequent pulse generation and thus bring into play information that might be stored on those synapses. The paper also introduced rules for synaptic adaptation during pulse transmission, and considered the interaction of wave-generating parts of the cortex that were rapidly firing (local bursting) with the surrounding, quiescent, and wave-transmitting cortex, so as to promote overall stability, rather than epileptic runaway excitation.

In the second of these two late papers


I showed that that at equilibrium of signal exchange, all excitatory cell pulses would approach zero-lag correlation, while excitatory and inhibitory cell pulses would also be locally correlated, but would exhibit 1/4 cycle lag-correlation at greater distances of separation, conforming to experimental results obtained by Freeman. These simulation findings also accorded with more recent work from his group, on the propagation of waves out of, and into, foci of wave generation.

It thus became possible to conceive of information transfer in the active cortex as a series of punctuated equilibria – equilibria reached repeatedly, with perturbations of the neural activity away from equilibrium caused by exogenous inputs and by foci of pulse-bursting, the whole forming a possible basis of cognitive sequences. Synchronous states could thus be considered as attractors, in the same sense as the point attractors in the more abstract attractor neural networks. Retrograde propagation of action potentials into the dendritic trees would release stored information into the wave medium, and interactions between foci, with synaptic adaptations, and subcortical inhibitory feedbacks, would enable the cortex to be self-stabilising at a level of excitation set by subcortical input. This would permit the release of pulse activity in complex sequences. The simulations did not simultaneously produce all these phenomena at once – but they showed that, thus assembled, a set of synergic effects could hold the activated cortex near the threshold for transition between bursting and quiescent wave transmission.

On the other hand, our models gave no suggestion that the transmission of pulse/wave activity in the electrocortical waves could carry large information content at any time, nor offer a communication channel of high storage capacity. More work would be needed to understand how the seemingly simple electrocortical waves could be associated with so many different pulse combinations, and yet both pulses and waves be mediators of information transfer.

For Adrienne and me, our time in Australia had come to an end. The commercial applications of our work had led to the formation of a company launched on the stock exchange, and led by Evian as CEO. To remain in this endeavour I would need to stay, and probably live, in Sydney. Family problems were a mounting source of anxiety to us.

Friends back in Auckland (Allen Fraser, Rob Kydd and Wayne Miles) arranged a clinical job for me, linked to an honorary professorship, with generous time for continuing research.

Home again.
Reticular Activation and the Dynamics of Neuronal Networks

J. J. Wright
Department of Psychiatry and Behavioural Science, School of Medicine, University of Auckland, Auckland, New Zealand

Abstract. It is postulated that during arousal the cortical system is driven by a spatially and temporally noisy signal arising from non-specific reticulo-cortical pathways. An elementary unit of cortical neuroanatomy is assumed, which permits non-linear dynamics to be represented by stochastic linear equations. Under these assumptions the resonant modes of the system of cortical dendrites approach thermodynamic equilibrium. Specific sensory signals perturb the dendritic system about equilibrium, generate low frequency, linear, non-dispersive waves corresponding to the EEG, which in turn regulate action potential sequences, and instantiate internal inputs to the dendritic field. A large and distributed memory capacity in axo-synaptic couplings, resistance to interference between functionally separate logical operations, and a very large next-state function set emerge as properties of the network. The model is able to explain the close association of the EEG with cognition, the channel of low capacity corresponding to the field of immediate attention, the low overall correlation of action potentials with EEG, and specificity of action potentials in some neurons during particular cognitive activity. Predictions made from hypothesis include features of thermal equilibrium in EEG (determinable by autoregression) and expectation that the cortical evoked response can be accounted for as the response to a sensory impulse of specific time characteristics.

1 Introduction

Arousal, alerting and subjective consciousness have long been known to be mediated by reticulo-cortical afferent pathways, which concurrently influence the electroencephalogram (EEG) (Moruzzi and Magoun 1949). The mechanisms underlying these associations are only partly understood. The cellular origin of the EEG signal from superficial cortical dendrites is now widely accepted (Elul 1972a, b; Mitzdorf 1988), and many precise and sensitive associations of both EEG spectral properties, and EEG evoked potentials, with attentional and cognitive states have been determined (Pockberger et al. 1988; Squires and Donchin 1976; Gevins 1988). However, a failure to obtain correlations between EEG and cortical action potentials in alert states has frequently been used to argue that the EEG is an epiphenomenon without direct relevance to cortical information processing. Against this interpretation, there is recent evidence of coupling of single cell activities to EEG frequencies in the visual cortex (Stryker 1989). There is also almost total absence of spatial coherence of cortical dendritic slow potentials, and the recorded EEG appears to reflect only minor and evanescent correlation among dendritic membrane potentials (Elul 1972a, b; Bullock 1988).

The anatomy of reticulo-cortical pathways is exceedingly complicated. Their anatomy and transmitter properties indicate that reticulo-cortical afferents (including the catechol and indolaminergic afferents) regulate the overall excitatory/inhibitory balance of cortical, thalamic and striatal neurones (Scheibel and Scheibel 1967; Anden et al. 1966; Robbins and Everitt 1982). Reticulo-cortical projections are largely polysynaptic, divergent, and diffuse in termination (Scheibel and Scheibel 1967), and their topology contrasts markedly with that of the specific sensory pathways. Their capacity to induce arousal is little influenced by major damage within their nuclei of origin (Adametz 1959). The lack of functional dependence upon structured afferents has suggested an essential absence of structured signal in the input of non-specific pathways to the telencephalon (Nauta and Feirtag 1979). This contrasts to the theoretical ascription to reticular efferents of selective and widespread co-ordination of cortical neuronal interactions (Crick 1984).

Part of the problem of understanding the reticulo-cortical interactions arises from the lack of a success-
fully tested model for local cortical information processing at the "semi-micro" level (Bullock 1988). Despite progress in the study of single units and groups of units in the cortex (Hubel and Wiesel 1977; Mountcastle 1978; Stryker 1989), there is no definitive consensus of opinion regarding the mechanisms of interneuronal couplings (Bullock 1988; Adey 1975), nor the levels of intrinsic noise in neuronal firings, nor the extent to which cortical neuronal interactions may be chaotic (Alspector et al. 1987; Carpenter and Grossberg 1987; Babcock and Westerveldt 1987; Mpitso 1987; Rossler 1983). The identification problem for electrocortical waves has not been solved, although evidence for their group linearity (Nunez 1981; Wright et al. 1987), and conversely, their low dimensional chaotic dynamics in the waking state, have been advanced (Watt and Hammeroff 1988; Freeman and van Dijk 1987).

In the presence of so many uncertainties, it appears to the author that qualitative modelling issues need resolution, before detailed specific modelling of cerebral neuron interactions can be undertaken. The present paper offers an hypothesis concerning reticulo-cortical interactions developed from earlier work (Wright et al. 1985), and proposes tests of that hypothesis. Reticulo-cortical input is assumed to be spatially and temporally noisy at a macroscopic level, and cortical interneuronal interactions to be non-linear and noise-free at cellular level. Cortical anatomy is represented as a system of recurrent, coupled, near identical units, for which the anatomical unit is not specified, but is assumed broadly consistent with Mountcastle’s (1978) proposal. The conclusions reached depend upon use of stochastic second-order differential equations, whose solution can approximate those of non-linear differential equations. These can be conveniently manipulated under the condition of near-independence of the second-order parameters, and this condition is held to apply because of the effect of reticulo-cortical input. General properties of the lowest frequencies of group resonant properties in the brain are emphasised and it is argued that signal/noise considerations make the lowest frequencies of the greatest informational importance. This approach is shown able to explain EEG activity as both an index of, and a component of, control processes within the brain, as follows:

2 Description of Model

2.1 Descriptive Methods

The cortical anatomy presents units which are recurrent and symmetrical at a number of scales, including the scales of cortical columns, supercolumns, or still larger units. Let the space-average potential of the superficial dendritic network overlying the appropriate unit be a state-variable \( x_i(t) \). Then a general coupled system of \( n \) units with input \( u_i(t) + x_{000}(t) \) (the uncorrelated, and cross-correlated parts of the input respectively) can be approximated as a stochastic linear system

\[
\begin{align*}
\dot{x}_i + D_i(t) \dot{x}_i + N_i^2(t) x_i &= u_i(t) + x_{000}(t) \\
+ \sum_{j=1}^{n} K_{ij}(t) (x_j + u_j + x_{000}) \\
(i &= 1 \ldots n, j = 1 \ldots n, j \neq i) \tag{1}
\end{align*}
\]

and

\[
D_i(t) = \sum_{j=1}^{n} F_{ij}(t) (x_j + u_j + x_{000}) \\
(i = 1 \ldots n, j = 1 \ldots n). \tag{2}
\]

Time-varying additive and multiplicative unit coupling parameters \( \{K_{ij}, F_{ij}\} \) are to represent axonal coupling between the state-variables, while \( \{N_i\} \) and \( \{D_i\} \) may be interpreted as the dominant intrinsic frequencies and dampings of the units, respectively. The stochastic parameters in (1) and (2) are used to represent implicit non-linear interactions, given by

\[
\begin{align*}
K_{ij}(t) &= g_i(x_1, \ldots, x_n, \dot{x}_i, \dot{x}_j, u_i, \ldots, u_n, \dot{u}_i, \ldots, \dot{u}_n, x_{000}, \dot{x}_{000}), \\
F_{ij}(t) &= h_i(x_1, \ldots, x_n, \dot{x}_i, \dot{x}_j, u_i, \ldots, u_n, \dot{u}_i, \ldots, \dot{u}_n, x_{000}, \dot{x}_{000}), \\
N_i(t) &= l_i(x_1, \ldots, x_n, \dot{x}_i, \dot{x}_j, u_i, \ldots, u_n, \dot{u}_i, \ldots, \dot{u}_n, x_{000}, \dot{x}_{000}) \\
(3)
\end{align*}
\]

for which the sensitivities \( \psi_{K_{ij}}, \psi_{F_{ij}} \) (gradients of each \( K_{ij}, F_{ij} \) in the state space) have maxima associated with the local dendritic potentials near threshold of transition for action potentials, and in general, each coupling presents maximum sensitivity to a unique state variable combination. Equations (3) may include time-lag operators describing effects of conduction times in axons, temporal summation in dendrites, etc.

Consistent with the findings of Elul (1972a, b) and Nunez (1981) the recorded EEG is here regarded as some linear transform of the space-averaged, low temporal-frequency components of \( \{x_i(t) + u_i(t) + x_{000}(t)\} \), such that the correlated variables dominate in the recorded time-series.

2.2 Model Assumptions

2.2.1 There exists one or more scales of cortical symmetry for which, during conditions of strong cortical activation \( \{K_{ij}, F_{ij}, N_i\} \) has all members finite and stationary over sequential epochs of duration \( T \) (up to several seconds), or zero at all times (where there
are no couplings between units). This assumption is mathematically equivalent to assuming all \(x_i(t)\) continuous, differentiable, bounded in time, and stationary over sufficient epochs. Physiologically, the description may correspond to a system of low-pass, smoothing components (dendrites) coupled by non-linear axo-synaptic (and other) processes.

2.2.2 When \(u_i \gg x_i\) for all \(i\), all parameters in \(\{K_{ij}, F_{ij}, N_{ij}\}\) approach stochastic independence during each epoch. This implies the coupling strengths are chaotic, or each perturbed by an extrinsic noise, as defined in 2.2.3.

In alert states \(u_i \gg x_i\) for all \(i\)

2.2.3 Reticulo-cortical input signals to cortical neurones are uncorrelated in space and time during any epoch, and form \(\{u_i\}\). Specific sensory pathways introduce transient signals which are spatially and temporally correlated during each epoch and form \(\{x_{00}\}\). Whether this is true of reticulo-cortical input is at present speculation. For specific sensory pathways, the assumption merely reflects the topological order of the fields of projection of these fibre systems.

2.2.4 All units have similar intracerebral efferents and afferents, are symmetrically and weakly coupled to many but not most others, and the intrinsic frequencies of units are high compared to EEG frequencies. This assumption is consistent with known small-unit physiology (Jack et al. 1975).

2.3 State Equation and Group Parameter Properties

Equations (1) and (2) are equivalent to

\[
\dot{z} = \mathcal{A} z
\]

(4)

where

\[
z = x + u + x_0
\]

i.e.,

\[
z = [x_1, \ldots, x_n, x_0] + [u_1, \ldots, u_n, 0]^T
\]

[\[x_{0(1)}, \ldots, x_{0(n)}\] + [0, \ldots, 0, 0]^T,

\[
\dot{z} = \frac{dz}{dt}
\]

\[
\mathcal{A}(t) =
\begin{bmatrix}
0 & 1 & 0 & 0 & \cdots \\
-N_1 & D_1 & K_{1,2} & 0 & K_{1,3} \\
0 & 0 & 0 & 1 & \cdots \\
K_{2,1} & 0 & -N_2 & D_2 & \cdots \\
0 & \vdots & \vdots & \vdots & \ddots \\
K_{n,1} & \cdots & 0 & \cdots & \cdots \\
0 & \cdots & 0 & \cdots & \cdots
\end{bmatrix}
\]

and

\[
D = Bz_d
\]

where

\[
D(t) = [0, D_1, 0, D_2, \ldots, D_n]^T,
\]

\[
B(t) =
\begin{bmatrix}
0 & 0 & 0 & \cdots & 0 & F_{n,1} \\
0 & 0 & 0 & \cdots & 0 & F_{n,2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & F_{n,n}
\end{bmatrix}
\]

\[
z_d = [\dot{x}_1, \ldots, \dot{x}_n, x_n] + [0, u_1, \ldots, 0, u_n]^T
\]

[\[0, x_{0(1)}, \ldots, 0, x_{0(n)}\]^T,

\[
\mathcal{A} = \mathcal{A}_o + \mathcal{A}_f,
\]

(6)

where the elements of \(\mathcal{A}_o\) are without cross-correlation or autocorrelation and \(\mathcal{A}_f\) is the matrix of autocorrelated parts of the elements of \(\mathcal{A}\). In Sects. 2.3.1 to 2.6 \(\mathcal{A}\) is considered a null matrix, in accord with assumption 2.2.2.

The system characteristic equation \(C(\lambda)\) relates unit parameters \(\{K_{ij}, F_{ij}, N_{ij}\}\) to group parameters \(\{M_i, \mathcal{D}_i\}\)

\[
C(\lambda) = a_m \lambda^m + a_{m-1} \lambda^{m-1} + \cdots + a_0 = 0
\]

\[
= [\lambda I - \mathcal{A}]^{n} = \prod (\lambda^2 + \mathcal{D}_i \lambda + M_i)
\]

(7)

where \(I\) is an \(m \times m\) identity matrix, \(m = 2n\), and \(\lambda\) are eigenvalues, \(\mathcal{D}_i\) is the damping coefficient of a resonant mode, and \(M_i\) is the corresponding natural frequency.

2.3.1 Unit Damping Coefficients and Damping of Resonant Modes. Assumption 2.2.4 applied to Eq. (5) requires the subset of \(\{F_{ij}\}\) in each row of \(B\) to be similar in all rows. Since \(u_i \gg x_i\) for all \(i\), and \(u_i\) are random variables, all \(D_i\) are random about a single mean.

In Appendix 1 it is shown that at all \(t\), for every member of the set \(\{D_i\}\) there is an equal member of the set \(\{\mathcal{D}_i\}\).

Therefore all \(\mathcal{D}_i\) are random about \(\mathcal{D}_i\) (the average damping factor) and \(\mathcal{D}_i \propto \bar{u}_i\) (the average value of dendritic potentials).

i.e. Unit and resonant mode damping coefficients are perturbed about a single parameter value set by the balance of excitatory and inhibitory influences.

2.3.2 Unit Natural Frequencies and Mode Frequencies. Again in Appendix 1 it is shown that

\[
\sum \{ \frac{1}{\sqrt{\prod N_i(z) \delta_i}} \} = \sum \{ \prod M_i \}
\]

(8)
for \( r = 1 \ldots n \) where \( \delta_r \) are terms arising from a subset of \( \mathbf{K}_{ij} \) (Appendix 1, (1.7) and (1.8)) and the summation indicated is that of the \( \binom{n}{r} \) possible products of combinations of \( r \) members of the sets \( \{ N_i \} \) or \( \{ M_i \} \).

Also
\[
\sum_{i=1}^{n} \sigma^2_{M_i} = n \sigma^2_{\gamma_M}
\]
and
\[
\prod_{i=1}^{n} \frac{M_i^2}{N_i^2} = \frac{1}{\prod_{i=1}^{n} (\gamma_{M_i} - 1)}
\]
(9)

where
\[
\gamma_{M_i} = 1 + \frac{\sigma^2_{M_i}}{M_i^2}, \quad \gamma_{N_i} = 1 + \frac{\sigma^2_{N_i}}{N_i^2}.
\]

\( \bar{M_i^2}, \sigma^2_{M_i}, \bar{N_i^2}, \sigma^2_{N_i} \) are the means squared, and variances of the squares of \( M_i \) and \( N_i \), respectively.

These relations require that variance of the low frequency modes \( \rightarrow 0 \) if \( n \) is very large.

**i.e. The EEG reflects low frequency resonances with natural frequencies which are effectively constant over time.**

### 2.4 Additivity of Group Waves and Dispersion Relations

A spatial transfer function \( H(s, T)_{ab} \) may be defined within a brief epoch \( T \), for any two points \( a, b \) by summing transfer functions over all possible pathways (\( q \) in number) by which signals may pass from \( a \) to \( b \) via \( p \) intervening units, where \( p = 1, 2 \ldots \) for different paths. Where \( K_{ij}, N_i \) here refer only to the subset of \( \{ N_i, K_{ij} \} \) which describe those units and pathways connecting \( a \) to \( b \), then

\[
H(s, T)_{ab} = q \left( \frac{1}{\prod_{i=1}^{n} \left( s^2 + D_i(t)s + N_i^2(t) \right)} \cdot K_{ij}(t) \right),
\]
(10)

where \( s \) is the Laplace operator.

Since \( D_i, N_i \) and \( K_{ij} \) are stationary and independent, where \( q \) is very large, \( H(s, T)_{ab} \) tends to constancy with respect to time, and is non-zero when (by assumption 2.2.4) points \( a \) and \( b \) are connected by multiple parallel pathways of similar conduction lag and sign of coupling (i.e. predominantly excitatory or inhibitory). Wave fronts propagating through the ensemble (group waves), since they do so with constant spatial transfer relations, must exhibit superposition.

The transfer from \( a \) to \( b \) approaches non-dispersive wave propagation when \( D_i \) is low, but depending on heterogeneity in the range and strength of the couplings described by \( \{ K_{ij} \} \), the dispersion relation may be branched.

**i.e. low frequency EEG waves are group linear waves.**

### 2.5 Equilibrium and Perturbations

Since EEG activity approaches conditions for linear nondispersive waves, and the cortical surface is asymmetrical with invariant boundaries, the lower frequency modes must have non-degenerate eigenfunctions, and their internal energies may be approximated as those of simple harmonic oscillators. In the simplified steady-state case where both \( x_0 \) and \( u \) are spatially and temporally wide-spectrum noise, then the internal energy \( U_i \) of each mode (Feynman 1963) is

\[
U_i \propto \frac{\left[ u(\omega) + x_0(\omega) \right]^2}{4 \omega_i} \tan^{-1} \left( \frac{\omega_i W}{M_i^2 - W^2} \right),
\]
(11)

where \( \left[ u(\omega) + x_0(\omega) \right] \) is the \( \omega^0 \) Fourier component modulus (equal for all \( \omega, \omega = 0 \ldots \), \( W \) of \( x_0 + u \), and \( 0 - W \) is the temporal bandwidth of the driving signals.

Each low frequency mode has entropy, \( S_i \),

\[
S_i = \left( 1 + \frac{U_i}{\epsilon} \right) \log \left( 1 + \frac{U_i}{\epsilon} \right) - \left( \frac{U_i}{\epsilon} \right) \log \left( \frac{U_i}{\epsilon} \right),
\]
(12)

where \( \epsilon \) is a suitable energy unit, \( \epsilon < U_i \), and the ensemble free energy \( \psi \), is given by the Gibbs relation

\[
\psi = \sum_{i=1}^{n} U_i - \tau \sum_{i=1}^{n} S_i,
\]
(13)

where \( \tau \) is the average energy of the modes.

Where all \( \omega_i \approx \omega \) and \( W \gg M_i \) for all \( i \), then the steady state corresponds to thermal equilibrium, since \( \sum_{i=1}^{n} \frac{U_i}{\epsilon} \rightarrow \text{a maximum} \) and \( \psi \rightarrow \text{a minimum} \).

Actually all \( \omega_i \) are perturbed about \( \omega \) (Sect. 2.3.1), and \( W \gg M_i \) for low frequency modes. Further \( u(\omega) \rightarrow 0 \) for all \( \omega \), since \( \{ U_i \} \) are uncorrelated noise inputs in both space and time, acting over the entire system. Therefore transients in \( x_0(t) \) perturb mode energies away from a low energy equilibrium state.

**i.e. the specific sensory signals generate propagating EEG waves, while brain-stem input maintains an equilibrium condition, equivalent to thermodynamic equilibrium with minimal wave generation.**

### 2.6 Effects of Slow Potentials on Local Coupling Processes

Linear resonant modes have amplitude proportional to \( \sqrt{U_i / M_i^2} \) and weakly dispersive waves diminish in amplitude with distance of propagation, \( R \), approxi-
mately as $e^{-\beta_i R M_i}$. Where all $U_i \to \tau$, the contribution of a single resonant mode to the variance of $x_R$ (a unit dendritic potential located at distance $R$ from a source of input) is therefore

$$\sigma^2(x_R) \propto \frac{e^{-2 \beta_i R M_i}}{M_i^2}$$  \hspace{1cm} (14)$$

and

$$\frac{\partial \sigma^2(x_R)}{\partial \beta_i} \propto -2 \frac{R}{M_i} e^{-2 \beta_i R M_i}.$$  \hspace{1cm} (15)$$

$$\frac{\partial \sigma^2(x_R)}{\partial R} \propto -2 \frac{\beta_i}{M_i} e^{-2 \beta_i R M_i}.$$  \hspace{1cm} (16)$$

Equations (14) indicate that the lowest frequency modes exert the greatest influence upon dendritic activity, with increasing range, and are most sensitive to variation of the average damping factor. The single parameter, $\beta_i$, can thus regulate the total covariance of $\{x_i\}$, and as a corollary, $\mathcal{A}(t)$ is a function of $\beta_i$. i.e. EEG waves can co-ordinate sequences of action potentials among widely separated units, in the presence of largely random neuronal firings among the intervening units.

2.7 Nonlinearity and Memory

In Sects. 2.3.1-2.6, $\mathcal{A}$ was treated as a null matrix. Actually $\mathcal{A} = 0$ only at equilibrium. Rewriting (4) as

$$z = \mathcal{A} z + \mathcal{A} (x)$$  \hspace{1cm} (17)$$

shows $\mathcal{A} z$ is equivalent to an input signal, but this signal has been internally generated by weakly cross-correlated action potential sequences, themselves determined by cross- and auto-correlated dendritic potentials. This equivalence requires the capacity for storage of information in axo-synaptic couplings to have a maximum dictated by the channel capacity of the near-equilibrium system, approximated as

$$z = \mathcal{A} (x).$$

2.7.1 Information Storage. Near equilibrium, if $\frac{x_i}{U_i}$ is the signal/noise ratio for dendrites, and if $\overline{x}_i$ is the mean number of couplings between units $i$ and $j$, the sequential flow of information through the dendrites and axons in series, defines

$$\chi = \frac{x_i}{U_i} \beta_i \eta$$

for which $\chi$ may be interpreted as the number of couplings needed to transmit $\{x_i\}$ alone. There are therefore

$$\chi = e^{-\sum_{x=1}^{n} \left( \frac{n x_i}{\chi} \right)}$$

ways to array signals in the axonal network, for which each array can be considered an independent channel, where $q < n \overline{x}_i$, and $q$ is set by the maximum dendritic signal/noise ratio. Applying the theorem of channel capacity (Shannon and Weaver 1949),

$$C_{ax} = \sum_{x=1}^{n} \left[ \left( \frac{n x_i}{\chi} \right) \cdot 2 W_i \log_2 \left( 1 + \frac{x}{n \overline{x}_i} \right) \right].$$

where $C_{ax}$ is the maximum possible information storage capacity of axonal couplings in $\{K_{ij}, F_{ij}\}$ and $W_i$ is the Nyquist frequency for all signals carried by axonal couplings. The actual information storage capacity achieved depends upon the degree of redundancy implicit in the array and sensitivities of unit couplings, but is not affected by the mixing of signal and noise in the coupling’s transmissions.

2.7.2 Information Access. Only a portion of the total information stored is accessible for a particular value of $\beta_i$, and strength of reticular noise input, since $\mathcal{A} z$ is a function of these two parameters. Access is also limited by the dendritic channel capacity, given by

$$C_{den}(\beta_i, \mathcal{A}(u_i)) = 2 W_i \log_2 \left( 1 + \sum_{x=1}^{n} \frac{\sigma(M_i)}{\sigma(u_i)} \right).$$

where $\sigma(M_i)$ is the rms amplitude of the $i$th mode's eigenfunction, $\sigma(u_i)$ is the corresponding noise amplitude, and $\sum_{x=1}^{n} \sigma(M_i)$ is the square root of summed mean square amplitudes of the modes. $C_{den}$ is the sum of two sub-channel capacities – the channel capacity utilised by sensory inputs and the channel capacity available to internally generated signals. Since $u_i > x_i$ for all $i$, and $\sigma(M_i)$ is inversely proportional to $M_i$ near equilibrium, then $C_{den}$ and its sub-channels are of low capacity compared to $C_{ax}$, and their capacity is principally attributable to the low frequency resonant modes.

Similar reasoning to that in 2.5 indicates that only the correlated components in $\mathcal{A} z$ will significantly perturb the system from equilibrium, and generate electrocortical waves.

i.e. EEG waves carry little information, but can continuously instantiate inputs to themselves from a large information store formed of axo-synaptic connections.

3 Discussion

3.1 Summary of Implications

Provided only that a restricted class of stochastic differential equations can be used to approximate the extreme non-linearities of neurones anatomically organised with translational symmetry, then a strong
noise component introduced to the cortical dendritic activity may exert a number of effects greatly influencing the brain's properties. By random perturbation of interneuronal interactions the noise input gives rise to an equilibrium state which permits both the generation of linear waves by specific sensory input signals, and the release of information stored within interneuronal connections. This in turn brings about the linear superposition of input signals and internally generated signals. Variation of the strength of the non-specific noisy input and the temporal damping for dendritic waves (the excitatory/inhibitory balance) can adjust access to selected small portions of a potentially very large internal axo-synaptic memory. In this view EEG waves are not epiphenomenal to cortical information processing. The longer propagation range and higher signal-to-noise ratio of low frequencies permits their preeminence as regulators of widespread neuronal activity.

3.2 Relation to Physiological, Psychological, and EEG Analysis Findings

The low spatial coherence of cortical dendritic micropotentials, the near-zero correlation of cortical action potentials with EEG in the alert state, and cognitive and behavioural dependencies of EEG, are consistent with of the present model. Uncorrelated “background” action potentials reflect the low signal/noise ratio of the activated state, while conversely, the changes in firing rates of single neurones seen in specific attentional and behavioural states may provide evidence of critical involvement of these cells in particular correlated action patterns (e.g. Stryker 1989).

The mediation of immediate cognitive processes by dendritic slow waves of low information capacity offers a correlate for the content of immediate attention. The concept of a low capacity information channel identical to the focus of current consciousness is employed in both modern cognitive psychology (Broadbent 1971) and traditional descriptive phenomenology (Jaspers 1963, Kydd and Wright 1986).

The expected linear wave dynamics of low frequency resonant modes accord with reports of alpha activity (Lehmann 1971; Lehmann et al. 1987; Nunez 1981) showing bidirectional non-interfering wavefronts, and standing wave activity. The lateral hypothalamic pathways (principal through-ways of the reticulo-cortical afferents)(Scheibel and Scheibel 1967) exert effects on EEG compatible with the regulation of time-varying damping in a system with constant frequencies of resonance (Wright et al. 1987; Franaszczuk and Blinowska 1985; Mitraszewski et al. 1987).

Conversely, there is no contradiction with the reports of low dimensional non-linear behaviour in deep sleep EEG, and in epileptic states (Babloyantz et al. 1985; Babloyantz and Destexhe 1986), as chaos might be expected in the autonomous behaviour of coupled highly non-linear elements (Rossler 1983). The withdrawal of reticulo-cortical drive, and strong local excitation within epileptic foci are both autonomous states of this model.

Reports of low dimensional attractors demonstrable in EEG in the awake state (Watt and Hammeroff 1988; Freeman and van Dijk 1987) appear in conflict with hypothesis of a noise-driven system, and this problem will require further consideration.

3.3 Predictive Tests of Hypothesis

Two quantitative tests of hypothesis emerge despite the qualitative nature of the model.

3.3.1 Near Equilibrium Properties of EEG Waves. Hypothesis requires that the alert EEG reflect linear modes with “thermal” properties, so damping coefficients will be randomly perturbed over time, and not frequency dependent. Experimental test must allow for the limited signal resolution possible with macro-electrode techniques. The EEG may be approximated by autoregression, using models of comparatively low order, to obtain initial estimates of mode damping coefficients (Franaszczuk and Blinowska 1985). It can be shown (Wright et al. 1990) that the parametric estimates of damping obtained by use of a low order approximate model must also exhibit randomness and frequency independence, if the “underlying” system has the predicted properties. As a corollary, in conditions removed from high alertness, EEG should depart from the equilibrium state and parametric estimates should then exhibit intermodulal dependencies.

3.3.2 Evoked Potentials and Cortical Impulse Function. Hypothesis requires that alert EEG reflects linear waves driven by specific sensory afferents. In conditions in which the EEG approaches stationarity, with repeated brief sensory stimuli (as in the evoked potential paradigm) the input of signals to cortex should be determinable by elementary transfer function calculations. Analysis of the EEG (again applying autoregression) in the pre-stimulus epochs would yield the (current) linear transfer function for dendritic fields driven by background sensory inputs. The post-stimulus evoked potential is then the impulse response induced in a system of known transfer function, and the time course of the input to cortex can therefore be directly computed. The input computed must correspond to the times of arrival of sensory signals to cortex, established by conventional physiological means (Shaw 1988), to support hypothesis.

3.3.3 Other Tests and Preliminary Results. As indicated above, earlier tests offer provisional support of hypo-
thesis (Wright et al. 1987). Studies of EEG wave properties using multielectrode arrays should reveal weakly dispersive or non-dispersive linear waves, with or without a multiply branched frequency/wave number relation.

The tests described in 3.3.1 and 3.3.2 have been partly carried out by members of the author's group (Wright et al. 1990) and results are supportive of hypothesis. Detailed data will be reported later. Assumptions 2.2.1, 2.2.2, and 2.2.3, are themselves subject to more direct test, by appropriate microelectrode techniques.

3.4 Some Relations to Other Brain and Neural Models

The present hypothesis is quite unlike the Crick (1984) hypothesis for reticular activation which postulates selective reticulo-cortical inputs. There is some resemblance to parallel processing in artificial neural networks, since network behaviour would converge toward an attractor of minimum energy. This resemblance does not require real neurons to share other dynamic properties with PDP networks, Hopfield networks, or Boltzman machines (McClelland et al. 1986a).

Neither does the present hypothesis fully correspond to assumptions of any specific neuronal model for cortical EEG yet advanced. It is not congruent with the Nunez model (1981) for alpha rhythm. The rationale for linearity of the wave medium is quite different. More importantly, the Nunez model predicts strong dependence of cerebral natural frequencies on axonal conduction lags. The present hypothesis indicates a dependence of wave number, but not natural frequency, upon conduction lag. However, the Nunez model is intended for conditions of arousal lower than full alerting. The Lopes da Silva (1974) model for alpha rhythm is not compatible with this hypothesis, nor is any other model attributing major cerebral rhythms to local circuit oscillations alone, rather than whole-brain resonances.

The use of assumed properties of a small, recurrent, sub-set of neurons is in accord with the K set technique of the Freeman model for olfactory bulb dynamics (Freeman 1987; Skarda and Freeman 1987).

Application of the K set technique to the cerebral cortex, would require a different consideration of the effects of reticular input, in contrast to that required in the case of the olfactory bulb (Freeman, personal communication).

Low frequency modes may provide means of initialising and co-ordinating local and highly non-linear information processing by neural arrays, however these arrays may be conceived and ultimately demonstrated to operate (Cowan 1968; Hopfield and Tank 1986; Heizel and Salverston 1988; Marder 1988; Carpenter and Grossberg 1987; Ingber 1986; Ingber and Nunez 1990). In the present context, high frequency local activity is little considered, but must be crucial to the release of stored information in the postulated memory mechanism (Sect. 2.7). A feature of possible importance is that the presence of high sensitivities of the interunit couplings within only a small domain of the total state space, would permit modification of memory without inevitable changes in “cross-talk” (McClelland et al. 1986b).

A complete model would also include cortico-reticular (centripetal) controls. Frontal, limbic, and hippocampal cerebral sub-systems (MacLean 1955) may act as mediators between the entire cortex and reticulo-cortical pathways – a process which would permit cortical self-regulation, and might incidentally explain the powerful motivational effects mediated by sub-components of reticulo-cortical afferents, which are closely associated with general arousal (Rolls 1974; Wright and Craggs 1977; Wise 1978).

Appendix 1

Mean and Variance of Unit and Group Parameters

The following conventions are defined:

(i) Where \( \prod \{ \bar{\sum} \}_{i=1}^{r} \{ \bar{\sum} \} \) represents the sum of the set of all possible \( \prod \{ \bar{\sum} \} \).

(ii) To apply the above convention to expansion products of \( |\lambda - a| \), a square bracket convention is introduced. As

\[
|\lambda - a| = \begin{vmatrix}
\lambda & -1 & 0 & 0 & 0 \\
N_{1} & \lambda + D_{1} & -K_{1,2} & 0 & -K_{1,3} & 0 & -K_{1,4} & 0 & \ldots \\
0 & 0 & \lambda & -1 & 0 & 0 \\
-K_{2,1} & 0 & N_{2} & \lambda + D_{2} & 0 & 0 & 0 & \lambda \\
0 & 0 & 0 & 0 & \lambda & 0 \\
-K_{3,2} & 0 & 0 & \lambda \\
0 & \ldots & & & & & & &
\end{vmatrix}
\]
\[ \prod \left[ X \right] \prod \left[ Y \right] \] indicates the sum of all possible products formed with regard to \( r-i \) elements from the symbol sub-classes, \( X \), and \( i \) elements from the symbol sub-classes \( Y \), where \( X \) and \( Y \) are exclusive.

(iv) Where ratios \( \frac{\prod \left[ K_0 \right]}{\prod \left[ N_{00} \right]} \) are used the subscripts \( ii(j) \) are to indicate that the element subscripted \( ii(j) \) comes from the same row as the element subscripted \( j \).

Mean and Variance of Damping

The characteristic equation with damping can be obtained from the expansions

\[ C(\lambda) = \sum_{r=0}^{\infty} \left( \prod \left[ K_{0i} \right] \prod \left[ N_{22} \right] \right) \]

(1.1)

and

\[ C(\lambda) = \prod \left[ (\lambda^2 + \beta_1 \lambda + \eta) \right] \]

(1.2)

Both (1.1) and (1.2) can be written as equivalent series of factorized products in which \( \lambda^2 \) and \( \lambda \) can be substituted as distinct variables, as follows:

Because the \( -1 \) elements in \( [i-\omega] \) are the only non-zero elements common to the row and column of each pair of elements \( \lambda, \lambda + D_i \) then any non-zero product in the form defined by (1.1), having an order in \( \lambda > 0 \), must have \( (\lambda^2 + \beta_1 \lambda + \eta) \) as one or more of its factors.

Therefore with \( \lambda^2 = b, \lambda = \sigma \) substituted in the factors \( (\lambda^2 + \beta_1 \lambda + \eta) \), the sums of all terms of orders \( c^{n-1}b \) and \( c^n \) in (1.1) are

\[ \sum_{r=0}^{\infty} \left( \prod \left[ K_{0i} \right] \right) \cdot c^{n-1}b \quad \text{and} \quad \sum_{r=0}^{\infty} \left( \prod \left[ K_{0i} \right] \right) c^n. \]

Equation (1.2) can be arranged as

\[ C(\lambda) = \sum_{r=0}^{\infty} \left( \prod \left[ \alpha_{il} \right] \right) \cdot c^{n-1}b \quad \text{and} \quad \sum_{r=0}^{\infty} \left( \prod \left[ \alpha_{il} \right] \right) c^n. \]

Equation (1.4) require that for every member of \( \{ D_i \} \) a member of the set \( \{ \aleph_i \} \) of equal magnitude.

Means and Variances of Natural Frequencies

In the case where all \( D_i = 0 \), and therefore all \( \aleph_i = 0 \)

\[ C(\lambda) = \prod \left[ \lambda \right] \prod \left[ \lambda, \lambda, \lambda, N_{22}, 1, 0 \right] \]

(1.5)

and

\[ C(\lambda) = \prod \left[ (\lambda^2 + M^2) \right] \]

(1.6)

where \( \alpha_{2n-2} \) is the coefficient of \( \lambda^{2n-2} \) in \( C(\lambda) \) then from (1.5) by listing of all possible combination products which have \( \lambda^{2n-2} \) as a factor, yields

\[ a_{2n-2} = \frac{\prod \left[ K_0 \right]}{\prod \left[ K_0 \right]} \times \frac{\prod \left[ N_{22} \right] \cdot \prod \left[ K_0 \right]}{\prod \left[ K_0 \right]} \times \frac{\prod \left[ K_0 \right]}{\prod \left[ K_0 \right]} \quad \text{etc.} \]

(1.7)

and from (1.6)

\[ a_{2n-2} = \frac{\prod \left[ M^2 \right]}{\prod \left[ M^2 \right]} \]

(1.8)

Since variances of sums and products of stochastic variables are always positive, then applying the standard form for the product of independent variables, from (1.8)

\[ \text{var} a_{2n-2} = \frac{\prod \left[ (M^2 + \sigma^2) \right]}{\prod \left[ (M^2) \right]} \]

(1.9)

(\( M^2 \) is, strictly, the squared mean squared \( M^2 \). By rearranging (1.7) in accordance with convention (iv) defined above, and obtaining the variance for products of stochastic variables and the sums of products, gives

\[ \text{var} a_{2n-2} = \frac{\prod \left[ (N^2 + \sigma^2) \right]}{\prod \left[ (M^2) \right]} \]

(1.10)

When

\[ \frac{\prod \left[ K_0 \right]}{\prod \left[ N_{22} \right]} \rightarrow \text{zero} \quad \text{for all} \quad l = 1 \ldots r \]

\[ \prod \left[ K_0 \right] \prod \left[ N_{22} \right] \]
then
\[ \var{r} \sum (r M^2) \rightarrow \var{r} \sum (r N^2) \] for all \( r = 1 \ldots n. \) \( (1.9) \)

This limit is approached as \( K_u \) and \( \sigma_{U^2} < 1 \) and \( N_{iq} \geq K_u \) for most \( N_{iq} \). These are the conditions describing weak couplings between each unit, and many, but not most others, with high intrinsic frequencies for unit self-excitation or resonance (Assumption 2.2.4). Near this limit, for \( r = 1 \)

\[ \sum \sigma_{M^i} = \sum \sigma_{N^i} \]

and when \( r = n \), with the substitutions \( \gamma_{Mi} = 1 + \frac{\sigma_{M^i}}{M^2} \gamma_{Ni} = 1 + \frac{\sigma_{N^i}}{N^2} \)

and rearranging yields

\[ \frac{1}{\sum M^2} = \frac{1}{\sum N^2} \]

\[ \frac{1}{\sum N^2} = \frac{1}{\sum M^2} \]

References

Adamezt JH (1959) Rate of recovery of functioning in cats with rostral reticular lesions. J Neurosurg 16:85–97


Moruzzi G, Magoun HW (1949) Brain stem reticular formation and activation of the EEG. Electroencephalogr Clin Neurophysiol 1:455–473


Wright JJ, Kydd RR, Sergejew AA (1990) Autoregression models of EEG. Results compared with expectations for a high order multilinear near-equilibrium biophysical process. Biol Cybern 62:201–210

Received: June 19, 1989
Accepted in revised form: October 19, 1989

Prof J. J. Wright
Department of Psychiatry and Behavioural Science
University of Auckland
School of Medicine
Private Bag
Auckland
New Zealand
Autoregression Models of EEG*

Results Compared with Expectations for a Multilinear Near-Equilibrium Biophysical Process

J. J. Wright, R. R. Kydd, and A. A. Sergejew
Department of Psychiatry and Behavioural Science, School of Medicine, University of Auckland, Auckland, New Zealand

Abstract. This paper considers the properties of parameters (natural frequencies and damping coefficients) obtained from segment-by-segment autoregression analysis of ECoG of rat. The use of a reference signal as control for parameter estimate errors, and multiple regression analyses indicate that the dependencies among parameters calculated from ECoG in the alert (desynchronised) state are of a form consistent with imposition of time-invariance assumptions (implicit in autoregression) on an inherently non-stationary, multimodal, linear and near-equilibrium "thermal" process.

1 Introduction

Characterisation of electrocorticogram (ECoG) or electroencephalogram (EEG) by autoregression (AR) and autoregression/moving average (ARMA) techniques has become widely applied, (eg. Al-Nashi 1986; Cerutti et al. 1986; Heinz and Kunkel 1984a, b; Ning and Bronzino 1987; Paarmann et al. 1987). In many circumstances these techniques are applied with intent only to compress data, rather than to obtain direct estimates of physically meaningful parameters (eg. Akaike 1974; Cerrutti et al. 1987; Crowell et al. 1977; Gath et al. 1983; Gundel 1983; Pomalaza et al. 1979; Simonsen et al. 1987; Vaz et al. 1987). The simplest physical interpretation that can be applied to autoregression coefficients is valid for low-order, time invariant linear filter processes driven by white noise, in which case the natural frequencies and damping coefficients of the system's resonant modes are calculable from solution of the characteristic equation defined by the autoregression coefficients (Gersh and Yonemoto 1977; Gersh et al. 1977; Freeman 1975; Franaszczuk and Biwowska 1985). There are excellent reasons for rejecting the most simple physical interpretation. Firstly, the system may be of very high order. Order optimising criteria such as that of Akaike (1974), yield the order which is most statistically efficient for forward-prediction. This need not correspond to the system order of the underlying physical process. Secondly, the EEG is actually non-stationary. Interpreting the physical meaning of the AR parameters therefore requires clarification of the class of dynamics applicable, as a prelude to system identification.

There is some reason to believe that the EEG can be treated as a multi-linear process. One set of arguments and experiments in favour of this view have been advanced by us (Wright and Kydd 1984a–c; Wright et al. 1984, 1985a–c; Wright 1990). We contend that the cooperative processes among cortical and subcortical neurons during conditions of strong "activation" (desynchronisation) lead to a group state which may be characterised as multilinear, and near-equilibrium in the thermodynamic sense. This implies superposition of group wave processes, and perturbation about an equilibrium point of maximum entropy and minimum free energy. Such an equilibrium requires near-equality of temporal damping for each of the multitude of possible group linear resonant modes (Wright 1990). These proposals regarding gross EEG activity in the alert state are derived from considerations of the extremely non-linear, chaotic and/or noisy properties of interactions between real neurons. This approach commits us to a "top down" analysis of EEG to determine relevant group properties, intended to complement the "bottom up" approach of models based upon specific simplifications of the neuronal elements (eg. Freeman 1975, 1987; Karwahrea 1980; Lopez da Silva et al. 1974; Nakagawa and Ohashi 1980; Nunez 1981; Wilson and Cowan 1973; Zhadin 1984). Earlier critical tests were contrived from measuring changes in relative amplitude and phase of
wave activity at laterally symmetrical electrocortical sites, recorded before and after unilateral lesion of activation pathways which relay through the lateral hypothalamus (Wright et al. 1985a–c).

The present paper follows the methods of Franaszczuk and Blinowska (1985) in applying AR to the EEG time-series. These authors have offered a parametrisation of the EEG signal as the superposition of variably damped resonant modes, and have also applied this parametrisation in an attempt to seek physical description, as well as for compact data reduction (Mitraszewski et al. 1987; Penczek et al. 1987). We attempt to extend this approach, considering further whether the results obtained are consistent with multilinearity and near-equilibrium dynamics in the underlying process, when a low-order, time-invariant statistical model is imposed on the signal. Use is made of an incidental interference signal (50 Hz mains interference) as a control signal to detect systematic error covariances among descriptive parameters, and other means are introduced to analyse parameter covariances which may not be implicit in the underlying ECoG process.

2 Expected Model Properties

2.1 The Underlying Process

The hypothesised underlying EEG process falls within a large class of models of the form

\[ \sum_{j=0}^{2p} a_j(t)x_{n-j} = e_n, \]

where \( x_n \) is the value of the sampled signal at \( t = n \cdot \Delta t \), \( n = 0, 1 \ldots 2p \) is the number of resonant modes and \( e_n \) is a noise process. Equation (1) has a continuous time form with a characteristic equation

\[ \prod_{i=1}^{2p} \left( s^2 + \mathcal{D}_i s + M_i^2 \right) = 0, \]

where \( M_i \) and \( \mathcal{D}_i \) are time-varying parameters.

Our hypothesis (Wright 1990) restricts the general model of (1) in the following ways:

(a) Model order is high, since the EEG may reflect hundreds or thousands of resonant modes.

(b) \( M_i \) are constants with respect to time.

(c) \( \mathcal{D}_i(t), (i=1 \ldots p) \) are independent of each other and of \( M_i \).

(d) \( \mathcal{D}_i(t) \) have mean values which are equal during an epoch of cortical "activation" (EEG desynchronisation) as a consequence of the approach of the electrocortical system to conditions comparable to thermal equilibrium.

2.2 The Statistical Model Process

The AR model of an EEG segment at an even order (Frasanzczuk and Blinowska 1985) is defined by

\[ \sum_{j=0}^{2q} a_j x_{n-j} = e_n, \]

where \( a_j \) are constants, \( j = 0 \ldots 2q \), \( a_0 = 1 \), \( e_n \) has zero mean. The set \( \{ a_j \} \) defines a discrete-time characteristic Eq. with roots \( z_j \) from which are obtained

\[ \omega_d(j) = \arg \left( \frac{z_j}{\Delta \tau} \right), \quad \beta(j) = \frac{\ln(\text{modulus } z_j)}{\Delta \tau}, \]

\[ \omega_d(j) = (\omega_d^2(j) + \beta^2(j))^{1/2} \]

from which we define the parameters \( \{ U_p, F_j \} \) for a continuous-time characteristic equation,

\[ \prod_{j=1}^{2p} \left( s^2 + \frac{s}{U_j} + F_j^2 \right) = 0, \]

where \( U_j = 1/\beta(j) \) and \( F_j = \omega_d(j) \). The change of symbols from conventional usage is for later convenience.

2.3 Model Equivalences

When an AR process of time-invariant form is used to approximate a time-varying process then the roots of the approximating characteristic polynomial represent the root loci of the actual polynomial, according to a criterion of minimized time-series prediction error – i.e., when \( s \) are poles approximating the pole loci of the underlying system characteristic polynomial in the epoch \( 0 \ldots \tau \) and \( s_m \) are the poles of the model system polynomial, then

\[ \prod_{i=1}^{2p} \left( s^2 + \mathcal{D}_i s + M_i^2 \right) = \sum_{j=1}^{2p} \left( s_m^2 + \frac{s_m}{U_j} + F_j \right)^{K_j} + E(s_m), \]

where \( \sum K_j = p \) and all \( K_j, U_p, F_j \) are single parameter values such that \( E(s_m) \) is approximate to zero.

2.4 Sequential Observations and Stochastic Dependencies of Parameters

Consider AR models obtained from a sequence of epochs of ECoG recording. Obtaining the optimum fit of (5) for each epoch is equivalent to finding optimum single poles, each to represent one cluster of a set of \( q \) clusters of roots of the underlying process. Each optimum is found when

\[ \left( s_m^2 + \frac{s_m}{U_j} + F_j \right) \rightarrow \sqrt{\prod K(s^2 + \mathcal{D}_i s + M_i^2)} \]

for each of the \( q \) clusters.

Variation of optimum model order and stochastic dependencies in the set of model parameters (of
constant model order) obtained for the sequence of epochs will arise even if the underlying process is essentially a near-equilibrium and multilinear process of the type hypothesised in Sect. 2.1, for the following reasons:

2.4.1 Optimum Model Order by Akaike Criterion. If the poles of (2) are tightly clustered in the complex plane, the AIC optimum order must tend to q, the number of clusters of poles, while conversely, for conditions with only weak clustering, optimum order will be variably determined for sequential epochs even though the underlying system does not vary in order.

2.4.2 Parameter Dependencies in the Model System. If a parameter set of order 2q (a fixed model order) is obtained from r non-overlapping epochs of similarly sampled ECoG we define \( \{ F_0, U_j \}_{j=1}^q \) to be the ordered set of all parameters obtained. Although \( \{ F_0 \} \) and \( \{ U_j \} \) are least biased estimators of \( \{ M_j \} \) and \( \{ 1/2 \theta_j \} \) respectively, (by a ready extension of the argument in Sect. 2.4), all members of \( \{ F_0, U_j \} \) must be stochastically dependent if the r samples are samples of an underlying multilinear process with time-varying damping, since parameters associated with adjacent frequencies will exhibit unique (type III) dependencies upon each other, consequent to sensitive interactions of \( K_p, P_p, U_j \) near convergence to a model optimum in each of the r data segments. This is because unique interactions in a cluster modelling problem of minimized distances to q cluster centers necessarily involves choices of members for adjacent clusters.

2.5 Expectations

2.5.1 Parameters from Optimum Order EEG Models. These tests are to show that the EEG is akin to filtered "thermal" signals in distribution of parameters.

2.5.1.1 AIC optimum model order should differ from that of noise driving of the recording system, and should exhibit more variable order.

2.5.1.2 Estimates of centre frequencies for EEG should be clustered about certain frequencies, and differ from expectation for the recording system alone.

2.5.1.3 Estimates of damping coefficients for the EEG should qualitatively resemble a thermal distribution.

2.5.2 Multiple Regressions among the Parameter Set Obtained with a Fixed Model Order. These tests are intended to show dependencies to be artefacts of modelling.

We formally define a set of multiple regressions (by a forward model with standardised variance) of the form \( x_t = R(\{ A_i \} | \{ B_j \}) \) applied to \( \{ F_0, U_j \}_{j=1}^q \), where \( x_t \in \{ F_1, ..., F_q, U_1, ..., U_q \} \) \( (F_j, U_j) \) ordered pairs by frequency of \( F_j \), \( \{ A_i \} \) and \( \{ B_j \} \) are subsets of \( \{ F_1, ..., F_q, U_1, ..., U_q \} \). \( \{ B_j \} \) are those parameters whose variance has been removed initially. The stepwise partitions of variance so defined can then be used to obtain:

(i) Variance in each parameter associated with a reference signal, as a test and control for structural error variance.

(ii) Individual error variance, to test and control for individual experimental subject factors such as electrode placement.

(iii) Variance attributable to direct interaction among \( \{ U_j \} \), after removal of other covariances, i.e. where \( \{ B_j \} \) includes covariance with reference signal, individual variance, and covariance with \( \{ F_j \} \), then only direct covariances among \( \{ U_j \} \) are modelled. This gives the least biased test of independence of the damping coefficients in the underlying process. Similarly, direct interactions among \( \{ F_j \} \) are a least biased test of independence of the underlying natural frequencies.

(iv) Unique (last out) interactions among all parameters. These are expected to be small, but where significant should exhibit only interactions between parameters of adjacent frequencies and their associated damping factors, as expected from Sect. 2.4.2.

2.5.3 Canonical Correlations among Parameters. If dependencies in \( \{ F_j \} \) and \( \{ U_j \} \) are purely factitious consequences of application of a model of low order, then canonical correlations of \( \{ F_j \} \) and \( \{ U_j \} \) should exhibit dependencies upon model order which are comparable for EEG and filtered noise. Wilks' Lambda, the likelihood ratio for the hypothesis that the set of canonical correlations between \( \{ F_j \} \) and \( \{ U_j \} \) is zero, permits the simplest comparison.

3 Methods

3.1 Animal Recordings

Electrocartiogram (ECoG) was obtained from a total of 24 male Wistar rats in the 300–400 gm range, with implanted bilateral bipolar platinum ball extra-dural recording electrodes, 0.5 mm in diameter. The electrode pairs were implanted at P7.5, L5.1, and P9.5, L5.1, from Bregma zero. After 10–14 days post-operative recovery, recordings were obtained in conditions of quiet rest (animals sitting or gently restrained). ECoG was obtained in 32 s blocks via unity-gain FET pre-amplifiers, Tektronics AM502 amplifiers and Butterworth filters, with a total gain of 5 K, and 3 dB points at 1 Hz (12 dB/octave high pass) and 35 Hz (48 dB/octave low pass). A/D converters digitised signals at 8 ms intervals. ECoG was subsequently selected by inspection of each timeseries, and those exhibiting high amplitude low frequency components.
in the theta or alpha range were excluded. Thus selection was for “desynchronised” EEG epochs, by traditional description. This selection was made because the hypothesis of multilinearity and near-equilibrium dynamics (Wright 1990) is not considered to apply to ECoG dynamics outside highly alert conditions. Data were not selected with regard to frequency content in any other way than slow-wave exclusion. Between 1 and 20 blocks were selected from each of the various animals, yielding 3104 two-s epochs per channel (left and right) for later AR analysis.

3.2 Noise Recordings

Nine data blocks per channel of identical format to the animal recordings were generated by feeding the output of a Hewlett-Packard Model 3722A low-frequency broadband noise generator through the same pre-amplifier/amplifier/filter system used for ECoG recordings. The generator provided random near-Gaussian input, effectively of flat-spectrum over the range of interest.

3.3 Autoregression

All data blocks (for both left and right channels) were divided into 2-s epochs (256 points per epoch). The procedure estimated autoregression coefficients by full unconditional maximum likelihood for each two-second epoch, by an algorithm essentially that of Harvey and Phillips (1975) (SAS/ETS Proc. AutoReg (SAS 1984)). A subset of data (3 blocks of data from each of twelve rats) underwent AR analysis for all even numbered model orders from 2 to 30. Subsequently the entire data set was analysed at order 12. All noise blocks underwent the complete analysis at all orders.

3.4 Estimation of Signal Parameters

By numerical solution of the characteristic equation defined by the autoregression coefficients, an equivalent set of natural frequencies and damping coefficients was computed from the analysis of each epoch at each order (Franaszczuk and Blinowska 1985).

3.5 Subsequent Analysis

3.5.1 Optimum Order Parameters. From the data subset with analysis available at all orders, the parameter set at each epoch at the optimum AIC order (first local minimum with ascending order of AIC) was selected for graphical purposes (see results).

3.5.2 Multiple Regressions among Fixed Order Parameters. The parameter values at order 12 for the entire data set (of each channel) then entered minimum least-squares multiple regression modelling. The six natural frequencies (in ascending rank of frequency) and the reciprocal of the six damping coefficients (each given the rank of their associated natural frequency) formed the parameter sets for multiple regressions between the resulting twelve parameters, as required by Sect. 2.4.2. The presence of 50 Hz interference provided a convenient reference signal (see results).

Reciprocal damping was chosen to parametrize damping, since the reciprocal of damping better met the parametrical statistical assumptions of distribution of variables required in multiple-regression modelling.

3.5.3 Canonical Correlation. This was separately performed on parameter sets obtained at each model order, for EEG and noise respectively (SAS Proc. CanCorr (SAS 1984)). A subset of the rat data available was chosen randomly, to balance the size of the animal recording and noise data sets in this procedure. Canonical correlations between \( \{F_i\} \) and \( \{U_j\} \) were then compared by Wilks’ Lambda.

3.6 Replications and Controls

Split-half replications were undertaken for all the analyses reported. These comprised between-channel replication, split-half by animal replication, and split-half by alternate EEG epoch, within the left-channel data. Multiple regression analyses omitting particular stages of the regression were performed to be sure that the conclusions reached were not critically dependent upon each early step in analysis. Noise-driven data was subjected to comparable regression analysis to that applied to the ECoG as a check that non-parametric distributions, numerical errors or other artefact had not contributed in an unexpected way to the results obtained.

These replications and controls supported the conclusions drawn, and are not further reported. Essentially, all conclusions were fairly robust even when individual error and reference signal errors were not partialled out (see results). Regression analyses of filtered noise parameters were in accordance with expectation.

4 Results

All results given in graphs and figures are for signals (ECoG or noise) input to the left amplifier/filter channel.

4.1 Parameter Distributions for Akaike Optimum Order Analyses

Figure 1 shows the average Akaike Information Criterion (AIC) as a function of model order, for both ECoG and noise input. Figure 2 indicates the probability that an AIC minimum will be identified at a particular order for a given data epoch. These figures confirm that ECoG has a higher average model order, with greater
Fig. 1. Aikake Information Criterion for different model orders. Continuous line — results for electrocorticogram of alert rats. Dashed line — for noise passed through identical amplifier/filter stages to the electrocorticogram.

Fig. 2. Probability that the optimum model order will occur at a particular order. Results from optimas of all epochs in the same data set used in Fig. 1. Continuous line — ECoG. Dashed line — filtered noise.

Variance than can be accounted for by sampling characteristics, recording and filter properties alone.

Figure 3 shows the density of distribution of natural frequencies obtained from the optimum order data set including a band attributed by us to 50 Hz interference. Such 50 Hz interference was not apparent in the ECoG time series and was minimal on spectral analysis of ECoG. No such interference was found in the noise-driven data sets. ECoG natural frequencies appear distinct from expectation for similarly filtered white noise, but are imprecisely separated or defined, particularly at frequencies near the filter cut off at 35 Hz.

Figure 4 shows the distribution of damping coefficients, matching the natural frequencies. The group of damping values corresponding to natural frequencies in the 50 Hz interference band are separately indicated. ECoG contains activity of low damping compared to expectation for similarly filtered noise, although the type of distribution of this parameter class appears similar for noise, ECoG, and interference, i.e., all are unimodal and skewed.
4.2 Multiple Regression Results

On the basis of the above, the analysis proceeded to systematic exploration of parameters obtained at order 12 from the total animal ECoG data set. This order was selected because (i) it was the median order found as described in Sect. 4.1, (ii) the more limited parameter set simplified linear regression modelling, compared to the order 14 data set (the mode order).

Figure 5 shows a random selection of order 12 parameters, plotted without regard to temporal order. Each epoch has contributed one and only one data point to each cluster of paired natural frequencies and damping estimates there displayed. Correction has been made for co-variance with the 50 Hz interference. Frequency estimates appear clustered at values near those expected from the optimum models (Fig. 3). The distributions of values of damping at each frequency differ significantly but the cluster means approach equality, as expected.

Tables 1–3 reveal patterns of correlation and covariance within the order 12 data set, after designation of the natural frequencies by rank (F1 to F6) in ascending order of frequency (F6 being associated with 50 Hz activity) and the reciprocal of damping (U1 to U6) in corresponding rank, by association with natural frequency.

Table 1 shows significant correlations are present between all parameters, including those associated with external interference.

<table>
<thead>
<tr>
<th>U2</th>
<th>U3</th>
<th>U4</th>
<th>U5</th>
<th>U6</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.16839</td>
<td>0.05242</td>
<td>0.04199</td>
<td>0.03294</td>
<td>0.17155</td>
<td>-0.0872</td>
<td>0.35282</td>
<td>0.22434</td>
<td>0.12682</td>
<td>0.11506</td>
<td>0.10664</td>
</tr>
<tr>
<td>0.0001</td>
<td>0.00035</td>
<td>0.0193</td>
<td>0.0665</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.37089</td>
<td>0.20406</td>
<td>0.16148</td>
<td>0.15279</td>
<td>-0.43695</td>
<td>-0.08154</td>
<td>0.30782</td>
<td>0.20689</td>
<td>0.17485</td>
<td>0.27102</td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.44400</td>
<td>0.31255</td>
<td>0.18825</td>
<td>-0.25282</td>
<td>-0.32837</td>
<td>-0.00307</td>
<td>0.38509</td>
<td>0.27550</td>
<td>0.36510</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.42123</td>
<td>0.24325</td>
<td>-0.18463</td>
<td>-0.16399</td>
<td>-0.20914</td>
<td>0.09694</td>
<td>0.43639</td>
<td>0.36745</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.34462</td>
<td>-0.19100</td>
<td>-0.09890</td>
<td>-0.09797</td>
<td>-0.22775</td>
<td>0.06232</td>
<td>0.45331</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>-0.11189</td>
<td>-0.08431</td>
<td>-0.07327</td>
<td>-0.04917</td>
<td>-0.00784</td>
<td>0.0625</td>
<td>0.26575</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>-0.08989</td>
<td>-0.16707</td>
<td>-0.13015</td>
<td>-0.12846</td>
<td>-0.35345</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.22714</td>
<td>-0.06733</td>
<td>-0.12957</td>
<td>-0.16794</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.30652</td>
<td>0.09066</td>
<td>-0.06162</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.29124</td>
<td>0.02374</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.1861</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.17219</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table 1. Matrix of correlations and associated levels of significance (below each correlation) for the parameters F (natural frequencies) and U (reciprocal damping coefficients) obtained from 12th order AR analysis of 2 second ECoG epochs. F1 to F6 are in ascending order of frequency. U1 to U6 are the associated damping estimates. F6 and U6 are associated with 50 Hz interference.
<table>
<thead>
<tr>
<th>Parameter predicted</th>
<th>$r^2$ Statistic for multiple regression</th>
<th>From reference signal (U6, F6)</th>
<th>From individual variation (Class variable main effect)</th>
<th>From all other parameters of same class</th>
<th>First out effect</th>
<th>Last out effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1</td>
<td>0.0334</td>
<td>0.1471</td>
<td></td>
<td>0.0184</td>
<td>0.1790</td>
<td>0.0041 0.1578</td>
</tr>
<tr>
<td>U2</td>
<td>0.0805</td>
<td>0.0718</td>
<td></td>
<td>0.0681</td>
<td>0.2053</td>
<td>0.0139 0.1431</td>
</tr>
<tr>
<td>U3</td>
<td>0.1423</td>
<td>0.0732</td>
<td></td>
<td>0.0829</td>
<td>0.2191</td>
<td>0.0175 0.1383</td>
</tr>
<tr>
<td>U4</td>
<td>0.1578</td>
<td>0.0880</td>
<td></td>
<td>0.0496</td>
<td>0.1694</td>
<td>0.0105 0.1581</td>
</tr>
<tr>
<td>U5</td>
<td>0.2596</td>
<td>0.1101</td>
<td></td>
<td>0.0170</td>
<td>0.0814</td>
<td>0.0146 0.0810</td>
</tr>
<tr>
<td>F1</td>
<td>0.1253</td>
<td>0.1574</td>
<td></td>
<td>0.1166</td>
<td>0.0611</td>
<td>0.0720 0.0227</td>
</tr>
<tr>
<td>F2</td>
<td>0.0299</td>
<td>0.1464</td>
<td></td>
<td>0.2367</td>
<td>0.0736</td>
<td>0.2086 0.0524</td>
</tr>
<tr>
<td>F3</td>
<td>0.0073</td>
<td>0.1221</td>
<td></td>
<td>0.2299</td>
<td>0.0853</td>
<td>0.1639 0.0282</td>
</tr>
<tr>
<td>F4</td>
<td>0.0039</td>
<td>0.0800</td>
<td></td>
<td>0.3020</td>
<td>0.1198</td>
<td>0.2058 0.0285</td>
</tr>
<tr>
<td>F5</td>
<td>0.0327</td>
<td>0.0582</td>
<td></td>
<td>0.1874</td>
<td>0.0680</td>
<td>0.1298 0.0177</td>
</tr>
</tbody>
</table>

Table 3. Values for unique predictions (last-out in forward model) for multiple regression among all ECoG 12th order parameters

<table>
<thead>
<tr>
<th>Parameter predicted</th>
<th>U1</th>
<th>U2</th>
<th>U3</th>
<th>U4</th>
<th>U5</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1</td>
<td>10.76</td>
<td>13.95</td>
<td>1.63</td>
<td>1.36</td>
<td>6.70</td>
<td>542.27</td>
<td>28.01</td>
<td>4.17</td>
<td>48.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0011</td>
<td>0.0002</td>
<td>0.2013</td>
<td>0.2429</td>
<td>0.0097</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0411</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>U2</td>
<td>27.99</td>
<td>37.88</td>
<td>0.40</td>
<td>42.97</td>
<td>360.07</td>
<td>79.43</td>
<td>289.08</td>
<td>0.57</td>
<td>4.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.5297</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.4489</td>
<td>0.0385</td>
<td></td>
</tr>
<tr>
<td>U3</td>
<td>15.56</td>
<td>64.68</td>
<td>17.34</td>
<td>0.69</td>
<td>0.25</td>
<td>292.73</td>
<td>38.20</td>
<td>425.46</td>
<td>5.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.4053</td>
<td>0.6160</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0187</td>
</tr>
<tr>
<td>U4</td>
<td>0.03</td>
<td>6.91</td>
<td>37.61</td>
<td>4.04</td>
<td>3.30</td>
<td>1.47</td>
<td>333.16</td>
<td>0.25</td>
<td>638.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0086</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0446</td>
<td>0.0694</td>
<td>0.2259</td>
<td>0.0001</td>
<td>0.6175</td>
<td>0.0001</td>
<td>0.0187</td>
</tr>
<tr>
<td>U5</td>
<td>0.08</td>
<td>2.25</td>
<td>44.16</td>
<td>40.64</td>
<td>0.59</td>
<td>2.56</td>
<td>0.87</td>
<td>459.90</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1338</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.4423</td>
<td>0.1098</td>
<td>0.3519</td>
<td>0.0001</td>
<td>0.7553</td>
<td>0.0001</td>
<td>0.0187</td>
</tr>
<tr>
<td>F1</td>
<td>1.29</td>
<td>420.42</td>
<td>0.03</td>
<td>1.00</td>
<td>3.98</td>
<td>129.27</td>
<td>0.87</td>
<td>9.24</td>
<td>5.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2560</td>
<td>0.0001</td>
<td>0.8653</td>
<td>0.3182</td>
<td>0.0463</td>
<td>0.0001</td>
<td>0.3513</td>
<td>0.0024</td>
<td>0.0151</td>
<td>0.0187</td>
</tr>
<tr>
<td>F2</td>
<td>763.35</td>
<td>58.09</td>
<td>344.66</td>
<td>1.45</td>
<td>0.04</td>
<td>327.33</td>
<td>10.09</td>
<td>7.28</td>
<td>31.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.8430</td>
<td>0.0001</td>
<td>0.0015</td>
<td>0.0070</td>
<td>0.0001</td>
<td>0.0187</td>
<td>0.0187</td>
</tr>
<tr>
<td>F3</td>
<td>96.62</td>
<td>346.07</td>
<td>36.25</td>
<td>352.84</td>
<td>1.77</td>
<td>29.26</td>
<td>0.39</td>
<td>111.59</td>
<td>34.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.1836</td>
<td>0.0001</td>
<td>0.5338</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0187</td>
</tr>
<tr>
<td>F4</td>
<td>36.39</td>
<td>0.73</td>
<td>522.83</td>
<td>0.01</td>
<td>700.27</td>
<td>16.49</td>
<td>28.73</td>
<td>92.72</td>
<td>41.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.3921</td>
<td>0.0001</td>
<td>0.9113</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0187</td>
</tr>
<tr>
<td>F5</td>
<td>40.89</td>
<td>1.07</td>
<td>0.15</td>
<td>527.05</td>
<td>103.42</td>
<td>0.09</td>
<td>44.83</td>
<td>24.08</td>
<td>27.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.3001</td>
<td>0.6949</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.7676</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0187</td>
</tr>
</tbody>
</table>
Table 2 indicates the amount of variance accounted for in each parameter by covariance with artefact, individual variation and interactions among the $U$ and $F$ parameter classes. These patterns of regression indicate that individual members of each of the two classes of parameter (the class of parameters $F$ and parameters $U$) have little direct dependence upon members of their own class when other sources of covariance are removed.

Table 3 shows that in most instances, the unique dependencies between all parameters, after removal of all other indirect covariance is generally between parameters associated with adjacent frequencies.

4.3 Canonical Correlation Results

Figure 6 shows the decline in likelihood of independence between $\{F_i\}$ and $\{U_i\}$ to zero for values equal to or greater than optimum model order, in both EEG and filtered noise.

5 Conclusions

Although the dynamic class of the EEG signal is not demonstrated in a deductive fashion by these experiments, it is shown that the electrocorticogram of alert rat has properties which conform to expectation for a multilinear near-equilibrium process. The tests of hypothesis are not exhaustive. However, results obtained at model order 12 are believed to be representative of parameter interdependencies at all orders, as is suggested by the results of the canonical correlation analysis.

The centre frequencies for the system are distinct from those expected from noise-driving of the amplifier-filter system, although in itself this is a trivial consequence of the difference in model order of EEG and similarly filtered noise. A pattern of harmonics at multiples of about 7 Hz is suggested by the present data. The estimates of the centre frequencies found represent compromise of a number of factors (particularly analog filter characteristics) with the ECoG frequencies. For any inferences to be drawn regarding center frequencies in the unfiltered EEG, allowance would need to be made for filter characteristics and for the effects of model order. An interesting possibility would be to compare the behaviour of natural frequencies and damping parameters of ECoG with a stationary filtered noise of identical spectrum to the average ECoG, thus permitting more precise comparison of natural frequency values and variation of damping factors with those of the EEG.

Again, comparison with noise-driven input, and with parameters estimating a 50 Hz imposed signal, suggests that ECoG has a distribution of damping coefficients comparable to that of a thermal process. While the 50 Hz signal does not exhibit strict stationarity in these estimates, variance of both noise and ECoG damping estimates are greater than those of the 50 Hz signal, and mean damping appears lower for ECoG than for filtered noise. This variability may be a consequence of the differing optimal order of EEG and the filtered noise signal.

When consideration is given to the stochastic dependencies between parameters estimating ECoG activity at fixed model order over many epochs, it is seen that strong dependencies are present, but these do not indicate that the underlying ECoG is similarly structured. All the dependencies, including the ECoG dependence with an extrinsic reference signal, conform to a predictable pattern expected in approximation of a higher order, unstructured underlying signal. Despite the high statistical significance, the low percentage of total model variance accounted for by direct interaction between damping coefficients alone, or natural frequencies alone, is a strong finding in favour of hypothesis. Canonical correlation analysis shows the dependencies between parameters to be little different for ECoG and noise, when allowance is made for differences in optimal model order.

Conversely, these findings do not exclude the presence of transient time-dependencies among parameters of the underlying signal. Time-dependencies among the damping coefficients are necessarily lost in the multiple regression analysis. We contend that specific sensory input signals organise time depen-
dencies among the mode energies in the near-equilibrium alert state, so that the maximum-entropy condition is a reference attractor state (Wright 1990).

These findings limit the classes of alternate dynamic theories of EEG in certain ways. Any model in which highly non-linear limit-cycles are involved is excluded, unless these are non-linearities in which no direct dependencies exist between the real parts or the moduli of the system eigenvalues. Although the results appear to favor a stochastic interpretation of the EEG, they are not necessarily in contradiction to evidence that the ECoG is a chaotic process (Freeman and Van Dijk 1987). It is possible that the signals driving the EEG "generators" may be a chaotic process of broad spectrum rather than the "white noise" assumed in AR modelling.

Finally, it should again be emphasised that we expect that ECoG obtained in conditions other than "desynchronisation" will obey other dynamic rules and produce results different to those reported here.

Acknowledgements. The authors are grateful to J. A. West and N. Hawthorne for technical services.

References


146 WORK TOWARD A THEORY OF BRAIN FUNCTION

Wright JJ, Kydd RR, Lees GJ (1985a) Quantification of a mass action of dopaminergic neurones regulating temporal damping of linear electrocortical waves. Biol Cybern 52:281–290


Wright JJ, Kydd RR, Lees GJ (1985c) State-changes in the brain viewed as linear steady-states and non-linear transitions between steady-states. Biol Cybern 53:11–17


Received: September 22, 1988
Accepted in revised form: September 13, 1989

Professor J. J. Wright
Department of Psychiatry and Behavioural Science
University of Auckland School of Medicine
Private Bag
Auckland
New Zealand
Inverse Filter Computation of the Neural Impulse Giving Rise to the Auditory Evoked Potential

James J. Wright*, Alexei A. Sergejew*, and Hans G. Stempfer **

Summary: An impulse response hypothesis for evoked potentials is tested. The auditory evoked potential (AEP) is shown to be the consequence of an impulse (the arrival of sensory signals in cortex) giving rise to an impulse response (the resonance of electrocortical activity in the form of group linear waves). To demonstrate this, pre- and post-stimulus EEG activity was recorded from subjects engaged in performance of an auditory oddball experiment. For each stimulus, the impulse required to account for the single auditory evoked potential (AEP) as a linear impulse response, was computed by use of the inverse of a filter obtained by autoregression analysis of the pre-stimulus EEG epoch. Single estimations of the impulse were then averaged. The average impulse exhibits a time course and topology consistent with the arrival of neural volleys in the cortex. The physical validity of the hypothesis is supported by a high lag correlation of the following values of the AEP to the average impulse. A further test calculation supports the linear additivity assumptions of the hypothesis.

Key words: EP; EEG; Impulse response; Wave linearity; Additivity; Averaging.

Introduction

We have better understanding of the physiological basis of the early components of evoked potentials than of the later components (Shaw 1988). Taking auditory evoked potentials (AEP) as an example, the cortical auditory evoked potential (CAEP) is associated with the earliest arrival of neural volleys in the auditory cortex (Hall and Barbely 1970; Shaw 1988). However, potentials in the middle auditory evoked potential (MAEP) - the P wave and related events - are less clearly understood, as to both the sites of their generation, and the pathways critically involved (Picton et al. 1974; Knight et al. 1985; Shaw 1988). Late evoked potential (EP) components, including N100, P300 etc. are even less well localised and understood, although their close correlation to cognitive and behavioural events has been extensively studied (e.g., Sutton et al. 1965; Pritchard 1986).

A factor complicating physiological analysis of the EP is uncertainty as to the most appropriate signal detection paradigm (Jervis et al. 1983). The conventional averaging method assumes a stationary noisy background EEG activity and an invariant additive signal produced by each stimulus. There is no reason to believe either assumption is correct in most situations (Sayers 1974; Squires and Donchin 1976; Van der Tweel et al. 1980; Jervis et al. 1983; Stempfer and Basar 1988). An alternative set of assumptions has been advanced by Sayers (1974), who suggested that the late components of the EP were not separable from the background EEG at all, but represented an influence of the input signal in resetting the phases of the background EEG rhythms. This approach has in turn been criticised because an additive signal against background noise would also exhibit the phase-locking described by Sayers, as well as explaining the changes observed in the energy content of the EEG spectrum when the pre- and post-stimulus epochs are compared (Jervis et al. 1983). There remains a significant unexplained feature - the similarity of spectral content in EP and background EEG (Stempfer and Basar 1988) - which may provide an important clue as to the type of additive model properly applicable to EEG, and EP.

Three quite different additivity assumptions need to be distinguished (McGillem and Aunon 1987). Figure 1
illustrates each of these models and indicates that certain physical assumptions have been made about EEG, partly to fit practical application. Firstly, in the "classic" paradigm used in signal averaging, it is assumed that the signal is completely invariant, and that both signal and noise pass through time-invariant filters. Next, in the more general Kalman filter and related approaches, both the signal and the background EEG are treated as time-varying additive signals (Childers et al. 1970, 1987; Gevins and Morgan 1986; Heinze et al. 1988; Ohshio 1981; Walter 1969). Finally, in the model introduced by Salomon and Barford (1977), the EEG and EP are treated as the output of a single time-varying filter, driven by noise, plus a time-locked input. It is contended here that the third (Salomon and Barford) alternative is physically valid, and incidentally, reconciles the different interpretations of the Sayers and Jervis findings. A biophysical justification for this approach is implied in at least three independent theoretical approaches to the EEG. The first of these (Lopes da Silva et al. 1974) focuses upon origin of alpha rhythm from resonance of local cortical circuitry. The second (Nunez 1981) emphasizes cortical resonance induced by correlated cortico-cortical association-fibre activity. The third (Wright 1990) emphasizes a permissive role of diffuse noise input from the reticular formation in generating rhythmic cortical wave activity from specific sensory input. While these three theoretical approaches differ significantly from each other, they all predict that electrocortical activity is a linear wave phenomenon generated by group activity of cortical neurones, and imply that time-locked sensory inputs may be treated as an impulse, that subsequently generates an impulse-response, within the electrocortical resonant system. If this view is correct, then by appropriate means, the EP (and in particular the AEP) should permit us to compute correctly the impulse signal arriving at cortical level from specific sensory inputs, by an application of the Salomon and Barford model, using an inverse filter based upon autoregression (Lopes da Silva and Mars 1987; Jazwinski 1970).

This paper describes a preliminary calculation of the impulse signal, and critical tests for the linear impulse response hypothesis. These tests show that the AEP exhibits a "following" response to a specific neural impulse, and that wave additivity assumptions are justifiable, within the impulse-response model.

Methods and Materials

(1) Method overview

To test the linear impulse-response model, we used the "predictive filter" technique (Jazwinski 1970), an inverse filter method which permits iterative forward computation of the impulse starting from the immediate pre-stimulus filter state. (Section (4) below). It will be shown that the AEP must exhibit lag correlation with the preceding values of the impulse, if the model is valid. We then reverse the calculations, and use the averaged input, and pre-stimulus filter characteristics, to recalculate the averaged AEP. Comparison of this recalculated AEP and the real AEP, tests assumptions of signal linearity. Brief consideration is given to the topology and context dependence of the impulse.

(a) Comparisons of autoregression models

Autoregression (AR) fits the EEG signal to a linear-filter model of the form.

\[ X(t) = \sum_{j=1}^{p} a_j X(t-j) + E(t) \]

where \( E(t) \) is a white noise input signal (plus a measurement error component) at time \( t \), \( X(t) \) is the EEG signal, \( p \) is the order of the model, and \( a_j \) are the AR coefficients. Application of this procedure can describe the EEG as a resonant process, driven by a white noise (Zetterberg 1969; Lopes da Silva and Mars 1987; Franaszczuk and Blinowska 1985), in which case it defines the transfer function corresponding to the EEG state prior to stimulation without ambiguity, in all of the three additive models. Averaging values of \( X(t) \) at a given value of \( t \), from many pre-stimulus epochs Eq (1) leads to \( \overline{X(t)} = 0 \) for all pre-stimulus \( t \).

It is useful to contrast the appropriateness of a post-stimulus AR model, in the "classical", and "linear response" models of Figure 1. For this purpose the classical averaging method can be used to represent properties of the Kalman type of model, also.

For the linear response model, post-stimulus,

\[ X(t) = A(t) - E(t) = \sum_{j=1}^{p} a_j X(t-j) \]

where \( A(t) \) is a time-locked impulse, delivered as a driving signal to the resonant system. Whereas, for the classical averaging model, post-stimulus
\[ X(t) - A(t) = \sum_{j=1}^{p} a_j [X(t-j) - A(t-j)] \]

since \( A(t) \) is not a driving signal of the same system as the noise. (For simplicity in comparing the two models, a filter operation on \( A(t) \), required to maintain strict comparability with Figure 1, is omitted in Eq (3) and in subsequent discussion of the classical model.) The average post-stimulus model of the linear response type is given by

\[
\bar{X}(t) - \bar{A}(t) = \sum_{j=1}^{p} a_j \bar{X}(t-j)
\]

and correspondingly for the classical type, the average is simply the time-locked input \( \bar{A}(t) \), at all post-stimulus \( t \), since in all cases random variables tend to zero, as the sample number tends large.

(b) Test for model validity

\[
\bar{X}(t) = \sum_{n=0}^{p} \prod_{n=0}^{n} w_j A(t-n)
\]

where \( w_j \) are weights derived from AR coefficients (see Appendix). This relation describes the way the impulse response "follows" the sequential discrete input steps, starting from \( t=0 \). Thus for a stable dissipative signal (such as EEG), the average computed input \( \bar{A}(t) \) is correlated with the evoked potential, \( \bar{X}(t+\Delta t) \) at a later (lagged) time, \( \Delta t \). The other additive models do not imply this relationship. If the linear impulse model is invalidly applied to parallel independent additive signals, then the invalid model cast in form analogous to Eq (5) is given by

\[
\bar{X}(t) = \sum_{n=0}^{p} \prod_{n=0}^{n} w_j B(t-n)
\]

Figure 1: Additive techniques for computation of the signal-locked component in the EEG. \( A(t) \) is a time-locked signal, \( E(t) \) a white noise. \( X(t) \) is the output recorded as EEG. Assumptions about signal transformation within the brain are represented as filter processes, \( F \). (a) The classical ERP averaging model (b) The most general additive model (c) The linear impulse model.
where

\[ B(t) = A(t) - a_1 A(t-1) \]
i.e., in this case, since \( A(t) \) is not a driving signal, but a parallel signal, \( X(t) \) does not "follow" the input, and \( B(t) \) cannot be lag correlated with later values of \( X(t) \). Thus, the presence of a strong lag correlation between values of the average computed impulse (computed iteratively from antecedent data) and the following AEP, is a test of the physical validity of the linear impulse response model. Where \( p \) the model order, is low the lag of correlation should be correspondingly short.

(c) A subsequent check on wave linearity

Given the average value of the impulse \( \overline{A(t)} \), and the pre-stimulus AR coefficients, Eq (2) can be used to recalculate a "noise-free" value of the EP for each recording sweep. These can in turn be averaged, and compared to the real AEP.

i.e. since

\[ \overline{A(t)} = \sum_{j=0}^{p} a_j X(\text{obs})(t-j) \]

and

\[ \overline{A(t)} = \sum_{j=0}^{p} a_j X(\text{calc})(t-j) \]

where \( X(\text{obs})(t) \) and \( X(\text{calc})(t) \) are the real EP and recalculated EP values respectively, then \( \overline{X(\text{obs})(t)} = \overline{X(\text{calc})(t)} \) only if the AR coefficients were errorlessly calculated and \( \overline{X(\text{obs})(t)} \) is a linear sum of underlying wave processes. It is useful to distinguish static non-linearities (in which the waves summate at all times according to the same non-linear function of total voltage) from dynamic non-linearities, in which the interaction of waves is also a function of time. The ratio \( \frac{\overline{X(\text{calc})(t)}}{\overline{X(\text{obs})(t)}} \) would act as a direct measure of the error attributable to a static non-linearity of waves, but the present method does not permit discrimination from errors due to a dynamic non-linearity. However a close match of real and recalculated EPs would confirm that the assumption of linear signal superposition upon which the modelling has been based, is justified.

(2) Subjects

These were five normal volunteers, aged between 19 and 26 years. Four were male, one female. All were right-handed. All gave fully informed consent.

(3) Recordings and experimental task

Subjects were seated on a comfortable chair in a darkened soundproof room. Auditory stimuli consisting of 50 msec tone bursts were delivered through head phones at a comfortable 70dB above individual subject hearing threshold (previously measured). Two different tones were used: 1550 Hz "target" tones and 1500 Hz "non-target" tones, each of which was delivered by gating in the appropriate one of two free-running oscillators by a monophase pulse.

Subjects were asked to count mentally the number of target tones and to ignore the non-target tones. This task was preferred for the present experiments because it was felt that attentional distractions, extraneous motor acts, etc. would be likely to disturb the stationary EEG state required for forward prediction, using stationary-state linear models (see Discussion). Short recording runs were preferred, again (hopefully) to optimise stationarity. Target tones were randomly interspersed with non-target tones. The inter-stimulus interval varied randomly between 1.54 - 8.0 seconds. Each experiment consisted of 80 stimulus presentations of which up to 30% were target tones. Blocks of stimuli were each delivered to either the right or the left ear. Subjects underwent both left and right testing, and all correctly completed the target tone count. A Sieman's source derivation EEG machine was used to obtain data from 19 electrodes placed in accordance with the 10-20 system for electrode placement. This machine computed the EEG at each channel by calculation of a nearest neighbour Laplacian derivative (Hjorth technique), as an embodiment in hardware. Linked earlobes were used as reference. The analogue filter bandwidth was set to 0.1-70 Hz. Each "sweep" (single stimulus epoch) of 256 pre- and 256 post-stimulus samples of data was acquired at a sampling rate of 3 msec.

Real time microprocessor controlled artefact rejection based on voltage limits was carried out by software during every sampling interval. Artefact contaminated sweeps were rejected and flagged on a VDU. The timing
Inverse Filter Computation of the Neural Impulse

and identity of presented stimuli was charted in parallel with eye movement activity recorded continuously during each experiment. Final off-line rejection of artefact contaminated sweeps was carried out by visual inspection of hard copy single sweep data and associated paper chart records of eye movement. In addition to this selection, data associated with some non-target tones was discarded, to balance the sampling. In every instance this was done by rejection of non-target data furthest separated in time from a target tone, leaving a total of 277 “sweeps” approximately equally divided between subjects, side of stimulus, and target condition. All data were stored in computer core memory as single sweeps, then transferred to magnetic media, and ultimately underwent analysis on a DEC Microvax II computer.

(4) Computing methods

Analysis was carried out on all nine “inner” channels of EEG. AR coefficients were estimated by the Maximum Entropy Method of Burg (1967) described by Andersen (1974). The optimum model order was found (for each pre-stimulus epoch) by application of the Schwartz Information Criterion (Schwartz 1978) after analysis at orders 2-30. The optimum model (typically of order 4 or 6) was used for all subsequent calculations. These results accord closely with those found for EEG using Akaike criterion (Lopes da Silva and Mars 1987).

Given the (pre-stimulus) computed values of $a_i$, and the $p$ last values of $X(t)$ in each “sweep’s” pre-stimulus epoch to initialise the computation, the impulse (noise plus time-locked input) for the post-stimulus epoch was computed by forward iteration of Eq (2) (the “discrete prediction filter”) (Jazwinski 1970). Each post-stimulus impulse thus computed was then averaged, over all sweeps from a single subject, channel, target condition, and side of stimulus. These averages (subsequently called “average impulse”), could then be pooled to demonstrate common features of side of stimulus, etc., as well as for computing the tests of hypothesis given in Methods 1(b) and 1(c).

Results

The averaged AEP results showed typical changes associated with target/non-target conditions, including P300 accentuation to target signals, as described elsewhere (Sutton et al. 1965; Stumper and Basar 1988). We here concentrate on those results directly bearing upon hypothesis. Figure 2 shows the grand average evoked potential, for channel CZ - all subjects, in both target and non-target conditions, and from stimuli delivered to either the left or right ear. CZ was selected as the most representative channel, because of the lateralisation effects apparent with side of stimulus - see below. Superimposed upon this trace, at a scale of twenty times the amplitude of the AEP, is the corresponding grand average impulse. It is seen that inflections of the average impulse precede the inflections of the average AEP, throughout the time course of the events. The average impulse magnitude most closely approaches that of the AEP during the earliest part of the post stimulus period, and is relatively attenuated in events after 50 msec post-stimulus. A lag correlation analysis showed that maximum correlation between the impulse and evoked response is found at a lag of 18 msec. for which $r = 0.8865$. The lag of 6 times the sampling interval of the recording, matches the low order of the optimum pre-stimulus AR models, as expected. Tables 1-3 show the results of lag correlation by channel and for differing subjects, side of stimulus and target condition. Figures 3(a) and 3(b) show the average result over all subjects, displayed for each channel for selected stimulus conditions - left-sided target tones have been chosen as representative of the
Table 1: Matrix of lag correlation maxima for AEP versus average impulse. Data averaged by channel. (Rig is measured by msec lead of impulse over AEP at maximum correlation. *r* is the correlation coefficient at that lag, and *n* is the number of “sweep” samples). Correlation was performed using data points 3-600 msec, and lags of 3-300 msec.

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>Z</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>r = 0.8562, lag = 15, n = 277</td>
<td>r = 0.8436, lag = 18</td>
<td>r = 0.8657, lag = 18</td>
</tr>
<tr>
<td>C</td>
<td>r = 0.9121, lag = 15</td>
<td>r = 0.8865, lag = 18</td>
<td>r = 0.9028, lag = 18</td>
</tr>
<tr>
<td>P</td>
<td>r = 0.8959, lag = 15</td>
<td>r = 0.8620, lag = 15</td>
<td>r = 0.8747, lag = 15</td>
</tr>
</tbody>
</table>

Table 2: Lag correlation maxima for channel CZ, by subject.

<table>
<thead>
<tr>
<th>Subject</th>
<th>r</th>
<th>lag</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.6455</td>
<td>12</td>
<td>44</td>
</tr>
<tr>
<td>2</td>
<td>0.8117</td>
<td>27</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>0.7972</td>
<td>12</td>
<td>62</td>
</tr>
<tr>
<td>4</td>
<td>0.7608</td>
<td>18</td>
<td>78</td>
</tr>
<tr>
<td>5</td>
<td>0.7550</td>
<td>15</td>
<td>53</td>
</tr>
</tbody>
</table>

Table 3: Lag correlation maxima for data averaged over all subjects/channels, by side of stimulus and target condition.

<table>
<thead>
<tr>
<th></th>
<th>Target</th>
<th>Non-target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>r = 0.7700, lag = 9, n = 51</td>
<td>r = 0.8042, lag = 15, n = 68</td>
</tr>
<tr>
<td>Right</td>
<td>r = 0.7826, lag = 15, n = 70</td>
<td>r = 0.7708, lag = 21, n = 88</td>
</tr>
</tbody>
</table>

principal features of importance. Figure 4 shows frontolateral leads over all conditions. It is seen from Figures 3 and 4 that while the computed impulse generally precedes the evoked potential response, the positive inflections of the impulse at 6-9 and 18-27 msec. show a sensitivity to side of stimulation and target condition. The net effect is that the fronto-lateral leads opposite the side of stimulation best show anticipation of evoked response by the computed impulse, during target-tone responses, for the earliest time intervals. A more widespread association of impulse and AEP is apparent for later events. These trends were apparent in the data of all individuals.

Figure 5 shows the grand average of the recalculated AEP superimposed upon the true grand average AEP, pooled in this case over all subjects, channels, and target conditions. Numerical errors in the recalculation could account for less than .01% difference. Errors in the estimates of AR coefficients will have contributed to the total error. A high degree of static non-linearity in the EEG and EP is not indicated by these results, for which the upper bound of error is 9.1% at t = 108 msec.

Discussion

Although the sample of subjects and data is small in this study, intersubject and interchannel consistency and significance appears high. Therefore the Salomon and Barford, or linear impulse, hypothesis is supported. The computed impulse precedes the AEP in the fashion expected for a following response, generated in a resonant wave system. The early components in the average impulse clearly precede wave forms within the first 50 msec. of the AEP, which we take to be the CAEP and MAEP, though the wave forms are imperfectly resolved at the sampling rate and sites, of recording applied. This we take as evidence of the arrival of sensory pathway neural volleys at special sensory cortex, and the shift of early impulse inflections with change of side of stimulus supports this. The relatively high-voltage N100, P300, and other peaks in the AEP are seen to be accounted for by quite small after-oscillations in the late phases of the impulse. These late phases of the impulse are amplified considerably in the AEP, presumably by the developing resonant response in the extended cortex. The match of the reversed computation of the AEP to the averaged impulse, supports the validity of assuming electrocortical signal linearity, with background “white” noise driving. These findings do not exclude the possibility of
Figures 3(a) and (b): The average AEP and average impulse for left ear target tones, pooled over subjects, for each channel. For clarity, the same data are represented on a restricted time-base, (Figure 3(a)) and a more extended one, (Figure 3(b)). Figure 3(a) - Impulse solid, AEP dotted line Figure 3(b) - Impulse jagged, AEP smooth line. In each instance the impulse is shown at two times its true amplitude to the AEP and the voltage scale is that of the AEP.
distortions in the computed impulse, attributable to time-variation in filter state from pre-stimulus, to post-stimulus epochs.

All three linear models for electrocortical activity cited earlier (Lopes da Silva et al. 1974; Nunez 1981; Wright 1990) are supported by these findings, which meet an explicit prediction of the third model. The linear impulse model also reconciles the Sayers (1974) and Jervis et al. (1983) viewpoints, by indicating that EPs fit a type of additive model in which input signals act to organise amplitude and phase relations in ongoing EEG activity, and this, in turn raises questions about the mechanisms of endogenous EEG generation. The ongoing EEG may be best conceived of as a sum of internally and externally generated linear impulse responses (Franaszczuk and Blinowska 1985; Wright 1990).

The linear impulse technique of analysis of the ERP has a number of consequences for further work, if it is found to be generally valid. It permits the specific sensory input (impulse) to be distinguished from the resonant response characterising the brain's state at the time the signal is delivered. Both the impulse, and the resonant response appear to be set dependent, as suggested provisionally by the very early differences seen in the computed impulse for target and non-target tones, in the present data. This may permit useful distinction of a global brain set, from a state of input filter set exerted upon the sensory pathways. Further, the present models could be extended to include consideration of time-variation in electrocortical filter state. We suspect that EP task paradigms involving marked shift of attention or arousal during the task will require application of time-varying models (Gersch 1987), to account for the EP phenomena then observed.

Finally, the present technique has implications for the analysis of P300 abnormalities in conditions such as schizophrenia. These will be reported elsewhere.

Appendix

Consider a special case of Eq(2), the response function for unit impulse:

\[ X(t) = \sum_{j=1}^{p} a_j X(t-j) + A(t) + E(t) \]

where \( E(t) = 0 \) for all \( t \), \( A(0) = 1 \), \( A(t) = 0 \) for \( t \neq 0 \)
Then by definition,

\[ X(0) = a_0 A(0) = A(0) \]

\[ X(1) = a_1 A(0) \]

\[ X(2) = a_1 a_2 A(0) + a_2 A(0) \]

\[ X(3) = a_1 a_2 a_3 A(0) + a_1 a_2 A(0) + a_2 a_1 A(0) + a_3 A(0) \]

etc.

That is, each \( X(t) \) is equal to \( A(0) \), multiplied by a different linear combination of products of the AR coefficients.

Also, for \( t = n, n = 0 \ldots p, X(t) \) may be represented as

\[
X(n) = \prod_{j=0}^{n} w_j A(j)
\]

where \( w_0 = 1 \), and \( w_j = X(j) / X(j-1) \), for \( j = 1 \ldots n \), for the unit impulse case.

Thus \( w_j \) are ratios of linear combinations of products of AR coefficients, and are independent of the particular input.

It follows from the principle of superposition in linear systems, that for the general input \( A(t) \),

\[
X(t) = \sum_{n=0}^{p} \prod_{j=0}^{n} w_j A(t-n)
\]

and the time-locked average for signals in which \( E(t) = 0 \) is given by Eq (5).

References


Squires, K.C. and Donchin, E. Beyond averaging : the use of
Radial coherence, wave velocity and damping of electrocortical waves

J.J. Wright and A.A. Sergejew

Department of Psychiatry and Behavioural Science, University of Auckland School of Medicine, Auckland (New Zealand)

(Accepted for publication: 26 March 1991)

Summary
Mean squared coherence was calculated as a function of frequency (1–32 Hz) and electrode separation (1–8 mm) from 64-channel extradural arrays on occipito-parietal association cortex of cats. A two-parameter theoretical function was then fitted to sets of pooled estimates. The theoretical function described coherences between recording sites of small separation for linear, non-dispersive, dissipative waves moving on an infinite homogeneous plane medium, and driven by spatio-temporally noisy inputs. Residuals of fit were then plotted as a function of frequency and distance, and were found to show no systematic trends with frequency, but an irregular and generally increasing relation to distance. This was the result predicted for linear non-dispersive waves on a surface actually folded, and with significant additional wave action generated between electrodes. Further recordings of coherence from more widely separated electrodes indicated that boundary conditions were absorbing or remote, rather than closed or reentrant. The phase velocity for electrocortical waves obtained from autoregression estimates of temporal damping and parameters of fit to coherence, was found in the range 0.1–0.29 m/sec, and appeared independent of the direction of electrode alignment. This compares with the velocity of 0.33 m/sec for alpha waves earlier found by Lopes da Silva and Storm van Leeuwen.

Key words: Coherence; Electrocortical wave; Linearity; Wave velocity; Damping; Boundary conditions; EEG model

Controversy surrounds the origins and significance of cerebral rhythmic activities and electrocortical waves, even as to whether intrinsic rhythmic properties of single neurones are essential for the genesis of brain waves of different frequencies or whether such rhythms basically emerge from synaptic interactions of large neuronal pools. Steriade et al. (1990) have undertaken an eclectic review of the topic, and show the intrinsic rhythmicities of certain neurones to be of importance in sleep spindles, the generation of some alpha activity at least, and in certain slow waves. The generation of waves in alert, 'desynchronised' states is apparently a more complicated matter.

It appears established that EEG activity covaries closely with cognitive processes (Gevins 1988), and arises from a small cross-correlated component in cortical dendritic potentials (Elul 1972a,b; Bullock 1988; Mitzdorff 1988). It remains a matter of controversy whether the EEG is integral to the processing of information in the cortex, or is an epiphenomenon of slight functional significance compared to the details of the exchange of action potentials among cortical neurones (Stryker 1989). Some resolution of this problem is being achieved in studies of visual cortex information processing (Eckhorn et al. 1988; Grey et al. 1989) and this resolution would be assisted by clearer understanding of the dynamic laws of the EEG. Unfortunately, dynamic properties of electrocortical waves are hard to establish from multi-electrode recordings, because of the transformations due to volume conduction, the complicated cortical surface topology, the multiplicity of possible wave sources, and the risk of spatial aliasing (Nunez 1981). Without knowing either the initial conditions, boundary conditions or wave dynamics, little can be inferred from topological descriptions. An important exception is the measurement of the velocity of alpha activity in dogs spreading from cortical epicentres over distances of up to 4 mm travelling at about 0.33 m/sec (Lopes da Silva and Storm van Leeuwen 1977).

In contrast to topological analysis, other workers have used estimates of coherence to try to define functional relationships between cortical areas. This measure is subject to criticism since it is a linear correlation technique applied to what may be nonlinear signal relationships (Salzberg et al. 1986). Also, the use of common ground references has been shown to yield unreliable estimates of coherence when applied to scalp EEG (Fein et al. 1988). A recent study (Bullock and McClune 1989) defined some general properties of cortical coherence, and showed that coherence falls off with both frequency and distance, over small distances (0.5–10 mm) on the cortical surface of rats and rabbits. Sharp discontinuities of coher-
ence were found even at small radial distances, and over larger distances, values tended to zero. However, other studies have shown quite high coherences between widely separated sites — for example, at symmetrical sites, where callosal connections make left/right interaction strong (Tenhouten et al. 1987) and as a reflection of short and long cortico-cortical fibre connections (Thatcher et al. 1986).

Explanation of these millimetric scale electrocortical interactions is problematic. It is unclear whether EEG waves are true waves — i.e., activity propagated from site to site in the cortex, thus implying lawful interaction, or pseudowaves — i.e., activity arising purely from co-activation of sites on a non-transmitting medium (Franaszczuk et al. 1985). There are apparently contradictory views concerning whether (at one extreme) they are linear superposition waves (Frasnaczuk and Blinowska 1985) or (at the other extreme) they are highly non-linear chaotic events (Freeman and van Dijk 1987). Detailed cellular models have been proposed to attempt linkage of wave dynamics to individual neuronal activity, but relatively few tests of hypothesis have yet been undertaken (Lopes da Silva et al. 1974; Nunez 1981, 1989; Van Rotterdam et al. 1982; Ingber and Nunez 1990).

In an attempt to help bridge the gap between theoretical cellular modelling and macroscopic EEG findings we have proposed a qualitative model for the dynamics of cortical neuronal networks which directs attention to the role of the reticular activating system in maintaining the condition for wakeful consciousness. It is assumed that the reticulo-cortical inputs provide a strong spatially and temporally noisy input to cortical columns. Stochastic linear differential equations can then be used to represent the dynamics of the superficial cortical dendritic networks, although the underlying axo-synaptic and other neuronal interactions are accepted as being highly non-linear and possibly chaotic (Wright 1990). At the lowest frequencies of resonance (corresponding to the EEG range) and at the macroscopic scale the cortical dendritic system is predicted to exhibit linear wave superposition and non-dispersive transmission, i.e., transmission of waves of all frequencies at a common velocity. In this model the non-dispersive nature of the waves is a simple consequence of conduction from point-to-point in cortex by modulated axonal transmission. Further consequences include wave generation by coherent sensory inputs, as well as the internal generation of coherent signals by coactivated neurones at widely separated sites. The EEG is thus envisaged as reflecting processes which link together neuronal interactions at long ranges, not despite, but because of, a high degree of noise and chaos in the detailed cell firings.

Tests of predictions made from the above model have shown that the auditory evoked potential represents the following response of the electrocortical resonances to specific sensory input (i.e., that the evoked potential is a linear impulse response) rather than a signal independent of the ongoing EEG, as is conventionally assumed (Wright et al. 1990b), and that in the desynchronised state the EEG approaches a condition analogous to thermodynamic equilibrium, with random perturbation of the energy of resonant modes about a state of energy equipartition (Wright et al. 1990a).

The above model could contribute to the development of specific cellular models if it can be used to provide quantitative measures such as wave velocity, so long as the range of conditions in which it is applicable can also be determined. The present paper describes the fitting of a theoretical equation to data on coherence derived from electrode sites with millimetric separations. We have been able to provide a specific test for the hypothesis that waves in the desynchronised EEG are linear and non-dispersive, by examining the distribution of residuals in the model fit, while also obtaining an estimate of EEG wave velocity.

**Coherence in an idealised plane-wave model**

Arguing from properties of coupled second order stochastic systems of high dimension, our 1990 model gives reasons to expect EEG in the desynchronised state to have wave properties which are equivalent to waves in an extended system of symmetrically coupled identical simple harmonic oscillators with uniform damping. We can further idealise the model to that of a 2-dimensional sheet of coupled elements with absorbing bounds, or bounds at infinity, on which waves arise from white or coloured spatially and temporally noisy input signals. Waves of frequency $\omega$ will sum at any point $a$, to a single wave of amplitude $A(t_0)$ at time $t_0$, which then moves on at constant velocity toward some other point, $b$, dissipating energy both from temporal damping, D, and from dissipation due to incomplete transmission of internal energy from excited harmonic elements to their coupled neighbours during each wave cycle.

For simplicity of argument we first consider plane waves moving in the line $ab$, then generalise the case to multi-directional waves.

Allowing for temporal damping (i.e., frictional loss) only, the wave amplitude $A(t_0)$ at $a$, will have attenuated to $Ae^{D/2}$ when passing point $b$ at time $t$ where $b$ is at a distance of $R = ct$ from $a$, where $c$ is the wave velocity. Thus temporal damping alone attenuates the wave amplitude as $e^{-DR^2/2c}$.

Allowing for spatial dissipation only, if $1/K$ is the fraction per unit energy of internal energy remaining in each excited oscillator after one cycle, i.e., the fraction of energy not immediately propagated to its coupled
neighbours lying in the direction of wave propagation, then the wave amplitude $A(t_0)$ at $a$, will attenuate to $A e^{-n/2K}$ in the component reaching $b$ at $t$, where $n$ is the number of cycles gone through as the wave moves from $a$ to $b$. Since $n = R/λ$, where $λ$ is the wavelength, and $λ = c/ω$, then $A e^{-n/2K}$ may be rewritten as $e^{-SωR/2}$, where $S = 1/Kc$.

The total attenuation of amplitude from temporal damping and spatial dissipation is $e^{-(DR/c + SωR)/2}$.

The bidirectional wave case can be generalised by considering all waves crossing the line $ab$ at an angle $θ$, where $0 ≤ θ ≤ π$ for waves from half the field, crossing $a$ to $b$.

Wavefronts cross $a$, then $b$, in a time $R \sin θ/c$, having travelled a distance $R \sin θ$. By integration and application of the mean value theorem, the average time of transit is

$$t_{ab} = \frac{R}{c} \left( \frac{1}{π} \int_0^{π/2} \sin θ \cdot dθ = \frac{2}{π} \right) R$$

and the average distance travelled is

$$r_{ab} = \frac{2R}{π} \left( \frac{1}{π} \int_0^{π/2} \sin θ \cdot dθ = \frac{2R}{π} \right)$$

The total attenuation of amplitude from temporal damping and spatial dissipation is therefore approximated by $e^{-β/2}$, where

$$β = \frac{2}{π} (DR/c + SωR)$$

We can initially neglect waves originating between electrodes of small separation, so if $A_{ab}$ is the root mean square value of the amplitudes of waves passing $a$, travelling from $a$ to $b$, and $A_{ba}$, the corresponding value for waves moving from $b$ to $a$, and we make the further provisional assumption that all wavelengths are large compared to the electrode separation, so that phase lags between $a$ and $b$ are small for all frequencies and angles of incidence, then the cross-spectral power of signals at $a$ and $b$ is

$$σ_{ab}^2 = \left( A_{ab} \cdot A_{ab} e^{-β/2} + A_{ba} A_{ba} e^{-β/2} \right)^2$$

The signal powers at $a$ and $b$, $σ_a^2$, $σ_b^2$ are

$$σ_a^2 = A_{ab}^2 + A_{ba}^2 e^{-β}$$

$$σ_b^2 = A_{ba}^2 + A_{ab}^2 e^{-β}$$

The squared coherence between $a$ and $b$ is

$$r^2 = \frac{σ_{ab}^2}{σ_a^2 σ_b^2}$$

$A_{ba}$ for spatially random driving signals, so by substitution and cancellation

$$r^2 = \frac{4e^{-β}}{(1 + e^{-β})^2} \quad (1)$$

By fitting of Eq. (1) to experimental mean squared coherence measured as a function of frequency and electrode separation on the cortical surface, we can obtain estimates of the parameters $D/c$ and $S$.

If mean squared coherence $r^2$ is to be used to estimate $c$ and $S$, then $D$ must be independently obtained. This can be achieved from autoregression analysis of recordings from each of the single channels, since the mean damping of the surface elements can be shown to be equal to the mean of the damping coefficients obtained from the system eigenvalues (Frañaszcuk and Blinowska 1985; Wright 1990; Wright et al. 1990a).

Test of hypothesis, and consequences of failures of assumption

Eq. (1) fitted to experimental measures of squared coherence as a function of frequency and distance permits a strong test of the hypothesis of linear non-dispersive dissipative wave propagation, as follows:

Eq. (1) can be re-written as $r^2 = f(\omega)$, where $g = R/c (D + ω/K)$ and $K, c$ and $D$ are constants for linear non-dispersive waves in conditions of stable temporal damping.

The residuals of fit, given by

Residual $r^2 = \text{experimental } r^2 - \text{theoretical } r^2$

will be systematically distributed with respect to either distance of electrode separation, or frequency, or both, depending upon various limitations (i.e., failures of assumption of the model). Topological distortions of the ideal plane, i.e., surface folding, and wave motion arising between the electrodes will cause the estimates of $R$ measured simply from the recording electrode separations to be in error, with proportional error in the estimates of $c$ (and $S$), and systematic distribution of residual $r^2$ with respect to $R$. However, for a given electrode separation the error in the apparent value of $R$ will apply equally at all frequencies. Therefore systematic residuals with respect to $ω$ at any value of electrode separation are not expected. Random residuals with regard to frequency at all electrode separations is thus a strong test that $c$ is, in fact, a constant independent of $ω$, i.e., that the waves are both linear and non-dispersive.

The following effects of failures of particular assumptions need to be considered:

(i) If $A_{ba} ≠ A_{ab}$. This would occur if electrocortical waves had a substantial, and continuing, directional component, e.g., evoked potentials arising from a specific locus being large compared to other sources. In this case $r^2$ would tend to 1, for all affected electrode pairs, over relevant frequencies. Multiple wave sources which were not adequately approximated as noise,
could produce (in principle) any pattern of residuals to the model fit whatsoever.

(ii) If waves crossing a to b were coherent with waves crossing b to a. This would occur if boundary conditions were reflective, or closed, in such a way as to permit significant 're-entry' of waves. In this case persistent standing waves would develop, so that \( r^2 \) would be high at widely separated cortical sites.

(iii) As the wave medium is not a simple plane, but a folded cortical surface, then idiosyncratic relations of residual \( r^2 \) to distance but not frequency, should arise for electrode pairs of particular position and orientation, particularly where sulci have been crossed.

(iv) As significant wave generation may be expected from sites between electrode pairs then coherence would be greater than predicted from Eq. (1). This effect is liable to be greater for more widely separated pairs than those closely positioned, so that residual \( r^2 \) should increase with electrode separation, but not frequency.

Departures of experimental \( r^2 \) from Eq. (1) of the types (iii) and (iv) above, are therefore to be expected, while limitations of types (i), and (ii) might disturb the test of hypothesis to a variable degree. These matters are again taken up in the conclusion.

**Method**

**Animals**

These were 4 young adult female cats, subsequently referred to by their 'in-house' names.

**Sixty-four-channel electrode implants**

These were pre-constructed for surgical implantation to cats, and in 3 cases, constituted an \( 8 \times 8 \) regular square array of platinum ball cortical electrodes, of diameter 0.4–0.5 mm each, and separation distances as described below. These distances and ball sizes were aimed at minimising spatial aliasing for corollary studies of Fourier spectra in space and time (Nunez 1988). The platinum electrodes were held within a gelatine film which could be applied to the cortex, and each electrode was soldered to a miniature connector, the wiring being secured within an insulating synthetic rubber extrusion. In one cat's case (Simone) the electrode separation was 1 mm between the edges of the recording balls, and 0.86 mm in the case of two others (Germaine and Kate). The fourth cat (Sandra) had electrodes sited in a \( 4 \times 11 \) array, flanked by 2 columns of \( 1 \times 11 \), moulded to the full extent of the dorsal cortical surface accessible surgically. The separation of adjacent electrodes in Sandra's case was 2.3 mm in the antero-posterior and 2.25 mm in the lateral direction.

**Surgery**

All procedures were approved by the University of Auckland Medical School Animal Ethical Committee. Each cat underwent implantation under general anaesthesia, with skull fixation in a stereotaxic frame. Scalp and rostral temporalis muscle were reflected, and vault of skull removed to the right side of the midline (in the cases of Simone and Kate), so as to admit the recording array to the dural surface, overlying the occipitoparietal cortex. In Germaine's case the array was implanted over the left cortex, laterally symmetric to the former 2 cats' arrays. Lateral placement and orientation of the array on the dural surface varied slightly between subjects. The medial row of electrodes was placed within 2 mm of the medial edge of cortex adjacent to the falx cerebri, and thus the array overlapped the prominent superior sulus of the cat, which traversed the array variably between subjects, in a generally antero-posterior direction. The array was centred on the most superior and mid-portion of the cortical surface, so that the electrodes overlay association cortex rather than special sensory areas. Sandra's array was applied over the entire accessible right cortical surface. Skull screws were introduced elsewhere on the cranial vault, and formed the reference earth. The electrode assembly and screws were then embedded in dental cement, and scalp skin replaced about the implant connector. Post-mortem confirmation of absence of damage to the cortical surface has been obtained for all animals, and showed that electrode placement was as intended.

**Recording method**

In sessions beginning at least 3 weeks after surgery, the cat was seated comfortably on an experimenter's knee, and subsequently kept (by the experimenter's interaction with the cat), in a state of wakeful alertness, without other external distraction. A connector containing the pre-amplifier stage of the recording system, was gently inserted by the use of counter-screws to the electrode implant plug. During the acquisition of EEG the cat was not moved by the experimenter, did not move itself, and was in eyes-open condition. Thirty-second continuous acquisitions of EEG were then obtained using the system now described.

**Sixty-four-channel recording system**

The pre-amplifiers (operational amplifier buffers) were pseudo-differential with the ground reference connected to a skull screw, and their frequency response was effectively flat, with negligible cable artefact and channel cross-talk. The frequency response of each amplifier was set by a high-pass filter of 12 dB/oct slope and 3 dB point at 0.234 Hz, and a low pass filter of 48 dB/oct slope and 3 dB point at 34 Hz. The amplifiers were tuned to match frequency re-
sponse, so that the complex phasor gain difference was less than 5% between 1 and 30 Hz for any pair of amplifiers. All outputs were sampled simultaneously (via sample and hold arrays) at 8 msec intervals, digitised by a 12-bit A-D converter, and transferred to magnetic media.

Data selection

Each of the 64 channels of each acquisition file was first computer-checked for the presence of gross artefact of the following type — DC segments, signal clipping by A-D converters, and the presence of voltage measurements beyond statistical expectation for the time-series (i.e., number of data points beyond the mean ± 3 S.D.). Data not fulfilling these criteria were rejected (there was actually little apparent artefact) and 10 acquisition files were arbitrarily selected for each cat, from the acceptable files.

Electrode configurations for analysis

For mean squared coherence computations, electrodes in each animal’s array were treated as members of pairs in each of 4 configurations, and each configuration was subsequently analysed independently, viz.:

Rows — all electrode pairs oriented in (roughly) the transverse direction. Columns — all electrode pairs oriented in the antero-posterior direction. Diagonal 1 — all electrode pairs on the main diagonal slope (left upper to right lower electrode, viewed from above, facing forwards), and parallel to this across the array. Diagonal 2 — all electrode pairs arranged at right angles to diagonal 1.

Coherence analysis

For all electrode pairs of the configuration, the mean squared coherence was computed for the frequency range 1–32 Hz, at 1 Hz intervals, using 1 sec epochs of the acquisition. Time series were trend corrected by classical least squares and tapered by a 10% split cosine bell window, prior to estimation of the unsmoothed spectrum by a 512-point Fast Fourier transform (the last 384 points set to zero). Smoothing in the frequency domain was performed by a triangular window of width $2\pi/32$, which gave a moderate degree of smoothing with a 1 Hz frequency interval. Identical windowing and smoothing were used in the estimation of univariate- and cross-spectra.

Mean squared coherence was then plotted as a function of frequency, and of electrode separation, for each configuration (i.e., rows or columns or diagonals) within each cat’s array, for the averaged results of 1 sec epochs from each 30 sec acquisition. In each of the cases of Simone and Kate, a single channel was omitted from analysis, as these were suspected of defect during implant construction.

Estimates of temporal damping

Each channel of each acquisition underwent autoregression analysis based upon 1 sec epochs, as for coherence. The Maximum Entropy method (Burg 1967; Andersen 1974) was applied at orders 2–32, and the optimum order ascertained by Schwartz criterion (Schwartz 1978). AR coefficients from each optimum order model were then used to compute the resonant mode damping coefficients (i.e., the real parts of system eigenvalues associated with natural frequencies in the 1–32 Hz range) (Franaszczuk and Blinowska 1985) and all values were then averaged over all epochs, and all channels of the 30 sec acquisition, to provide a measure of $D$, the temporal damping. Inspection of the distribution of damping coefficients associated with activity at different frequencies in the 1–32 Hz band, showed damping values to be independent of associated frequency values in accord with our earlier finding of apparent equipartition of energy across modes, as reported in Wright et al. (1990a). This we took both as justification of our averaging their values to obtain a single estimate, and a part criterion of the presence of a ‘desynchronised’ state, as defined by the assumptions of Wright (1990). Fast Fourier average normalised power spectra calculated in association with the coherence estimates showed a typical parabolic decline of power with frequency, as also expected for equipartition of energy over all frequencies. Specific tests for signal stationarity were not applied.

Surface-fitting

Eq. (1) was surface-fitted by least-squares minimisation, to the estimates of average coherence (averaged over 1 sec epochs of a 30 sec continuous acquisition as described above) in each configuration, to yield estimates of $D/c$ and $S$. The residuals to be minimised (the root mean-squared deviations of $r^2$) were weighted by the square root of the number of electrode pairs at any given separation distance with the intent of achieving a fit optimal for total standard error. Minima were robustly found from a wide range of starting values, and run times were short on the Silicon Graphics 4D/240S mainframe used. For further purposes (see Fig. 1 and Table II) we also performed surface fits for averaged 1 sec coherence over larger total times of acquisition.

Analysis of residuals

With the values of the parameters $D/c$ and $S$ obtained from surface-fitting, theoretical surfaces of best fit were plotted, as were residuals of each fit, calculated by subtracting the theoretical $r^2$ surface from the experimental $r^2$ surface.

The residuals were then surface-fitted themselves, by a plane of form

$$\text{Residual } r^2 = A\omega + BR + C$$
where \( \omega \) was the frequency in radians/sec and \( R \) the interelectrode distance in metres, with the parameters \( A, B, C \). This permitted a concise estimate of the model's trends of fit in the frequency (parameter \( A \)) and distance (parameter \( B \)) dimensions respectively.

**Results**

Fig. 1 shows representative examples of the experimentally derived plots of squared coherence versus frequency and distance, the theoretical surfaces of best fit corresponding to each data set, and the distribution of residuals in each case. It should be noted that the illustrations are based on pooled data from larger subsets of recordings than 30 sec, and that a representative example has been chosen from each of the 3 cats, at differing electrode orientation in each case. It is seen that the characteristic result is a good correspondence of theoretical surface to data at all frequencies, but with residuals increasing in the positive sense with increasing distance.

Data from all electrode configurations, from all 3 animals with closely spaced electrodes, differed principally in the distribution of residuals at the greatest distances, but in every case the type of residual distribution was invariant over different data sets based on 30 sec acquisition, obtained from the same cat and electrode configuration, over all epochs recorded. Associated plots of standard error for the coherence estimates showed all trends of residuals to be highly significantly different between most configurations (both within and between animals) at distances greater than 5 mm.

Plots of coherence versus frequency and distance for cat Sandra showed a pattern similar to that of the 3 other cats at comparable distances, with a fall to low coherences at the highest electrode separations (see Fig. 2).

Table I lists the parameters obtained from curve-fitting from each of cats Simone, Kate and Germaine. In some instances a satisfactory global minimum could not be found during curve-fitting (16 out of 120 curve fits performed) and these instances were found to be due to the presence of residuals in the fit, at the higher electrode separations, these distortions being characteristic of a unique electrode configuration, as described above. High discrepancies of this sort led to negative estimates of \( S \). In these circumstances we chose to exclude the discrepant data from those listed in Table I, and in the interests of conservative estimates, we excluded also the results found from the same recordings analysed according to different configurations of electrode pairs. The effect of these exclusions on the final estimates of parameters (over all animals and electrode configurations used) was as follows (median value is given, followed by range in brackets):

Without exclusions: \( c \) (m/sec) 0.124 (0.01–0.40), \( D \) (sec^{-1}) 37.04 (17.69–50.97), \( S \) (sec/m) 0.422 (0–4.37).

With exclusions: \( c \) (m/sec) 0.115 (0.07–0.40), \( D \) (sec^{-1}) 35.69 (17.69–45.83), \( S \) (sec/m) 0.498 (0.020–0.893).

(Fig. 1 likewise reflects this exclusion, yet this also produced little change in the resulting figures from graphs made without exclusion of any data set.) There is no consistent difference in these estimates as a function of the direction of orientation of the electrode assemblies in general, but different orientations in each cat show individually consistent tendencies over the epochs analysed.

Table II lists the results of fitting a plane to the residuals of fit, obtained from the pooled data with exclusions, of which Fig. 1 is illustrative. It is seen that A is close to zero in all cases, implying less than a 1% difference of phase velocity over the 1–32 Hz band. The positive values of \( B \) and negative values of \( C \) indicate progressive failure of the model with increasing distance, with associated distortion of the fitting parameters. It thus appears that the data excluded from tabulation, above, were probably only indicative of extreme instances of similar model-fitting limitations in all the data.

**Conclusion**

The hypothesis of linear non-dispersive wave transmission is supported. The non-systematic (random) distribution of residuals at all frequencies in the 1–32 Hz band at all distances of electrode separation confirms hypothesis, while residuals as a function of distance rise systematically, consistent with the increasing intrusion of sources of wave generation between electrodes with increasing distance, as anticipated in item (iv) in the introduction. Idiosyncratic distortions of residuals with distance peculiar to particular electrode configurations, and presumably indicating the effects of surface folding are also seen, as anticipated in item (iii). Strong effects from failures of other assumptions are not apparent. Estimates of wave velocity are consistent with each other, from electrode configurations of orthogonal orientation and indicate a wave velocity between 0.1 and 0.29 m/sec (the range of median estimates in Table I) equal for all frequencies 1–32 Hz. This agrees reasonably well with Lopes da Silva and Storm van Leeuwen's (1977) estimate of 0.33 m/sec for alpha waves, despite the difference in species used. Possible errors in our estimate are introduced via the estimate of temporal damping from autoregression modelling, which may underestimate true damping in
Fig. 1. Plots of squared coherence versus distance (in millimetres) and frequency (in Hertz) shown with the theoretical surface obtained by fitting the 2-parameter model, and the residual of the fitted surface. a–c: cat Simone. All electrode pairs oriented in the antero-posterior direction. Average found from 300 one sec recording epochs. d–f: cat Kate. All electrode pairs oriented in the transverse direction. Average found from 120 one sec recording epochs. g–i: cat Germaine. All electrode pairs on the main diagonal, viewed from above. Average found from 240 one sec recording epochs.

the short epochs analysed (and thus underestimate c), while fitting a model with an implicit problem of rising residuals with distance, introduces error toward overestimate of the true velocity. Further sources of error from extrinsic noise or the presence of directional wave movement cannot be excluded. Future studies might improve estimates by using smaller arrays confined more precisely to smooth cortex, and by utilising longer periods of analysis, so long as signal stationarity was satisfactory.

The continuing decline of coherence with increasing distance seen in data from cat Sandra argues against the presence of much standing wave formation of the type to be expected in the presence of reflective boundaries. The wave velocity is high enough to imply that the lowest EEG frequencies (delta waves) would reflect wave activity of wavelength approaching or greater than the circumference of the cerebral hemisphere of a cat. Even activity in the beta range would be centimetric in wavelength. The provisional assumption of wavelengths long compared to electrode separation, made in the introduction, thus appears consistent.

Overall, these results support an interpretation of electrocortical waves in the 1–32 Hz range as linear, non-dispersive, and uninfluenced by boundary conditions of a reflective, or strongly closed (reentrant) type.

The estimate of wave velocity obtained from surface-fitting is much lower than the velocity of conduction for myelinated or unmyelinated axons, and indicates that wave conduction probably depends upon comparatively short range fibre connections, thus introducing many synaptic delays between signals travelling...
millimetric distances, as is indicated in the model for cortical alpha waves of Van Rotterdam et al. (1982) which is, in turn, consistent with the model linking effects of combination of long and short range neocortical interactions advanced by Nunez (1989). A significant discrepancy is that the solutions of spatio-temporal transfer functions in these theories predict branched dispersion relations, not non-dispersive waves. The finding is consistent with a still wider range of possible models, each differing as to the details of cortical resonance or rhythmicity on the microscopic scale, but agreeing with regard to the velocity of propagation across the cortical field. Suggestions made by Freeman (1991) on the existence of excitatory/inhibitory recurrent loops in neocortex with a self-excitation cycle rate of 40 Hz offer a possible model within this class with relevance for local information processing (Eckhorn et al. 1988; Grey et al. 1989). The estimates of parameter S could assist in the further development of such models by providing information of the strength of additive coupling between columnar elements, while it now also appears justified to use the autoregression based estimate of D as a measure of decay of reverberation in local circuits.

Since the wave model applied (Wright 1990) does not imply linear and non-dispersive wave properties for high frequencies (rather the reverse), reports of chaotic dynamics of high frequency EEG components (Freeman and van Dijk 1987) may be consistent with the present results.

We expect breakdown of coherence from predictions made from this homogeneous conduction model

\[ \text{Experimental } r^2 \]

\[ \text{Fig. 2. Cat Sandra. Mean squared coherence as a function of frequency and distance, for more widely separated electrodes than the results shown in Fig. 1a, d and g. All electrode pairs oriented in the antero-posterior direction. Average from 300 one sec recording epochs.} \]
at medium ranges as a consequence of known inhomogeneous projections (Thatcher et al. 1986), and between laminae of cortex (Bullock and McClune 1989).

A commonly observed feature of evoked potentials – their short lag intervals for various phases of the late potentials at widely separated sites – is in contradiction with slow propagation of wave motion from the specific sensory areas, and for our interpretation to remain consistent we must argue that the widespread synchrony of the late evoked potential reflects either volume conduction, or wide co-activation of cortex by non-classical sensory afferents, or both. We note that the transient semi-stationary ‘landscapes’ revealed by topographic EEG analysis (Lehmann et al. 1987), may be interpreted as interference patterns of quasi-randomly generated waves passing through each other at comparatively low velocities.

More generally, these results further justify the provisional model of electrocortical organisation upon which they were based (Wright 1990) and delineate something of the behavioural conditions, the frequency range and distance scale on which linear assumptions regarding the EEG may apply.

The authors thank John West, Nicholas Hawthorn and Robert Kydd for assistance.

Supported by the New Zealand Medical Research Council and the Auckland Medical Research Foundation.

References


REVIEW ARTICLE

The electroencephalogram and cortical neural networks

J J Wright† and R R Kydd
Department of Psychiatry and Behavioural Science, School of Medicine, University of Auckland, PB 92019, Auckland, New Zealand

Received 15 June 1992

Abstract. The character of the EEG, its cellular sources, and its relationship to cognitive events are outlined. Then four theories of the EEG are discussed—the Amsterdam group's model of alpha activity, the Nunez model of global resonance, Freeman's model of oscillation in the cortical minicolumn, and the New Zealand group's stochastic model of EEG at millimetric scale. Experiments supporting these theories are outlined, including spatial and temporal characteristics of the alpha rhythm, velocities of EEG wave propagation, and phase relations of cell action potentials with EEG near 40 Hz.

These theories are mutually consistent, differing only with regard to the scale of phenomenon accounted for. They imply that real cortical dynamic properties bear analogy to those of Hopfield networks, Boltzmann machines, and Amit probabilistic attractor networks. The cortex may be described as a system with a single instantaneous basin of attraction, the locus of the basin being subject to adiabatic control by brain-stem afferents to the cortex.

1. Introduction

This review considers some aspects of research into the electroencephalogram (EEG), by which is meant the electric field recorded from the cortex by recording electrodes placed on the scalp, on the cortical surface, or among the cortical cells themselves. Findings in this field may help the artificial neural network modelling of brain function on scales between those of discrete, small assemblies of cells, and the gross, whole-brain level.

Research into the EEG has a difficult and conflicted history. Berger, the discoverer of the human EEG (Gloor 1969), had grand hopes for his work as a high road to insight into cognitive processes, but later workers became disenchanted by the difficulties posed, particularly in relating the EEG to simple correlates of single-neurone activity. Indeed, there is at present a widespread belief among general physiologists that the EEG is an epiphemomenon of little consequence to the information processing activity of the cortex (Eccles 1952). While doubt as to the EEG's exact significance is still justified, understanding of the complex relationships between cortical dynamics and the EEG appears to be presently increasing, as will be discussed below.

We will contend in section 2, that the EEG reflects cooperative processes in the brain of importance to information processing and outline briefly the biophysics and anatomy of the EEG. In section 3, we will outline some theoretical treatments chosen
especially for their physical nature, emphasizing wave processes in large neuronal pools. In section 4, our emphasis will be on developments which have analogy to ‘classical’ attractor neural network theory.

2. The electroencephalogram

2.1. Origin of the recorded signal

The biophysics of the recorded EEG have been described in detail by Nunez (1981, 1992). The relationship between electrical events at differing scales, and the types of electrical field recorded, are shown in figure 1.

![Electroencephalogram Diagram](image)

Figure 1. A schematic diagram showing how the electrical fields of cortical neurones are sensed by recording electrodes at various scales. Microelectrodes within the neurone sense directly the fluctuations of dendritic potentials (“neuronal waves”) created by synapses (IPSP and EPSP) afferent to the cell. These fluctuations then generate current flow in the surrounding media, which sum locally as the local field potentials, at larger scale as the electrocorticogram (for electrodes on the cortex) and, at the largest scale, as the scalp electroencephalogram. Low-pass filtering by the neural components largely excludes the relatively powerful short-range fields generated by action potentials from inclusion in the LFP, ECoG and EEG, which thus represent activity particularly in pyramidal cell dendrites and soma.

Essentially, the EEG, and the electrocorticogram (ECoG), are time-varying differential voltages, recorded between some point on the scalp, or the brain’s surface, and a more remote site on the subject’s body. The head and body are inhomogeneous conducting media. Thus potential differences, generated by bioelectrical events at diverse points in the brain and elsewhere, generate currents which sum linearly in
accord with Poisson's law, at the electrode. The voltage recorded is then generated in accord with Ohm's law. Only ideally is the reference electrode remote enough to be considered independent of the same sources. The complex media of the brain, dura, bone, scalp and skin all act to various degrees as low-pass temporal filters, and the size of the electrode determines the degree of low-pass spatial filtering.

The signal component with which we are concerned principally is that arising from the cortical mantle. With sufficiently small electrodes, properly positioned, it is possible to record the local field potentials (LFP) arising from electrical activity in the dendrites of one or a few nearby cells. These signals range in frequency from less than 1 Hz, to more than 100 Hz—action potential components requiring recording in the kilohertz range are thus largely excluded. Conversely, larger and more remote electrodes lead to the recording of events spatially summed over larger and larger areas.

The dendritic sources of the LFPs are thought to be fluctuations of the membrane potential in the millivolt range, which have the same spectrum as the LFPs and can be recorded with intraneuronal electrodes as 'neuronal waves' (Elul 1972). These intraneuronal fluctuations may be taken to arise from the effects of the bombardment of the neurone with afferent signals (excitatory and inhibitory post synaptic potentials—EPSP and IPSP) at synaptic junctions.

The net EEG is in the microvolt range, and current/source density studies indicate that the cells which are the principal source in cortex are the pyramidal cells of layers III and V (Mitzdorf 1985, 1987). These cells are so oriented as to generate greater net dipoles with regard to a surface electrode, and are also larger and more numerous cells than the inhibitory cells which surround them.

2.2. Relation of the EEG to cognitive events

It is clear that in one sense the EEG is an epiphenomenon, since it is merely a weak electric field created by current flows of a complicated type, and is well below the field strength of the neural events themselves. While even this conservative assertion has been challenged (Adey 1988) it does not enable us to tell a priori, whether the EEG is a reliable mirror of cognitive events. All available evidence indicates that it is such a mirror.

The following lines of evidence are some of the more important.

1. The EEG time series (or power spectrum) is an indicator of wakefulness (Moruzzi and Magoun 1949, Walter et al 1967). During intense alertness the EEG approaches a state called 'desynchronization', in which its spectrum resembles that of 1/f noise (Freeman and van Dijk 1987). During drowsiness, a powerful peak of activity is apparent in the spectrum between 8 and 14 Hz—the alpha band—and further changes characterize sleep (Walter et al 1967).

2. Components in the evoked potential (the signal-locked events in the EEG after a specific event) from about 100 ms after a stimulus to a second later, are exquisitely sensitive to the attentional set of the subject, subject expectation of the stimulus, and stimulus properties (Picton and Hillyard 1988).

WORK TOWARD A THEORY OF BRAIN FUNCTION

4. Topographic and spectral characteristics of EEG can be used to separate groups of individuals with cognitive disabilities such as dyslexia, schizophrenia, etc (John et al 1977). Similarly psychotropic drugs exert differential effects upon the EEG (Fink 1978).

5. Patterns of EEG activity in the olfactory bulb can be shown to be not merely odour specific, but specific to conditioned expectation, or stimulus significance (Freeman 1981, 1983, Freeman and Skarda 1985).

6. The input to the brain of a very weak microwave carrier signal, amplitude modulated at EEG frequencies, disrupts cognitive processes at energy levels without thermal effects (Adey 1977, 1979, 1981).

Collectively, these findings make it likely that the EEG is an informationally significant transform of dendritic activity.

A larger question, which remains unanswered, is how much of the total information being processed in the cortex at a given time can be extracted from the EEG? Although technical considerations obviously set a limit to the extent to which detailed resolution over extensive areas of cortex can be achieved, the spatial and temporal bandwidth necessary to specify the brain’s present state might be much less than the brain’s enormous connectivity at first suggests. Psychological experiments and common experience (Broadbent 1958) indicate that at any moment the brain/mind is utilizing a channel of limited information capacity—the field of immediate attention. The long wavelengths and low frequencies of the EEG might suffice to represent the information of current consciousness, or a substantial part of this information.

2.3. Dendritic events versus action potentials

A substantial stumbling block to progress in understanding of the EEG has been the failure to demonstrate any reproducible correlation between LFPs and density of action potentials, in most circumstances (Burns 1958). Expectation that such a relationship would be found, and disappointment that this was not apparent, except in certain conditions such as discharge of an epileptic focus or strong slow wave EEG activity, seems to have led to the disparagement of the EEG as an ‘epiphenomenon’.

In view of present knowledge of the physics of EEG recording (Nunez 1981), the complexity of neuronal connectivity (Braitenberg and Schüz 1991, Korn and Faber 1979, Schmitt 1979, Shaw et al 1982) and the nonlinearity of neuronal interactions, it now seems less surprising that no simple correlation exists.

It has recently been recognized that at 40+ Hz, the LFP and single-unit firing are correlated, and further, that spike densities at separated sites are sometimes cross-correlated near this frequency (Eckhorn et al 1988, Gray et al 1989). The matter will be taken up further below, in reference to Freeman’s chaotic theory of LFPs.

2.4. Cerebral anatomy, subcortical control, and the EEG

A complete theory for the EEG would need to consider the organization of the entire brain, in some simplified manner. Important aspects which require incorporation include:

1. The anatomy of the cortex. No unit of cortical anatomy is perfectly modular, although considerable homogeneity exists across widely separated areas (Braitenberg 1978). A sufficient functional unit for present purposes is the minicolumn (Szentagothai 1978b, 1979, Mountcastle 1979). This unit is equivalent to a single excitatory
pyramidal cell, with its cell body located in one of the several layers of dense cellular matter in the cortex, plus its surround of largely inhibitory cells. Within cortex the excitatory cells are laterally coupled over distances up to 3 mm by 'recurrent collateral' fibres (Szentagothai 1978a). Within the minicolumn inhibitory cells are coupled to each other, to the reference pyramidal cell, to the column's neighbouring pyramidal cells, and to other less common cells, over much shorter lateral ranges than the recurrent collaterals, typically about $3 \times 10^{-2}$ mm, which range defines the minicolumn.

2. Cortico-cortical connections. Nunez (1981, 1989, 1992) has emphasized the importance of these long-range fibre connections linking pyramidal cells over distances much greater than recurrent collaterals. In humans these fibres form the massive subcortical white matter, which connects cells over ranges up to many centimetres. When this long-range fibre system is compared to the intracortical recurrent collateral system, it is apparent that for cells up to a few millimetres apart, recurrent collaterals mediate the bulk of the interaction, and cortico-cortical fibres, those between cells at greater range.

3. Subcortical/cortical interactions. Cortical afferents form a numerically small part of the connectivity of the cortex, but are functionally vital.

We may exclude from immediate consideration such vital subsystems as the specific nuclei of the thalamus, the basal ganglia, the hippocampus and related circuitry, noting only that these systems are all situated so as to act as to-and-fro relay systems in interaction with the cortex itself (Nauta and Feirtag 1979, Nauta and Karten 1970).

Of principal concern is the reticular activating system—a complex and diffuse system projecting from the brain-stem to the cortex, which provides both inhibitory and excitatory input. This subsystem is itself under control from the cortex, as well as from collaterals of sensory pathways (Nauta and Feirtag 1979). The activating system and its cognates, including catecholaminergic cells (Lindvall and Björklund 1974, 1982) is most crucially concerned with maintenance of the waking state, desynchronization of the EEG, direction of attention and governance of motivation (Clavier and Routtenberg 1974, Lindsley 1982, Marshall 1979, Olds 1981, Wright 1981, Wright et al 1987).

A large body of research has identified rhythmic EEG processes with rhythmic driving from subcortical sites (Steriade et al 1990). Almost all this literature concerns processes associated with sleep or the onset of sleep and will not be further considered here. Although there is no direct proof that rhythmic driving of cortex is a factor in the alert state, nor any evidence defining the spatial and temporal spectra of reticular inputs to cortex during wakefulness, all the models which follow assume nonetheless that reticolocortical input is spatially and temporally a white noise.

3. Network theories of EEG

We now outline four theories (or models) of the generation of cortical EEG rhythms. These have been selected for two principle reasons, from a larger literature which has been well summarized in Lagerlund and Sharbrough (1988). Firstly, each of the models selected makes comparatively clear predictions about EEG neocortical characteristics, which have been experimentally tested. Secondly, the four models appear complementary.
The common feature of all the theories is an emphasis on oscillatory activity generated by each pyramidal cell and its inhibitory surround, in interaction with coupled neighbours. Variations include the following differences in basic assumptions:

1. The way in which the summation of postsynaptic potentials is linearized, or considered as a static or dynamic nonlinearity.
2. Range and strength of intercellular couplings.
3. Boundary conditions.
4. The spatial scale of the recorded signal, which is the experimental variable being explained.

Figure 2 shows schematically features of the latter three of these four theories, emphasizing their relationships as function of scale.

3.1. Spatial-temporal characteristics of alpha rhythm, modelled as a local cortical phenomenon

The first theory to be considered arises from the work of the Amsterdam group—specifically that of van Rotterdam et al. (1982). This model is a generalization of the earlier 'lumped' model of Lopes da Silva et al. (1974, 1976) which was intended only to account for the time domain properties of cortical alpha waves. We describe both the 'lumped' and 'local' versions of this model in some detail, because its mathematical rigour makes it a classic of the field. It will be seen that this rigour is achieved at some expense, in terms of required simplification of the complexity of the brain circuitry. All the following theories might be regarded as attempts to overcome this limitation.

The 'lumped' model worked as follows. Two interacting populations of cells were considered—the pyramidal cells and the local interneurons. These subpopulations were characterized by two transfer functions $h_2(t)$ representing the EPSPs, and $h_3(t)$ for the IPSPs, two static nonlinearities, $f_2(V)$ and $f_3(V)$ relating the dendritic membrane potentials to their input pulse density, and $C_1$ and $C_2$ which were interconnectivity constants representing the total number of synaptic connections between each pyramidal cell and local interneuron, and vice versa. The principal signal to be observed was $V_r(t)$, the potential of a pyramidal cell's dendrites, to which the recorded EEG was assumed proportional.

Figure 2. An 'exploded' diagram of levels of organization within the brain, and the putative physical characteristics of the electrical fields associated with wave, or wave-like, activity at three scales. The top third of the diagram represents the microscopic level. Pyramidal cells (white triangles) and their inhibitory surround cells (black circles) engage in local nonlinear oscillation, generating local field potentials with a principal periodicity perturbed about a value of about $\frac{1}{40}$ s. The middle third of the diagram represents the millimetric scale. Coupled interactions among many cells generate collective waves which are effectively linear, non-dispersive, superposition waves, with a velocity of about 0.2 m s$^{-1}$. Patches of excited cells at microscopic scale act as localized signals driving the collective waves. The lower third of the diagram represents the whole-brain scale. Closed boundary conditions impose preferential modes akin to spherical harmonics. At this scale larger range cortico-cortical fibres may become dominant in effect over the intracortical connections dominant at millimetric scale. Wave velocities (in the human) are 7-9 m s$^{-1}$. Interaction of cortex with brain-stem (shown by the descending arrow and return cell projections) constitutes a major determinant upon behaviour of the system, acting to change the focus of the system basis of attraction on a time scale slower than interactions among the cortical cells themselves.
The nonlinearity coefficients $f_2(V)$ and $f_3(V)$ were converted to linearity, by assuming that the dendritic membrane is operating in the linear range with amplification factors $q_{e1}$ and $q_{i1}$. Then $h_e(t)$ and $h_i(t)$ have the time course and amplitude experimentally obtained for EPSP and IPSP, and they can be associated in the frequency domain with

$$H_e(jw) = Aa/(jw + a)^2$$

$$H_i(jw) = Bb/(jw + b)^2$$

(1)

where $A = 3.25$ mV, $B = 22$ mV, $a = 100$ s$^{-1}$, $b = 50$ s$^{-1}$.

This leads to the transfer function

$$V_e(jw) = \frac{Aa(jw + b)^2}{(jw + a)^2(jw + b)^2 + K^- P(jw)}$$

(2)

where $P(jw)$ represents a spectral component of $p(t)$, an input signal, and $K^- = C_1 C_2 Aa Bb q_{e1} q_{i1}$.

With the 'proper choice' of $C_1$, $C_2$, $q_{e1}$ and $q_{i1}$, and $p(t)$ an uncorrelated white noise, the power spectrum of $V_e(t)$ has a peak in the alpha range, and the amplitude, peak frequency and width of this peak can be set comparable to the experimental case by choosing $K^- = 0.7 \times 10^8$ s$^{-4}$. This is a case of close coupling. At low strength of coupling, the spectrum of $V_e(t)$ approaches the noise spectrum of the input.

To extend the lumped model to a spatially distributed case, van Rotterdam et al (1982) assumed an infinite one-dimensional chain of pyramidal cells and interneurons, interconnected by means of spatially distributed collaterals and inhibitory fibres. The interconnections were described by $C_e(k, l)$ denoting the number of excitatory synapses from the $k$th pyramidal cell projecting on the $l$th interneuron, and $C_i(k, l)$ denoting the number of inhibitory synapses from the $k$th interneuron to the $l$th pyramidal cell. They then obtain a transfer function for the temporal frequency $\omega$ and spatial frequency $\lambda$,

$$U(j\lambda, jw) = \frac{H_e(jw) P(j\lambda, jw)}{1 + q_{e1} q_{i1} H_e(jw) H_i(jw) C_e(j\lambda) C_i(j\lambda)}$$

(3)

When the neuronal chain is driven by a spatially and temporally noisy signal, with similar choice of parameters to the lumped model, dispersive waves propagate in the chain, with preferential propagation of power at the alpha frequency, and the phase velocity of these waves is dependent on $C_e(k, l)$ and $C_i(k, l)$. When $C_e(k, l)$ is chosen to correspond to the distance and approximate density of spread of intracortical pyramidal axons---i.e. over up to 3 mm (Szentagothai 1978a)---and $C_i(k, l)$ is similarly chosen to correspond to the spread of basket inhibitory cells---i.e. 500--1000 µm (Tombol 1978)---then a phase velocity for alpha waves of 0.3 m s$^{-1}$ is found. This velocity corresponds to that found in the visual cortex of the dog, by phase-difference estimates obtained with millimetric electrodes at adjacent points a few millimetres apart (Lopes da Silva and Storm van Leeuwen 1978).

In summary, this model achieves a good fit between the temporal and spatial properties of alpha rhythm at millimetric scale, and allied physiological and anatomical data, at the price of linearizing the system, imposing a simplified anatomical structure, and introducing a strength of coupling which is somewhat arbitrary. Boundary conditions are open, and waves outside the alpha range are not considered explicitly, but are dispersive.
3.2 Cortical global resonance

We turn now to the work of Nunez (1981, 1989, 1992), who has developed, partly in association with Katznelson (1981) and later Ingber (Ingber and Nunez 1990), a theory of origin of the EEG which goes beyond the Amsterdam model in important ways. The principal extensions are:


2. Consideration of boundary conditions. Since the cortical surface can be considered a deformed oblate spheroid, wave activity of long range will be influenced by rather complicated periodic boundary conditions.

3. Explicit consideration of the scaling effects of electrode size and situation upon the type of EEG activity recorded.

Development of Nunez's theory has been stepwise, considering first an infinite one-dimensional brain (Nunez 1974), a closed one-dimensional brain, a sphere, and an oblate spheroid, under differing assumptions of connectivity and external input, and this work is not yet completed. The crucial notion is, to quote Nunez (1989), that 'the equations linking synaptic activity to action potentials \( g(r,t) \ldots \) follow directly from the basic idea that synaptic action in a mass of neural tissue is due to action potentials fired from some other location, perhaps at earlier times'. That is,

\[
\begin{align*}
    h_E(r,t) &= u_E(r,t) + \int_0^\infty dv \int_{r_1} R_E(r,r_1,v) g \left( r_1, t - \frac{|r - r_1|}{v} \right) d^2 r_1 \\
    h_I(r,t) &= u_I(r,t) + \int_{r_1} R_I(r,r_1,v) g(r_1,t) d^2 r_1
\end{align*}
\]

where \( h_E(r,t) \) and \( h_I(r,t) \) are excitatory and inhibitory synaptic action densities (number of active synapses per unit volume, at time \( t \)), and \( u(r,t) \) are excitatory and inhibitory inputs to the cortex. The distribution functions \( R_E(r,r_1,v) \) and \( R_I(r,r_1,v) \) describe the number of cortico-cortical fibres and short inhibitory fibres connecting locations \( r \) and \( r_1 \) and their action potential velocity, \( v \).

Nunez's general technique is to take \( R_E, R_I \) as functions based on the range of long and short excitatory and inhibitory fibres, set \( u(r,t) = u_x \), for consistent inhibitory tone, assume a linear response of dendritic potentials to synaptic inputs (similarly to the van Rotterdam et al model) so that the synaptic action densities may be treated as having dynamics identical to the local field potentials at \( r, t \), and then consider the standing waves which will be generated when periodic boundary conditions force a restriction of possible wavenumbers for the standing waves.

The results of this approach are complicated, and under differing constraints dispersive and non-dispersive waves are variously predicted. Certain general features emerge which are of theoretical and experimental importance.

(a) The Amsterdam model can be incorporated as a local wave description, embedded in the global processes (Nunez 1989).

(b) Alpha rhythm for brains of human size is predicted to occur also as a global (large-scale) phenomenon, as a standing wave—the first harmonic of the global system—{}with a frequency dependent on head size. This relation is found experimentally (Nunez 1981).
Phase velocity for alpha waves of the global type is predicted to be near the conduction velocity of cortico-cortical fibres, at $6-9 \text{ m s}^{-1}$. Alpha activity has been measured travelling at $4-20 \text{ m s}^{-1}$ (Thatcher et al 1986) and $7 \text{ m s}^{-1}$ by phase shift of MEG (magnetoencephalogram) during driving by sinusoidally modulated light (Silberstein et al 1992).

Halothane rhythm—a cortical rhythm found in anaesthesia—behaves like a global resonance which varies in frequency with depth of anaesthesia, and hence velocity of axonal conduction. The frequency of this rhythm is appropriately different in dogs and humans, according to brain size (Nunez 1981).

The frequencies of other major cerebral rhythms correspond approximately to those predicted for resonant modes of an oblate spheroid (Nunez 1992).

3.3. Local chaos in cortex

The two preceding models account well and consistently for activity in the alpha band. Both take steps to regularize description of the anatomy of cortical connections, and linearize dendritic summations in an assumed narrow operating range. In neither case has explicit detailed consideration been given to the extent to which predictions of the models may be sensitive to limitations in the basic linearization assumptions, although Nunez in particular has been at pains to indicate that microscopically highly nonlinear activities might smooth on the macroscopic scale to linear waves (Nunez 1992). However, it is not immediately clear that this would be the case. Analogies from general physics seem disconcerting rather than reassuring. Although molecular motion is chaotic, sound waves are predictably deterministic, linear waves, in common conditions. Yet weather is chaotic. What, then, is the case in the brain?

In this section we summarize recent proposals of Freeman on the microscopic dynamics of cortical columns. In the following section a link between nonlinear microscopic and linear macroscopic dynamics is described.

In the past 25 years Freeman and his co-workers (Freeman 1964, 1972, 1975, 1979, 1987a,b, 1988) have systematically developed a 'bottom-up' model of olfaction, starting from consideration of the anatomical configuration of olfactory neurones, measurements of neuronal nonlinearity, and open loop gains. They have achieved simulation of olfactory bulb EEG in inhalation and resting states, unifying these findings into a preliminary model of perception. In a recent paper (Freeman 1991) this body of information is brought to bear to make predictions concerning neocortical dynamics.

3.3.1. The basic unit of oscillation. Freeman rejects intrinsically rhythmic cells of cortex, and points to Burns' (1958) finding of electrical silence in isolated intact neocortical slabs, as evidence that without external input, cortical cells reach a point equilibrium.

From reasoning related to the basic nature of the Hodgkin-Huxley equations, Freeman considers that similar values to his estimate of the 'open-loop gain' for paleocortical cells will be found in neocortex and will be equal for both pyramidal cells and inhibitory surround cells. The 'open-loop gain' parameters, obtained from the anaesthetized animal, determine, among other things, the time of transmission of a change of level of dendritic depolarization through a single neurone to near neighbours, where the time of axonal transmission may be neglected. The time of transmission is 5.8 ms, composed of 1.3 ms of synaptic and dendritic cable delay.

WORK TOWARD A THEORY OF BRAIN FUNCTION | 177
and 4.5 ms due to passive electrical properties of the dendritic membrane. Within a cortical minicolumn (in which a pyramidal cell interacts with its inhibitory surround) a cycle of sequential steps may develop, composed of pyramidal excitation, excitation of the inhibitory cell, inhibition of the pyramidal cell, disinhibition of the pyramidal cell, etc.

The cycle would impose a crudely sinusoidal or saw-toothed periodicity of 23.2 ms, (or a fundamental frequency of 43.1 Hz) upon local field potentials associated with pyramidal cell dendrites. As remarked earlier, coherence has been found experimentally between neuronal firing and the EEG near this frequency (Eckhorn et al 1988, Gray et al 1989). Also, a quarter-cycle phase-lag in the periodicity of cell firing should develop between the excitatory, and the inhibitory cells involved in the loops. This relation has also been found wherever it has been sought, including all parts of the central olfactory system (Freeman 1975) the entorhinal cortex (Eeckman and Freeman 1990) and the hippocampus (Horowitz 1972).

3.3.2. Locally chaotic dynamics. Reticulo-cortical afferents are considered as regulators of the overall level of excitation of the cells, necessary to initiate oscillation as follows.

While oscillation depends on the excitatory/inhibitory loops, the other types of cellular interactions must increase the complexity of local interactions considerably. The sigmoid response and the absolute refractory time of neurones guarantees overall stability, but the sharp nonlinearity of neuronal threshold confers a sensitive dependence of action potential trains, upon both initial conditions and ongoing perturbation. Consequently, both the firing rate and field potential of any pyramidal cell raised above a critical level of excitation, may well be chaotic. Below a critical level of excitation, the local system would sink to the point attractor of electrical silence.

Freeman points out that the near 1/f spectrum of LFPs, and the Poisson distribution of axon potential intervals accord with chaos, not entrained limit cycle oscillations. In the present context it is irrelevant whether the dynamics of the interaction is strictly chaotic, or some limit cycle strongly perturbed by noise. There is also evidence of chaos (loosely defined) in ECoG in the 25–40 Hz gamma range (Freeman and van Dijk 1987). His argument is thus largely independent of the controversy which surrounds experimental measures of chaotic dimension from EEG (Layne et al 1986, Rössler and Hudson 1989).

Two supportive details from other findings may be added to Freeman’s account. Firstly, that the minicolumn is of such dimension that the action potential of the pyramidal cell would generate sufficient field to coordinate feedback responses from the surrounding inhibitory cells (Abeles 1982). Secondly, during wakefulness most cells must not be excited to such a level that oscillation occurs in more than a minority of cells at a time—this follows, since the observed mean firing rate of cells is low (as low as 10 or 20 spikes s⁻¹), and cells are more likely silent than active at any given time (Fuster 1973, Niki 1974, Abeles et al 1990, cited in Amit and Tsodyks 1991). During their normal activity most cells must thus be near the sharply nonlinear threshold—an effect at first sight seemingly drastic for the models of Nunez and the Amsterdam group.

3.4. Cortical interactions and coupled stochastic oscillators

It should now be apparent that problems are posed in developing a theory which allows adequately for both the extreme nonlinearity of neural behaviour at the cellular
level, while accounting in a tractable manner for macroscopic EEG phenomena. A really adequate solution (if one exists) may yet be found by formal application of techniques such as group renormalization. The work of Ingber (1982, 1991) represents the most rigorous attempts in this direction. With analogous motivation, Frolov and Medvedev (1986) have attempted, by simulation, the prediction of EEG spectral properties in large scale and at a point approximation, from specific classes of possible microscopic models.

Our own group in New Zealand has attacked the problem by constructing and testing a theory of events on an intermediate scale (Wright 1990, Wright et al 1990a,b, Wright and Sergejew 1991). Our concern was to show that without linear element response, conditions in the cortex permit the generation of linear wave motion at macroscopic scale.

We consider the cortex as a system of coupled nonlinear oscillators, deferring for the present any precise definition of the oscillators, other than that they are small systems of cells coupled by axons. We will later discuss the consequences of identifying the unit oscillator with Freeman's conception of oscillation at the minicolumn level.

A coupled system can be described generally by second-order stochastic differential equations of dimension \( n \), thus

\[
\dot{X}_i + D_i \dot{X}_i + N_i^2 X_i = \sum_{j=1}^{n} K_{ij} X_j + \sum_{i=1}^{n} U_i \quad (i \neq j)
\]

(5)

where \( \{X_i\} \) are the local field potentials associated with each unit oscillator, \( \{N_i, K_{ij}, F_{ij}\} \) are stochastic parameters, and \( \{U_i\} \) are input signals. By definition \( \{K_{ij}\} \) represent the instantaneous additive coupling of the \( j \)-th to the \( i \)-th unit, \( \{N_i\} \) represent the instantaneous frequencies of oscillation, and \( \{F_{ij}\} \) describe instantaneous multiplicative couplings.

The parameters \( \{N_i, K_{ij}, F_{ij}\} \) are implicitly functions of \( \{X_i\} \) describing local interactions which are nonlinear and may be chaotic. During cortical 'activation' (when the subject is fully alert and the EEG becomes 'desynchronized'), the inputs \( \{U_i\} \) (arising from all afferents to the cortex) may approach white noise, with super-added variable 'coloured' noise. Consequently \( \{N_i, K_{ij}, F_{ij}\} \) can all be provisionally assumed stationary and stochastically independent in the large of each other, and of \( \{X_i, X_j\} \) if conditions of input remain constant. Several conclusions emerge, as follows.

3.4.1. Linear and non-dispersive wave transmission. At a scale sufficiently large compared to the unit oscillators, the 'instantaneous' transfer function over a brief epoch of duration \( \Theta \) between two cortical points \( a \) and \( b \), is given by

\[
H(j\omega, \Theta, t) = q \frac{(p)}{(\Pi)} \frac{K_{ij}}{jD_i \omega + (N_i^2 - \omega^2)}
\]

(6)

where \( q \) is the number of possible paths over which signals may pass from \( a \) to \( b \), \( K_{ij}, D_i, N_i^2 \) represent that subset of parameters describing the involved units and their connections along one of the \( q \) paths, and \( (p) \) is a short notation indicating taking the product of the unit transfer functions along any particular path.
The electroencephalogram and cortical neural networks

Where \( q \) is large and \( K_{ij} \) less than unity, \( H(j\omega, A, t) \) tends invariant with time. Thus group waves propagating through the mass will exhibit linear superposition.

For \( D \ll N \), the phase shifts in transition through units are minimized, and phase velocity largely reflects lag times in the couplings. Since axons transmit all frequencies at a common velocity, the waves are non-dispersive. Recalling that the coupled system may have active elements, and is assumed driven by a diffuse white noise, then, in regions where cortical anatomy is homogeneous over a considerable area, wave motion will occur as though arising from multicentric sites on a non-dispersive plane medium.

The linear superposition character of the waves in alert conditions has now been established by several methods, including modelling of the evoked response by linear filter methods (Wright et al 1980b) and by comparing linear with nonlinear forecasting of the EEG time series (Blinowska and Malinowski 1991).

Non-dispersive linear wave characteristics of the electrocorticogram at millimetric distances and at frequencies from 1–30 Hz have been demonstrated in the occipitoparietal cortex of the cat (Wright and Sergejew 1991). This was done by recording the ECoG from 64 sites, each arrayed in an eight-by-eight grid, and each separated by less than a millimetre. Coherence at each frequency (the squared cross-spectra power, divided by the product of spectral powers), could then be computed from ECoG recording from any pair of sites and averaged, to provide a surface which is a function of distance, \( R \) and frequency \( \omega \). Theoretically, when non-dispersive linear waves of multicentric generation move on a plane, coherence \( r^2 \) is given by

\[
    r^2 = \frac{4e^{-\beta}}{(1 + \beta)^2} \tag{7}
\]

with

\[
    \beta = \frac{2}{\pi} \left( \frac{D R}{c} + S \omega R \right)
\]

where \( c \) is the phase velocity of the waves and \( S \) is a constant related to onward transmission of energy by additive coupling. \( D \) (the temporal damping) can be measured independently, using autocorrelation analysis of the EEG (see below).

With \( c \) and \( S \) as free parameters in the fit of (7) to experimental \( r^2 \), the fit is robust for \( r^2 \) versus \( \omega \), independent of distance (indicating that the waves are non-dispersive), but deteriorates in \( r^2 \) versus \( R \), with increasing \( R \), as expected for the folded and inhomogeneous cortex—a plane other than ideal. The parameter \( S \) finds a value from which it can be calculated that the total additive coupling from one unit to all its neighbours is close to unity, implying \( K_{ij} \) is fractional, as expected.

The velocity, \( c \), is found to be 0.1–0.3 m s\(^{-1}\).

3.4.2. Near equilibrium group dynamics. A further feature of the stochastic model is of relevance for the arguments in section 4. A series of studies show that the control of \( D_i \) is mediated in the brain by inputs to cortex from the transhypothalamic pathways mediating arousal (Wright et al 1984, 1985a,b, 1990a). It can be seen from equation (5) that the mean value of \( \{X^T_i, X_j\} \) determines temporal damping for the units, so this sensitivity to diffuse tonic regulation of the cortex is expected. It can be further shown (Wright 1990) that when all values of \( D_i \) are equal, the system has reached a condition of equipartition of energy, i.e. thermal equilibrium.
Values of $D_i$ can be obtained from autoregression (AR) analysis of the EEG, as they are equal to the mean of temporal damping estimates obtained from the roots of an optimal-order AR model for a short segment of EEG (Wright 1990, Franaszczuk and Blinowska 1985). Results from the EEG of alert rats show resonances to be associated with similar mean and distribution of damping coefficients at all frequencies, thus implying perturbation of system energy about a set point, analogous to thermal equilibrium (Wright et al 1990a).

3.5. Integration of models

The models outlined above were initially developed to explain quite different aspects of the EEG, yet they are not incompatible with each other, and show certain quantitative agreement.

Firstly, Freeman's description of the cortical minicolumn is consistent with the stochastic differential equation description of the New Zealand group. It seems reasonable to identify the chaotic oscillation of the minicolumn described in the former, with the unit oscillator of the latter. Likewise, transverse interactions of the minicolumns via pyramidal axons and dendrites may be described by stochastic additive and multiplicative coupled interactions.

Freeman's estimates of open-loop gain indicate that the time required for a displacement of dendritic potential to be transmitted from one pyramidal cell to another, should be 5.8 ms, when the axonal delay is negligible. Taken in association with the mean range of intracortical (recurrent collateral) axons connecting pyramidal cells transversely (Szentagothai 1978a) this leads to a predicted average velocity of about 0.2 m s$^{-1}$ for waves on the millimetric scale. This is the velocity found by us from electrocorticall coherence (vide supra). Further, the velocity found is similar to that predicted, and found, for local alpha activity, by the Amsterdam workers, and that velocity, in turn, actually depends on estimates of synaptic rise times and delays similar to Freeman's open-loop gain estimates.

Secondly, if local EEG activity at millimetric scale and frequencies below 30 Hz, can be adequately described as a linear wave phenomenon, then this justifies treating longer-range coupling parameters in terms of their average values, and for linearizing local wave summations even although dendrites may not be individually operating in the linear range. Thus Nunez's equation (4) can be consistently applied with Freeman's and our own account. Interestingly, this implies that non-dispersive waves at millimetric scale may be matched in frequency to waves which may be dispersive at a larger scale. In recent work conducted collaboratively between the New Zealand group, the Nunez group, and the Silverstein group at the Swinburne Institute, Melbourne, Australia, we have found evidence for the simultaneous presence of short-wavelength cortical activity and long-wavelength activity (apparently due to a global phenomenon), at the same temporal frequency. Additionally, specific spatial Fourier components attributable to the input of active pyramidal cells, which are driving the resonant modes and local waves, can be detected by frequency/wavenumber analysis. These findings, utilizing driving of human scalp EEG and cat EEG by sinusoidal visual flicker, will be described in detail elsewhere.

It should be noted that the local wave activity predicted by von Rotterdam et al is strictly intended to explain the alpha band activity. A contradiction between the Amsterdam model and the New Zealand model is apparent. The coupling required for the Amsterdam model to function is strong coupling, whereas the stochastic model indicates weak coupling. For the Amsterdam model to be included in the same
framework as the other three, we would have to assume some special circumstance pertains when alpha power is high compared to broad-band activity. It is interesting to note that a similar crucial effect of strength of coupling in converting 'synchronized' to 'desynchronized' EEG has been identified in the simulations of Frolov and Medvedev (1986) cited in section 3.4.

4. Comparison of stochastic and chaotic network theories of EEG to artificial neural networks

If we accept that, together, the models presented above give a valid semiquantitative picture of the dynamics of the cerebral cortex, does any consequence follow which casts light on the information processing capability of the cortex? In particular, does analogy to processes recognized in artificial network theory emerge?

Such analogy does emerge in two ways, the first at the gross level of cortical organization, where the whole cortex can be compared to an enormous attractor neural network with a capacity for self-control, mediated by interaction of cortex and brain-stem. The second analogy is apparent at the microscopic level, in the match to model networks developed by Amit and Tsodyks (1990, 1991).

4.1. Analogy to attractor neural networks at a macroscopic level

In a state of alerting the energy of the cortical oscillators can be described by a Hamiltonian function, viz:

$$E = \sum_{i=1}^{n} \sum_{j=1}^{n} K_{ij} X_i X_j + \frac{1}{2} \sum_{i=1}^{n} \left( \frac{\sum_{j=1}^{n} K_{ij} X_j}{\sum_{j=1}^{n} N_j} - \frac{\sum_{j=1}^{n} F_{ij}}{\sum_{j=1}^{n} F_{ij}} \right)^2$$

which is derived by summing 'potential' and 'kinetic' energy terms from equation (5) (Wright et al 1992). We take this energy function to represent in a generic way, all four of the above EEG theories. The stochastic variables and parameters can then be assigned either mean values or signs, viz: $N_j$ should be about 40+ Hz for all oscillators, $\{X_i, X_j\}$ are all positive and have values in the range of dendritic potentials for pyramidal cells, $\{K_{ij}\}$ have all members less than unity, and positive (reflecting excitatory coupling between minicolumns) and $\{F_{ij}\}$ can be either positive or negative in sign, depending on the levels of mutual excitation pertaining between minicolumns. The mean values of $X_i, X_j$ will be significantly influenced by reticular afferents. Equation (8) may describe either an energy-gaining, or energy-losing system, depending on the net effect of the instantaneous signs of $\{F_{ij}\}$. This determines whether the state vector flow is toward a system energy minimum, or maximum, at a given instant.

From its form, equation (8) can be compared to the energy function of a Hopfield network, which in turn, is homologous to that of a spin glass (Hopfield 1982), i.e.

$$E = -\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} S_i S_j + \sum_{i=1}^{n} \Theta_i S_i \quad (i \neq j)$$

where $S_i, S_j$ are analogous to $X_i, X_j$, where $w_{ij}$ are analogous weights to $\overline{K_{ij}}$, and, since $\overline{X_i}$ is proportional to $X_i$ for oscillators, equation (8)'s analogue of $\Theta_i$ is
a function of the state variables, rather than being a constant-element threshold, as is usual in equation (9).

Through the analogy of (8) and (9), an inference regarding the interaction of brain and reticular formation can be drawn. Since all $K_{ij}$ are at all times positive, the analogous magnetic example (an Ising model with only positive couplings) is a ferromagnet. Thus the system would have at any instant only a single energy minimum, and since the analogue of $\Theta_x$ is a function of the state variables, then the position of the energy minimum in state space is a function of the present state (Sherrington 1991, Amit 1990). That is, the cortex may be compared to a ferromagnet under perturbation by an external magnetic field, with the proviso that the external field is some function of the ferromagnet's internal fields at an earlier moment.

Alternatively, comparison can be made to neural networks in which couplings are by both 'fast' and 'slow' synapses, resulting in a mobile basin of attraction, or 'adiabatic landscape' (Amit 1990, Nebenzahl 1987, Dehaene et al 1987). See figure 3.

![Figure 3](image)

**Figure 3.** An adiabatic landscape in simplified three-dimensional form, viz: energy versus two local field potentials at (a) initial time and (b) late time. System energy, shown schematically by the arrowhead, tends at all times to a basin of attraction, but the locus of the basin is a function of the state vector.

A fairly obvious anatomical and physiological parallel to the controlled external field, or the 'slow' synapses, is apparent. This is provided by the centripetal effects of the cortex on the reticular activating system, and the reticular system's return influence upon the cortex. The lag terms in Nunez's treatment (equation (4)) permit us to account for such slow interactions, if (4) was explicitly introduced into (8), to describe the movement of signals across the cortex, down descending pathways and finally via cortical re-entrant pathways, as indicated in figure 2. As foreshadowed in section 3.4.2, the presence of this means of self-control within the brain is supported by a series of studies we have conducted (Wright and Craggs 1977, Wright et al 1984, Wright and Kudd 1984, Wright et al 1985a,b) on ECoG power spectral changes subsequent to unilateral lesions of fibre pathways ascending through the lateral hypothalamus. These pathways are functionally involved in the maintenance of arousal and the mediation of motivation and affect (Hess 1936, Scheibel and Scheibel 1967, Wise 1978). Lesion-induced changes in the ECoG spectrum are consistent with changes in both the temporal damping and strength of driving signals within a linear wave medium. The changes in temporal damping are essentially a movement away from a sustained equipartition of energy described in section 3.4.2, toward lower damping...
at the lower frequencies, thus increasing power in the slow-wave range. It seems plausible then, that smaller, physiological variation in the strength of input of these pathways could perturb the basin of attraction of the cortical system.

It can be noticed also, that with the presence of noise, or chaos at the microscopic level, that the cortical system is also analogous to a Boltzmann machine (Ackley et al 1985). This might improve reliability of the vector flow in state space rather than degrade it.

4.2. Analogy to Amit’s model of the cortical column

A model for neural network dynamics in pools of $10^5$ or so neurons, has been advanced by Amit and Tsodyks (1990, 1991). These workers have addressed, as their primary problem, storage and retrieval in real biological networks. They sought to explain why, if real neurone activity bears analogy to the attractor dynamics of Hopfield networks (Hopfield 1984a,b), persisting high firing rates are not encountered in real neurons in the absence of persisting stimulus. Instead cortical cells exhibit low mean firing rates (as discussed in section 3.3.2). To overcome the difficulties posed by earlier solutions to this problem, they have introduced a series of highly biological features to their artificial neurones, viz:

1. Neurons are excitatory and symmetrically coupled and immersed in a ‘bath’ of recurrent inhibitory cells, which are also explicitly realistically modelled.

2. They have an absolute refractory period of 2 ms.

3. Membrane potentials have continuous dynamics except at spike emission, and realistic membrane decay constants and synaptic rise times.

4. They are immersed in random continuous afferent input, as from surrounding cell activity, with short non-random input also occurring as the input to be ‘classified’.

5. Random asynchronous updating is essential, to prevent fluctuations from eliminating currently stored patterns.

As a consequence all neurons in the network have continuous, coherent, intranetwork postsynaptic potentials, below threshold, and spikes are emitted primarily due to noisy continuous afferent, but specific neurones spike more rapidly during a pattern recognition—i.e. when the assembly is centred on an attractor. Firing rates remain low.

With membrane decay constants in the 6–8 ms range the modelled network oscillates with 13 ms periodicity, in certain conditions which have not yet been systematically analysed, but are included under two general conditions. These are relatively high excitatory tone, and persisting, weak, non-random afferents.

This type of network can store and retrieve patterns as for a Hopfield network, but point attractors are replaced by non-ergodic domains. As in the case of other attractor networks (Caianiello et al 1967, Amari 1972, Little and Shaw 1975, Peretto and Niez 1986), the introduction of asymmetric connectivity would enable short temporal sequences of stimuli to be stored and retrieved. Additionally, Amit and co-workers have recently shown (Griñasty et al 1992) that their networks can convert temporal correlations between stimuli to spatial correlations between attractors, thus mirroring the experimental findings of Miyashita and co-workers (Miyashita and Chang 1988, Miyashita 1988, Sakai and Miyashita 1991).
It can be seen that the general anatomical structure, the dendritic time constants, and period of oscillation in Amit's models resemble those of Freeman's description of cortex, while the ambient noise, and asynchronous random updating make stochastic coupling between units, as in the work of our own group, applicable.

5. Conclusions

Although no simple relationship exists between neurone action potentials and the EEG, it is now sufficiently clear that the EEG cannot be dismissed as a meaningless epiphenomenon, as has sometimes been done in the past. Nor is it necessary to ascribe to the EEG some direct role in mediating brain processes for its relevance to the description of brain processes to be apparent. Instead, study of the EEG raises qualitative issues of the network scale appropriate for the description of cognitive processes. It may not be satisfactory to analyse cognitive subsystems as though each were separable, localizable and discrete. To do so may be to revisit yet again the errors of phrenology.

It is reasonable to suppose that the partial theories described above might be more formally integrated into a multiscale description suitable for simulation. Successful development and testing of such a supermodel must depend in part on whether a match can be found between parameters describing anatomical connectivity and nonlinearities, and wave characteristics at all intermediate scales. A full description would enable insight into how effects of activity at different scales in the brain may coordinate local network activities and vice versa.

We have raised the possibility that brain self-control may be achieved by frontolimbic regulation of the reticular formation thus exerting a return control on the attractor basin for the entire cortex. This is not incompatible with the autocorrelator role ascribed to the hippocampus by Rolls (Rolls 1990, Rolls and Treves 1990).

More generally, a shift of emphasis concerning input and output relationships of local neuronal models is suggested. Input to local networks should not be considered exclusively in terms of patterned input from some narrowly defined site elsewhere, but may require additional consideration of the transient biasing of all the cells in the local network by the passage of group waves through the immediate locale. Likewise, output from local networks might be thought of not so much as producing discrete output for relay to some further highly discrete site, but as producing information required to determine the future vector of the cortical system. As we have implied above, the vector of temporal dampings for the EEG may act as a guide to the present focus of the basin of attraction.

Finally, at a more philosophical level, we consider that this concept of the brain's self-control may provide some physiological basis for certain phenomenological accounts of conscious process—particularly those associated with Jaspers (Kydd and Wright 1986). Jaspers' (1963) classical descriptions of the 'present focus of consciousness' and the 'determining tendency (of cognition)' might be associated (to the extent that such a mental/somatic identification is plausible, even in principle) with the present state vector, and the present locus of the basin of attraction, respectively. In this view, the involvement of reticulo-cortical pathways in alerting, attention, and affective tone may be accounted for as operations upon the guidance of ongoing cognition, in parallel to phenomenological descriptions.
The electroencephalogram and cortical neural networks

References

Abeles M 1982 Local Cortical Circuits (New York: Springer)
Abeles M, Vaadia E and Bergman H 1990 Firing patterns of single units in the prefrontal cortex and neural network models Network 1 13-25
Adey W R 1977 The sensorium and the modulation of cerebral states: tonic environmental influences on limbic and related systems Amo. NY Acad. Sci. 290 396-420
— 1979 Neurophysiological effects of radio frequency and microwave radiation Bull. NY Acad. Med. 55 1079-93
— 1981 Tissue interactions with nonionizing electromagnetic fields Physiol. Rev. 61 435-514
— 1988 Do EEG-like processes influence brain function at a physiological level? Dynamics of Sensory and Cognitive Processing by the Brain ed E Basar (Berlin: Springer) pp 362-7
Blaworska K J and Malinowski M 1991 Nonlinear and linear forecasting of the EEG times series Biol. Cybern. 66 159-65
Bratenberg V 1978 Cortical architectonics: general and areal Architectonics of the Cerebral Cortex ed M A B Brazer and H Petsche (New York: Raven)
Broadbent D F 1958 Perception and Communication (Oxford: Pergamon)
Burns B D 1958 The Mammalian Cortex (Baltimore, MD: Williams and Wilkins)
Clavier R M and Routtenberg A 1974 Ascending monoamine-containing fibre pathways related to intracranial self-stimulation: histochemical fluorescence study Brain Res. 72 25-40
Eekman F H and Freeman W J 1990 Correlations between unit firing and EEG in the rat olfactory system Brain Res. 528 238-44
Elai R 1972 The genesis of the EEG Int. Rev. Neurobiol. 15 227-72
Fink M 1978 EEG and psychopharmacology Contemporary Clinical Neurophysiology suppl. 4, ed W A Cobb and H Van Duijn (Amsterdam: Elsevier) pp 41-56
Fraas C P and Bilowowska K J 1985 Linear model of brain electrical activity—EEG as a superposition of damped oscillatory modes Biol. Cybern. 53 19-25
Freeman W J 1964 A linear distributed feedback model for prepyriform cortex Exp. Neurology 10 525-47
— 1972 Measurement of open-loop responses to electrical stimulation in olfactory bulb of cat J. Neurophysiol. 35 745-61
— 1983 The physiological basis of mental images Biol. Psychiatry 18 1167-25
— 1987a Techniques used in the search for the physiological basis of the EEG Handbook of Electroencephalography and Clinical Neurophysiology vol 3A, ed S S Gevins and A Remond (Amsterdam: Elsevier) pp 583-664
— 1987b Simulation of chaotic EEG patterns with dynamic model of the olfactory system Biol. Cybern. 56 139-50
WORK TOWARD A THEORY OF BRAIN FUNCTION

J J Wright and R R Kydd

— 1991 Predictions on neocortical dynamics derived from studies in paleocortex Induced Rhythms of the Brain ed E Basar and T H Bullock (Basel: Birkhauser)
Freeman W J and Skarda C A 1985 Spatial EEG patterns, nonlinear dynamics and perception: the neo-Sherringtonian view Brain Res. Rev. 10 147–75
Freeman W J and van Dijk B 1987 Spatial patterns of visual cortical fast EEG during conditioned reflex in thalamic monkey Brain Res. 422 267–75
Frolov A A and Medvedev A V 1986 Substantiation of the ‘point approximation’ for describing the total electrical activity of the brain with use of a simulation model Biophysics 31 332–7
Futter J M 1973 Behavioural electrophysiology of the prefrontal cortex I. Neurophysiol. 36 61–78
Gevins A S 1988 Recent advances in neurocognitive pattern analysis Dynamics of Sensory and Cognitive Processing by the Brain ed E Basar (Berlin: Springer) pp 80–102
Hess W R 1936 Hypothalamus und die zentrales autonomes neurensystems Physiologie Archiv. für Psychiatrie und Nervenkrankheiten (Berlin) 104 548–57
— 1984a Neurons with graded response have collective computational properties like those of two state neurons Proc. Natl Acad. Sci. USA 81 3088–92
— 1984b Collective processing and neural states Modelling and Analysis in Biomedicine ed C Nicollini (Singapore: World Scientific)
Jaspers K 1963 General Psychopathology (Manchester: Manchester University Press)
John E R et al 1977 Neuronomics Science 196 1393–403
Kydd R R and Weight J J 1986 Mental phenomena as changes of state in a finite state machine Austral. New Zealand J. Psychiatry 20 158–65
Lindvall O and Björklund A 1974 The organisation of the ascending catecholamine neurone systems in the rat brain as revealed by the glycine acid fluorescence method Physiol. Scand. Suppl. 412 1–48
Little W A and Shaw G L 1975 A statistical theory of short and long term memory Behav Biol. 14 115–33
Lopes da Silva F H and Storm van Leeuwen W 1978 The cortical alpha rhythm in dog: the depth and surface profile of phase Architectonics of the Cerebral Cortex ed M A B Brazier and H Petsche (New York: Raven) pp 319–33
Marshall J F 1979 Somatosensory inattention after dopamine depleting intracerebral 6-OHDA injections: spontaneous recovery and pharmacological control Brain Res. 177 311–24
Mitsufuji U 1985 Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena Physiol. Rev. 65 37–100
Miyashita Y 1988 Neuronal correlate of visual associative long-term memory in the primate temporal cortex Nature 335 817–20
Miyashita Y and Chang H S 1988 Neural correlate of pictorial short-term memory in the primate temporal cortex Nature 331 68–70
Nauta W J and Feirtag M 1979 The organisation of the brain Sci. Amer. 241 78–105
Niki H 1974 Prefrontal unit activity during delayed alternation in the monkey Brain Res. 68 185–197
—— 1989 Generation of human EEG by a combination of long and short range neocortical interactions Brain Topography 1 199–215
—— 1992 Neurological Dynamics and Human EEG Rhythms (Oxford: Oxford University Press) to be published
Rolls E T and Treves A 1990 The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain Network 1 407–21
Rössler R O and Hudson J L 1989 Self-similarity in hyperchaotic data Brain Dynamics ed E Basar and T H Bullock (Berlin: Springer)
Sakai K and Miyashita Y 1991 Neural organisation for the long-term memory of paired associates Nature 354 152
Scheibel M E and Scheibel A B 1967 Structural organisation of nonspcific thalamic nuclei and their projection toward cortex Brain Res. 6 60–94

The electroencephalogram and cortical neural networks

361

WORK TOWARD A THEORY OF BRAIN FUNCTION

188
WORK TOWARD A THEORY OF BRAIN FUNCTION
Intracortical connectivity of pyramidal and stellate cells: estimates of synaptic densities and coupling symmetry

D T J Liley and J J Wright

Department of Psychiatry and Behavioural Science, School of Medicine,
University of Auckland, FB 92019, Auckland, New Zealand

Received 30 November 1993

Abstract. A method is outlined for estimating the the average number of synapses forming between cortical neurons as a function of their intercellular separation and the geometry of their dendritic and axonal arborization. Consideration is confined to the formation of local intracortical connections and to the case where the distribution of axonal and dendritic fibres has spherical symmetry. Parameters are deduced from quantitative anatomical studies in neurone cortex. It is demonstrated that the majority of local connections forming within a given volume of isotropic cortex can be accounted for on the assumption that local connections between neurons form randomly.

From these computations the symmetry of connection between neurons, the likely position for synapse formation on the dendritic tree and the relative synaptic densities attributable to long- and short-range interaction between excitatory and inhibitory neural subsets is determined. Local intracortical couplings appear to be highly asymmetric, and account for about 3200 synapses forming on pyramidal and stellate cells.

1. Introduction

For the purposes of simulating cortical neural networks, it would be desirable to have general analytical relationships describing the density of interaction of neurocellular components at the different scales found in the cerebral cortex. Defining this anatomical interaction has implications for the magnitudes of feedforward and feedback gains in both lumped and discrete models of cortical electro-rhythmogenesis (see following paper [21]) and also specifies the spatial transformation of activity to be expected in discrete cellular models.

The neocortex is complex. Areal variations exist in the cortex, both in terms of vertical (the familiar laminae of heterotypical cortex) and horizontal (e.g. Brodmann’s areas) organization. However, if we are prepared to make a number of simplifying assumptions, progress can be made towards defining some sort of regularity and order. While upwards of 30 morphologically distinct neuronal subtypes have been identified [16, 22], it is possible to divide the cortical neuronal population broadly into two groups based on the geometry and distribution of axonal and dendritic ramification. These are the pyramidal and stellate cells. This morphological classification also corresponds to their functionality: pyramidal cells are thought to be exclusively excitatory and the stellate cells inhibitory [4].

† The current addresses for both authors are: Mental Health Research Institute, Royal Park Hospital, Parkville, Victoria 3052, Australia and Swinburne Centre for Applied Neuroscience, Swinburne University, Hawthorne, Victoria 3122, Australia. Correspondence may be addressed to either author at either address or sent by email to dt@brain.physics.swin.edu.au or jjw@brain.physics.swin.edu.au.
Interaction between neurons occurs at a number of different spatial scales: short range (local intracortical interactions), medium range (the surface-parallel conduction fibres in layer I) and long range (the anisotropic and isotropic cortico-cortical efferents and afferents).

Anatomical data suggests that lawful relationships can be found to describe the geometric organization of branching in pyramidal and stellate cells in the visual and motor cortices of the cat [14]. Regrettably there is a deficit of similar data for axonal ramification, in part because of technical difficulties imposed by the narrow diameter of axons and, in the case of pyramidal cells, their wide area of distribution. This has necessitated our attempt to describe axonal branching using very schematized models of axonal growth that ignore many features of development.

For the purposes of this paper we confine ourselves to the investigation of local intracortical connectivity between pyramidal and stellate cells, developing a simple 'stochastic' method for determining the expected number of synapses between two cells separated by a distance \( r \). This method was motivated by the work of Uttley [19].

However, Uttley's treatment never dealt with the more general cases, had a fatal theoretical error and, due to the paucity of anatomical data at the time, was never applied to estimating total synaptic numbers or coupling symmetry.

2. A method for calculating connectivity

2.1. Basic assumptions

The following assumptions apply throughout. Other specific assumptions and simplifications are introduced as needed.

- The basal dendritic system of a neuron has approximately spherical symmetry. This implies that the expected distribution of axonal and dendritic branches crossing any arbitrary concentric sphere about the cell body is uniform.
- The apical dendritic tree can be ignored for calculation of intracortical connectivity as it usually ramifies in the uppermost layers and is believed not to be a significant recipient of local intracortical axons.
- Only axo-dendritic connections need be accounted for.
- As the volume of axonal and dendritic branching per cell is small compared with the volume in which the fibres distribute themselves the probability of a synapse forming between any two cells, at any particular point in the field of their overlap, is small.
- Systematic variations in dendritic radius and horizontal and vertical anisotropies in cortical organization can be ignored for the purposes of simplicity.
- Axonal fibre density for all neurons can be described by a radially homogenous exponential distribution of the form \( a_{nx}e^{-r/r_0} \), where \( r_0 \) is the axonal space constant, and \( a_{nx} \) is the 'density' of the axonal tree at \( r = 0 \).

Our approach uses certain established quantitative anatomical measurements as primary to our calculations. Their specific values are introduced as needed, while both primary and derived statistics are presented in tabular form (table 1).

2.2. Formulation

We choose to treat each fibre tree as composed of small straight segments. These segments may correspond physically to the sections between bifurcations or fibre deviations. Because the growth of the dendritic and axonal trees is complex, we ignore correlations between the
Cortical connectivity of pyramidal and stellate cells

Table 1. Tabular outline of all the empirical data used to calculate derived estimates of cortical connectivity. Also shown, for comparison, is the correspondence between quantities which can be estimated both empirically and theoretically.

<table>
<thead>
<tr>
<th>Quantity Description</th>
<th>Empirical value</th>
<th>Derived value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Reference</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Equation</td>
</tr>
<tr>
<td>$r_{ax}$ radius of an axonal fibre ($\mu$m)</td>
<td>0.15 [3]</td>
<td>—</td>
</tr>
<tr>
<td>$r_d$ radius of a dendritic fibre ($\mu$m)</td>
<td>0.45 [3]</td>
<td>—</td>
</tr>
<tr>
<td>$\rho$ cell density (mm$^{-3}$)</td>
<td>$9 \times 10^4$ [3]</td>
<td>—</td>
</tr>
<tr>
<td>$\rho_r$ relative pyramidal cell density</td>
<td>0.85 [3]</td>
<td>—</td>
</tr>
<tr>
<td>$\rho_s$ relative stellate cell density</td>
<td>0.15 [3]</td>
<td>—</td>
</tr>
<tr>
<td>$L_D$ stellate cell dendrite length (mm)</td>
<td>2.16 [3]</td>
<td>—</td>
</tr>
<tr>
<td>$L_D$ total pyramidal cell dendrite length (mm)</td>
<td>3.08 [3]</td>
<td>—</td>
</tr>
<tr>
<td>$L_A$ stellate cell axon length (mm)</td>
<td>20 [3]</td>
<td>—</td>
</tr>
<tr>
<td>$L_A$ total local pyramidal axon length (mm)</td>
<td>20 [3]</td>
<td>—</td>
</tr>
<tr>
<td>$a$, $r_0$ parameters describing branching of pyramidal basal and stellate dendrite ($\mu$m$^{-3}$, $\mu$m)</td>
<td>0.0028, 31.25 [15,14,15]</td>
<td>—</td>
</tr>
<tr>
<td>$a_{ax}$, $r_2$ parameters describing branching of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>local pyramidal cell axon</td>
<td>—</td>
<td>0.8, 3.17 (15)</td>
</tr>
<tr>
<td>stellate cell axon ($r_{ax}^2$, $r_0$)</td>
<td>—</td>
<td>2.73, 2.12 (15)</td>
</tr>
<tr>
<td>$d_A$ stellate axonal intersynapse spacing ($\mu$m)</td>
<td>1.25–5.0 [3]</td>
<td>4.45 $L_A^{N_{ax,ax}}$ (24)</td>
</tr>
<tr>
<td>$d_{Ax}$ pyramidal axonal intersynapse spacing ($\mu$m)</td>
<td>1.25–5.0 [3]</td>
<td>6.81 $L_A^{N_{ax,ax}}$ (24)</td>
</tr>
<tr>
<td>$d_D$ stellate dendritic intersynapse spacing ($\mu$m)</td>
<td>0.33 [3]</td>
<td>0.68 $L_D^{N_{ax,ax}}$ (24)</td>
</tr>
<tr>
<td>$d_{Dx}$ pyramidal dendritic intersynapse spacing ($\mu$m)</td>
<td>0.4</td>
<td>0.68 $L_D^{N_{ax,ax}}$ (24)</td>
</tr>
<tr>
<td>$N_{ax,ax}$ total number of synapses per stellate cell dendrite 6545</td>
<td>3169 (24)</td>
<td></td>
</tr>
<tr>
<td>$N_{ax,ax}$ total number of synapses per pyramidal cell basal dendrite 6545</td>
<td>3169 (24)</td>
<td></td>
</tr>
<tr>
<td>$N_{ax,ax}$ total number of synapses per stellate cell axon 4000–16000</td>
<td>4496 (24)</td>
<td></td>
</tr>
<tr>
<td>$N_{ax,ax}$ total number of synapses per pyramidal cell local axon 4000–16000</td>
<td>2935 (24)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Mean of all cortical areas of the mouse for all quantities except $a$ and $r_0$ measured in visual cortex of the cat.

positions and orientations of these segments. Therefore we are considering a large number of small straight axonal and dendritic segments, distributed randomly, but not necessarily uniformly, over space, orientation and length.

Further, we assume that these segments are so short that the probability that a given axonal segment will intersect the same dendritic tree more than once is negligible. Additionally we require that the orientations of the fibres are uniformly distributed over the unit sphere, and that location, orientation and length are independent random variables.

Let each segment of the axonal or dendritic tree be described by the location in space of its centre, $r$, its length $l$ and its orientation ($\theta$, $\phi$) (the familiar polar angles). Therefore the number of segments (axonal (A) or dendritic (D)) within $d^2r$ about $r$, with length between $l$ and $l + dl$ and orientation in $(\theta, \theta + d\theta) \times (\phi, \phi + d\phi)$ is given by

$$dN_k = \rho_k(r) f_k(l) \frac{\sin \theta}{4\pi} dl \, d\theta \, d\phi$$

where $k = A, D$

An axon segment with orientation ($\theta$, $\phi$) and length $l_A$, will intersect a dendritic segment with centre at $r_D$, orientation ($\theta$, $\phi$) and length $l_D$, if its centre, $r_A$, lies in a parallelepiped about $r_D$ (see figure 1). When this happens a synapse is said to have formed.
Here we have assumed that the diameters of the fibres are small compared with the linear dimensions of the region. We have also ignored the detailed shape of the sides of the region. The volume of this region is seen to be

\[ V_{\text{int}}(l_A, l_D, \theta, \phi, \eta, \psi) = 2\epsilon l_A l_D \sin \gamma(\theta, \phi, \eta, \psi) \]  

where \( \epsilon = r_{ax} + r_d \) (\( r_{ax}, r_d \) are the average axonal and dendritic radii respectively) and \( \sin \gamma(\theta, \phi, \eta, \psi) \) is the angle between an axonal and dendritic segment. If this region is small enough for the axonal segment density function to be regarded as constant within it, then the number of intersections of axons of this orientation and length with the dendritic segment is

\[ \rho_A(r_A)V_{\text{int}}. \]  

For the more general case, the total number of synapses forming, \( N_s \), is

\[ N_s = \int \cdots \int 2\epsilon l_A l_D \sin \gamma(\theta, \phi, \eta, \psi)\rho_A(r_A)\rho_D(r_D)f_A(l_A)f_D(l_D) \]

\[ \times \frac{\sin \theta \sin \eta}{4\pi} \frac{1}{4\pi} \, d^3r_A \, d^3r_D \, d\theta \, d\phi \, d\eta \, d\psi. \]

Defining

\[ \bar{I}_k = \int l f_k(l) \, dl \quad \text{for} \quad k = A, D \]

we are able to define the length of axonal or dendritic fibre per unit volume as

\[ A(r) = \bar{I}_A \rho_A(r) \quad D(r) = \bar{I}_D \rho_D(r). \]

Thus, using equations (5) and (6) equation (4), the expected number of synapses forming between a presynaptic cell \( i \) and a postsynaptic cell \( j \), can be written, after partial integration, as

\[ N_{ij} = \frac{\pi \epsilon}{2} \int \lambda_i(r)D_j(r) \, d^3r. \]

If we define either the axonal or dendritic fibre density function relative to the position of the respective cell body, then (7) can be rewritten as the convolution

\[ N_{ij}(s) = \frac{\pi \epsilon}{2} \int \lambda_i(r)D_j(r - s) \, d^3r \]
where \( s \) is the distance separating the pre- and postsynaptic cells.

The probability that a single axonal segment selected at random will intersect a dendrite somewhere in the region of co-arborization will be small. The number of such axonal segments will probably be large. Therefore the distribution of such connections or synapses will be reasonably well described by a Poisson distribution \([7]\). Hence the probability that a postsynaptic neuron will receive \( q \) connections from a presynaptic neuron is

\[
p(q, r) = \frac{\exp[-N_{ij}(r)] N_{ij}(r)^q}{q!}.
\]

\( \text{(9)} \)

2.3. Specifying the axonal and dendritic fibre densities

2.3.1. The dendritic tree. For the case of a spherically symmetrical dendritic tree the dendritic fibre density at a distance \( r \) from the soma, \( D \), can be determined from anatomical data as \([14]\)

\[
D = \frac{n}{(4\pi r^2 \cos \theta)}
\]

\( \text{(10)} \)

This is shown in figure 2.

![Figure 2](image)

Assuming that all dendritic fibres intersect this imaginary sphere at right angles \( \cos \theta = 1 \) (i.e. all fibres in a unit volume can be considered to be parallel to each other) then the length of fibre per unit volume is equal to the number of fibres crossing a unit area multiplied by a unit distance.

Sholl \([14]\) has measured the number of dendrites crossing concentric spheres of radius \( r \) centred on the perikaryon for the dendritic systems of stellate and pyramidial neurons in the striate and motor areas of the cat. Spherical symmetry holds approximately for the basal dendritic system, and Sholl \([14]\) has shown that \( D \) falls off exponentially with \( r \). Thus \( D \) is given empirically as

\[
D = ce^{-r/\theta}.
\]

\( \text{(11)} \)
The values of \(a\) and \(r_0\) vary little between the dendritic systems of stellate and pyramidal cells. Thus we have chosen, from Sholl [14] average regression estimates of \(a\) and \(r_0\) as 0.0028 \(\mu m^{-2}\) and 31.25 \(\mu m\). If we measure all distances in units of \(r_0\), the treatment becomes dimensionless making computation easier. Taking this step

\[
D = ae^{-\gamma}
\]

where

\[
a = 2.73 r_0^{-2}.
\]

Measurements of the radius of dendrites and axons can similarly be modified. Electron micrographs suggest [4] that \(r_d\) is of the order of 0.45 \(\mu m\) and \(r_{ac}\) is of the order of 0.15 \(\mu m\).

\[
r_d = 0.014r_0 \quad r_{ac} = 0.0048r_0.
\]

2.3.2. The axonal tree. Statistical descriptions of axonal systems analogous to those for dendrites appear not to exist. For what follows we assume that axonal fibre density for all neurons can be described by a radially homogenous exponential distribution of the form

\[
A = a_{ac}e^{-r/r_{c}}.
\]

Using anatomical information we can make a tentative attempt to estimate \(r_a\) and \(a_{ac}\) for pyramidal and stellate cells. These two parameters can be estimated by observing that:

- cell diameter \(2r_c \approx 20 \mu m\);
- only one axon emanates from the soma for pyramidal cells and \(a_{ac}\) is estimated as

\[
a_{ac} = \frac{1}{4\pi r_c^2}.
\]

The pattern of axonal ramifications for stellate cells is highly variable with dense branching near the cell body [4, 15]. Thus in the absence of contrary evidence, we assume that the axonal and dendritic branching near the soma resemble each other. This being the case, we take \(a_{ac}\) for stellate cells to be 2.73 \(r_0^{-2}\).

For both stellate and pyramidal cells we have taken the total length of local axon to be 20 \(mm\) (or 640 \(r_0^{-1}\)) based on anatomical values obtained from murine cerebral cortex [4]. These values appear to represent upper bounds.

Thus, as \(a_{ac}\) is known, we can calculate \(r_a\) by integrating (15) for a radially symmetric ‘exponential’ tree,

\[
r_a = \left(\frac{L}{8\pi a_{ac}}\right)^{1/3}
\]

where \(L\) is the total length of local intracortical axon. Solving equation (17) we obtain \(r_a = 3.17\) for pyramidal cells and \(r_a = 2.12\) for stellate cells.

2.4. Calculation of the number of synapses given or received

The fibre density distributions of the axonal and basal dendritic trees are approximated by equations (11) and (15). Equation (8), exploiting the symmetry inherent in the problem, can then be written as

\[
N_d(r) = \pi^2 e a_{ac} \int_0^{\infty} \int_{-\infty}^{\infty} y \exp \left(-\frac{[(x-r/2)^2 + y^2]^{1/2}}{r_a} - [(x+r/2)^2 + y^2]^{1/2}\right) dx dy.
\]
Cortical connectivity of pyramidal and stellate cells

This can be integrated, by changing to elliptic coordinates, to yield

\[ N_{ij}(r) = 2\pi^2(r_a + r_{xx})a_{xx}Q_{ij}(r) \]  

(19)

where

\[ Q_{ij}(r) = \frac{1}{rf^2d^3}[e^{-r/r_0}(d^2(f'r - 2) + f'^2(dr + 2)) + e^{-r/r_0}(d^2(f'r + 2) - f'^2(dr + 2))] \]  

(20)

and

\[ N_{ij}(0) = \frac{4\pi^2(r_a + r_{xx})a_{xx}}{d^3} \]  

(21)

\[ f = \frac{(r_a - 1)}{r_a} \]  

(22)

\[ d = \frac{(r_a + 1)}{r_a} \]  

(23)

Figure 3 shows the connectivity profile for two pyramidal cells obtained by substituting (19) with the parameter values obtained in section 2.3. Figure 4 shows the connectivity profile for a presynaptic stellate cell and a postsynaptic pyramidal cell.

Figure 3. The probability (p) of n synapses forming between two pyramidal cells separated by a radial distance r (in units of r0).

Figure 4. The probability (p) of n synapses forming between a presynaptic stellate cell and a postsynaptic pyramidal cell separated by a radial distance r (in units of r0).
3. Some consequences for cortical connectivity

We can now estimate:

- the total number of synapses a stellate cell is likely to give and receive;
- the number of local intracortical synapses a pyramidal cell will receive and project;
- the likely distribution of synapses within the dendritic tree;
- coupling symmetry.

3.1. Total synaptic number

The total number of synapses given by a presynaptic cell of type \( k \) located at \( r_0 \), \( N_{\text{total},k}^A(r_0) \), is

\[
N_{\text{total},k}^A(r_0) = \sum_i \int_R d^3r \, N_{kl}^A(r - r_0) \rho_k(r)
\]  

(24)

where \( N_{kl}^A(r) \) is the "connectivity" function for the \( k \)th presynaptic cell type and the \( l \)th postsynaptic cell type, \( \rho_k(r) \) is the \( l \)th cell density for the postsynaptic cell of the pair. A similar expression exists for \( N_{\text{total},j}^D \).

The estimators for the mean and variance of the total number of synapses, are, assuming constant cell density,

\[
\mu_k = N_{\text{total},k}^A
\]  

(25)

\[
\sigma_k^2 = \sum_i \rho_i N_{\text{total},kl}^A
\]  

(26)

where \( N_{\text{total},kl}^A \) is the total number of synapses forming between a presynaptic cell of type \( k \) and all other postsynaptic cells of type \( l \).

When (19) is substituted into (24), all parameters to calculate \( N_{\text{total},k}^A \) are available. For constant cell density, \( 9\times10^4 \text{ mm}^{-3} \) [4], over the region of integration,

\[
N_{\text{total},k}^A = 16\pi^3 \rho_0 \sigma_{\text{ax}} \sigma (r_a + r_{ax}) Q
\]  

(27)

where

\[
Q = \frac{r_a^5}{(r_a - 1)^2(r_a + 1)}
\]  

(28)

Table 1 summarizes values of \( N_{\text{total}} \) for derived axonal space constants for stellate and pyramidal cells on the assumption that 85% of neurons are pyramidal and 15% stellate and that cells are mixed randomly [4].

3.2. The likely position for synapse formation

Equations (8) and (11) also allow us to predict the radial distance from the soma of a postsynaptic neuron at which a synapse from a given presynaptic cell is most likely to form. The mean radial position of a synapse is

\[
\langle |s| \rangle_r = \frac{\pi e \int |s| A(s) D(r - s) d^3s}{2N(r)}
\]  

(29)

Figure 5 shows \( \langle |s| \rangle_r \) as a function of intercellular distance and axonal space constant.
3.3. Estimates of coupling symmetry

Figures 6 and 7 show a number of estimates of the symmetry of connection between neurons, based on

$$P_{ij}(n, r_{ij}) = p_{ij}(n, r_{ij})p_{ji}(n, r_{ji})$$

(30)

where $P_{ij}$ is the probability that two cells $i$ and $j$, separated by an intercellular distance $r_{ij}$, will give each other exactly $n$ synapses. $p_{ij}$ and $p_{ji}$ are calculated from equations (8) and (9).

We define two measures of symmetry. The first we shall call 'strong' symmetry. By this we mean that two cells will give each other exactly $n$ synapses.

The second measure we shall call 'weak' symmetry; this is the probability that two cells sharing a total of $n$ presynaptic contacts will have these synapses distributed as $n - m$ to one cell and $m$ to the other, where $0 < m < n$.

$$P_{ij}^{s}(r_{ij}) = \sum_{n=1}^{\infty} p_{ij}(n, r_{ij})p_{ji}(n, r_{ji}).$$

(31)

For the case of weak symmetry we have the discrete convolution

$$P_{ij}^{w}(r_{ij}) = \sum_{n=2}^{\infty} \sum_{m=1}^{n-1} p_{ij}(m, r_{ij})p_{ji}(n-m, r_{ji}).$$

(32)
Combining these measures of symmetry defines a quantity called the ‘expected symmetry’ on the range $[0,1]$ (0 being the case where no connections are shared, 1 being the case when all connections are reciprocal), which varies with intercellular separation

$$\Phi_{ij}(r_{ij}) = 1 - E \left( \frac{m-n}{m+n} \right) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \left( 1 - \frac{m-n}{m+n} \right) p_{ij}(m, r_{ij}) p_{ij}(n, r_{ij}).$$ (33)

Figures 8 and 9 show $\Phi_{ij}(r_{ij})$, $P^p_{ij}(r_{ij})$ and $P^s_{ij}(r_{ij})$ for a pair of pyramidal cells and a stellate and pyramidal cell.

3.4. A general schema for connectivity in homotypical cortex

Figure 10 shows derived estimates for axo-synaptic couplings in an arbitrary volume of homotypical cortex. We define the axo-synaptic coupling density between the various cell groups to be the fraction of the total number of synapses per unit volume involved in this interaction. The figures were derived on the following assumptions.

- The basal dendritic tree and the dendritic tree of the stellate cell are of approximately the same length and the variation of this dendritic fibre density with respect to the distance
from the cell body is described by the same relationship (viz equation (11)) [4, 16].

- Termination of cortico-cortical afferents is assumed to be distributed uniformly over all layers. This has been demonstrated to be so for prefrontal cortico-cortical efferents in rhesus monkeys [9]. There is, however, great variability and specificity in the laminar distribution of such efferents as well as anisotropy of connectivity of ipsi- and contralateral cortico-cortical projections [11].

- For a given volume of cortical tissue the number density of incoming cortico-cortical fibres equals the number of outgoing cortico-cortical fibres [2, 16]

- Non-specific subcortical efferents and afferents have been conservatively calculated at 1% of the total number of cortical efferents and afferents [2]

Under these assumptions the fraction of the total number of synapses per unit volume involved in the local intracortical interaction between the i-th presynaptic cell group and the j-th postsynaptic cell group is

\[
\beta_{ij} = \frac{\rho_i \rho_j \int d^r r \, N_{ij}}{N_{\text{total}}}
\]

where \(N_{ij}\) is defined by equation (19), \(\rho_i, \rho_j\) are the respective cell densities and the \(\beta_{ij}\) are the local intracortical anatomical coupling coefficients between the given neuronal groups.

\[
N_{\text{total}} = \sum_k \rho_k (N_{\text{total}, e}^\sigma + N_{\text{total}, i}^\sigma + N_{\text{total}, k}^\mu)
\]

where the addends on the right side of (35) are the number of synapses on dendrites due to local, cortico-cortical and subcortical interactions respectively.

Similarly the fraction of the total number of synapses involved in long-range interactions between excitatory cells and between excitatory and inhibitory cells is

\[
\alpha_{cc} = \frac{\rho_e \rho_e N_{\text{total}, e}^\sigma}{N_{\text{total}}} \quad \alpha_{ci} = \frac{\rho_i \rho_e N_{\text{total}, i}^\sigma}{N_{\text{total}}}
\]

where \(\rho_e, \rho_i\) are the densities of pyramidal and stellate cells respectively, and \(N_{\text{total}, e}^\sigma, N_{\text{total}, i}^\sigma\) are the number of synapses per excitatory cell and inhibitory cell, respectively, that are due to long-range afferents. Because stellate cells are not believed to give rise to long-range projections \(\alpha_{ei}\) and \(\alpha_{ii}\) are necessarily zero.
The $\mu_{\lambda}$, the relative synaptic densities attributable to subcortical fibres, are determined in a similar manner to the anatomical coupling coefficients in equation (36). $N_{\text{total},e}$ and $N_{\text{total},i}$ can be estimated by noting that

$$N_{\text{total},e} + N_{\text{total},e}^\mu = \frac{L_p^D}{d_p^D} - N_{\text{total},e}^D$$

$$N_{\text{total},i} + N_{\text{total},i}^\mu = \frac{L_i^D}{d_i^D} - N_{\text{total},i}^D$$

(37)

where $L_p^D$ is the total length of pyramidal cell dendrite, $L_i^D$ is the total length of stellate cell dendrite and $d_p^D$, $d_i^D$ are the anatomically determined intersynapse spacing on pyramidal and stellate cell dendrite respectively, and where

$$N_{\text{total},e}^\mu = \sum_p N_{\text{total},e}^{\mu,uv} \quad N_{\text{total},i}^\mu = \sum_p N_{\text{total},i}^{\mu,uv}$$

(38)

where the sums are over all afferent subcortical fibre systems, $v$.

If the isotropic cortico-cortical and subcortical fibre systems can be assumed to contribute cortical synapses in direct proportion to the relative numbers of each fibre type entering an arbitrary volume of cortex we can make progress towards defining values for the quantities on the left-hand side of (37). Anatomical evidence suggests that the numbers of fibres entering cortex from subcortical and distant cortical areas are in the ratio 1:100 [2, 3, 12]. Denoting this ratio by $\kappa$, from (37) we have

$$N_{\text{total},e}^\mu = \kappa N_{\text{total},e}^\alpha \quad N_{\text{total},i}^\mu = \kappa N_{\text{total},i}^\alpha$$

(39)

and

$$N_{\text{total},e}^\alpha = \frac{1}{\kappa + 1} \left( \frac{L_p^D}{d_p^D} - N_{\text{total},e}^D \right)$$

$$N_{\text{total},i}^\alpha = \frac{1}{\kappa + 1} \left( \frac{L_i^D}{d_i^D} - N_{\text{total},i}^D \right)$$

(40)

A putative estimate of the magnitude of the anatomical coupling coefficients in homotypical murine neocortex is shown in figure 10.

4. Conclusions

This model is an a posteriori model. No attempt has been made to incorporate the effects of growth and development. Nonetheless, this simple formulation is able to account for the observed pre- and postsynaptic synaptic densities seen in adult murine cortex, within the limits of experimental uncertainty. We have calculated the mean total number of synapses (from local fibre distribution only) for the basal dendritic tree of a pyramidal cell and the dendritic tree of the stellate cell to be $\sim 3200$. From this we deduced that the expected number of synapses on a pyramidal cell, due to long-range isotropic cortico-cortical fibres and subcortical fibres, is of the order of 4500. Thus the total number of synapses a pyramidal cell receives is about 7700. From this we conclude that the total number of synapses per mm$^3$ is $7.4 \times 10^9$. Braitenberg and Schutz [4] have measured, using detailed electron-micrographs, a mean value of $7-9 \times 10^9$ synapses per mm$^3$. Based on this they estimate that, on average, a pyramidal cell in the murine cortex would expect to form about 8000 synapses. Table 1 summarizes all the empirical data used to determine the theoretical values outlined above, together with estimates of these values.
In section 3.4, we showed that, in terms of the numbers of synapses contributed from 'non-local' and local fibre systems, the effect of short-range pyramidal axons is about half as strong as that due to long-range afferents, in the case of the mouse.

In our calculations the ratio of thalamic-cortical to cortico-cortical afferents (\( \kappa \) in equation (39)) has been treated very conservatively. Thalamic-cortical afferents in the mouse may be as much as 10-fold greater than the 1% figure cited for the human work, but no clear estimates for the mouse are presently available. Larger values of \( \kappa \) would lower the estimates for the cortico-cortical connectivity determined by our method.

This formulation can be generalized to take into account any arbitrary distribution of axonal or dendritic fibre density. For instance, the distribution of pyramidal cell axons may be better described by a functional form other than a radially homogenous exponential axonal field.

In this treatment we have not accounted for the axo-somatic synapses formed exclusively by stellate cells on pyramidal cells [4]. For a stochastic mechanism to describe the formation of these synapses, it would require axonal growth and synaptogenesis in stellate cells to precede axonal and dendritic development in pyramidal cells. Alternatively, or in addition, chemo-affinic mechanisms may need to be considered [10].

It has been assumed that a synapse forms only when an axonal fibre and a dendritic fibre are in direct physical contact. However, the effect, as yet incompletely verified, of soluble trophic metabolic gradients in initiating and facilitating synapse formation has not been included. The effect of this may be to modify \( \epsilon \) such that two fibres do not need to be in physical apposition for a synapse to form. In principle such a mechanism can be easily incorporated in our formulation, as can areal variations in cell density and branching.

We have ignored, for the present, the role spines may have in promoting and modifying connectivity. Spines are thought to have a role in increasing the surface area of a dendrite such that a greater number of synaptic contacts can be accommodated [5]. The latter assertion is difficult to demonstrate due to the difficulty in obtaining quantitative anatomical evidence. Despite this, certain general conclusions can be arrived at. If the number of synapses per unit length of smooth dendrite is compared with the number of synapses per unit length of spiny dendrite, the figures are essentially equivalent – about 3 synapses/\( \mu \)m [4]. Also, it is rare to find a case in which a spine does not have an associated synapse.
simple calculation, which we shall omit, shows that if spine development preceded synapse formation, then for the case of adult murine cortex, one would expect only 13% of all spines in the cortex to have an associated synapse. Thus, the role of spines in facilitating the formation of connections in the cortex remains doubtful. For these and other anatomical reasons [13] we have ignored their effect.

Autosynaptic and autoinhibitory synapse formation appears to be significant, in contrast to some expectation [8]. Autoinhibitory synapses are more likely as the stellate cell axon arborizes within the volume of distribution of dendritic ramification. Referring to figure 4 at \( r = 0 \), and as the branching of the pyramidal cell basal dendritic tree is approximately the same as that of the dendritic tree of the stellate cell, we see that 1 to 3 autoinhibitory synapses per stellate cell are likely to form.

One important aspect of our conclusions is the high asymmetry of cortical neuronal connections which our calculations imply. We discuss this in more detail elsewhere [20]. Amit [1] has investigated the case for networks in which asymmetry has been introduced into a symmetric network having \( N^2 \) synapses, by deleting a fraction \( \gamma \) (typically defined on \( 0 < \gamma \leq 1 \)) of \( \frac{1}{2}N^2 \) synapses. For the case of 100% dilution this implies that every pair of neurons has probability 0.25 of being reciprocally connected (i.e. \( \gamma^2/4 - \gamma + 1 \)). Although the analogy with our calculations of asymmetry is incomplete, from our estimates, for strong asymmetry, the likelihood of reciprocally connected pyramidal cells is 0.16, for weak asymmetry this rises to 0.44 (for two pyramidal cells separated by one dendritic space constant). Thus a real neural network, by analogy with a spin glass, is 120% diluted for the case of strong symmetry and 67% diluted for the case of weak symmetry. At 100% asymmetric dilution spin glass effects are totally absent. However, Derrida et al [6] have shown, for the case of extreme asymmetric dilution, that such networks may perform as well as their symmetric counterparts. In fact, the storage capacity per synapse is higher than that for a fully connected network. At this extreme dilution almost all feedback loops are eliminated, i.e. the network forms a branched graph.

Additional tests of the plausibility of this model of cortical connectivity might be found in experiments in which spike-triggered averaging of postsynaptic potentials is used to identify local circuit connections between neurons [17, 18].

As a partial test of the approximate validity of these results – in particular the schema for connectivity in homotypical connectivity – the following paper [21] introduces our calculations of relative synaptic density into a model of the origin of the electrocorticogram, in the form of coupling coefficients among lumped aggregates of excitatory and inhibitory cells.

Acknowledgments

The authors would like to thank Mr Mike Pemberton of the Department of Mathematics at the University of Queensland and Dr Peter Cadusch of the Department of Physics, Swinburne University of Technology, for many helpful comments and suggestions. This work was supported by grants from the Health Research Council of New Zealand and the Auckland Medical Research Foundation.

References


Cortical connectivity of pyramidal and stellate cells


[5] Collinser M 1968 Synaptic patterns on different cell types in the different laminae of the cat visual cortex. An electron microscope study Brain Res. 9 268–87


[16] Sholl D A 1956 The Organization of the Cerebral Cortex (New York: Wiley)


[22] y Cajal S Ramon 1911 Histologie du systeme nerveux de l’homme et des vertebres trans L Azoulay (Madrid: Consejo superior de investigaciones cientifica, Instituto Ramon y Cajal)
A millimetric-scale simulation of electrocortical wave dynamics based on anatomical estimates of cortical synaptic density

J J Wright and D T J Liley†

Department of Psychiatry and Behavioural Science, School of Medicine, University of Auckland, PB 92019, Auckland, New Zealand

Received 30 November 1993

Abstract. Estimates of the connectivity and synaptic densities of pyramidal and stellate cells of the cerebral cortex obtained earlier permit simulation of a lumped cellular network in which synaptic densities are translated into network coupling coefficients. The simulation reproduces spectral, autoregression and frequency/wavenumber properties of real electrocortical waves, including all major cerebral rhythms. The system control parameter is the mean strength of non-specific afferent inputs to the cortex, by analogy with non-specific cortical activation by the reticular formaion and associated pathways. This control regulates total power, spectral density, spatial dumping and wave velocity. The range of velocities also corresponds to physiological measurement.

1. Introduction

In the preceding paper [9] we described estimates of the asymmetry and synaptic density of connections among pyramidal and stellate cells of the cerebral cortex. In the present paper we use the estimates of synaptic density as coupling parameters for a model of electrocortical activity at millimetric scale. This permits indirect test of the calculations of synaptic density, by comparison of the simulation’s performance with physiological data.

Our wider motivation is to link network simulations potentially capable of logical operations to a gross observable of cognitive function in real brains — specifically, the electrocorticogram (EEG). In an earlier review [20] we have outlined problems of this field, and the models which address them.

The lumped neural network reported on here depends only upon synaptic densities, very simplified dendritic properties, and an assumed Gaussian distribution of action potential thresholds with respect to the local field potential. A single control parameter is used. The model’s only free parameter (mean synaptic gain) is optimized so as to enable the control parameter to operate in the physiological range.

The anatomical range and distribution of fibres are those obtained by Brairtenberg and Shiz [3] from the mouse cortex, and those deriving from our further calculations which are also based upon mouse statistics, although the simulations are compared for experimental

† The current addresses for both authors are: Mental Health Research Institute, Royal Park Hospital, Parkville, Victoria 3052, Australia, and Swinburne Centre for Applied Neuroscience, Swinburne University, Hawthorn, Victoria 3122, Australia. Correspondence may be addressed to either author at either address, or sent by email to dl@brain.physics.swin.oz.au or jjw@brain.physics.swin.oz.au.
veracity with ECoG from the cat. At the scale involved there is anatomical reason to believe that such density and range estimates correspond roughly across species [4, 16, 18].

Comparison with earlier EEG models and neural networks is deferred to the conclusion (section 5).

2. Basis of the simulation

We make three assumptions:

(i) At millimetric scale synaptic coupling densities depend upon random cell contacts, under the constraint of the stochastic geometry of the axonal and dendritic trees.

(ii) Cell threshold potentials have a Gaussian distribution with respect to local field potential.

(iii) The local field potential of a volume of cortical tissue is proportional to the lagged action density of excitatory synapto-dendritic transmissions.

Assumption (i) was the basis for calculation of connectivity and coupling strengths in the preceding paper [9]. Assumption (ii) is implicit in assumption (i) and assumption (iii) is a consequence of Poisson's equation [12] applied to summation of a uniformly distributed population of current sources and sinks in a conductive medium homogeneous at millimetric scale. The negligible contribution to the ECoG made by non-pyramidal cells [11] means that the ECoG may be considered as the product of excitatory processes only.

Further simplifications are intended only as interim at this stage of modelling, and are introduced in context below.

2.1. Unit volume

The use of a rectangular recording array to obtain ECoG in our animal experiments sets a scale limit on the ECoG space frequencies which can be resolved, and thus on the smallest volume of cerebral cortex which can enter as a unit volume into a lumped simulation. The unit volume is given by \( \pi r^2 \theta \) where \( 1/r \) is the highest spatial wavenumber which can be resolved by the array, and \( \theta \) is the cortical thickness. We have selected an array size such that within unit volume intracortical connections form a substantially closed system and cortico-cortical fibres form the principal mode of connection between unit volumes. This condition is met by electrode separations of 0.86 mm, resolving cortex electrotopography to approximately 1.6 mm, a distance greater than the range of intracortical fibres and much less than the cortico-cortical fibre range.

2.2. Pulse density

We define pulse density, \( Q \), as the number of cells in unit volume which are above threshold, normalized to the range zero to one. Pulse partial densities of excitatory and inhibitory components are distinguished from section 2.6 onward. For the \( N \) cells in unit volume

\[
Q = \left( \frac{1}{N} \right) \sum_{i=1}^{N} q_i
\]

where \( q_1 \ldots q_N \) is the state vector of the cells and \( q_i = 0 \) or 1 depending on whether the individual cell is sub- or suprathreshold.
2.3. Threshold distribution and local field potential

The local field potential, \( V \), and pulse density, \( Q \), under assumptions (ii) and (iii), can be related approximately by the logistic function

\[
Q = (1 + e^{\alpha V})^{-1}
\]  

(2)

where \( \alpha = -9.919 \) is an arbitrary value chosen to ensure that almost all cell threshold potentials lie in the range \( V = -1 \) to \( +1 \). Thus, for \( V = -1 \) to \( +1 \), \( Q \) ranges from 0.000202 to 0.999898. For values of \( V \) between \(-1\) and \(-0.5\), \( Q \) is a small fraction and the \( Q/V \) relation is effectively linear.

\[
dQ/dV \text{ is proportional to the density of cells 'near threshold', where 'near threshold' means that input to one further excitatory synapse would induce action potentials in the postsynaptic cell.}
\]

2.4. Threshold sensitivity gain

Since afferent fibres reaching the unit volume from all sources are assumed to make random synaptic contact within the volume, pulses in these afferents raise cells above threshold with effect proportional to the number of cells presently near threshold. Thus, for small perturbations about a negligibly-low base-firing rate, a cell mass has a gain factor, \( T \), the efferent/afferent pulse ratio, which is proportional to \( dQ/dV \), and from equation (2)

\[
dQ/dV = -(1 + e^{\alpha V})^{-2} \alpha e^{\alpha V}.
\]  

(3)

2.5. Synaptic density

Synapses within a unit volume, classified by the origin and termination of axonal fibres, have a characteristic density for each class, as calculated in the preceding paper and shown in figure 10 of that paper [9]. This figure is reproduced here as figure 1. The synaptic densities are defined as follows:

- \( \alpha_{ec}, \alpha_{ee} \): Synaptic densities of cortico-cortical fibres afferent to inhibitory and excitatory cells respectively. For a typical unit volume at cortical coordinates (0,0), partial densities can be defined by their axonal origins from surrounding volumes at \((x, y)\), as \( \alpha_{ee}(x, y) \), etc.
- \( \beta_{ec}, \beta_{ei}, \beta_{ee}, \beta_{ei} \): Synaptic densities of intracortical connections, excitatory to excitatory, excitatory to inhibitory, etc.
- \( M_{ee} \): Synaptic densities of specific cortical afferents, for those special cortical volumes which receive such afferents.
- \( \mu_{ee}, \mu_{ei} \): Synaptic densities of non-specific cortical afferents: values treated as uniform for all unit volumes.

For convenience, the results reported below were calculated with the following synaptic densities (being the values obtained from a preliminary calculation during the preparation of the companion paper [9]): \( \alpha_{ee} = 47.50, \alpha_{ei} = 0.45, \beta_{ee} = 37, \beta_{ei} = 13.50, \beta_{ee} = 13, \beta_{ei} = 0.47, \mu_{ee} = 0.4, \mu_{ei} = 0.07 \). These variations were later shown to produce small numerical changes in the results as compared with the final estimates of synaptic density, and in no way affect the conclusions drawn.
2.6. Dendritic delay and state-transition equations

Neglecting axonal delay for short connections, system delays are imposed only by the lag time of transmission from afferent pulse to efferent pulse via the cortical dendrites. In accord with the findings of Freeman for olfactory neurons [6, 7, 8] and his reasoning regarding the transferability of his estimates of olfactory measurements of synapto-dendritic delays to neocortex [7], we introduce these delays in maximally simplified form. Dendritic lags are imposed by the rise of dendritic potential in response to afferent pulses at time zero, to a maximum at 5 ms, with a symmetrical fall. Efferent action potentials are generated at a rate proportional to the current sum of dendritic potentials, $V$, and the threshold sensitivity gain, $T$.

Thus, simplified state-transition equations for each unit volume are given by

$$Q_e = gV_e \frac{dQ_e}{dV_e} \quad Q_i = gV_i \frac{dQ_i}{dV_i}$$

Here, $g$ is a constant of proportionality which is a measure of individual synaptic gain, and is the model's only free parameter. $Q_e$ and $Q_i$ are the unit-volume partial pulse densities of excitatory and inhibitory cells, respectively. $V_e$ and $V_i$ are the associated dendritic potentials, calculated as follows

$$U_e = \sum_{j=1}^{n} w_j Q_{ae}(t-j) \quad U_i = \sum_{j=1}^{n} w_j Q_{ai}(t-j)$$

The $w_j$ are discretized lag weights approximating the rise and fall of dendritic potential in response to a unit-amplitude pulse input. The $w_j$, $j = 1 \ldots n$ form an isosceles triangular function, with maximum at $j = n/2$ corresponding to the 5 ms lag weight. For $j = 1$ and $n$, $w_j = 0$, and $\sum w_j = 1$.

$U_e$ and $U_i$ are used to compute $V_e$ and $V_i$ by substituting $U$ for $Q$ in equation (2) and solving for $V$. This manipulation exploits the near linearity of the $V/Q$ relation for low values of $Q$, and avoids the difficulty of defining an absolute zero for values of $V$, arising from the asymptotic nature of equation (2). The values of $dQ_e/dV_e$ and $dQ_i/dV_i$ are then computable using equation (3). $Q_{ae}$ and $Q_{ai}$ represent the summed and synaptic-density-weighted pulse partial densities for all afferents to the cells in the unit volume, arising
intrinsic to and extrinsic to the unit volume; thus they are given by

\[ Q_{an} = \beta_{ae} Q_e - \beta_{ae} Q_i + M_{ae} Q_a + \mu_{ae} Q_{an} + \sum_{(x,y)} \alpha_{ae}(x, y) Q_e(x, y) \]  

(6)

for excitatory cells and

\[ Q_{in} = \beta_{ai} Q_e - \beta_{ai} Q_i + M_{ai} Q_i + \mu_{ai} Q_{in} + \sum_{(x,y)} \alpha_{ai}(x, y) Q_e(x, y) \]  

(7)

for inhibitory cells.

\[ M_{ae}, M_{ai}, \mu_{ae}, \mu_{ai}, \alpha_{ae}, \alpha_{ai} \]

and are the synaptic-density-weighted specific and non-specific afferent pulse densities.

At the present level of model development we include only excitatory non-specific projections and further confine specific projections to excitatory projections reaching excitatory cells.

\[ Q_{an} \]

is the system control parameter.

2.7. Coupling of unit volumes to simulate an extended cortical matrix

An extended area of cortex can now be simulated by arranging unit volumes in a matrix with each volume interconnected by efferent couplings arranged to simulate short cortico-cortical fibres. Following Braitenberg and Schuz [3] efferent-projection synaptic densities can be roughly approximated as a two-dimensional Gaussian function, with mean centred on each unit volume and range 7 mm at two standard deviations. By symmetry, the total synaptic gain of efferents from each unit volume is equal to the total synaptic gain of its external afferents. When the range of cortico-cortical fibres is corrected for the radius of the unit volume, couplings between unit volumes on the matrix conform to anatomical range and synaptic density.

2.8. Matrix size and boundary conditions

In the simulations which follow matrices of size 10 x 10, 20 x 20 and 40 x 40 were used. Boundary conditions for the matrix were 'zero', 'cylindrical', or 'toroidal' depending on whether or not cortico-cortical connections were reflected to units on the opposite matrix edge. Variations of the matrix size and the boundary conditions were used to explore the sensitivity of the simulation to these conditions, rather than in imitation of real boundary conditions. For the results which follow, these factors had no effect. The results shown are for zero bounds.

2.9. Input, output and control parameters

Inputs simulating specific cortical afferents were introduced as zero-mean white noise or sine-wave trains of input pulses introduced to the excitatory cell dendrites of the unit volumes on one edge of the matrix, thus imitating special sensory cortex. Synaptic densities for specific afferents \( M_{ae} \) were therefore not needed.

Output, simulating the EEG, was the vector time series \( V_e(t) \). A recording array was imitated by confining consideration to an 8 x 8 square subset of the complete matrix. Two positions of this 'recording' matrix were used: a position one row in from the row simulating specific sensory cortex, and a position as far away from specific sensory cortex as possible.

The single variable control parameter, the pulse density \( Q_{an} \), was applied uniformly to the non-specific synaptic couplings of excitatory and inhibitory cell components in the entire matrix, including specific sensory cortex.
2.10. Initial conditions, updating, and data acquisition

All $Q_s$, $Q_t$, were initialized at zero, and the selected value of $Q_{ns}$ was applied from time zero. Synchronous updating at 0.1 ms intervals was then performed for 0.2 s, until values of $Q_s$ reached a steady state. Specific inputs were then introduced and equilibration was again allowed for 0.2 s. The amplitude of specific inputs was always less than the mean value of $Q_s$ to avoid clipping of the input. Output data then entered temporary file storage in 1 s epochs.

2.11. Effect of the free parameter, $g$

In equation (4), it can be seen that $g$ and $dQ/dV$ both act as simple multipliers of system gain. With small perturbations of input, it is the input mean which regulates $dQ/dV$, and it is by this operation upon $dQ/dV$ that $Q_{ns}$ acts as the control parameter. Thus, by a choice of $g$ such that $Q_{ns}$ regulates gain in the physiological range, the free parameter $g$ can be optimized.

We set $g = 2.75$ since at this value the mean pulse rate in the simulation most closely approximated to low physiological rate, and also remained in keeping with the simplified state-transition equations applicable to low base-rate firing, as mentioned in section 2.4 (see also section 4.1). At higher values of $g$ the simulation was unstable, and the small perturbation approximations were clearly not applicable.

3. Matched animal experiments

Unless otherwise stated, the results obtained in simulation were compared with recordings obtained from 64-channel cat ECOG in experiments reported elsewhere [22, 23]. The animals had undergone chronic implantation of extradural 0.5 mm diameter platinum-ball recording electrodes arranged in an $8 \times 8$ array at approximately 0.86 mm separation. These electrode dimensions eliminate risk of spatial aliasing. The arrays were situated over posterior association cortex, with variation among individuals. Recordings were made during comfortable but alert rest for the animals, under two visual-stimulus conditions. These were (a) normal lighting within a laboratory, and (b) with imposed visual sinusoidal flicker near the mid-field of the animal's gaze, in surrounding partial darkness. The recorded signals were obtained in 30 s continuous runs, digitized at 8 ms intervals, and low-pass filtered at 30 Hz. From these recording epochs, 1 s segments were used to calculate the power spectrum, frequency/wavenumber spectrum, and coherence as a function of distance and frequency. The 1 s results were then averaged. Output from the simulation was similarly treated.

4. Results

4.1. Base pulse density

With $Q_{mw}$ at its maximum of 1 and $Q_s = 0$, the values of $Q_s$ reached equilibrium at 0.007. The low value of mean $Q_s$ corresponds to the low firing rate during alert conditions of real cortical cells [1, 2]. With $Q_s \neq 0$, $Q_s$ was perturbed about the 0.007 mean.
4.2. Power spectra

4.2.1. Noise input. When the simulation was driven by making $Q_n$ a synchronous white noise, the average power spectrum of the output signals approached that of $1/f$ noise, with peaks of power at approximately 5, 14, 25, and 40 Hz. Thus, the spectrum is similar to that seen in the desynchronized (alert) state, with realistic components in the theta, alpha, beta, and gamma ranges. When the simulated recording array was located further from the driven matrix column, the lower frequencies were relatively pronounced in spectral density. Variation of $Q_m$ over the range 0.1 to 1 produced an increase in total output power, and an increase in relative spectral density at the lowest frequencies (see figure 2).

![Figure 2. Spectrum of simulated eCoG with variation of control parameter $Q_m$.](image1)

![Figure 3. Logarithmic power spectra of real cat eCoG and simulated eCoG during driving with a sinusoidal input to specific (visual) pathways, showing typical occurrence of harmonics.](image2)

4.2.2. Autoregression characteristics. Maximum-entropy autoregression analysis was performed by analogy with the application of this technique to real eCoG [21]. Optimum model order and distribution of autoregression coefficients was comparable with that of real eCoG, when allowance was made for the effect of low-pass analog filtering of the real eCoG. Variation of $Q_m$ did not affect optimum model order, the range of autoregression coefficients, or the system eigenvalues which can be derived from these. The principal effect was upon the amplitude of the zero-mean uncorrelated component of the signal, thus...
accounting for the sensitivity of total spectral power to $Q_m$.

4.2.3. Sine-wave inputs. With the addition of sine-wave components to the white noise driving to the 'specific cortical' matrix column, the output showed the appearance of power at the fundamental and integer harmonics of the driving signal (see figure 3). The example shown is for 10 Hz input, but equivalent harmonic responses were produced for all sinusoidal frequencies. This reproduces the harmonic response of cortex during visual driving by sinusoidally flickering light [15, 23]. The harmonic frequencies were independent of the amplitude of input signals, but the relative distribution of power to higher harmonics increased with higher input amplitude and with increase in $Q_m$.

4.3. Frequency/wavenumber spectra

These were computed for 5, 10, 20, and 40 Hz, and for conditions of white noise driving with and without added sine-wave inputs. The wavenumber plots showed an initially unexpected result, namely an orthogonal pattern of power at wavenumbers which were invariant with the temporal frequency analysed. The same pattern is seen in real ECoG [23] and we have previously shown that this pattern of power distribution is attributable to a small-scale resonant interaction of each unit volume with surrounding volumes. The wavenumbers at which power is characterized by the frequency/wavenumber method are not influenced by the control parameter, and activity at these wavenumbers contains only a small fraction of the total ECoG power.

Representative wavenumber plots are shown in figure 4.

4.4. Travelling-wave velocities

The majority of spectral power is associated with travelling waves of wavelength too great to be estimated by the frequency/wavenumber method, at the simulated or comparable real ECoG electrode separations. Phase velocity and attenuation of amplitude for the travelling waves was estimated from the Fourier components of individual output channels in the simulation. It was found that both phase velocity and spatial damping of the travelling waves varied with $Q_m$. Examples of this variation are shown in figure 5.

The range of velocities determined in the simulation accord approximately with our earlier estimations of wave velocity at millimetric scale and in the frequency band 1–30 Hz, in real ECoG [22]. We inferred from estimates of coherence that real millimetric ECoG waves are essentially non-dispersive and have a velocity between 0.1 and 0.3 m s$^{-1}$. Lopes da Silva and Storm van Leeuwen [10] have measured the velocity in millimetric alpha activity at 0.3 m s$^{-1}$, with some estimates as high as 1.2 m s$^{-1}$.

4.5. Sensitivity considerations

A detailed consideration of the simulation’s parameter sensitivities is beyond the scope of the present paper, but the following general features were demonstrated:

- The distribution of power at the wavenumbers identified by frequency/wavenumber analysis is sensitive to variation of both the range and centrality of the distribution of cortico-cortical fibres, and matches the real ECoG in near optimal fashion at the distributions we have imposed from anatomical considerations.
Electrocortical wave dynamics at millimetric scale

Figure 4. Representative frequency/wavenumber results. Upper figures: power at wavenumber components \( K_y, K_x \) for real and simulated ECoG at 10 Hz (wavenumbers associated with wavelengths greater than grid dimensions excluded). The wavenumbers estimated do not vary with the frequency analysed. Lower figures: cross-sections through the \( K_y \) axes of the full wavenumber spectrum at 10 Hz. The marked difference in the range of power in real and simulated ECoG reflects the use of normalized excitations in the simulation.

- Variation of the time of peak lag weight simulating the rise and fall of dendritic postsynaptic potentials shifts the peaks of the average power spectrum away from their physiological values.
- Decrease in the total cortico-cortical synaptic densities, or increase in the synaptic density of intracortical inhibitory-to-excitatory synaptic densities produces progressive flattening of the power spectrum associated with increased spectral density at 40 Hz.

5. Conclusion

The properties of our simulation appear to correspond to those of real ECoG. We believe this to be the first simulation of neocortex at millimetric scale reproducing this range of features, and the only one based upon quantitative estimates of synaptic density. The present model is, in effect, without free parameters and can be readily generalized beyond the small perturbation case.

The results obtained conform to much preexisting theory. Most prominently, they approach expectations of Freeman [7] regarding neocortex. Freeman's asymmetric sigmoid relation between LFP and action potential probability [5] is implicit in this simulation. With
the introduction of strong excitatory–inhibitory feedback, spectral properties for olfactory bulb during inhalation [6] are approximated.

A further conceptual similarity to the global electrocortical model of Nunez [12, 14] is apparent. State equations for the simulation resemble Nunez’s basic equations [13]. The scale of electrocortical events is the principal difference, since Nunez’s work is intended to apply to wave action at centimetric level.

The simulation is also in conceptual keeping with our own work [19] which approaches electrocortical dynamics by the use of second-order stochastic differential equations, applied to local cortical circuitry. The near-equilibrium (1/f spectral distribution) and near-linear-superposition properties exhibited by the present simulation accord with our earlier predictions.

It is not yet clear whether the present simulation is compatible with or in contradiction to the model of Rotterdam et al [17] for cortical alpha, to which it is indebted in terms of technical style.

While it is apparent from the match with experimental results that the synaptic connectivities and densities taken from the preceding paper are plausible extensions of empirical anatomical data, it does not follow that our findings exclude non-random short-range connection selectivities. However, it does appear that such selectivities are not required to model dynamics at millimetric scale. We believe it would be of interest to consider the modelling of memory mechanisms in neural networks with just this degree of stochasticity and dilution, embedded in input signals which have the spectral composition of EOG waves. This might be achieved by suitable modification of networks resembling those described by Amit and Tsodyks [1, 2]. By this suggestion, we wish to imply that organized
signal interchange among cortical networks of the scale of our 'unit volume' may proceed under laws identical to those that govern the generation of the EEG. Notably, for critics who have considered the EEG an epiphenomenon, it is apparent that EEG waves have high spatial bandwidth, as shown in the above results, and therefore do not offer an implausibly narrow information channel for the exchange of signals.

The present simulation is suitable for development in a number of different ways, at both cellular and microscopic scales. These include increased resolution of local anatomical detail, to incorporate a wider range of cellular types, details of cortical layering, detailed modelling of the specific and non-specific inputs, and consideration of anisotropy of longer-range cortico-cortical connections. It remains to be seen whether with progressive improvement in anatomical veracity, the simulation proves to mimic real EEG dynamics ever more closely.

Acknowledgments

The authors would like to thank Mr Nicholas Hawthorn for technical assistance. This work was supported by grants from the Health Research Council of New Zealand and the Auckland Medical Research Foundation.

References

Dynamics of the brain at global and microscopic scales: Neural networks and the EEG

J. J. Wright and D. T. J. Liley
Mental Health Research Institute, Parkville, Victoria 3052, and Swinburne Center for Applied Neuroscience, Hawthorne, Victoria 3122, Melbourne, Australia
Electronic mail: jlw@cortex.mhri.edu.au

Abstract: There is some complementarity of models for the origin of the electroencephalogram (EEG) and neural network models for information storage in brainlike systems. From the EEG models of Freeman, of Nunez, and of the authors' group we argue that the wavelike processes revealed in the EEG exhibit linear and near-equilibrium dynamics at macroscopic scale, despite extremely nonlinear—probably chaotic—dynamics at microscopic scale. Simulations of cortical neuronal interactions at global and microscopic scales are then presented. The simulations depend on anatomical and physiological estimates of synaptic densities, coupling symmetries, synaptic gain, dendritic time constants, and axonal delays. It is shown that the frequency content, wave velocities, frequency/wavenumber spectra and response to cortical activation of the electrocorticogram (EOcG) can be reproduced by a "lumped" simulation treating small cortical areas as single-function units. The corresponding cellular neural network simulation has properties that include those of attractor neural networks proposed by Amit and by Parisi. Within the simulations at both scales, sharp transitions occur between low and high cell firing rates. These transitions may form a basis for neural interactions across scale. To maintain overall cortical dynamics in the normal low firing-rate range, interactions between the cortex and the subcortical systems are required to prevent runaway global excitation. Thus, the interaction of cortex and subcortex via corticostriatal and related pathways may partly regulate global dynamics by a principle analogous to adiabatic control of artificial neural networks.

Keywords: chaos; EEG simulation; electrocorticogram; neocortex; network symmetry; neurodynamics

1. Introduction

Although we now know much of the structure and functional attributes of the brain and can understand some of its mechanisms of information processing, we lack certainty regarding the overall dynamical properties of the brain. That is, we have little idea how the subcomponents of the brain are integrated into a functional whole. Recognizing that the problem of the brain's overall integration raises poorly posed but important questions, Churchland (1986) discussed the need for the systematic arrangement of neuroscientific theories within hierarchies, so that reduction of some theories to the level of others can take place.

Two areas requiring such comparative reduction are those of attractor neural networks (ANN) and of mechanisms giving rise to the electroencephalogram (EEG). The ANN field has been largely motivated by the need to reproduce brainlike logical operations. In those studies in which ANN properties are compared to activity patterns of real neurons, the emphasis is generally on comparison to single- or multiple-unit recordings (Amit 1990). [See also Amit: "The Hebbian Paradigm Reintegrated" BBS 18(4) 1995.] Yet since the EEG is an observable property of large fields of real neurons that are engaged in cortical information processing, the ANN field might benefit from considering the dynamics of the EEG as a means to cast light on the mechanisms of cognitive processes at larger scales, up to that of the entire brain.

We shall argue that a preliminary integration of neocortical EEG theory with ANN principles is now practicable, and that this unification has consequences for both fields as well as for studies in single-unit neurophysiology.

To demonstrate the utility of our arguments we present accounts of the simulation of both macroscopic EEG and microscopic cellular interactions, basing both types of simulation upon calculations and measurements of relevant anatomical and physiological parameters. Thus we seek to avoid "explanations" depending on simulations that utilise "free" (arbitrary) parameters. The results of these simulations raise further questions about global control processes in the brain. In our attempts to address these additional questions we found we needed to consider not only the dynamics of neuronal interaction at intercellular and global levels but also the interactions of the cortex with the brain stem.

We have not attempted an exhaustive review of findings and theory in either the EEG or the ANN fields. The EEG models upon which our work is drawn most strongly are those proposed by Freeman (1991) for neocortex, and the global model for EEG developed by Nunez (1989) — each a remarkable work for experimental or theoretical scope, but
at first examination seemingly in mutual contradiction, as we shall describe later. For ANN work that exhibits properties of relevance to EEG modeling we have drawn principally on the work of Amit and Tsodyks (1990; 1991) and Parisi (1986a; 1986b) — ANN authorities who have aimed their work as much as possible toward meeting the requirements of neurobiology. To link these disparate accounts together we have added some description of the earlier work of our own research group, and the concurrent work of others with related concerns.

2. Freeman's nonlinear model for neocortex

In the past 25 years Freeman and his coworkers have systematically developed a model of perceptual processing in the olfactory bulb (Freeman 1964; 1972; 1975; 1979; 1987a; 1987b; 1988; Freeman & Skarda 1985). [See also Skarda & Freeman: “How Brains Make Chaos in Order to Make Sense of the World.” BBS 10(2) 1987.] In a recent paper (Freeman 1991) this body of information was used to make predictions concerning neocortical dynamics. Similar considerations appear to motivate the ongoing work of his group.

2.1. The basic unit of oscillation. Freeman notes that pyramidal cells occur in multilayered, loosely columnar structures with recurrent inhibition provided by inhibitory surrounds. Within this extended structure Freeman finds it convenient to recognize a subunit configuration called a KII set. KII sets are “lumped” approximations of interacting aggregates of cells. The most common KII set is one of interacting aggregates of excitatory and inhibitory cells, with interactions including all combinations of feedforward and feedback relationships. Interactions between KII sets in neocortex are provided by longer range transverse couplings — the excitatory axonal fields of the pyramidal cells. He predicts that properties similar to those his group has found for neuronal aggregates in the olfactory system will be found in neocortex.

The first key property on which Freeman's model depends is the lagged time response of dendrites to synaptic input. His estimates indicate that the maximum dendritic depolarization after input occurs 5.8 msec after action potentials reach the afferent synapses — a delay composed of 1.3 msec of synaptic delay and 4.5 msec due to passive electrical properties of the dendritic membrane. In conditions of sufficient excitatory drive to all the neurons in the local cortical network, these delays impose a tendency to oscillation, since excitatory cells interact with surrounding inhibitory cells, which in turn provide recurrent connections to the excitatory cells. The oscillation, Freeman predicts, should have a fundamental frequency of 40+ Hz, imposed by the times required for surges of excitation, recurrent inhibition, disexcitation, and disinhibition, to complete a cycle. These surges of firing will impose a similar cyclicity on the local field potentials (LFP). Coherence between neuronal firing and the EEG near this frequency has been experimentally demonstrated (Eckhorn et al. 1988; Gray et al. 1989). In contrast, EEG at other frequencies may reflect more complicated interactions and need not be simply correlated with local cell firing.

The absence of broad-band correlation of EEG with action potentials has long been the grounds upon which the EEG has been trivialized as an epiphenomenon by many physiologists (Streyer 1989). The laying to rest of this trivialization is a small part of Freeman's concern.

In his 1991 account Freeman does not deal with excitatory-excitatory and inhibitory-inhibitory interactions, although he recognizes this deficit and expects these interactions to complicate the picture considerably. Such interactions would make correlation of 40 Hz ECoG and action potentials rather improbable, unless observations were made in states (such as anaesthesia) in which local activity may be relatively decoupled from its surrounds. This may account in part for the controversy that has surrounded the original reports (e.g., Tovee & Rolls 1992).

2.2. Nonlinear wave-pulse relations. The second key property Freeman proposes will be found in neocortex, analogous to that measured in the olfactory cortex (Eeckman & Freeman 1991), is the one described by the asymmetric sigmoid curve. This function relates the normalized pulse probability for an individual neuron to the concurrent EEG wave amplitude. Freeman takes its form to describe indirectly the way the pulse density rises with progressive depolarization of a small mass of cells in response to afferent bombardment, and he predicts that the form of the curve is dependent on the state of nonspecific activation.

An important part of Freeman's (1979) theoretical derivation of the exact form of the asymmetric sigmoid curve is embodied in the asymptotic limit of the curve. The limit, designated in Freeman's terminology as Qm, is the maximum pulse density that can be sustained by a group of neurons. Qm itself rises with increasing nonspecific cortical activation and is considered by Freeman to impose a stabilizing upper bound on the perturbations induced by other inputs to the KII set. It is notable that the asymmetric sigmoid curve is as of yet a theoretical construct. Freeman's experimental results do not demonstrate convincingly the presence of the state-dependent asymptote since, at least as commonly as an asymptotic limit is apparent, pulse density seems to rise exponentially with rising wave amplitude.

2.3 Locally chaotic dynamics. Freeman remarks that although both the asymmetric sigmoid curve and the absolute refractory time of neurons guarantee overall stability, the sharp nonlinearity of neuronal threshold confers a sensitive dependence on both initial conditions and ongoing perturbation. Consequently, both the firing rate and the field potential of any pyramidal cell raised above a critical level of excitation would, according to Freeman, be chaotic. Freeman and Jakubith (1993) have embodied the principles of Freeman (1975; 1979; 1991) in two simulations that reproduce the predicted chaotic dynamic properties.

Incontrovertible proof that EEG reflects any simple chaotic process is generally lacking. There are grounds for reservation concerning reports of the dimensionality of EEG from direct measurement. Fundamental difficulties lie in the applicability of estimation algorithms to EEG data because of limitation in the size of data sets, noise contamination, and lack of signal stationarity (Ruelle 1994). These difficulties do appear to be overcome in certain situations in which return-map techniques can be applied to controlled cellular interactions (Schiff et al. 1994). Without depending on problematic direct demonstration in all circumstances, the likelihood that cellular interactions are commonly chaotic is made further plausible by Freeman's argument that
the Poisson distribution of cell-firing rates and the "1/f" nature of the EEG spectrum associated with this, are incompatible with coupled nonlinear limit-cycle oscillations – the only possible alternative. However, this argument depends on an imprecise use of the term 'chaos'. Although the absence of periodic determinate behaviors suggest that local neuronal interactions may rapidly lose information about their initial conditions, this leaves undecided the autonomous dynamic characteristics of local neuronal interactions in the absence of ongoing perturbation. Because any small pool of neurons is under heavy perturbation by afferents, rapid loss of information on initial conditions is to be expected.

2.4. Globally chaotic dynamics. If local neuronal interactions are accepted as chaotic, is the global system then necessarily chaotic? An important class of simulation studies suggests that this must be the case ( Kaneko 1990; 1992). These studies concern one-dimensional chaotic maps (that is, chaotic numerical subprocesses of considerable generality) that are globally coupled, each to all others. Such coupled maps exhibit global chaos and appear to escape from the law of large numbers and the central limit theorem. The analogy to masses of coupled neurons is evident ( Tsuda 1992; 1994). However, the escape from the law of large numbers does not occur in the presence of noise ( Kaneko 1990). We will argue in section 4 that this effect of noise is crucial to the type of dynamics observable at macroscopic scale in the ECoG.

3. Nunez's linear model for neocortical EEG

The concept that the EEG is an essentially chaotic process at global scale strikes a strong contrast with a well-developed as well as very general model advanced by Nunez (1981; 1995). In this model, EEG is treated as a linear wave process, and the dynamics of the entire telencephalon is treated as a problem of mass action of coupled neuronlike elements. In application of mass action laws, the applicability of the theorems of central tendency (central limit theorem and law of large numbers) is implicit, but inescapable.

Nunez places emphasis on a number of factors commonly ignored in models of neuronal interaction. These are:

2. The quantitative predominance of long-range interactions mediated by the cortico-cortical fibres.
3. Boundary conditions – the global cortical system is a closed medium for the propagation of traveling waves, implying the possibility of global resonance and standing waves. The alpha rhythm is considered by Nunez to arise as a global standing wave.
4. The likelihood that EEG waves will obey a law of linear superposition, because they are based on a measure (average depolarization of the cortical cells consequent to afferent bombardment) that Nunez expects to be nearly linear over the relevant operating range, namely, directly proportional to synaptic action density.

Nunez's basic dynamic equations have been solved for spherical, spheroidal, and other cases approximating the brain. Principal results include a confirmed prediction (Burkitt 1994; Thatcher et al. 1986) of a wave propagation velocity of 7–11 m/sec for human alpha waves, and a number of other findings in qualitative agreement with the theory (Nunez 1981; 1995).

A strength of Nunez's work has been the consideration of widely differing classes of solution to his fundamental equations, with variation in the key structural parameters. A weakness has been uncertainty as to exactly what ranges of these parameters actually apply in mammalian brains. Nunez has recognized explicitly that both local and global interactions may be relevant to the origin of EEG (Nunez 1989; Nunez & Srinivasan 1993) and together with Inger (Inger & Nunez 1990) has explored methods for the application of group renormalization techniques to neuronal dynamics.

4. A stochastic model of electrocortical activity

Our group has attempted to bridge the microscopic and global scales by providing a concept that recognizes nonlinear dynamics at microscopic scale and emergent linearity of macroscopic electrocortical waves (Wright 1990; Wright et al. 1990a; 1990b; Wright & Sergejew 1991).

4.1. Basic assumptions. A unit structure in the cortex capable of oscillation was postulated – one that was loosely defined but considered to be a mass of excitatory and inhibitory cells roughly equivalent to a minicolumn. Each oscillator was considered symmetrically and reciprocally coupled to many neighbors by essentially excitatory couplings. The system variables were the local field potentials associated with each unit oscillator.

For conditions pertaining to the ECoG, such a system can be described in a highly general way by stochastic second-order differential equations, in which the parameters of the state equations (representing the analogs of natural frequencies, damping coefficients, additive and multiplicative couplings of simple harmonic oscillators) are all considered to vary freely to whatever values are needed to fit the behavior of the system variables. The time variation of these free parameters thus describes deviations from the simple harmonic case arising from the extreme nonlinearity of the underlying neuronal properties and the ongoing influence of noisy input from the reticular activating system. We then assumed that because of the idiosyncrasy, complexity, and noisiness of local cellular interactions, the time-varying parameters were stochastically independent in the large.

4.2. Consequences. It follows (Wright 1990) that despite the extreme nonlinearity of the elements, the macroscopic wave motions at low frequencies in such a system tend to approximate the linear dynamics of coupled simple harmonic oscillators, for which the free parameters are constants. That is, waves obey a superposition principle. The lowest resonant mode frequencies tend to constant values. Energy is equipartitioned among the resonant modes, yielding a "1/f" type of average spectrum. Traveling wave velocities at different frequencies tend to constancy and are governed by axonal and dendritic delays.

These statistical tendencies hold more closely the lower the frequency is and thus the longer the wavelength considered. Some of these tendencies reverse at high frequencies, but high-frequency activity is greatly attenuated in its contribution to total amplitude of the macroscopic ECoG.
4.3. Tests. Three indirect tests for the presence of wave motion of the predicted type in the ECoG have yielded results in favor of hypothesis. Each of these tests applied linear analysis methods to the ECoG and then critically considered the physical adequacy (rather than the numerical efficiency) of the linear descriptors. Specifically, these descriptors show that:

1. The distribution of damping factors in autoregression (AR) analysis of the ECoG is that expected in a near-equilibrium process, with equipartition of energy among resonant modes (Wright et al. 1990a).

2. Linear inverse filter methods can be applied to the averaged evoked response to determine the impulse giving rise to the evoked response. The forward filter can then be applied in such a way that the likely linearity of the wave process can be demonstrated. (Wright et al. 1990b).

3. An equation describing coherence at separated points on a medium traversed by multidirectional linear non-dispersive waves can be fitted to coherence measures from multichannel ECoG. The distribution of the residuals of fit supports the physical validity of the equation (Wright & Sergejew 1991).

5. Complementary aspects of the theoretical models

From the above, it seems reasonable to view interactions among cortical cells as a two-scale phenomenon, with interaction across scales. If we consider the interactions of single neurons in a small locale (ignoring for the moment perturbations from surrounding cells) then the free parameters utilized in a stochastic description cannot be expected to be independent. Thus laws of central tendency may be violated and chaotic interactions are to be expected. But consider the synaptic activity impinging upon this small locale, arising from the sea of surrounding cells. Linear wave attributes may apply to the macroscopic ECoG fields correlated with the sources of this synaptic bombardment, and the action potential flux in the afferent volleys will reflect the space average of LFPs in all the efferent fields even though the afferent bombardments of individual cells may be poorly time correlated.

The abstract second-order oscillator of the stochastic model may be taken to correspond to the KII set of Freeman. The stochastic model can be consistent with Freeman’s formulation, if interaction among individual cells involves noisily perturbed chaotic activity, particularly if this chaotic activity is of high Lyapunov exponent. The presence of global chaos in the coupled chaotic maps studied by Kaneko (1990) may not be relevant in the cortex, when ongoing noisy perturbation is provided by the reticular formation, in conditions of cortical activation. By extension, the linearized, mass action treatment applied by Nunez within his general coupling equations would also be justified, at least as first approximation. But we are left at this point with little idea as to the ways in which local cellular dynamics might override their merely stochastic participation in the global dynamics – a requirement, surely, since it is a neurophysiological commonplace that individual cells reliably exhibit transient bursts of activity in response to selective and remote stimuli. Indeed, without such a capacity for relatively autonomous activity, how are neurons to exhibit any form of local ANN behaviors?

We therefore turned to simulation studies at both macroscopic and microscopic levels, to see to what extent consistency in such simulations could be maintained while describing essential experimental data at both levels.

6. Computer simulations

Before simulation could be attempted, an adequate description of connectivity in the cortex had to be provided and simulations devoid of arbitrary parameters had to be devised. Further, the simulations at macroscopic and microscopic levels had to share a consistent parameterization as fully as possible.

6.1. Cortical connectivity. To simulate realistic cortical networks, general rules in the form of equations describing the density of interaction of neurocellular components were needed. We have applied a modification of a method initiated by Uttley (1956) to compute the needed connectivities. A detailed treatment is reported in Liley and Wright (1994). The assumptions made in our calculations are:

1. The basal dendritic system and intracortical axonal trees of a neuron are distributed with average fibre densities that decline exponentially with distance from the cell body and with spherical symmetry.

2. The apical dendritic tree can be ignored for calculation of intracortical connectivity, since it is involved mainly in cortico-cortical synapses.

3. Only axo-dendritic connections need be accounted for.

4. An axon and dendrite meeting at random in a "connection space" (defined by the radii of intracortical axons and dendrites) form one synapse (equivalently, when a dendrite and an axon touch, a synapse forms).

On the basis of these assumptions, and using the anatomical data and initial calculations of Sholl (1953) and of Brantenberg and Schuz (1991), we were able to calculate:

1. The percent density of synapses in unit cortical volume attributable to cortico-cortical excitatory connections, intracortical excitatory-inhibitory, inhibitory-excitatory and inhibitory-inhibitory connections, and nonspecific afferents; and
2. The probability that n synapses are formed between any two cells in the cortex as a function of their separation by a small distance, r.

We have subsequently generalized these calculations so that they can be applied to mammals of variable size, by corrections based on relative cortical cell densities (Wright & Liley 1995). The most recent resultant estimates of synaptic densities for mouse, cat, and human are given in Table 1. Figure 1 shows the probability of synaptic connections between individual cells.

If, as we have assumed, connectivity between individual cells can be appropriately considered stochastic at a local intracortical level, then an important conclusion can be drawn concerning these connections: the connectivity between individual cells is highly asymmetric. Conversely, despite the high asymmetry of individual cell connections, absolute synaptic densities of coupling between pools of cells can be considered roughly symmetric. These estimates form the basis for the simulations that follow, which thus share a common parameterization of connectivity.

6.2. Macroscopic simulation of the ECoG. A full account of the macroscopic simulation of human and feline ECoG is
Table 1. Approximate synaptic densities for mouse, cat, and human (percentage of synapses in average cortical unit volume)

<table>
<thead>
<tr>
<th>Synaptic type</th>
<th>Mouse</th>
<th>Cat</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortico-cortical (ex-ex)</td>
<td>45.1</td>
<td>70.0</td>
<td>78.0</td>
</tr>
<tr>
<td>Cortico-cortical (ex-inh)</td>
<td>5.6</td>
<td>10.0</td>
<td>11.5</td>
</tr>
<tr>
<td>Intracortical (ex-ex)</td>
<td>35.13</td>
<td>14.4</td>
<td>7.3</td>
</tr>
<tr>
<td>Intracortical (ex-inh)</td>
<td>6.26</td>
<td>2.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Intracortical (inh-ex)</td>
<td>6.26</td>
<td>2.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Intracortical (inh-inh)</td>
<td>1.12</td>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>Nonspecific afferents</td>
<td>0.53</td>
<td>0.8</td>
<td>0.8</td>
</tr>
</tbody>
</table>

*Type of cells connected: ex = excitatory (pyramidal); inh = inhibitory (stellate).

given in Wright and Liley (1995), with related aspects elsewhere (Wright et al. 1994; Wright & Liley 1994). Each functional unit in the simulation was a “lumped” subsystem, equivalent to a KII set, or the “unit oscillator” of the stochastic model.

The state equations then applied were a specific instantiation of the general equations of Nunez. As such, there was implicit application of the principles of mass action, the law of large numbers, and the central limit theorem – as justified by the arguments of section 4. The state equations described the local field potentials and action potential densities as vector fields over the cortex, with structural parameters introduced to take account of:

1. Dendritic delays. These were applied to each set of excitatory cells and inhibitory cells within a lumped unit, using a simplified approximation of Freeman's estimates for dendritic transmission delays, namely, a lag of about 5 msec to peak dendritic response after afferent bombardment.

2. Axonal delays, applied to the long-range connections. These were equivalent to an action potential propagation velocity of 9 m/sec (again in accord with the figures used by Nunez) and a distribution and maximum range of cortico-cortical connections in accord with anatomical estimates.

3. Coupling strengths within and between all lumped component groups of inhibitory and excitatory cells that were proportional to the synaptic densities of their connections.

4. A Gaussian distribution of cell action potential threshold, with respect to somatic membrane potential or local field potential, giving rise to a sigmoidal pulse-density/LFP relationship for each of the excitatory and inhibitory subcomponents of the lumped units.

5. Synaptic gain (see below).

Boundary conditions considered for the long-range couplings were the toroidal (closed) and zero-bound (open) conditions – the former providing a closed geometry sufficient for present purposes, and the zero-bound conditions being used only for comparison.

Specific sensory input was imitated by input of white noise to a restricted portion of the simulated matrix of lumped cortical elements, and nonspecific activation by diffuse input throughout the matrix, weighted by the relative density of nonspecific synapses. The simulation successfully reproduced the following EEG characteristics, some of which are displayed in Figures 2, 3, and 4.

1. The spectral properties of the output included peak densities at the frequencies of the major cerebral rhythms, with a “1/f” spectral envelope (see Fig. 2).

Figure 1. A diagrammatic representation of the variation of intracortical synaptic coupling symmetry between two “typical” neocortical pyramidal cells as a function of their intercellular separation. The measure of coupling symmetry (or asymmetry) is the probability that two cells share n presynaptic contacts. Inter- cellular distances are measured in units of $r_p$, where $r_p$ is 31.25 μm (the space constant for the basal dendritic tree; Liley & Wright 1994).

Figure 2. Normalized power spectra (1.22 Hz bins) obtained at steady state from the simulation of human and cat electrocorticogram, for a variety of values of nonspecific activation. “$Q_{0.5}$“ is a measure increasing with nonspecific excitatory tone (Wright & Liley 1995).
2. A spectral “shift to the right” with increasing total power occurred as nonspecific cortical activation was increased (also Fig. 2).

3. Steady-state traveling waves with a phase velocity of 5–7 m/sec in the human simulation, and less than 1 m/sec in that of the cat (Fig. 3) were demonstrated. Both velocities are in accord with experimental estimates from the appropriate species (Thatcher et al. 1986; Burkitt 1994; Lopes da Silva & Storm van Leeuwen 1978; Wright & Sergejew 1991).

4. Frequency/wavenumber results (at millimetric scale) revealed an unusual pattern — apparently that of bidirectional and orthogonal standing waves, with a periodic wavenumber spectrum, and wavenumbers independent of the temporal frequency of analysis. This unusual pattern closely matched results from real cat ECoG (Wright et al. 1994). It may be in part artifactual, arising because of intrinsic limitations of the frequency/wavenumber technique. The frequency/wavenumber technique revealed little of the traveling wave activity, most of which lay at wavelengths outside the range estimable by this method (see Fig. 4).

5. Harmonic responses to strong sinusoidal driving analogous to those generated in cortex in response to sinusoidal flicker (Wright et al. 1995).

In these simulations only one parameter — the synaptic gain — was not established a priori from anatomical and physiological data. The value of synaptic gain (g) applied was not arbitrary, although the available means to estimate g was qualitative only. We chose a value of g calculated from the low-mean firing rate of cortical cells — approximately 10 spikes/sec (Amit & Tsodyks 1991) — and an estimate of the maximum sustainable firing rate of cortex which does not lead to runaway global excitation. The latter was estimated at only 20 spikes/sec, averaged over all cells in the cortex (Wright & Liley 1995).

There are important implications in this choice of g, which will be further considered later. But it can be immediately noted that this choice implies that high firing rates commonly seen for brief periods in single-unit recordings, would, if freely propagated in cortex, lead to runaway excitation, analogous to epileptic discharge. It is also argued in our 1995 paper that the value of g we applied is such that saturation afferent synaptic input to dendritic membranes is just capable of completely depolarizing the membrane. Additional support for the approximate accuracy of this choice of the value of g arose in our microscopic simulation.

6.3. Microscopic simulation. We are currently engaged in the study of a cell-by-cell simulation of local cortical dynamics, making use of the “Genesis” simulation package (Bower & Beeman 1995). This work will be reported in detail elsewhere. Its intention is to study dynamic interactions within locales of individual cells, which were “lumped” under mass action rules, as a component group in the macroscopic simulation.

Up to a thousand or more cortical cells with closely packed cell bodies are represented, with proportions of inhibitory and excitatory cells equivalent to that of real
cortex. Intracortical connections are stochastic and highly asymmetrical, in accord with the connectivity rule shown in Figure 1, appropriately modified for the pyramidal and stellate cell types.

The parameters of this simulation (connectivities, threshold distributions, etc.) were chosen so as to correspond to those of the macroscopic simulation as closely as possible. Some features not explicit in the macroscopic simulation are inevitably present, namely:

1. The location of synaptic inputs on the dendritic tree are modelled, and account taken of the relative distribution of inhibitory and excitatory synapses on the trees.

2. Each neuron is in effect a leaky integrate-and-fire process, with 4–8 dendritic compartments, spike threshold discrimination at the soma, and absolute and relative refractory periods.

3. Transmitter kinetics, passive dendritic properties, and synaptic delay are separately modeled.

4. There is no explicit dependence of the model on a synaptic gain term. Instead, this is implicit in the postsynaptic response to each spike, the parameters of which are based on single-cell recordings and patch-clamp experiments. Thus, all free parameters are involved.

We have so far studied the dynamics of this network under a time-invariant input condition and differing initial conditions. That is, a uniform and standing direct current stimulus is injected to the dendritic membranes from time zero, so that cellular interactions are enabled to evolve in the net without any time-varying external perturbation. We chose this condition because it is the simplest one in which to study intrinsic local network properties. The direct current injection can be related to an equivalent average afferent pulse density, as though the network were being bombarded by a steady pulse stream from all cells in the surrounding cortex. The initial conditions of the individual cells can be varied to obtain some insight into the trajectories followed.

The following properties are apparent:

1. Two distinct dynamic states emerge, depending on the level of injected current (or external synaptic bombardment) applied during the simulation run. A sharply defined transition between these types of network response occurs at a level of injected current equivalent to a low rate of input of external action potentials. The transition depends upon introducing a level of excitation of the network's component cells that is sufficient to raise the aggregate gain in the intercellular interactions to unity or above.

2. Below this transition, animations reveal a turbulent pattern of cellular interaction, manifested in the membrane potentials. The pattern is sensitive to small changes in the depolarizing bias applied to each cell and is associated with low intrinsic firing rates in the network. Time variation of the membrane potentials dies away after initiation of the input bias.

3. Above the transition level, runaway excitation in the network supervenes. This terminates in the entry of the network to cyclic activity, in which mean firing rates for the cells are high and include bursting behaviors. Preliminary results indicate that the network has many different limit-cycle or nearly limit-cycle attractors, and the trajectories followed from different initial conditions can diverge or converge from each other in complicated ways, prior to their ultimate termination in different or similar attractors.

4. The membrane potentials, or local field potential, associated with each of these classes of dynamic behavior exhibit spectral properties similar to EEG.

5. The high firing rate states can only be terminated by the application of a strong inhibitory tone to the entire network.

6.4. Comparison of the simulations with the antecedent theoretical EEG models. We conclude from the results of these simulations that our speculative comments in section 5 are supported in general. The macroscopic and microscopic simulations have been contrived as much as possible so that the one is the generalization under the principles of mass action of the other. Macroscopic wave motions are near-linear, relatively nondispersive, and near-equilibrium. A considerable body of experimental data is matched with regard to macroscopic wave motions, and at the microscopic level extremely complicated dynamics, with high sensitivity to initial conditions, are exhibited. We have not yet been able to define the Lyapunov exponents, or make other dimensional measures of the microscopic dynamics, but we believe these measures could have little meaning without considerable further classification of the basins of attraction present in the state space of any given ensemble of cells.

An effect not considered in our opening sections has become apparent—the sharp transition to excited states of neuronal activity, analogous to a change of phase, in the thermodynamic sense. This property is reminiscent of Freeman's (1964; Freeman & Skarda 1985) findings in the olfactory bulb, where contrasted dynamics associated with sniffing and resting conditions are apparent. Notably, the sharp transition is a feature of the microscopic model in line with the low maximum firing rate possible without runaway excitation at macroscopic level. Thus, the choice of synaptic gain imposed for other reasons in the macroscopic simulation appears independently justified. However, with this choice of synaptic gain, the simulations tend to become "stuck" at high firing rates.

Three further points of difference from the parent models from which this work began, emerge. First, no analog to the state-dependent asymptotic property of Freeman's asymmetric sigmoid function is present in either the macroscopic or microscopic simulations, and therefore no appeal to asymptotic limitation of cell-firing rate to stabilize cortical dynamics is applicable within the simulations. Instead, our simple assumption of a Gaussian distribution of action potential threshold for our cell populations imposes a static sigmoidal nonlinearity on wave-pulse relations.

Second, in the macroscopic model, boundary conditions do not exert the great effect on cortical dynamics predicted in most of Nunez's writings, even though the simulation reproduces experimental results against which the Nunez model has been previously tested. This appears to be due, at least in part, to Nunez's assumption that dendritic delays are negligible compared to axonal delays, which is not the case in the simulation. Thus, the simulation corresponds to an alternative, and relatively little explored, case within Nunez's general formulation.

Third, the appearance of harmonics in the simulation produced by sinusoidal driving shows that the macroscopic simulation is significantly nonlinear under perturbation of sufficient magnitude. The stochastic linear model is therefore only approximately applicable, even given free use of-
the central tendency theorems. The nonlinearity of response is attributable to the static sigmoid gain function, and the near linearity of wave motions depends on the fact that the macroscopic simulation runs in a stable fashion only at low average pulse densities, over which pulse-density/LFP relations are near linear.

We now turn to a brief description of types of dynamics seen in artificial neural networks, which were developed from a different perspective – principally that of the study of memory storage and retrieval.

7. Some aspects of attractor neural networks

7.1. Standard models. By "standard models" we mean networks with the following features:
1. Inputs to each neuron form a linear weighted sum with both positive and negative weight signs permitted.
2. Outputs of each neuron are two-state (Hopfield 1982) or graded (Hopfield 1984), with a threshold value for transition.
3. Connections are bidirectional and symmetric with small lags only.

Figure 5. Simulation of microscopic dynamics. The soma membrane potentials of 1600 cells (85% pyramidal cells and 15% stellate cells) arranged in a square grid, and coupled by intracortical connections in accord with physiologically and anatomically realistic synaptic and dendritic processes. Injected current is above the level of transition to high cell-firing rates. A single cycle of dynamic interactions, approximating a 13 Hz limit cycle, is seen in this case.

4. A system energy (Lyapunov energy) isomorphic with that of an Ising spin-glass can be defined as a function of the input weights, the current system states, and the unit thresholds.
5. Elements in the net change state asynchronously, and the system vector follows a trajectory to a point attractor – a static final state. At the attractor Lyapunov energy is at a local minimum.
6. Initial states leading to a common point attractor are thus classified identically by their shared basin of attraction. Because the trajectories to each point attractor converge, they are, by definition, not chaotic.
7. In the presence of high asymmetry of couplings, an energy is not defined and system convergence to a point attractor does not take place (Hopfield & Tank 1986).

7.2. Asymmetry and chaos. Amit (1990) draws attention to the limitations of standard models in so far as they may be considered models of real neuronal cooperativity. They are generally applied with the assumption that a network of $10^4$ elements or so is about the size of a realistic pool of completely interconnected neurons in the cortex. Their
dynamic properties are rather insensitive to symmetry of coupling strengths, until a point is reached at which the positions in state space of the basins of attraction are not fixed but are instead functions of the current state.

It may be that mobility of basins of attraction is crucial to the problem of storing and retrieving temporal sequences of inputs, since this property confers the ability to output a sequence of different patterns of action potential rather than the single fixed pattern of activity provided at convergence to a point attractor in a standard network (Amari 1972; Caianiello et al. 1967; Dehaene et al. 1987; Little & Shaw 1975; Nebenzahl 1987; Peretto & Nizey 1986).

Dynamic properties of trajectories in the state space of asymmetric nets can include chaotic and limit-cycle behaviors. These offer interesting features when asymmetric nets are considered in connection with Hebbian learning (Hopfield & Tank 1986; Parisi 1986a; 1986b). In the presence of a “pulsesset” type of continuously operating synaptic modification in which recurrent activation increases synaptic weight whereas disuse reduces the weight, learning and forgetting occurs depending on the class of dynamics in which neurons are engaged (Nadal et al. 1986; Parisi 1986a). Chaotic dynamics in some synapses prevent storage whereas other synapses engaged in activity of a limit-cycle type store memory.

7.3. Adiabatic landscapes. Kleinfield (1986) considers the case of symmetric couplings of two types between neurons – fast synapses, and those that begin to operate after a delay. This also confers mobility of the attractor basins, with potential for temporal sequencing of the network’s behavior. The associated energy function (or “landscape”) is said to be adiabatic (i.e., slowly changing).

7.4. Amit’s model of the cortical column. A neural design advanced by Amit and Tsodyks (1990; 1991) is composed of excitatory cells with inhibitory surrounds. The cells have an absolute refractory period and continuous membrane dynamics with realistic decay constants and rise times. The network is immersed in random, continuous afferent input as if from surrounding cell activity, with short bursts of nonrandom input as the signal to be classified.

These features lead to continuous and coherent intranetwork dendritic potentials, usually below action potential threshold, with spikes mainly emitted because of the noisy, continuous afferent. Firing rates remain low, but specific neurons spike more rapidly when the state vector is centered on an attractor.

Because of the ongoing noise the attractor dynamics cannot be classified as chaotic, limit cycle, or point attractor; each attractor basin corresponds only to a nonergodic domain.

Amit and colleagues (Griniasty et al. 1993) have shown that their symmetric networks can convert temporal correlations between stimuli to spatial correlations between attractors, paralleling the experimental findings of Miyashita and colleagues (Miyashita 1988; Miyashita & Chang 1988; Sakai & Miyashita 1991). The model can, in principle, be applied in conditions of coupling asymmetry (Amit, personal communication), but it is unclear at what level of asymmetry the definition of basins of attraction would be drastically affected.


Our macroscopic and microscopic simulations of ECoG and local field potential properties can, to a considerable extent, also be understood as special cases of ANN dynamics, as outlined above.

8.1. Local (microscopic) dynamics. Since cell-to-cell connections appear to be highly asymmetric, the energy concepts applicable in standard models do not apply in our microscopic simulation. The Amit and Tsodyks (1990; 1991) model is not directly comparable, since this model operates only at low firing rates, and has not been tested at such levels of asymmetry. However, there appears to be a good match with regard to the appearance of coherent fluctuations of local field potentials. Amit and Tsodyks’s model appears to be equivalent to a standard model ANN, under noisy perturbation, but with dendritic properties derived from physiological experiment as in our microscopic simulation. The appearance of coherent dendritic potential fluctuations at low firing rates may therefore reflect the way in which dendritic potentials are transduced into action potentials and depend little on specific connectivities.

It is unclear whether the success of Griniasty et al. (1993) in using Hebbian learning to reproduce the experimental findings of Miyashita et al. (sect. 7.4, para. 4) could be reproduced in our simulation. The principle source of uncertainty arises from the asymmetry of connections involved in the microscopic simulation. It may be, of course, that the cells involved in the Miyashita experiments share fairly symmetrical couplings, but we know of no reason to believe (or disbelieve) this is the case.

On the other hand, the types of Hebbian learning envisaged by Parisi (1986a; 1986b) and Nadal et al. (1986) (see sect. 7.2) appear applicable. In our simulation the combination of extreme divergence and asymmetry of couplings implies the necessary existence of many closed (and thus potentially self-exciting) loops capable of generating limit cycles when the overall firing rate is high, and thus creating basins of attraction.

Our microscopic simulation therefore seems to approximate the behaviors of certain asymmetric nets with potential to classify, store, and recall temporal sequences – but also to share properties with a low firing rate model, which describes the behavior of neural networks surrounded in a sea of concurrent input from the rest of the brain.

8.2. Dynamics at global scale. For scales greater than the macrocolumn, up to the entire closed extent of the cortex, average couplings can be more appropriately considered symmetrical, as indicated by the general success of our symmetrically coupled macroscopic simulation. Therefore, concepts of system energy are applicable.

The coupled stochastic oscillator model implies, at face value, that a very simple sort of basin of attraction governs the trajectories followed in the cortical state space. Since all long-range couplings are excitatory (positive) couplings, this implies a basin of attraction with a single energy minimum analogous to the case of a ferromagnet in spin glass theory. If that were so the trajectories followed would be defined simply by the principle of least action. This interpretation does not take account of complications in the
form of the energy landscape that might emerge consequent to interactions among the stochastic parameters, each of which was assumed effectively stochastically independent of the others.

Moreover, this "face value" interpretation is belied by the further properties that emerge from the simulations, which are:

1. The sharp transition from low firing rate to high firing rate dynamics. The appearance of patches of excited activity in the cortex would violate the assumption of independence of parameters in the stochastic model.

2. The requirement, imposed by this sharp transition, for a stabilization mechanism to prevent runaway excitation.

The implications of these complicating effects are discussed below.

8.3. Sharp transitions to high firing rates and interactions across scale. The presence of sharp transitions to rapid firing rates implies that, given sufficient excitement, small pools of cells can temporarily assume the capacity to drive activity elsewhere in the brain. During this state, the involved neurons would be readily and preferentially identified by single-cell recording techniques. Because complicated high firing rate dynamics at specific loci are associated in single-cell studies with very specific stimulus or response properties, this appears to be a strength of the simulations. Conversely, macroscopic scale events must interact with the single-unit events by virtue of the space averaging over large neuronal fields implicit in the determination of the synaptic input to small pools of neurons. This seems to offer a plausible explanation for the widely established association of temporal and topographic EEG with cognitive states (Gevins et al. 1983; Picton & Hillyard 1988).

If this interpretation is correct, then local areas of highly excited cells must be capable of highly perturbing the global brain state. The sharp transitions of "phase" about unity aggregate gain may be expected to confer a sensitive dependence upon initial conditions to the EEG time-series, even though mass action principles apply generally, and whether or not attractor dynamics are truly chaotic above or below this transition level of excitation. Highly erratic perturbation about a state approaching equipartition of energy among resonant modes is apparent in activated ECoG (Wright et al. 1990a).

8.4. The demand for global stability: Cortical afferents and adiabatic control. As they are presently formulated, both classes of simulation undergo transition to runaway excitation and stay there, under all but minimal perturbation. In the case of the macroscopic simulation it is easy to compare this runaway excitation to epilepsy, albeit in a very crude way. Indeed, the need for a low epileptic threshold formed part of our argument in arriving at a value for synaptic gain. But in the microscopic simulation the high firing rate states appear to correspond to attractor states necessary for functional classification of inputs. It is probably impossible to appeal to any local inhibitory mechanisms that might stabilize this runaway excitation, because once local excitation has become sufficiently intense to activate any local control, long-range spread of excitation would have already occurred.

Similar instability was encountered in the now-classic work of Wilson and Cowan (1973). They found that to achieve global stability in a model genetically similar to our simulations, they required excitatory-inhibitory interactions to occur at longer range than excitatory-excitatory interactions. This is a condition that cannot be met by a model concerned with the cortex alone. The requirement for stability arrived at by Wilson and Cowan means that large rings of inhibition must be generated around all excited cortical zones - a requirement that appears possible only if cortical-subcortical interactions are brought into play.

There is no shortage of candidate pathways for such control. The cortico-atrial-thalamic pathways and the long catecholaminergic and indolaminergic fibres to cortex suggest themselves, but a detailed consideration of the possible pathways of interaction is beyond the scope of this paper.

Cortical-subcortical interaction is generally polysynaptic, topographically organized in parts and diffuse in others. All delays in the total cortico-fugal and centrifugal extents of the interaction are likely to be much greater than cortico-cortical delays. Therefore, rapid control may be expected to involve relatively fast pathways of cortical-subcortical interaction. Perhaps local changes in intracortical inhibitory processes more complex than those introduced to our simulations so far also play a part, albeit a subsidiary one.

It remains as a large future modeling exercise to determine whether such interactions would be sufficient to retain overall stability or whether this requirement exposes a flaw in our conception.

If cortical global stability is eventually proven to depend on cortical-subcortical interactions, then a further mechanism of cerebral self-control - analogous in character to ANN findings - is implicit in the stabilizing process. Cortical-subcortical-cortical circuits can be considered equivalent to slow synapses within cortex and could thus act as a form of adiabatic control on the energy landscape. This control would operate on the slope at all points on the global energy basin, and thus partially control the state trajectory for the entire brain.

9. Conclusion

We have sketched an outline of the possible organization of brain dynamics and have indicated that whereas events at cellular and global scales may be studied to some degree in isolation, the rules by which the brain constitutes a cognitive engine may depend on both local and global network properties, with interactions across scale. This sketch explains basic wave properties of the EEG and is partially unified with some ANN principles, which themselves appear to offer an account for important aspects of cortical information processing and storage. The future study of learning rules within the simulations we have presented appears practicable.

Our outline appears sufficiently precise to be subject to critical test, as increasingly accurate physiological measurements are made and the simulations are developed to higher levels of anatomical and physiological veracity. As instances in the macroscopic simulation, if cortico-cortical connectivity is modeled with greater anatomical realism, then the predicted estimates of wave velocities and fre-
The simulations we have presented lack detail. Most seriously, they lack specification of the pathways, distributions, and relevant time-constants for the stabilizing interactions with subcortical structures on which their plausibility as genuine models of cortical dynamics depends. We have made little attempt to describe in-depth the influences of organization within the cortex. No regard has been given to the macroscopic and microscopic inhomogeneities of cortico-cortical and intracortical connectivity or the effects of rhythmic subcortical processes and resonances between cortex and subcortex, which undoubtedly contribute to the EEG (Steriade et al. 1990). We have considered neither the effects of neuromodulators in any detail; nor any complications of synaptic transmission, such as hyperpolarizing afterpotentials; nor cortical cellular specificity greater than that of two cell types—pyramidal and stellate. But it would appear that neither these factors nor many others are essential to a parsimonious treatment of general dynamic principles.

It also appears likely that the principles we have outlined apply only under certain further restrictions, for example, to waking states. A key point in our argument has been the conclusion that microscopic chaos (in the broadest sense) is compatible with the operation of mass action effects at macroscopic level. This argument depended on the introduction of spatially and temporally noisy input from the reticular formation. It may be that the withdrawal of this input permits the global chaos described in coupled chaotic maps by Kaneko (1990) to emerge.

Although the most important reservation on the completeness of this outline of dynamic principles appears to us to lie with specification of the cortical-subcortical interactions, it does appear that such interactions might prove relatively simple to model, for the very reason that to sustain low-average firing rate stability, the return pathways to the cortex must exert fast and widespread effects. Whether simple or complicated, the effect such control might have on the competitive evolution of dynamic patterns of activity in the cortex should be very rich.

Acknowledgments

Parts of this paper have already appeared in *Psychology, 93.4.60 EEG-chaos. I. Wright.*
local module is taken seriously to be composed of a large number of excitatory and inhibitory cells, one finds that excitatory input into a module can create inhibition because the inhibitory cells in the module typically react more rapidly than the excitatory ones.

(9) This raises a perennial comment (see also Author's Response, Amit 1995): Effective units (oscillators or other, sect. 4.1) composed out of neurons in an unspecified way are problematic and unnecessary. They are problematic because they require an additional level of modeling to produce out of credible neurons and synapses the required properties of the effective units. They are unnecessary because they produce nothing that cannot be produced at a single level of modeling.

(10) The above comment connects to an apparent inconsistency: if the units are effective minicolumns (sect. 4.1), then what do the (very useful) synaptic data refer to?

To conclude: The logical sequence of W&L's target article is somewhat perplexing. On the one hand, a large part of it reads like a rather critical review of ANN models of cortical module dynamics. On the other hand, some of it reads like a strenuous effort to link itself to a well-understood body of knowledge concerning EEG spectra and synaptic anatomy.

Is the distribution of coherence a test of the model?

Theodore H. Bullock
University of California, San Diego, La Jolla, CA 92093-0201.
bullock@usc.edu

Abstract: Does the Wright & Lilly model predict: (1) that subdural and hippocampal EEGs coherence tend to rise and fall in parallel for many frequencies, (2) that it is locally high or low within 10mm and falls steeply on average or, (3) that it is in constant flux, mostly rising and falling within 5–15 sec?

Three features of the relatively fine structure of EEG coherence in space and time recently found or emphasized (Bullock & McClune 1989; Bullock et al. 1995a; in press) would seem to be unexpected on the basis of Wright & Lilly's (W&L's) model.

I refer to the pairwise comparison of any two simultaneous, 0.3-50 Hz EEG records, taken with rows or grids of subdural macroelectrodes 5–10mm apart on the pial surface or with needle-like probes having 8–12 contacts deep in the temporal lobe in epileptic patients monitored 24 hours a day for many days. Coherence is meant in the technical sense: a value between zero and one for each frequency, representing the fraction of energy which is at the same phase (any phase) in the two time series throughout the period analyzed. (The term is sometimes used for cross-correlation, which is not frequency specific, and is amplitude sensitive and requires near zero phase. It can be high while coherence is low and vice versa.)

1. As coherence fluctuates, up and down, both in time and in space, it tends to move in the same direction at the same time for all frequencies over several octaves.

2. Coherence for all frequencies falls, on the average, as a function of distance between electrodes of a pair such that it is ca. 0.5 at 10mm and indistinguishable from chance at ca. 20–30mm. Besides average decline, local differentiation is common: adjacent pairs 10mm apart may differ greatly.

3. Coherence is constantly changing in the resting, alert state and in stage 1 sleep as well as during seizures. Some drifts may last for 30 sec or more before reversing; more commonly 5–10 sec.

The first feature argues against the view of the EEG as a mixture of independent oscillations, as the Fourier analysis assumes. Although W&L do not explicitly claim that their unit oscillators are independent, they may be implicitly so understood. They doubtless expect interactions between oscillators, but not such that the 10 and 11 and 12 and 13 Hz and every frequency component over two or more octaves typically tend to change coherence in parallel, across both space and time. Of course, a wideband change in coherence is equivalent to a change in the proportion of wideband stochastic activity. In our data, this stochastic activity is intrinsic brain activity, not extrinsic noise.

The second feature indicates a high degree of structure, differentiation or heterogeneity in the millimeter domain, in respect to this dynamic cooperativity measure. The coherence decline with distance is not due to passive spread and varies widely, except in averages. EEG from scalp electrodes typically falls much more slowly for all frequencies. From intracortical microelectrodes it can fall to <0.5 in 1 or 2 mm. This seems quite compatible with W&L's model but it is not clear how one could predict the slower fall with distance in scalp records or the higher differentiation subdurally or deep in the temporal lobe and the still higher local structure with microelectrodes. It would be useful if their model predicted the effect of recording from smaller and larger volumes of tissue upon the coherence and the absolute EEG amplitude and its spectrum. Likewise for the bicoherence. Without going into more detail here, we find numerous peaks or mountain ranges of bicoherence in many 10–20sec samples of waking and sleeping EEG whereas in many other samples there is no bicoherence above the 95% confidence level.

The third feature, like the others, is not essentially contrary to the model but a test, insofar as it may not have been predictable. In our fairly wide sample of cerebral lobes, pairs of electrodes, epochs, brain states, and individual subjects, it is characteristic that coherence for any given pair and band is constantly in flux. Only when we compute it every half second (with a corresponding increase in the uncertainty of seeing a difference) do we observe little or no difference between successive estimates. The spectrum of fluctuations has little energy <0.5 Hz; one misses little by computing every 2–3sec. Most of the spectral energy is in the 0.07–0.2 Hz range – fluctuations lasting 5–15 sec. A weak tendency to fluctuate periodically, more than expected by chance, is sometimes evident but usually insignificant. Coherence seems to be sensitive to any form of effective stimulation or endogenous event.

In this range of frequencies (0.3–50 Hz) and of tissue volume or population size (several cubic millimeters), it is a mesoscopic measure of assembly cooperativity, at each frequency and at any phase relation, that could not be predicted from a large number of electrodes tuned to unit spike activity. This is probably the best measure of synchrony, which is otherwise an eyeball guess highly subject to error due to changes in power spectrum without change in synchrony. Synchrony may be one of the primary features of brain dynamics that has evolved, comparing the cerebral pallium in fish, amphibians, reptiles, birds and mammals. What would Wright & Lilly suggest as possible key differences to make their model fit nonmammalian data?

Levels, models, and brain activities: Neurodynamics is pluralistic

Péter Érdi
Department of Biophysics, KFKI Research Institute for Particle and Nuclear Physics of the Hungarian Academy of Sciences, H-1525 Budapest, P.O. Box 49, Hungary. erdi@rmki.kfki.hu

Abstract: Some dichotomies related to modeling electrocortical activities are analyzed. Attractor neural networks versus biologically motivated models, near-equilibrium versus nonequilibrium processes, linear and nonlinear dynamics, stochastic and chaotic patterns, local and global scale simulation of cortical activities are discussed.

In the intersection of dichotomies, Wright & Lilly (W&L) have attempted to develop their own models for describing cortical electrogenesis in the light of the results of some important previous modeling efforts. They mention, sometimes explicitly, sometimes implicitly, a few dichotomies which have apparently canalized the analysis of their models.
The most important dichotomies I found in the target article are the following: (1) the physicists’ attractor neural networks versus biologically motivated models for the origin of the EEG; (2) near-equilibrium versus nonequilibrium processes; (3) linear versus nonlinear dynamics; (4) stochastic versus deterministic chaotic time series; (5) local (microscopic) and global (macroscopic) scales. In this commentary I take up their line with some reflections on these dichotomies.

1. Attractor neural networks versus biologically motivated models. “Computation with attractors” became a paradigm which suggests that dynamic system theory is a proper conceptual framework for understanding computational mechanisms in self-organizing systems such as certain complex physical structures, computing devices, and neural networks. Its standard form has a few properties, including the following: (1) the attractors are fixed points; (2) a separate learning stage precedes the recall process whereby the former is described by a static “one-shot” rule; (3) the time-dependent inputs are neglected; (4) the mathematical objects to be classified are the static initial values: those which are allocated in the same basin evolve towards the same attractor and can recall the memory trace stored there. In its extended form, not only fixed points, but also limit cycles and strange attractors can be involved. A continuous learning rule may be adopted but, in this case, the basins of the attractors can be distorted, which may even lead to qualitative changes in the nature of the attractors. W&L cautiously suggest (in sect. 5) that biologically motivated models of the cortex can be interpreted as special cases of attractor neural networks. More realistic models, however, which take the continuous interaction with the environment explicitly into account, are nonautonomous in the mathematical sense (Aradi et al. 1995; Erdi et al. 1992). Such systems do not have attractors in the general case. Consequently, attractor neural network models cannot be considered general frameworks for cortical models.

2. Near-equilibrium versus nonequilibrium processes. W&L state (sect. 4.3) that the ECoG is interpreted as a near-equilibrium process. On the contrary, Freeman’s approach is motivated by the theory of nonequilibrium thermodynamics (Nicolis & Prigogine 1977) and nonequilibrium phase-transitions (Haken 1977). The term “near-equilibrium” does not have a precise definition, and has had a bad reputation in better circles of the kingdom of Thermodynamics (e.g., Truesdell 1969); it is really preferable to avoid it. From a purely physical point of view the brain is a thermodynamics-based nonthermodynamic system. It has been pointed out (Erdi 1983) how nonequilibrium thermodynamics can be extended to “higher than molecule” hierarchical levels. Such “near-equilibrium” phenomena as random fluctuations around equilibrium (or stationary states) could and should be treated as special cases of nonequilibrium (stochastic) processes.

3. Linear versus nonlinear dynamics. W&L state (sect. 4) that nonlinearity in the small and linearity in the large can be reconciled. “Linear” and “nonlinear” are mathematical concepts; they have multiple meanings, and we must be cautious in using them. In Freeman’s model, for example, (1) the static transfer function (wave-pulse relationship) is nonlinear; (2) the differential equations for describing the activity of lumped cell aggregates are also called nonlinear. Though 1 do not see the formal connection between the two levels (and it certainly cannot be the aim of the target article to resolve such technicalities), in principle the transition between linear and nonlinear equations is possible. To mention a relevant example, in the case of discrete state space stochastic processes, the differential equations for the absolute probability are linear, but nonlinear equations for the expectation can be derived from them. These derived equations govern situations qualified experimentally as “nonlinear phenomena.”

4. Stochastic versus deterministic chaotic time series. On the one hand, classical signal analysis considers EEG records as realizations of (often stationary) stochastic processes, and spectral analysis has been the conventional method for extracting the dominant frequencies and other parameters of the rhythms. On the other hand, the occurrence of chaotic temporal patterns has been reported at different hierarchical levels of neural organization. Chaotic patterns can be generated at the single neuron level because of the nonlinearity of voltage-dependent channel kinetics of the ionic currents, at the multicellular network level arising from the interactions among neurons, and globally from spatiotemporal integration.

Dynamic system theory offers a conceptual approach to EEG signal processing that is different from classical analysis. Time series, even irregular ones, are considered deterministic phenomena generated by nonlinear differential equations. Though the methodological difficulties in interpreting the calculated quantities (Lyapunov exponents, fractal dimensions) characteristic of neurological categories are now acknowledged, we cannot deny that dynamic system theory provided a breath of fresh air to the methodology of processing of neural signals.

W&L themselves gave a stochastic model for the generation of electrocortical activity but tried to accept the occurrence of chaotic signals too. They have another “in small and in large” principle, namely, chaos in the small and stochasticity in the large. Despite the existence of improved statistical methods to discriminate between noise and chaos, it is still hard to decide uniquely whether electrocortical recordings should be considered purely random or deterministic chaotic patterns.

5. Local (microscopic) and global (macroscopic) scales. The adjective “local” means pointlike here, and “global” may be interpreted as the whole. Obviously, local descriptions may have very different (and relative) meanings, even if we do not go below the single cell level. A single neuron, which is certainly a submicroscopic unit relative to Freeman’s microscopic model, may be described by a spatially extended (global) model by using a multicompartmental technique. But Freeman’s model for neuron populations is local, since the whole population is lumped into a point. Instead of speaking about “local” and “global”, we should remember the whole hierarchy of the neural organization. A step in this direction is the statistical theory of neural fields (Ventriglia 1988, 1990) for describing the global cerebral activities in terms of interacting sub-fields. The main methodological problem, still unsolved, is to integrate the relevant aspects of different levels of description, from single cells through networks to population. W&L have clearly recognized both the necessity and the absence of integrative frameworks. Even though their analysis has not led to the “final solution” which would need deeper conceptual and mathematical consideration, W&L’s target article will certainly be thought-provoking for commentators and readers.

Acknowledgment
Support of my work by the Hungarian Scientific Research Fund under Grant No. T 017784 is acknowledged.

Multiscale modeling of the brain should be validated in more detail against the biological data

Harry R. Erwin
Department of Computer Science, George Mason University, Fairfax, VA 22030. herwin@gmu.edu

Abstract: Wright & Liley provide an advance in addressing the interaction of multiscale processes in the brain. It should address in more detail the biological evidence that underlies the models it proposes to replace.

Over the last fifty years, two scientific research communities have extended the early work done by McCulloch and Pitts (1943) modeling computation in the brain. Engineers and physical scientists studying artificial neural networks (ANNs) have emphasized the computational capabilities of networks of very simple neurons,
while biologists and neuroscientists have been more concerned with simple biological neural networks (BNNs) of complex neurons in behaving animals. In recent years, the two fields have converged as computer capacities have grown and modeling techniques have improved. That development has encouraged increased theorizing in both communities about how the brain works (Calvin & Ojemann 1994; Churchland & Sejnowski 1994; Crick 1994; Dennett 1991; Freeman 1985; Penrose 1989, Shepherd 1994). The modeling and theorizing needs to be validated against the biology—a crucial model of how the brain works, validated against the biological evidence, is superior to a detailed model that lacks a basis in biology.

Wright & Liley (W&L) approach modeling the brain at two scales, the macroscopic scale of the electroencephalogram (EEG) and the microscopic scale of the neural network. This is a marked advance on past work, and is likely to be followed up by other workers.

At the macroscopic level, W&L provide cogent criticism of Freeman's (1975) model of perceptual processing in the olfactory bulb. However, Freeman (1995) makes it clear that his modeling was motivated to clarify the biology, not to stand alone as an independent model of perceptual processing. As a consequence, the work described in Freeman (1975) and Yao and Freeman (1990) was left incomplete in a number of areas where the biology was inadequately understood at the time. It is not surprising that better models can be found, especially with the advances in neuroscience in the last ten years. I would suggest that as part of their model validation, W&L need to address the experimental data that originally motivated Freeman's modeling—for example, the evidence for near-zero phase lag between the oscillations of widely separated neurons (Bressler et al. 1993) and multiscale interactions with the transfers of activity across hierarchical levels that are discussed in sections 3.4 and 5.6 of Freeman (1995). Discussion of these areas would strengthen W&L's contribution.

The modeling also needs to be validated against the biology at the microscopic level. The assumptions the authors make are: (1) a spherical model for the dendritic and intracortical axonal trees, (2) no role for the apical dendritic tree, (3) only axo-dendritic connections, and (4) random intracortical connections. These disagree with the biological data in a number of ways that appear to be relevant to how cortex processes its input both locally and globally. For example, there is now evidence (Whittington et al. 1995) that the “40-hz” gamma wave is produced by interactions between inhibitory (CABAergic) interneurons and excitatory primary neurons via metabotropic glutamate receptors. In the olfactory bulb, mitral axon collaterals often synapse on spines on mitral basal dendrites that are also the sites for reciprocal synapses with inhibitory granule cells (Shepherd 1994). Recent work (Alkon et al. 1992) suggests that if a CABAergic synapse is triggered when the postsynaptic side is depolarized, the synapse becomes excitatory, resulting in immediate spiking with longer-term membrane depolarization and increased excitability—hence supporting a form of short-term associative memory localized to those spines. These processes seem clearly related to the mechanisms addressed by Freeman, and should be addressed in modeling the interaction between local and global cortical dynamics.

Neural system stability

Walter J. Freeman
Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley CA 94720-3200. wfreeman@garnet.berkeley.edu

Abstract: Two hypotheses concerning nonlinear elements in complex systems are contrasted: that neurons, intrinsically unstable, are stabilized through embedding in networks and populations; and, conversely, that cortical networks are intrinsically stable, but are destabilized through embedding in cortical populations and corticostriatal feedback systems. Tests are made by piecewise linearization of nonlinear dynamics at nonequilibrium operating points, followed by linear stability analysis.

Wright & Liley's (W&Ls) is a welcome exploration of the utility of linear analysis as a powerful toolkit for describing the evoked and ongoing activity of neural systems ranging from single neurons to human brains. Its use does not imply any assumption that the system under study is linear; rather the performance of a linear model simulates the observed activity of the system in a delimited range of its function. The standard methods for verifying applicability depend on paired shock testing for superposition and the derivation of impulse responses and records of background activity under assumptions of time invariance and Gaussian “white noise” input. Tests are based on comparison of Fourier spectra from system and model outputs and the linear decomposition of the impulse responses for comparison of eigenvalues (poles and zeros) derived by use of autoregression (W&L) and root locus techniques (Freeman 1975; 1992) with closed loop rate coefficients from heuristic models.

Restriction of a biosystem to a near-linear or piecewise-linear domain combined with linearization of its nonlinear model about a putative operating point is the most effective way to understand and analyze the stability properties of the biosystem. This is seen implicitly in the incorporation of the linear passive membrane and cable models of nerve into the Hodgkin-Huxley and Fitzhugh-Nagumo models of nerve axon. An effective way to model the Poisson-like distributions of action potentials of cortical neurons is to assume that every neuron spends 99% of its time near equilibrium and below threshold, but embedded in neuronal and bombarded by uncorrelated discharges continually. This is also a good model for the “white noise” required by the W&L approach. By local and global stability properties have been shown to depend on saturation of population firing rates without the necessity for inhibitory interneurons (Chang & Freeman in press; Freeman 1975). Another example is the brilliant analysis by James Houk (1974), showing how the gamma efferents suffice to linearize the neuromusculoskeletal system.

Experimental demonstration of a linear domain is more than a convenience for stability analysis; it manifests an operating characteristic of a neuron, cortical column or forebrain, which calls for a synthesis of the role of its attractors (point, limit cycle, chaotic) in the larger system of which it is a part, and for analysis of the neural mechanisms by which local stability or bistability is established, maintained and terminated as needed. In this respect I disagree with W&L.

In my view, the stability of domains of background activity is maintained locally at all levels—from single neurons, networks (spinal cord), and populations (my KI, KII, and KIII sets), to cortical columns and areas—as a solid foundation for the contribution of microscopic elements to the next higher hierarchical (“global”) level. The key property at the axonal level is the inactivation of the sodium channel, producing bistability. At the population level is the sigmoid curve, giving rise to cortical bistability (“burst” and “interburst”). W&L do not accept (sect. 2.2, para. 2) my experimental evidence (Fig. 3.18 in Freeman 1975) or formal derivation (Freeman 1979) for the existence of the upper inflection of the sigmoid, and account for the lower inflection “with our simple assumption of a Gaussian distribution or action potential threshold” (sect. 6.4, para. 3). Their assertion that the distribution “imposes a static nonlinearity on wave-pulse relations” (loc. cit.) does not follow, because the threshold does not account for the upper asymptote. In any case, they invoke the action of nuclei in the striatum and brain stem for the mechanism to prevent runaway excitation and epileptiform activity in cortex. Burns (1988) showed that isolated cortices go to a point attractor, not into sustained activity, and Becker and Freeman (1988) showed that the isolated olfactory system maintains normal, waking activity, lacking only the changes associated with sleep and hunger. Hence, the mechanisms for stability are to be found locally, and the transitions through instabilities are to be ascribed to the interactions of systems in the larger and encompassing global system at the next higher level.

A great service provided by W&L is to clarify the complemen-
tarity of the local network approach with the global approach by
Nunez (1995). My students and I have now provided evidence that
coopetative events in necortices during perception last
only 75 to 150 msec (Freenan & Barrie 1994) and occupy
domains not exceeding 1–2 cm. in diameter (Freenan et al.
1995) in rabbits; our data on phase velocities are consistent with
W&Ls data in Figure 3. We have comparable results from electrocorticograms in awake neurosur gical patients (Menon et al.,
in press). Concomitant recordings from cortical “patches” (Freen-
an 1995) separated by more than 2 cm. fail to show more than
chance levels of correlation or coherence, yet we believe that
they are cooperating in some manner. We are baffled. I find that
the linear approaches offered by Nunez and by W&L are the
best means currently available for systematic exploration of this
unsolved problem.

One last point: I do not wish to be hung with the albatross of “40 Hz.” That is a modal frequency for cat cortex, but the modes vary
over time and state in differing cortices, and across vertebrate
species from below 30 Hz in pigs to above 100 Hz in mice, as
shown by Bressler and Freeman (1980), where we coined the term
“gamma” to designate the EEG spectral range above alpha and beta.

Modeling for modeling’s sake?
Valerie Gray Hardcastle
Department of Philosophy, Virginia Polytechnic Institute and State
University, Blacksburg, VA 24061-0126, valerie@vt.edu

Abstract: Although this is an impressive piece of modeling work, I worry that the two models that Wright & Liley have created do not yet provide us with useful empirical information regarding brain processing.

My initial impression of Wright & Liley’s (W&Ls) article was that it is an impressive piece of modeling work. I still feel that way. My second thought was: What are these simulations supposed to show? I still haven’t answered this question to my satisfaction.

First pass answer. At the end of their piece, W&L write that “the rules by which the brain constitutes a cognitive engine may depend on both local and global network properties, with interactions across scale.” (sect. 9, para. 1).

This can’t really be their conclusion, however, for in the body of their text they never discuss the brain as a cognitive engine, nor do they discuss whether cognitive engines require any network properties. Being cognitive does not enter the picture at all, as far as I can tell.

Second pass answer. Let us assume that this concluding summary was simply a misstatement. Throughout the article, W&L allude repeatedly to network interactions across scales. For example, they write that the “sharp transitions [which] occur between low and high cell firing rates . . . may form a basis for neural interactions across scale” (Abstract). The brain is currently analyzed at many different levels of organization. One of the higher analyses relies on ECoG recordings; one of the lower ones on single cell recordings and local field potentials. However, few worry about how these two levels fit together. If W&L could demonstrate how the two levels interact, then they would be filling in a significant piece of the puzzle. I take it that they see this interaction as key to explaining how “the subcomponents of the brain are integrated into a functional whole” (sect. 1, para. 1).

But again, I remain unpersuaded. First, I fail to understand exactly what W&L could mean by interaction itself. There is a long (and sometimes tedious) philosophical tradition that worries about exactly how the different levels of organization in the world fit together. I won’t rehearse any bits of that here; suffice it to say that it is dubious that the higher levels of the world exert any “downward” causation. Let us assume that our neurons are arranged so that we can detect organized patterns at a macrolevel. Should we therefore conclude that these higher level patterns now causally interact with lower level neuronal activity? We can say it supervenes on the activity, it is composed of the activity, maybe we can even say that it is caused by the activity. But surely it is a mistake simply to assert that “macroscopic scale events must interact with the single-unit events” (sect. 8.3, para. 1, emphasis mine) without some further philosophical warrant. Consider: the chair I am currently sitting in is composed of molecules (or perhaps field potentials) whose interactions determine (among other things) its solidity. But I don’t believe that the chair exerts any influence back on its constitutive molecules. That strikes me as a confused metaphysics. At least, it is one that needs to be argued for. The same holds for the relationship between molecular interactions and neuronal firing patterns, between neuronal firing patterns and EEG patterns, and between molecular interactions and EEG patterns. Would we want to say that my EEG patterns somehow interfere with the molecules that make up my brain?

I also fail to understand how this interaction would explain how the different subcomponents are integrated into a whole. At this stage in the game, it is a great leap of faith to decide that the important “functional whole” in the brain is measured by EEG waves. One of the great difficulties with brain research today is knowing what the important levels of organization are, much less how they fit together.

Third pass answer. As a smaller goal, W&L mention that they show how many properties of the EEG can be modeled by a “lumped” simulation. However, this should not be terribly surprising. Insofar as the EEG waveforms reflect a statistical aggre-
gate of the transmembrane currents of all active neurons at a particular location, we should expect that treating these areas (or parts of these areas) as units with singular output should give us the same EEG waveforms as does a real brain. That is, we should be able to appreciate the “lumped” properties of EEG waveforms without W&Ls simulation by simply meditating on what an EEG measures.

Fourth pass answer. So what did W&L do? They created computer models of two models of cortex. I wonder whether the objects in the world – the actual objects to be modeled – aren’t being lost in these models of models. Take Nunez’s work for example. His equations are solved for variations on spherical shapes. Although what he has accomplished is enormously impressive, I worry whether these shapes approximate the brain, with all of its surface convolutions, in the relevant way. These convolutions wreak havoc on attempted localizations of activity from EEG data. Modeling this model of cortex as a model of cortex is probably premature.

In addition, that this modeling project captures some important regularities of the brain is rendered suspect as W&L note that their models easily fall into a state of runaway excitation (under all conditions except “minimal perturbation”; sect. 8.4, para. 1). Perhaps this points to the need for extracortical mod-
ulation; perhaps this points to fatal flaws in W&Ls models. The fact that we can’t choose between the alternatives highlights the need to tie models of brains more closely to known facts about the brain, as only a study of neurophysiology is going to answer this question.

But for all my negative comments, my first impression still remains – this is an impressive piece of modeling. Perhaps, as we collect more data from actual brains, we will be able to put this work to better use to generate novel and surprising predictions about how brains behave at some level of organization.
Nonlinear nonequilibrium nonquantum nonchaotic statistical mechanics of neocortical interactions

Lester Ingber
Lester Ingber Research, P. O. Box 857, McLean, VA 22101.
ingber@alumni.caltech.edu

Abstract: The work in progress reported by Wright & Liley shows great promise, primarily because of their experimental and simulation paradigms. However, their tentative conclusion that macroscopic neocortex may be considered (approximately) a linear near-equilibrium system is premature and does not correspond to tentative conclusions drawn from other studies of neocortex.

Importance of approach. At this time, there exists an interdisciplinary multidimensional gradation of published studies on neocortex, with one primary dimension of mathematical physics represented by two extremes. At one extreme, there is much scientifically unsupported talk of chaos and quantum physics being responsible for many important macroscopic neocortical processes (involving many thousands to millions of neurons) (Wilczek 1994). At another extreme, many nonmathematically trained neuroscientists uncritically lump all neocortical mathematical theory into one pile, and consider only statistical averages of citations for opinions on the quality of that research (Nunez 1985). In this context, it is important to appreciate that Wright and Liley (W&L) report on their scientifically sound studies on macroscopic and microscopic, but not on simulation and a blend of sound theory and reproducible experiments.

However, their pioneering work, given the absence of much knowledge of neocortex at this time, is open to criticism, especially with respect to their present inferences and conclusions. Their conclusion that EEG data exhibit linear near-equilibrium dynamics may very well be true, but only in the sense of focusing only on one local minima, possibly with individual-specific and physiological-state dependent parameters, embedded in a more relevant global context. It certainly would be more extreme, and probably quite incorrect, to conclude from W&L's inference that macroscopic neocortical phenomena in general can be treated as linear near-equilibrium systems.

No chaos. As discussed by W&L, more evidence is being collected on the irrelevance of chaos to macroscopic neocortical interactions; this is similar to serious work undertaken in several fields, where the impulse to identify "chaos" in a complex system has often proved premature (Rapp et al. 1993). It is not supported by the facts, tentative as they are, because of sparse data. Similar caution should be exercised regarding neocortical interactions.

Many such studies demonstrating chaos are quite model dependent; it is hence fair to present only the models as exhibiting chaos, not necessarily the actual physical system. There are, however, reports of model-independent experimental observations of distinctive stimulus-dependent parameters with extreme sensitivity to initial conditions.

It has been widely noted that measures such as the correlation dimension of data are difficult to calculate (Abarbanel et al. 1993). Perhaps the correlation dimension is often not even a well-founded concept, since, for example, the EEC of event-related potentials is probably nonstationary and highly context- and subject-dependent. Its calculation, even when supplemented with other statistical tests and noise reduction techniques, may prove fruitful, but probably only as a sensitivity index relative to shifting context and complementary to other models of EEG data.

Early in the development of a model of the statistical mechanics of neocortical interactions (SMNI) (Ingber 1981; 1982), it was noted that a remarkable sensitivity of macroscopic phenomena would be required to render any chaos that might exist at macroscopic scales relevant at columnar or macroscopic scales (Ingber 1983).

In a recent study using a highly accurate algorithm to evolve multivariate probability distributions, chaos persisted in the presence of moderate noise in a Duffing analogue of EEG systems (Ingber et al., in press). It was tentatively concluded (more CPU time is needed for longer detailed runs) that adding moderate noise to models of neocortical systems that might exhibit (highly model-dependent) chaos washes out traces of chaos.

No quantum mechanics. Quite often, quantum mechanics is invoked to "explain" the linearity of macroscopic information processing under some set of hypothesized circumstances, for example, preparatory to invoking the quantum mechanical principle of superposition. However, there does not seem to be any reasonable role for specific quantum processes in explaining any macroscopic neocortical processes (certainly at the levels of phenomena giving rise to scalp EEG, consciousness, etc.; Wilczek 1994).

One often gets queries concerning associations between quantum mechanics and SMNI, mostly because of the similarities in their mathematical structures. It must be made clear that SMNI and quantum theory are only related to the extent that some modern methods of classical nonlinear nonequilibrium statistical mechanics have developed a modern calculus in the late 1970s (Graham 1978; Langouche et al. 1982) that turns out to have been likewise used for some phenomena (too small to measure) in quantum gravitational physics (DeWitt 1957); these phenomena also have relevance to some measurable phenomena in nuclear physics (Ingber 1984a). There is absolutely no other physical relationship between SMNI and quantum mechanics. For example, there are some subtle quantum mechanical processes in information processing that are specific to quantum mechanics that have no relation to any possible classical information processing (Bennett 1995).

In this context, it is noteworthy that the modern calculus, used for many other systems as well as for SMNI, does not even need the usual ensemble-averaging statistical mechanical arguments for its development, for example, as is implicit in Langevin-type differential stochastic rate equations in which "noise" reflects on finer degrees of freedom rather than the order parameters. Rather, the mathematically equivalent Fokker-Planck ("Schrodinger-type") and Lagrangian path-integral representations, which can be derived by ensemble averages of such noise, can be directly derived by nonlinear nonequilibrium thermodynamic arguments (Ruppiner 1995).

Nonlinear nonequilibrium. It is hard to understand why a claim is made for linearity across the range of phenomena measured by EEG data, even if those data have the requisite resolving power to tell us much about neocortical function. This last condition is not trivially established, and is of course a subject of intense research.

For example, the existence of multiple attractors must certainly be taken as evidence for nonlinearity, by definition. Short-term memory (STM) certainly exhibits multiple states (Miller 1956; Zhang & Simon 1985), and it seems reasonable to conclude that this is due to nonlinearity in neocortical processing (Ingber 1984b; 1985; 1994; 1995a; in press a; in press c; Ingber & Nunez 1995). It is possible, but unlikely, that multiple states of STM are due to a single attractor, driven to extreme nonequilibrium within tenths of seconds (to access multiple memories within these epochs, as observed) to new contexts (thereby at least contradicting any claims of near-equilibrium!!). For example, there is strong evidence for the existence of mechanisms that establish at least quasi-stationarity of multiple attractors (Ingber, in press a; Lisman & Kliart 1995).

The reference in section 3 of W&L to the confluence of global and local interactions (Ingber & Nunez 1990) in fact argues against linearity of EEG-measured mechanisms. A more specific study, directly related to that reference, demonstrated that EEG could be fit quite well by SMNI, using only values of synaptic and neuronal parameters lying within experimentally determined ranges. This required the use of a sophisticated global optimization technique, adaptive simulated annealing (ASA) (Ingber 1993), previously called very fast simulated annealing (VFSR).
Commentary/Wright & Liley: Dynamics of the brain at global and microscopic scales

(Ingber 1989) was used to fit the global optimum over several attractors, specific to each subject's P300 data, within the group of 49 subjects. It was certainly clear from that study that there is a great deal of nonequilibrium dynamics inherent in neocortex; for example, these attractors would radically shift if one attempted the same study using 100 epochs of EEG.

A good analogue exists to demonstrate just how EEG would have to look to be linear. The same mathematical development used for SMNI has been applied to financial markets, including specific trading scenarios (Ingber 1984c; 1990; 1995d; 1996b; in press b). The financial data are well-known to be quite nonlinear and nonstationary (in fact, they also are quite hard to distinguish visually from some EEG data!). Global optimization, such as using ASA (Ingber 1993), is required to fit these bivariate nonlinear stochastic models to data. However, if decisions are to be made on recognizing patterns in the data, that is, for the purposes of trading, then trading rules based on a kind of moving average of "canonical moments" indicators (derived from the Lagrangian model fit to a separate independent training set of data) can indeed take advantage of "inefficiencies" in the market, where inefficiencies here imply nonlinear nonequilibrium states of the system.

A current project uses similar constructs for canonical momentum indicators derived for EEG from the SMNI Lagrangian as previously fit to EEG data (Ingber 1991), to correlate physiological or behavioral states in human subjects (Ingber 1995b; 1996a). It is recognized, however, that this will require "customization" for each subject. Similar to the recursive optimization performed for the above finance calculations, there is an inner shell of optimization on the Lagrangian model to data, and an outer shell of optimization on the rules for pattern recognition - in this case - parameterized rules used by clinicians.

The point here is that the use of moving averages might seem to imply that the system is being viewed as linear and near-equilibrium. However, this would be quite myopic, looking only within a short window of data instead of seeing or appreciating the full context of the system. This myopic view, if implemented in trading, without the possibility of adjusting in a timely fashion to indications that the system is quite nonlinear and nonequilibrium, would quickly lead to financial ruin. I believe it is fair to state that a similar view of EEG would quickly lead to incorrect diagnoses of many physiological and behavioral states across many patients and subjects.

Questions. Some questions can be raised specific to W&L, with a view to constructive improvement of their methodology. They outline some of their basic assumptions in section 4.1: Are all the parameters and variables in their assumed stochastic second-order differential equations tightly bound to specific neuronal entities, within experimentally established ranges? This is not clear for all parameters, notwithstanding their presentation of Table 1. Otherwise, some other system may be being fit. (Piece-wise) Linear models of data can usually be expected to extract linear systematics. It would be more useful to develop some realistic nonlinear model, and then to see whether it could exhibit a specific range that strongly exhibits linearity.

The use of "white noise" for W&L's simulations suffers serious defects. It has long been recognized that many complex chemical and biological systems have complex noise structures (van Kampen 1976), typically arising from stochastic and statistics of variables and parameters arising in the "deterministic" structures. This certainly was true in SMNI development.

In this context, W&L have taken their parameters to be time-independent as well as noise independent, to facilitate implementation of their model. In many complex systems, one can often use time-independent parameters if one takes into account that nonlinear nonequilibrium structures are required to model temporal evolution. Attempts to approximate these systems with quasilinear models (e.g., ARMA-type models) requires including many more higher-order structures to model this time-dependence (Kishida 1982, 1984).

The point here is that for many systems, only after a truly nonlinear study is undertaken (one with a full nonlinear structure in the noise) does it make sense to assume that one has found a local minimum ("the needle in the haystack"); similarly for ranges of constant noise.

Most serious is the fact that a crucial and influential parameter in W&L's model, the synaptic gain of set of questionable origin, has no specific neocortical origin. Anything might be expected, and W&L seem to find quite a bit of activity by tweaking this parameter.

It is important to note that W&L have undertaken further studies using the Genesis simulation, which can generate a good deal of microscopic realistic neocortical modeling. It remains to be seen whether they can carry out this program of basing conclusions about macroscopic neocortex on sampling such microscopic simulations. Not only must columnar dynamics be properly modeled, with a thousand minicolumns per macrocolumn, but global effects from long-range firings are likewise critical in neocortical EEG studies. For example, the macroscopic SMNI model, derived from a progression of development at neuronal and columnar scales, requires specific optimization, path-integral, and parallel algorithms to treat its minicolumnar network model (Ingber 1992). This model has taken advantage of the particular class of nonlinear structures many complex systems develop at meso-scales (Gaussian-Markovian; only in short epochs) and it remains to be seen how other complex models can be accurately calculated.

Like other complex systems (even "simple" fluids), a system as complex as neocortex is best treated as an overlap of quite different models at different spatial-temporal resolutions. I think W&L basically agree with this. The scientific challenges are to establish strong theoretical and experimental bases for overlaps of these models at different scales, and to articulate and understand the precise nature of emergent variables, parameters, and structures across these scales. It is dubious that much will be learned by attempting a "grand simulation" of such systems, especially since such simulations and models typically require approximations too harsh for specific calculations and empirical tests.

W&L are presumably undertaking the challenge of modeling macroscopic neocortex, respecting the importance of multiple scales of neocortical processing, and the mode they are developing will presumably be useful for future EEG data that will have better spatial resolution than they do today. These future resolutions are likely to lead to better correlations of EEG with real and individual human physiological and behavioral states; but then these models will have an even greater need to address nonlinear and nonequilibrium processes of neocortical interactions.

Comparative reduction of theories - or over-simplification?

Edgar Koerner
Honda R&D Co. Ltd., Wako Research Center, 1-4-1-Chuo Wako-shi, Saitama 351, Japan. edgar.koerner@r.d.honda.co.jp

Abstract: To model the organization of levels of cortical dynamics, at least some general scheme for hierarchy, functional diversity, and proper intrinsic control must be provided. Rhythmic control forces the system to iterate its state by short trajectories, which makes it much more stable and predictable without discarding the desirable ability of chaotic systems to make rapid phase transitions. Rhythmic control provides a fundamentally different systems dynamics, one not provided by models that allow the emergence of continuous trajectories in the system's state space.

Wright & Liley's (W&L') target article attempts to integrate two different abstract levels of systems description in neural network modeling by comparative reduction of attractor neural networks and the mechanisms giving rise to the EEG. The goal of the paper is a better understanding of the overall dynamic properties of the
Commentary/Wright & Liley: Dynamics of the brain at global and microscopic scales

brain (sect. 1), and W&L claim to have sketched an outline of the organization of brain dynamics (sect. 8). After reading the paper, one may be able to understand better the general dynamic principles of a system some of the crucial characteristics of which are captured by the model — but it is doubtful that this system would have much similarity with the brain. I will focus on two crucial characteristics the model fails to incorporate, namely, functional architecture, and related intrinsic global control mechanisms.

There is both a basic functional architecture, and a dynamically instantiated subset of the heterarchies within that architecture; the latter represents the currently selected internal state, which provides predictive feedback against which any activation is evaluated. There is convincing phylogenetic evidence that the cortex has developed top-down and that at least a certain top-down preference of control seems to exist. Hence, the momentarily selected organization of the system may control the routing and evaluation of any activation in the system, providing massive constraints for the dynamic behavior of the system by simply rigorously restricting its state space dynamically. Hence the assumption that "time-varying parameters were stochastically independent in the large" (sect. 4.1) is not justified. For a complex system such as the brain, functional organization matters; it cannot be discarded without rendering the model obsolete.

Minsky has made the point precisely:

The brain has evolved as (and, hence, consists of) hundreds of specialized subsystems, any one mutually connected with a more or less selected subset in a more or less specific way. Such a complex system must have an incredible number of control mechanisms... that enables its coordination according to the changing environment. (Minsky 1991, p. 345)

Rhythms are among the important control mechanisms to consider in simulations which claim to reflect brain dynamics. Rhythms seem to be imposed on the brain not only by subsystems of cortical inhibitory networks that can maintain large-scale oscillations at various frequencies (Buzsáki & Chrobak 1985) but also by subcortical systems such as the thalamic intralaminar nucleus (Llinás & Bilany 1993) for the gamma and septum (Stewart & Fox 1990; Vinogradova et al. 1993) for the theta range. The impact of rhythmic control on systems dynamics is more crucial than merely reflecting synchronization processes. Rhythms seem to define a cognitive "time span" for which the basically asynchronous events of neural processing are evaluated as being one cognitive object (Joliot et al. 1994); the positive correlation of theta and gamma rhythm (Buzsáki & Chrobak 1995) support this view that rhythmic control provides not only synchronization, but also (especially in the case of the theta rhythm) a limited period for activation driven evolution of the system's trajectory.

Consider chaotic dynamics of complex systems to have a rhythmic control as the brain does, it may simply be irrelevant to observe continuous trajectories generated by long-lasting iterations of the system. Instead, in any iteration rhythmic control allows only short activity-related epochs of instability, resetting the system at any respective period effectivly, enforcing a restart of the system with updated (or the same) initial conditions. Consequently, for modeling brain-like systems, only short system trajectories are relevant where the relation of a system's state to the initial conditions at the start of the trajectory is definitely not yet lost. Starting at any period at any local functional unit of the system, a limited trajectory with the current initial conditions can be an effective means of rapidly harmonizing the many local trajectories towards a globally consistent state, since with any restart the same or modified local trajectories may appear, effectively adjusting the global system's trajectory piecewise. This phenomenon can be compared to a "limited breadth parallel search" process and should improve both stability and predictability of the systems behavior considerably without discarding the benefit of rapid phase transition immanent in chaotic systems. From that point of view, models of chaotic processes that do not obey that very fundamental constraint are of limited relevance to understanding brain dynamics.

So, if W&L do indeed feel ready for models of ambition enough to describe the general organization of brain dynamics, they must include at least some general scheme for hierarchy, functional diversity, and proper control. There is surely a manifold of connectivity in the cortex which can be described statistically. At any moment, however, only a subset of that connectivity may be relevant for the system's dynamics.

Assume we are trying to model the general dynamics of a circuit board, not caring about the wiring and architecture; we define a general model of a local unit (say, a typical functional cluster of transistors such as AND, NOR) that appears more or less everywhere on the board, and link it according to section 4.1, neglecting its functional organization completely. What understanding of the true behavior of the board do we gain from stimulating that model with noise? My guess: close to nothing. Nor do I see a reasonable argument as to why the increasing complexity of the process under consideration eliminates the need to make an adequate simplification, or to limit the interpretation to the level supported by the model, which would be one homogeneous cortical subsystem only. And even so, the superficial reproduction of the characteristic frequency distribution and other phenomena of the EEG (which are at least to some degree due to rhythmic and other control) by that simple model does not demonstrate that it captures crucial characteristics of the target process. Hence, W&L's simulations do not support general conclusions on brain dynamics.

Why does the human brain need to be a nonlinear system?

Zbigniew J. Kowalik,a Andrzej Wrobel,b and Andrzej Rydzb

a Institute of Experimental Audiology, Biomagnetism Center, University of Münster, D-48129 Münster, Germany; b Nancki Institute of Experimental Biology, PL-02-093 Warsaw, Poland, zk@bihf.urz.uni-muenster.de

Abstract: We focus on one aspect of Wright & Liley's target article: the linearity of the EEG. According to the authors, some nonlinear models of the cortex can be reduced (approximated) to the linear case at the millimetric scale. We argue here that the statement about the linear character of EEG is too strong and that EEG exhibits nonlinear features which cannot be ignored.

Wright & Liley (W&L) target article concerns one of the most intriguing questions posed by neuroscientists: Is it possible to build an integrated model of the electrical potentials recorded from the brain? Such a model might provide the needed link between the microscopic and macroscopic description of neural activity and solve the question of the origin of the EEG. In addition, W&L discuss the problem of the linearity of micro and macroscopic information processing in the brain. We found the described estimations and simulations difficult to follow because the paper does not contain all the requisite mathematical information (some of this can be found in the cited papers). In what follows we have therefore restricted our comments to questions concerning the (non)linearity of the brain.

1. Problem of a scale. We feel that the division of brain processes into micro and macro domains is an oversimplification of reality. W&L also recognize this, referring to "subcomponents of the brain" "roughly equivalent" to a column, and to mechanisms underlying the EEG as "unit oscillators" or "single" and "multiple" units of information storage. We later read that cognitive processes are realized by still larger structures made up of many subcomponents, up to the brain as a whole. At none of these levels does the brain tissue appear to be homogeneous; rather, the brain is known to be built up from functional blocks. Altogether, W&L's approach to the scale problem is not well defined.

2. Local and global processing and the origin of the EEG. This is not just a semantic issue. W&L use terms such as local field...
potential (LFP), electrocorticogram (ECoG), and electroencephalogram (EEG) "somewhat interchangeably" in the target article. If the ECoG already "lumps" activity at a global scale (sect. 4.1) then the model lacks a description of the EEG which measures the electric potential "on the skin." The EEG is considered to aggregate the postsynaptic potentials of millions of neurons whereas ECoG is more local, not only because of a smaller distance between electrode and measured object (up to zero) but also because of more direct potential distribution (less mixing) as compared with scalp measurements. By localizing the sources of brain activity using MEG techniques (magnetoencephalography), we can often pinpoint a small region of the cortex where this activity takes place. Should the activity associated with this cortical area be described as local or global? It is of course global from the point of view of neuronal activity but at the same time local considering the cortex as a whole or the extent of presumed attractor neural networks (ANN). Adding to the previous finding that the LFPs correlation profile decreases as a function of distance between recording sites (Eckhorn 1994), we have recently shown that a specific global pattern of cortical activity forms while recording from the behaving cat (Krukowska et al. 1995).

The above examples, both measured at the millimeter scale, point to the difficulties of treating the cortex or even its parts as functionally homogeneous tissue.

3. Non-linearity. Our main argument about the nonlinear character of the EEG/MEG (for a review, see Elbert et al. 1994) is based on our finding that the divergence measure, in the form of largest Lyapunov exponent (LLE), is positive, and a test for determinism (Kaplan & Glass 1992) that indicates the deterministic character of brain signals (Mühlhöbel et al. 1994). We recognize the limitations associated with physiological data that rarely meet the mathematical assumptions required to define a system as chaotic, especially when considering nonstationarity and limited time epoch. Because of these reservations we cannot positively define the brain as chaotic, but we can test its nonlinear (possibly transiently chaotic) character. Recently, using a high-resolution 37-channel DC-SQUID neuromagnetometer, we estimated the largest Lyapunov exponent (LLE) for all MEG-channels (Kowalski et al. 1983). The positive values we found in all channels for all collected trials means that the initial sensitivity to an infinitesimal disturbance defines, for the case of a stationary deterministic processes, the chaotic state (transitory in a nonstationary case).

This nonlinear property means in general that the brain produces information. Indeed, we hypothesize that the brain, when producing a new quality at the global scale, generates new values which are not just a random (linear) projection of an existing reality (similar to the W&L model, which produces harmonics in Fourier analysis). Though there is lack of evidence for a low-dimensional chaotic behavior, the nonlinear character of EEG leaves no room for doubt (Fritchard et al. 1995).

We are also concerned about the apparent circularity of the argument motivating the choice of a linear model as W&L also justify their choice based on the performance of other mathematical techniques (e.g., AR models), which again need to be verified experimentally and not just numerically. If this justification exists, it would strengthen W&Ls theoretical position.

In explaining Freeman's (1991; Freeman & Jakobith 1993) model of a chaotic brain (sect. 2.4), W&L cite Kaneko (1990) to the effect that a set of nonlinear elements must be nonlinear. This statement is not valid in general and the global property depends on the number of nonlinear elements, the noise introduced into the system, and the coupling between elements. W&L further claim that at the macroscale their model exhibits linear properties. Using mechanics as an example, the question arises as to how the nonlinear character of the macrosopic pendulum depends on the order of the microscopic structure of the material used for its construction. Once again, coping with the problem of scale will be the deciding factor in answering this question.

Another apparent misunderstanding results from a lack of differentiation between global spatial and global temporal structure. The EEG/MEG model is a global measure, characteristic of the activity of a "large" cortical area. Its nonlinear character does not necessarily imply a chaotic spatial distribution of the EEG-amplitude on the scalp. In addition, a nonlinear system does not always produce chaotic time-patterns. It should also be noted that filtering EEG/MEG signals linearizes observed patterns. An additional argument for nonlinearity of the brain processing is that the information transfer between cortical structures requires nonlinear mechanisms of synchronization and that this synchronization is a global phenomenon (Abeles et al. 1994).

Despite all these criticisms, Wright & Liley's model is a promising step toward a common description of experimental and numerical results, even though it does not allow us to include information about the anatomically and functionally described pattern of brain organization or the identified connections between cortical layers and different brain areas.

Neuromodulation can significantly change the dynamical state of cortical networks

Hans Liljenström
SANS-Studies of Artificial Neural Systems, Department of Numerical Analysis and Computing Science, Royal Institute of Technology, S-100 44 Stockholm, Sweden. hali@sans.kth.se

Abstract: We present simulation results of an olfactory cortex model complementing the results presented in Wright & Liley's target article. We show how the cortical dynamics as expressed in EEG can be regulated by neuromodulation and discuss how the system can attain global stability without cortical-subcortical interaction, as presumed necessary by Wright & Liley. Network structure is shown to be crucial.

In addition to the more specific question of relating artificial neural networks (ANN) to EEG, Wright and Liley (W&L) bring up some fundamental issues in neuroscience. I will focus on the problems of different scales, global control and stability, and how these might be linked.

One of the main points of the target article concerns the interaction of different types of dynamics at microscopic and macroscopic scale. At the microscopic scale of single cells, the dynamics are characterized as nonlinear and probably chaotic, whereas it is characterized as linear and near-equilibrium at a macroscopic scale (sect. 5). Yet it is unclear how such an interaction between scales actually would be realized.

Apparently, the global dynamics of a system can be quite different from those of its elements. A highly complex behavior can arise in a system of simple elements, and vice versa. The problem lies primarily in the interaction between the elements, and how information is transmitted within the system. This is highly dependent on structure and organization. Computer simulations that are meant to capture the significant dynamical features of any particular system and scale should also be concerned with the relevant structural features at that level.

Comparing simulations of neural network models of the olfactory cortex at both "microscopic" (Wilson & Bower 1992) and "macroscopic" (Liljenström 1991) scales, reveal that the characteristic dynamics of the system depend primarily on parameters such as network connectivity, conduction delays, time constants, and mean firing frequency. The microscopic simulations were done with an early version of "Genesis," the same simulator package used by W&L. The "lumped" model has essentially the same network connectivity as the microscopic model, but its functional units correspond to populations of neurons with a sigmoid input-output relation, as derived by Freeman (1979). The additional features of the detailed model do not seem to be necessary for most of the dynamics experimentally revealed in EEG and local field potential recordings (Liljenström 1991). In fact, simulations with both models confirm many of the results.
presented in sections 6.2 and 6.3 of the target article, including travelling waves of activity.

Another issue that W&L raise is that of global stability and how it may be attained (sect. 8.4). They presume that cortical-subcortical interactions are needed, unless there are some intracortical inhibitory processes other than those introduced in their model. W&L will investigate this possibility using a two-layer version of their macroscopic simulation.

Both the microscopic and macroscopic simulations of the three-layered olfactory cortex, however, have shown that local inhibition from “feedforward” and “feedback” network units is sufficient to stabilize the global dynamics so that runaway excitation can be avoided. An increase in the excitatory output also results in increased inhibition, through feedback connections. In addition, long-range excitatory-excitatory connections can serve to stabilize the system, resulting in more coherent activity in different parts of the network.

Stability is indeed an important property of living systems, but it should be balanced with flexibility and sensitivity. It is conceivable that such a balance can be regulated in response to internal and environmental demands. This would presumably be reflected in different dynamical states of the system.

We have demonstrated how cortical dynamics can be under neuromodulatory control, either through an increased gain in the input-output function (Wu & Liljenström 1994), or through cholinergic suppression of neuronal adaptation and synaptic transmission (Liljenström & Hasselmo 1995). The dynamical state of the system can change from point attractor to limit cycle to strange (“chaotic”) attractor behavior. This also has a functional significance that will be discussed shortly.

Except for loosely hinting at its role for learning and cognition, W&L do not discuss the functional significance of the microscopic and macroscopic dynamics. Actually, it seems very difficult to discern whether the erratic time series obtained with various experimental methods are of chaotic or stochastic origin, and whether that would make any difference. If there is chaos or noise only at the microscopic scale, would it have any significance for global behavior, such as pattern recognition? What role does the dynamics play for the computational power of the system? Since computation, in general, requires convergence and stability, chaos would seemingly not be useful for an efficient computational process. Noise is normally also regarded as a nuisance. However, with our simulations of the olfactory cortex we have shown how noise (Liljenström & Wu 1995), as well as “transient chaos” that converges to near-limit cycle attractor dynamics (Wu & Liljenström 1994), can contribute to efficient information processing in associative memory tasks.

An initial chaotic state makes the system sensitive to changes in the input. It could also result in a rapid search of state space. Oscillatory behavior would correspond to a limit cycle attractor memory. Neuromodulatory control of the network dynamics can shift the balance between sensitivity and stability and result in a transition between “noisy” and more ordered behavior, depending on demands.

Whatever degree of realism one attempts to capture with a computer model, simulations can never be conclusive. At best, they can be helpful in testing our ideas and finding possible structures and mechanisms for certain phenomena, such as ECG. W&L have shown convincingly that a computational approach to important neurobiological problems can indeed serve this purpose; this is also supported by our own simulations of the olfactory cortex.

**Empirical data base for simulation: Firing rates and axonal conduction velocity for cortical neurones**

Robert Miller

Department of Anatomy and Structural Biology, University of Otago, P.O. Box 913, Dunedin, New Zealand. anat09@otago.ac.nz

**Abstract:** Simulation of brain dynamics requires the use of accurate empirical data. This commentary points out major errors in some of the empirical data used in Wright & Liley's simulation. The simulation is quantitatively very different from the real cortex, and may also have important qualitative differences.

Although I am a nonmathematician, I am aware that there is an important role for mathematics and computer simulation in unravelling brain dynamics. However, I am also keenly aware that because of the great complexity of biological systems, extended chains of mathematical reasoning have less importance in biology than in physics, unless corroborated at every pertinent step by empirical data. Thus I draw attention to the fact that Wright & Liley's (W&L's) reference list contains few papers dealing with empirical aspects of the structure or function of the neurones which generate the EEC. What empirical papers there are are mainly about large scale electrophoretic behaviour, which, though important, is factually less solid because of the complexity of data processing required to produce results. In view of this general observation, my specific comments are about the accuracy of the empirical data used, either as premises for the simulation presented here, or to compare with the result derived from it.

(1) The best data on firing rates of cortical neurones in the waking animal are from Swadlow (1994 and other papers), in unanesthetized rabbits. Swadlow reports for several cortical regions that neurones in lamina II/III and VI have firing rates between 0.1 and 1.0 impulses per second, while those in lamina V have higher rates, of the order of 5 impulses per second. These values may be overestimates, because the occurrence of spontaneous discharge is a major criterion for detecting single units, so silent neurones may escape detection. Far from 10/sec being the mean firing rate for cortical neurones (sect. 6.2), many cortical neurones projecting to other neurones in the cortex (i.e., those located in laminae II, III, and VI) may be essentially silent for seconds at a time.

(2) Estimates of conduction velocity in cortico-cortical axons give a wide range, from 0.1 m/sec up to 10 m/sec (Swadlow 1994 and other papers, for rabbit; Miller 1975, for cat). These estimates are subject to a variety of biases, all favoring rapidly conducting axons (see Miller 1994, in press). I believe that conduction velocities for cortico-cortical axons of the order of 0.3 m/sec are common and of far greater importance than is generally recognized. This value is far smaller than assumed by W&L.

(3) Long corticocortical axons have a patchy distribution, not one declining exponentially with distance. Even at the informal level this has important consequences for cortical dynamics and information processing (Schüz 1994).

(4) There is no empirical basis for suggesting (W&L's Table 1) that in human cortex long corticocortical connections are ten times more numerous than local ones.

(5) Nowhere do W&L specify their assumptions about the crucial empirical value of synaptic gain (i.e., number of presynaptic impulses required to produce each post-synaptic impulse), though they once refer to "aggregate gain" as "at unity or above." Abeles (e.g., 1982) has assumed a value of up to 30 for synaptic manipulations in general; and I have assumed a value of 3-4 for maximally strengthened synapses (Miller 1989, 1993).

With regard to comparison of simulation results with empirical data, I note that the change of EEC phase velocity with frequency is very much less in the simulation (Fig. 3) than in the empirical study in humans by Thatcher et al. (1986, Fig. 8). Also, the sharp transition to runaway excitation found in the simulation does not occur in most conditions in the real cortex (a fact noted and discussed by W&L).
I agree that a model of cortical dynamics dominated by excitatory-excitatory links is more realistic than that of Freeman, where inhibitory links are critical. Possibly the stability of the cortex is a less severe problem than has usually been assumed if one takes the very low firing rates reported by Swallow as one's premise. In this case, most synaptic activity on cortico-cortical neurones would be subthreshold, though it would still contribute to the EEC: linear superposition of neuronal events to form the EEC would occur, by interaction of extracellular potentials from many adjacent (but often largely independent) neurones.

The errors in initial premises seem to be large. The "relative insensitivity [of the model] to whatever errors are present in our anatomical estimates" (sect. 6.2) probably means that the reasoning which relates premises to conclusions is to say the least, imprecise. It hardly identifies a model with close quantitative relation to reality, and there may be important qualitative differences between the model and the realities of cortical dynamics.

Chaos in induced rhythms of the brain – the value of ERP studies

Márk Molnár
Institute for Psychology, Hungarian Academy of Sciences, Budapest, Hungary. molnar@oqgpsyphy.hu

Abstract: Event-related potentials (ERPs) – neglected almost entirely by Wright & Liley – allow objective investigation of information processing in the brain. The application of chaos theory to such an analysis broadens this possibility. Through the use of the point correlation dimension (PD2) accurate dimensional analysis of different Event-Related Potential components such as the P3 wave is possible.

Wright & Liley's (W&Ls) target article confronts two areas of electrophysiological research (neural network simulations and analysis of the EEG) while almost completely disregarding ERP (Event-Related Potential) studies. This, despite the fact that W&L refer in their Introduction to "cortical information processing" and "cognitive processes," two (closely related) notions for which ERPs, especially their late, "endogenous" components such as the P3 wave, have proved to be extremely informative (Donchin 1981).

Although the experimental circumstances in which these late ERP components appear are rather clear, less is known about the physiological processes they represent (Molnár 1994). New mathematical tools such as chaos theory may give a new momentum to the analysis of time series such as the EEG. However, there is still much controversy and doubt in this area. As Freeman (1995) put it: "Two things are immediately clear. First, psychologists have an opportunity and an obligation to explore this new offering from physics. Second, we do not agree on what it means or how it should be used." Nevertheless, as is generally agreed, calculation of the correlation-dimension (D2) of a time-series, such as the EEG, allows inferences to be drawn concerning the complexity of the underlying generators and is one of the most widely used methods of chaos theory (Skarda & Freeman 1987; Skinner et al. 1994).

Apart from studies of serious pathological conditions producing robust clinical symptoms, only in a few studies have the effects of cognitive "effort" on the D2 of the EEG been investigated (Gregson 1992; Lutzenberger et al. 1993; Rapp et al. 1989). D2 was found to increase as a consequence of cognitive "load." Novel stimuli, compared to when they are habituated, produce a similar effect (Skarda & Freeman 1987; Skinner et al. 1992).

As W&L point out (sect. 2.3), there are reasons to doubt the results of dimensional analysis of the EEG. We suggest that this should not necessarily be the case. D2 was calculated by the Grassberger-Procaccia (1983) algorithm in most of the studies performed so far. This algorithm assumes stationarity in the analyzed time-series, a condition which is rarely seen in biological systems. It also requires infinitely long data epochs for the analysis to be accurate. A new version (point correlation dimension, PD2) of the Grassberger-Procaccia algorithm was developed by Skinner et al. (1992, 1994) to overcome these difficulties. This new procedure is able to track changes of dimensionality in the EEG in time and is more accurate than others for estimating its dimensional complexity. After the calculation of the PD2 of each EEG epoch, including the ERPs, all of these are averaged and displayed. Thus it becomes possible to see dimensional changes associated with the occurrence of certain ERP components having more or less clearly defined psychophysiological "functions." In our studies, a significant decrease of the PD2 was found to be related to the occurrence of the P3 ERP component (Molnár & Skinner 1995, Molnár et al. 1995).

Our conclusion, based on these findings, was that the occurrence of the P3 ERP component was associated with a "context closure" event, as suggested by Verleger (1988) and not by "context updating," as proposed by Donchin (1981). Thus, we suggest that ERPs are still in the foreground of the repertoire of methods of neuroscience and that their dimensional analysis may help to solve long standing problems not accessible before.

Multiscale neocortical dynamics, experimental EEG measures, and global facilitation of local cell assemblies

Paul L. Nunez
Brain Physics Group, Department of Biomedical Engineering, Tulane University, New Orleans, LA 70118. pnunez@mailhost.tcs.tulane.edu

Abstract: Multiscale dynamics, linear approximations, global boundary conditions, experimental verification, and global influences on local cell assemblies are considered in the context of Wright & Liley's work. W&L provide a nice introduction to these issues and a reasonable simulation of intermediate scale dynamics, but the model does not adequately simulate combined local and global processes.

Why create neural models? Many artificial neural networks and brain dynamic theories seem more motivated by attraction to mathematics and computer simulations than to genuine interest in brain function. Thus, many papers use free parameters and avoid experimental connections. To Wright & Liley's (W&Ls) credit, they have emphasized quantitative physiology in their simulations. A principal value of physiologically based theory and simulation in the brain sciences is to guide experimental methods. Non-physiological models may provide interesting metaphors, but are unlikely to make a major impact. Without physiologically based models, experimental brain science may be driven by the gurus of current fashion.

Dynamics at multiple scales. The influence of small-scale dynamics on macroscopic dynamic behavior is considered in the seminal papers of Ingber (1995). Another issue concerns the matching of simulated variables with experimental data. For example, W&L provide simulations for human and cat EEG. But cat data are recorded from the cortical surface and most human data are recorded from the scalp. The former may involve tissue containing anything between one and perhaps 10⁶ neurons (depending on electrode size and location), whereas human scalp EEG is space averaged over regions containing between 10⁶ and 10⁷ neurons (independent of electrode size, but dependent on electrode density and computer method used to improve spatial resolution, Nunez et al. 1994; Nunez 1995). Realistic theory and simulations should distinguish between data recorded at different scales as well as from different species. More relative power at higher frequencies (15–30 Hz) is observed on the human cortex, consistent with the existence of wave dispersion relations that imply low-pass temporal filtering as a byproduct of low-pass spatial filtering by the head volume conductor (Nunez 1981; 1995;
Proceedings/Wright & Liley: Dynamics of the brain at global and microscopic scales

Pfurtscheller & Cooper (1975). Some cortical EEG recorded with small electrodes is dominated by incoherent activity, but only the coherent part can be recorded on the scalp. It is this coherent part that is most likely influenced (perhaps dominated) by global boundary conditions.

The excellent work of Wilson and Cowan (1973) is also relevant to scaling issues. They coarse-grained (space-averaged) small scale variables to make predictions at intermediate scales appropriate for comparison with some cortical EEG, but their variables are still too small-scale for direct comparison with human scalp EEG, which provides nearly all of its power at wavelengths longer than 5 or 10 cm. W&L simulations are unlikely to provide realistic predictions of human scalp EEG without the addition of many more elements or clever analytic methods so that predicted temporal frequencies match the experimentally determined spatial wavelength range.

Linear approximations. W&L's characterization of global EEG theory (Nunez 1995) as a "linear wave process" may be misleading. The theory consists of two linear integral equations (global) coupled to a third equation for local dynamics which is generally nonlinear and dependent on postsynaptic rise and decay times. I have not emphasized such local dynamic processes because of the necessary introduction of additional physiological parameters (which may be poorly known) and because others have published studies based only on local effects (Freeman 1975; van der Putt (et al. 1982; Wilson & Cowan 1973; Zhudin 1984). Various linear approximations to the third (local) equation have demonstrated predictive value in EEG. This approach is consistent with successful studies of genuine complex systems (as opposed to mathematical abstractions). That is, linear approximations are often used in science and engineering because they are useful, not because of any delusion that such systems are linear (a point sometimes not appreciated by mathematicians or physiologists). Recently, I expanded analytic action variables as linear combinations of global eigenfunctions (consistent with periodic boundary conditions), and showed that different spatial modes may exhibit separate limit cycle (or perhaps chaotic) behavior as limiting cases (Nunez 1995), thereby extending the global theory into the nonlinear range. Such modes may be considered "order parameters" in the parlance of modern dynamical theories (Friedrichs et al. 1991; Fuchs et al. 1992). This extension comes at the expense of introducing new parameters. W&L have added nice to quantitative parameter estimates, but substantial uncertainty remains, especially with regard to local feedback gains (Silberstein 1995). Most theoretical studies conclude that these gains strongly affect dynamics in apparent contrast to the W&L simulations.

Global boundary conditions. W&L find that global boundary conditions have minimal influence on dynamical. This is probably an artifact of their simulation model, which is based on interactions of perhaps 10^8 minicolumn-like units compared with 10^8 minicolumns in neocortex. The local-global theory (Nunez 1995) involves a local control parameter dependent on feedback gains in local circuits and a global parameter determined by corticocortical (long range) interactions. Some brain states are dominated by local properties, other states are dominated by global properties. The following EEG predictions of the theory have been largely verified: (1) Coherent spatial modes with frequencies in the 10 Hz range; (2) negative correlation between brain size and alpha rhythm frequency; (3) high spatial frequencies associated with high temporal frequencies (above about 10 Hz); (4) measurable phase velocities in the 5-10 m/sec range, matching corticocortical propagation velocities; (5) apparent nodal lines of standing waves observed in alpha rhythm and steady-state visual evoked potentials; (6) phase patterns (over the entire scalp) in visual evoked potentials showing a sensitive dependence on driving frequency (suggesting that nodal lines are not determined by fixed anatomical features like fissures and sulci); (7) coherence as a sensitive function of driving frequency. Many other properties of EEG dynamics cannot be predicted by this crude quasilinear theory. However, both theory and experiment suggest that global boundary conditions strongly influence several EEG states.

Top-down, multiscale, neocortical dynamic plasticity. The local-global theory suggests that multiple (perhaps closely spaced) global modes can drive local modes with matched frequencies (which vary with location), thereby facilitating a top-down mechanism for establishing coherent oscillations in widely separated cell groups with no direct interconnections (Greg et al. 1989; Nunez 1995; Silberstein 1995). Such matching of time scales in multispatial scale dynamics may have profound implications for information processing (Ingher 1983; Nunez 1989; Ingher & Nunez 1995). If so, one may further conjecture that neural plasticity allows this to occur in the developing brain so that, for example, time constants in local cell assemblies evolve partly under the influence of global boundary conditions (Nunez 1995; Silberstein 1995).

Dynamics of the brain – from the statistical properties of neural signals to the development of representations.

Andrew Olivar

Medical Research Council, Cognitive Development Unit, London WC1H 0BT, England.

Andrew@edu.uea.ac.uk

Abstract: The unification of microscopic and macroscopic models of brain behaviour is of paramount importance and Wright & Liley's target article provides some important groundwork. In this commentary, I propose that a useful approach for the future is to incorporate a developmental perspective into such models. This may be an important constraint, providing a key to understanding the nature of macroscopic measures of brain function such as functional measures like ERP.

The importance of the research programme discussed in Wright & Liley's (W&L) target article is beyond doubt. At present, much neural network research occurs in a vacuum, divorced from the fundamental problem of cognitive neuroscience: how the neural substrate of the brain gives rise to the complex behaviour of the organism. This includes not just the relatively "low level" behaviour expressed in the general statistical properties of the EEG signal, but the meaning of these signals as we attempt to infer them from functional experiments such as event related potential (ERP) studies and other functional brain imaging methods. At the extreme, this dissociation between neural network research and brain research leads to studies in which neural networks are designed to simulate behavioural results when the algorithms used by these systems are highly unlikely to operate in the real brain. Recently, the study of brain-like networks has become much more feasible, and W&L's article represents a significant advance in this approach.

This commentary raises some general considerations that need to be addressed in W&L's project. The next stage of the research program should attempt to address the meaning of the neural signals that are simulated. This involves considering the nature of representations in the brain and how these representations are manifested in EEG, ERP, and other functional recording methods. Considering brain development, the development of representations in the brain, and the general cognitive development of the infant may provide further important constraints on this kind of model, and a key to understanding how macroscopic recordings of brain activity may reflect representations in the brain and representational change.

A property that is clearly missing from W&L's work is a discussion of the learning properties of the models. This is reasonable up to a point if it is accepted that this has no effect on the gross statistical properties of the EEG signal, but it clearly becomes more important if one wishes to understand functional measures such as the ERP. However, much more interest to development...
talists, and I believe an important source of constraint on models that attempt to unify macroscopic and microscopic brain behaviour, is a thorough consideration of the developmental dynamics of the brain and its representations.

A developmental perspective is likely to be useful because changes in the EEG or ERP during the course of development are related to changes in the constants of the neural substrate at the microscopic level, such as neuronal, synaptic, and dendritic densities, neuronal firing thresholds, and the degree of myelination in different parts of the cortex. In addition to these gross changes, changes also occur in the representational complexity of the brain, and in particular the neocortex, as development proceeds. This leads one to consider the possible sources of these changes in representation, and how these may alter both the nature of the microstructure of the brain and the form of the EEG. Recent work has indicated that an interacting set of processes may induce these changes.

These developmental processes concern principally: (1) the endogenous architecture of the neocortex, especially the nature of the subcortical to cortical pathways such as the afferents from the thalamus to the sensory regions and (2) the nature of the exogenous informational environment that the developing cortex is exposed to. These two processes have been explored by several workers, for example in the pioneering work of Miller et al. (1989) on the development of ocularity in the visual cortex. (3) There is also a sequence of progressive changes of synaptic and dendritic growth, followed by a regressive process of retraction of neural connections, (4) processes of competition and cooperation between neurons and synapses to form stable representations, and (most controversially) (5) dynamic changes in the plasticity of the cortex (in terms of its ability to form representations) which take the form of a wave of plasticity passing over the cortex during development.

A recent model due to Kerszberg et al. (1992) expanded by Shrag and Johnson (1995) and Oliver et al. (1995) attempts to address these processes of development in a unified way. Thatcher (1992) and Case (1992) have examined the physiological and psychological evidence for the dynamic wave of neural plasticity. Thatcher has also attempted to unify these results with changes to the EEG during development.

To developmentalists, these processes are of considerable importance in their own right. For example, much effort is expended in attempting to solve the “inverse mapping problem” to localise the source of ERP components; recently this project has been expanded to attempt to observe changes in the ERP during development. However, such effects also seem likely to play an increasingly important role in constraining models that attempt to unify the macroscopic and microscopic properties of the neocortex. The developmental processes, namely, endogenous and exogenous factors, progressive and regressive processes, competition and cooperation, and changes to cortical plasticity, eventually need to be built into such models, especially as the models move from representing the gross signal properties of the EEG to attempting to account for the nature and meaning of these signals in terms of the representations they reflect.

The developmental perspective may serve to both constrain the models and extend their range of applicability, allowing them to consider the meaning of the EEG signals rather than just their overall statistical properties. Clearly, this is work for the future. Wright & Liley's target article lays some important groundwork, but in the long run, modellers ignore the developmental perspective at their peril!

Is there chaos in the brain?

Hubert Preissl, Werner Lutzenberger, and Friedemann Pulvermüller
Institut fuer Medizinische Psychologie und Verhaltensneurobiologie, Universitaet Tuebingen, Gartenstrasse 29, 72074 Tuebingen, Germany.
hubert.preissl@uni-tuebingen.de

Abstract: For some years there has been a controversy about whether brain state variables such as EEG or neuronal spike trains exhibit chaotic behaviour. Wright & Liley claim that the local dynamics measured by spike trains or local field potentials exhibit chaotic behaviour, but global measures like EEG should be governed by linear dynamics. We propose a different scheme. Based on simulation studies and various experiments, we suggest that the pointwise dimension of EEG time series may provide some valuable information about underlying neuronal generators.

In the eighties, it was shown that a special case of chaotic systems, so called “dissipative” chaotic systems, can be described using different measures, such as non-integer, so called fractal dimension, that is, correlation dimension, pointwise dimension and other dimension measures (Farmer 1982; Packard et al. 1980), or the Lyapunov exponents (Vastano & Kostelich 1986). It is obvious that these measures characterize different aspects of a single dynamical system. In this commentary changes in pointwise dimension are considered changes in complexity. This interpretation is based on the idea that a change in pointwise dimension measure can be interpreted as a change in the number of independent processes generating the signal. As Wright & Liley (W&L) point out, there are severe difficulties both in calculating these measures from time series and in interpreting them. In earlier work we reported some evidence that single spike train data are not useful for deciding whether the underlying brain process can be characterized by a chaotic dynamical system (Preissl & Aertsen 1992). However, one can still argue that dimensional analysis of more global measures, such as EEG time series, may provide some information about neuronal dynamics (Lutzenberger et al. 1995).

To connect dimensionality measures to neuronal processes, some assumptions are necessary.

A theoretical approach to brain function is based on Hebb’s concept of cell assemblies (Braitenberg 1978; Hebb 1949; Pulvermüller et al. 1994). See also Amit “The Hebbian Paradigm Revisited,” BBS 18(4) 1985.) According to this view, strongly coupled neuronal groups are the functional units of the brain. According to Hebb’s conjecture, these cell assemblies can be considered “closed systems,” that is, the exchange of information between them is much smaller than the information flow within each assembly. If cell assemblies are to some degree autonomous processing units, it should be possible that two or more are active at the same time without significant interaction between them. In experiments using visual stimulation it was found that distinct sets of neurons oscillated synchronously at different frequencies, while no synchrony was present between the sets (Eckhorn et al. 1988; Gray et al. 1989). This can easily be explained by assuming that each set of synchronously oscillating neurons belongs to a cell assembly, and that these assemblies do not strongly influence each other (see Aertsen & Arndt, 1993, and Singer 1994, for discussion). These data suggest that in the working brain there may be not only one or two, but much larger numbers of cell assemblies oscillating synchronously at different frequencies. In this case, the number of cell assemblies activated can be considered an indicator of the complexity of neuronal computations in the brain.

Given that the complexity of brain-internal computation can be measured in the number of activated cell assemblies, it is possible that measures such as pointwise dimension convey reliable information about this number. In this case, it would be possible to state the following: the number of simultaneously active assemblies can be measured with nonlinear dynamics. In a previous paper (Lutzenberger et al. 1992), it was shown that pointwise dimension measures obtained from EEG signals differ between experimental tasks. When subjects are looking at a double-pendulum which
The EEG data indicate stochastic nonlinearity

Walter S. Pritchard
Psychophysiology Laboratory, Bowman Gray Technical Center 611-12, R. J. Reynolds Tobacco Company, Winston-Salem, NC 27102. wally@rjt.com

Abstract: Wright & Liley contrast their theory that the global dynamics of the EEG are linear with that of Ghanayem, who hypothesizes an EEG governed by (nonlinear) deterministic-chaotic dynamics. A "call for further discussion" on the part of the authors is made as to how either theory fits with experimental findings indicating that EEG dynamics are nonlinear but stochastic.

Wright & Liley (W&L) have presented an ambitious and quite interesting model of the EEG. A detailed critique of their target article would require a careful reading (or, personally more accurately, multiple careful readings) of the numerous technically oriented references they cite, a task which I freely admit having undertaken only incompletely. So rather than a critique per se, I present what may be termed some "calls for further discussion."

W&L contrast two basic views of the EEG. The first is that of Freeman, who holds that cortical EEG, like the EEG of the olfactory bulb, ought to have underlying chaotic dynamics. The other view (Wright and colleagues; Nunez) is that, although at a microscopic level neuronal dynamics are highly nonlinear, at the macroscopic level of the EEG, global linear dynamics emerge/preval/etc.

My main call for further discussion is this: When dimension-estimation algorithms such as Grassberger-Procaccia (1983) have been applied to scalp-recorded human EEG in conjunction with surrogate-data testing (based on phase-angle randomization; Pijn et al. 1991; Theiler et al. 1992), no evidence of low-dimensional, deterministic chaos of the strange-attractor variety has been obtained (Palus 1994; Pritchard et al. 1995a; 1995b; Theiler et al. 1992). This is also true of deep-electrode recording of non cortical EEG activity in rats (Pijn et al. 1991); a similar finding was obtained for normal, resting human EEG by Glass et al. (1993) using the Kaplan and Glass (1992) state space vector-assignment test for determinism. However, a significant element of nonlinearity has been detected in normal, resting human EEG, even when the possibility of a nonlinear measurement function is controlled for using random-Gaussian surrogates (Palus 1994; Pritchard et al. 1995a; 1995b; Theiler et al. 1992). The question then becomes how to reconcile these experimental findings of "nonlinear stochasticity" with one view that predicts deterministic chaos and another that predicts stochastic linearity?

Assuming Freeman to be right, could cortical EEG be chaotic but of dimension higher than what can currently be resolved (or, alternatively, could it switch rapidly among multiple chaotic attractors), thus only appearing to be a nonlinear stochastic process? Or, considering W&L to be right, could the EEG be subject to constant "perturbations of sufficient magnitude" (target article, sect. 6.4, para. 5) to maintain the system in a nonlinear realm without going into a state of runaway excitation? Does the cortex somehow walk a fine line between linearity and "epilepsy"? The notion that epilepsy is some form of "runaway nonlinearity" would seem to fit with evidence that epileptic EEG is chaotic (e.g., Pijn et al. 1991).

Two other items I think need discussion include: (1) On what basis is input from the reticular activating system best modeled as white noise? Are there physiological data supporting this, or was white noise necessary for the model to work? (2) The model produces a prominent peak in the alpha frequency band of the power spectrum (Fig. 2) that is quite characteristic of normal human EEG recorded under resting, eyes-closed conditions. Yet it also produces prominent peaks at 20 and 40 Hz that are not present in the power spectra of most normal humans. Also, how does the model account for alpha blocking (virtual disappearance of the alpha peak) upon opening of the eyes? In Figure 2, one sees a diminution of the "alpha peak" with increasing values of the parameter ("recticular" noise input), but not a disappearance.

W&L are careful to point out the limitations of their model at the end of their article. I hope that their consideration of the questions I have raised leads to further refinements of this quite exciting beginning.
The form of chaos in the noisy brain can manifest function

Ichiro Tsuda
Department of Mathematics, Hokkaido University, Sapporo 060 Japan.
tsuda@math.hokudai.ac.jp

Abstract: I would like to emphasize the significance of chaotic dynamics at both local and macroscopic levels in the cortex. The basic notions dealt with in this commentary will be noise-induced order, chaotic "itinerancy" and dissipative structure. Wright & Liley's theory would be partially misleading, since emergent nonlinearity rather than the linearity at even a macroscopic level can actually subserve cortical functions.

Wright & Liley (W&L) attempt in their target article to provide a "unified theory" over a full range in cortical space, provided "nonlinear dynamics at local scale and emergent linearity of macroscopic electrocortical waves" (sect. 4) are present. They interpret chaotic dynamics in local networks and linear wave motions in global networks (the AN model) in order to make a bridge between these two levels. Asymmetric and symmetric ANN models are taken to be relevant to local and global dynamics respectively, based on the effective couplings and the presence or absence of noisy input from the retinoinformation or other cortical areas (sect. 8).

In this commentary, I would like to point out that W&L's theory could be misleading. The point is based on the fact that noise may destabilize chaos by creating some ordered motion rather than increasing the randomness, the latter giving rise to the limit theorem (Matsumoto & Tsuda 1983; Tsuda & Matsumoto 1984), and also to the fact that chaotic "itinerancy" (Ikeda et al. 1989; Kaneko 1990a; Tsuda 1991a; 1991b) has been observed in the experiment conducted to detect cortical activities over spatially extended areas (the macroscopic scale) responsible for olfactory perception and recognition (Kay et al. 1995).

W&L state that spatially and temporally noisy input can permit the linearity at the macroscopic scale; thus, withdrawal of such an input may permit the global chaos which can appear in globally coupled chaotic elements (Kaneko 1990b) (sect. 9). However, there are other kinds of noise effects in chaotic dynamics. Among other factors, noise-induced order (Matsumoto & Tsuda 1983; Tsuda & Matsumoto 1984) is here taken into account.

Noise-induced order was first found in one-dimensional chaotic maps which have a nonuniform Markov partition. This notion can be extended to higher-dimensional maps and vector fields. "Nonuniform" chaos occurs when the nonuniformity of an invariant measure in phase space can be destabilized by noise, forming ordered states. By applying (uniform) noise to nonuniform chaos, one can find some order, though it is a bit noisy. Noise-induced order is characterized by the appearance of a sharp peak in the power spectrum, the transition of the Lyapunov exponent from positive to negative, an abrupt decrease of Kolmogorov-Sinai entropy, and localization of orbits.

Nonuniform chaos is often seen in biological systems rather than physical ones. This is because chaos in biological systems can be accompanied by destabilization of excitation. Actually, chaos in neural systems can be an object of noise-induced order, since it shows nonuniformity of the invariant measure. A crucial question arises if we are concerned with the real cause of ordered motions observed in neural systems, since what is observed in experiments is always contaminated by noise. The question could be: How can one distinguish the ordered motions created by noise applied to nonuniform chaos from the noisy limit cycles? One possible method to answer this question is studying the principal peak frequency according to the change of the system's bifurcation parameter. If the ordered motion is that of noise-induced order, the principal peak frequency is observed to shift continuously by a systematic change of the parameter, whereas it remains constant if the ordered motion is a noisy limit cycle (Matsumoto & Tsuda 1983). One can thereby distinguish these two ordered states.

Kaneko's (1990b; 1992) unusual ordered state recovers the usual statistics at macroscopic scales, that is, the limit theorem when applying noise (sects. 2.4, 5, and 9). In this case global chaos is apparently of a different type from what creates ordered motions via interplay with noise. Taking into account the excitability of neural systems at every scale, it would be much more plausible that order emerges out of chaos with the help of noise in both local and global cortical dynamics rather than that global chaos is wiped out in the noisy environment to give rise to linearity. This notion is also consistent with the self-organization principle in macroscopic and dissipative systems (Nicolis & Prigogine 1979; Bublyanty & Laurence 1994). The emergence of dissipative structure demonstrates the presence not of the thermodynamic but the far-from-equilibrium branch at macroscopic scale.

There also exists the class of dynamic structure in a macroscopic system with large degrees of freedom that cannot be reduced to the usual stochastic process in its description. The notion of chaotic itinerancy proposed by us in this context has also been implicated as serving as a dynamic link in memory (Tsuda 1987; 1992), as a concurrent process of learning and retrieval (Tsuda 1994), a key process of hermeneutic states (Tsuda 1991b), input-driven perception and recognition (Freeman 1994; 1995a), and information transmission in wide cortical areas (Kay et al. 1995).

Assume a multi-attractor system, where an attractor can be a fixed point, limit cycle, torus, or chaos. This multi-attractor state can be unstable, triggered by a change in the effective system size or the input. A new stable state is represented by a global attractor which makes the original basins of attraction connect to each other. The orbit is attracted to one of the original attractors and remains there for some time, but departs along an unstable temporarily created manifold toward another original attractor. Then the overall behavior is itinerant among quasi-attractors, and history-dependent. It should also be noted that this behavior can be created by noisy synapses in neural networks (Tsuda 1992).

Cortical activities are, in general, "nonstationary," because of continually variable external and internal inputs. This nonstationary state may be characterized by some subset of the time series of chaotic itinerancy; thus the system loses the invariance of its activities with external input (Freeman 1995b). On the other hand, the system can retain invariance with internal input if chaotic itinerancy works, for it forms a global attractor. A global state in the cortex could therefore be still highly nonlinear, contrary to W&L's assertion.

Rhythmicity in the EEG and global stabilization of the average level of excitation in the cerebral cortex

M. N. Zhadin
Laboratory of Neurocibermatics, Institute of Cell Biophysics, 142292 Pushchino, Moscow Region, Russia. zhadin@mars.ibioc.serpukhov.ru

Abstract: The network model of EEG formation has revealed a unified mechanism for various rhythmic EEG phenomena: for various reactions as well as for ontogenetic and phylogenetic differences. EEG rhythm was shown to be an essential manifestation of the functioning of the intracortical stabilizing system which provides normal informational operations in the cerebral cortex.

Wright & Liley (W&L) touch on the mechanisms of EEG generation and its functional significance, on the relations between microscopic and macroscopic scales in cortical bioelectric activity as well as between linear and nonlinear processes in it and on stabilization of the excitation level in the cerebral cortex. This overall combination of problems has been the focus of our laboratory's attention for more than 20 years. As a consequence of a succession of theoretical and experimental studies, we can provide some supplementary information and unified interpretation re-
Authors’ Response

Multiscale modeling of brain dynamics depends upon approximations at each scale

J. J. Wright and D. T. J. Liley
Mental Health Research Institute, Parkville, Victoria 3052 and Swinburne Center for Applied Neuroscience, Hawthorne, Victoria 3122, Melbourne, Australia. jjw@ cortex.mhri.edu.au

Abstract: We outline fresh findings that show that our macroscopic electrocorticographic (ECoG) simulations can account for synchronous multunit pulse oscillations at separate, simultaneously activated cortical sites and the associated gamma-band ECoG activity. We clarify our views on the approximations of dynamic class applicable to neural events at macroscopic and microscopic scales, and the analogies drawn to classes of ANN behaviour. We accept the need to introduce memory processes and detailed anatomical and physiological information into any future developments of our simulations. On the issue of intrinsic cortical stability and the role of extrinsic fibre systems in maintaining stability, we argue that this position is not in extreme contradiction to those of our commentators, and that the mechanisms implicit in our simulations’ properties imply rich computational possibilities. We discuss some of the reasons for and against the existence of significant global resonances in the brain and explain why such behaviour appears absent in our simulations. Last, we discuss other phenomena, such as rhythmic driving of the cortex, which have not yet been introduced into our models, and indicate lines for future development of the simulations.

R0. Introduction

We thank all the commentators for their careful consideration. We will try to respond to the most important issues, noting the individual commentators under these headings.

R1. Do our models offer fresh insights — e.g., into synchronous oscillation at a distance?

Two of our commentators, Hardcastle and Erwin, question whether our use of simulations of electrocortical activity add anything useful, above and beyond the empirical information they purport to explain. We will set aside temporarily Harcastle’s doubts as to whether our work has anything to do with the brain’s cognitive properties, hoping to make our position on this a little clearer in the remainder of the commentary.

Hardcastle notes our theme of interaction between microscopic and macroscopic scales in the brain, and indicates that if we could explain how these interact, we would have achieved something of consequence. Erwin has similar concerns, posing a more precise challenge by suggesting that as part of our model validation we address the evidence for near-zero phase lag between oscillations of widely separated, simultaneously activated, neurons. Erwin cites Bressler et al. (1993) as his specific example. The body of work to which he is directing our attention has also been summarised recently by Singer (1994). As the synchrony-at-distance effect is seen in multiunit action potentials, and is associated with electrocortical potentials in the gamma band (around 40 Hz), this is indeed an instance of interrela-
tion of brain events at macroscopic and microscopic scales. Furthermore, this phenomenon appears important in cognitive terms – thus his challenge enables us to focus on Hardcastle’s comments as well.

By way of a very brief summary of Singer (1994): when separate stimuli are delivered simultaneously to the cortex via sensory inputs, multuniunit action potentials at or near the sites of input exhibit synchronous oscillation at zero phase lag. This effect is variably associated with local field potential activity in the gamma band, that is, at about 20 to 80 Hz. Singer and others have commented on the possible significance of the phenomenon with regard to feature-binding. Simultaneously excited networks, even though spatially separate and differently stimulated, might thus, under a suitable learning rule, form basins of attraction that reflect the spatial and temporal contiguity of their inputs.

We reproduced conditions analogous to those in Singer’s experiments, using our macroscopic lumped simulation, in the variant developed for the cat cortex (Wright & Liley 1995). To mimic the input of simultaneous specific sensory inputs to cortical sites, we delivered a relatively high level of excitatory bias mixed with zero-mean noise to two of the unit cell volumes on the simulated cortical surface. In the critical case, the noise components to each site were white and uncorrelated with each other, although we also studied cases with a degree of correlation between the two inputs. All other unit cell volumes received only a small excitatory biasing tone; in later experiments we also added low levels of asynchronous noise to all other sites. We varied the separation of the two sites of strong input – from adjacent to each other to as far apart as the toroidal form of the simulated cortical surface permitted. The essential results are shown in Figure R1.

When the noise inputs to the two excited sites were completely uncorrelated, the simulated pulse density (Qe, in our terminology, analogous to multuniunit action potential frequency) from the sites of input showed only irregular levels of correlation with each other. But Qe at all sites adjacent to the sites of input exhibited highly significant correlations at zero phase lag once the system had reached steady state. The correlation coefficient at zero lag decreased with the separation distance between the sites under study, and with the relative amplitude of noise delivered to the less-excited sites.

When the level of excitatory tone delivered to the excited sites was sufficiently high, the zero-lag correlation of pulse density was accompanied by a shift of the simulated ECoG spectrum from the 1/f form described in our target article, to gamma-band activity.

The above results will be reported in detail elsewhere. The broader issues of interaction across scale raised by Hardeastle and Erwin will be further addressed below, in relation to Ingber’s commentary.

R2. Is our discussion of linearity, nonlinearity, dynamic equilibria, and so forth misplaced or too crudely approximate?

A number of commentators take us to task in various ways about our application of terms widely used and misused in dynamics. Perhaps this is most explicit in the commentaries of Erdi and of Kowalik et al. Erdi rightly identifies the set of dichotomies of concern to all who formulate models of cerebral dynamics, and Kowalik et al. point out a number of ways in which our formulation of linear and nonlinear domains and scales could be criticised. We agree entirely with these comments, but feel here that perhaps our commentators have missed our point somewhat. We do not intend to say that the brain is an absolutely linear, equilibrium system at macroscopic scale – merely that it is approximately so, despite its extremely nonlinear properties at microscopic scale. We wished to emphasise, focusing on the works of Freeman and Nunez as strong examples, that these major bodies of work are not necessarily in conflict: each are relevant for the scale at which they have been applied. Our own earlier work, cited in the target article, attempted to demonstrate experimentally that (macroscopic) ECoG waves are indeed almost linear, and almost at equilibrium. Pritchard raises a related point when he indicates that some measures based on EEG time-series or equivalent measures reveal significant nonlinearity. We
hope it is clear that this kind of nonlinearity is different from the linearity of waves in the sense that the moving waves exhibit near obedience to the superposition principle. We expect the ongoing perturbations produced by interactions across scale to produce nonlinearities of the type to which Pritchard refers.

As Erdi and Kowalik et al. recognise, scale-dependent dynamics occur elsewhere in physics – e.g., the chaotic motion of air molecules, the linearity of sound waves, and the chaotic motions of the weather – but this particular transformation in approximate dynamic type as a function of scale is not logically inevitable. The important point for us is that the particular dynamic transformations with scale greatly constrain the classes of models that might be developed.

R3. Are ANN properties comparable to brain dynamics? The importance of learning rules

Amit concentrates on our possible misreadings of the ANN literature [See also Amit: “The Hebbian Paradigm Reintegrated” BBS 18(4) 1995.]

There are four substantive issues:

1. The local ANN program (i.e., the work of Amit and colleagues) is concerned with attractor storage, not the need to reproduce brainlike operations. We understand this. Our intent was to show that the putative classes of dynamic events in the brain show analogies with ANNs that are already relatively well developed, namely asymmetric networks with variable threshold, integrating and firing elements, and ANNs with fast and slow synapses – particularly those with capacities for recall in temporal sequence. This encourages the belief that the emerging models of cerebral dynamics may turn out to be models of cognitive mechanism without their construction having been forced by abstract, purely cognitive considerations. The comparisons are as yet imprecise, and are rather confounded with other issues of dynamic behaviour, as reflected at different scales.

2. The issue of global and local stability in networks with local inhibition and long-range excitatory couplings is not as clear-cut as we have made it seem; global stability is possible without long-range inhibition. We will respond to this crucial criticism below, in connection with Freeman’s commentary.

3. Learning as expressed in the synaptic matrix is of central explanatory importance (for cognition). We agree. However, the issue with which we are concerned – the particular nature of cerebral dynamics – requires resolution in its own right. It is our working assumption that a full account of dynamics would only require the appropriate physiological principles (e.g., LTP) to be introduced, to produce a much more complete account of cognition. This must be the case, unless it is disputed that very rapid synaptic adaptations change cerebral dynamic properties on the same time-scale(s) as are relevant to any type of cognition. This controversy is open at present, but our working assumption seems to be at least worthy of testing until it fails.

4. Lumped models contribute little or nothing over cellular models. With this we disagree, on grounds of practicality. An ultimate simulation might have every neuron in the brain specifically modelled, but this is not practical. Meantime, it appears that many of the brain’s properties may be understood only by considering events at both local and global scales, and possibly at intermediate scales too (see the commentaries of Ingher and Nunez). Lumped simulations seem the only way to approximate events at meso and macroscopic scales for now.

R4. Is a general dynamical account of brain activity ever likely to be a model of cognition without extremely detailed specifics of connectivity? The developmental perspective

Oliver encourages us to include a developmental perspective in later work and to use the ERP as a measure of subtler aspects of cognition. By this he is saying, in part, that later models need to be much more realistic than our very limited sketches, which thus far address some basic principles only. We agree.

Koerner is more explicitly critical along similar lines. He is unhappy with a central assumption – “time-varying parameters are stochastically independent in the large” – that enabled us to rationalize the emergence of near-equilibrium, near-linear, global wave phenomena. We feel he has not noticed that this assumption can sometimes be valid when applied to very complicated phenomena, even though these phenomena are actually deterministic. Quoting Minsky (1991), he emphasises that the brain is immensely complex and hierarchically organised, and indicates that little can be gained without such detail. Other aspects of his commentary are discussed below.

We feel that the controversy over the importance of specific connectivities and developmental detail reflects only different groups’ choice of starting point. Does one come to an understanding of a petrol engine by thinking of all the components separately or by grasping the principle of the four-stroke cycle? Ultimately, only by doing both.

Our simulations appear to imitate some general dynamic aspects of the organisation of real brains. They, we hope, reflect a class of systems from which evolution has drawn in selecting cognitive engines that actually interact with the real world in survival-favouring ways. Perhaps they can be made more brainlike by the progressive addition of detail.

It is not clear just how exact this added detail would need to be. In some parts of the brain, specifics of connectivity are less important than elsewhere. The rapid evolutionary expansion of the cerebral cortex in hominids and the immense plasticity of cortical function following damage both indicate that relatively general principles have been exploited in the cortex. This may be contrasted with the exact and slowly evolved pathways that dominate much of the anatomy of the brainstem.

R5. Do our simple models at different scale deal at all adequately with the effects of scale upon dynamics? What is the appropriate total dynamic description?

These questions are primary for Ingher and of relevance to the commentaries of Nunez, Erwin, Hardcastle, Amit, Freeman, and Liljenstrom.

The problem raised is a more powerful version of that addressed in section R2. The stronger form follows from considerations addressed in Ingher’s SMNI. This theoreti-
cal approach uses methods for describing events in very complex systems (neural events specifically) and relates these events to observable properties at many scales and over different durations. Models of this form may well prove ultimately to yield all-embracing descriptions of cerebral dynamics. We believe Ingber has rightly perceived our more limited intent, however.

In contrast to the full formalism of SMNI, we have broken cerebral events into two spatial scales and concerned ourself with events over about a second, as observed in the power spectrum. We have made no attempt to formulate a model that could account for the state-space trajectory of the brain, even within this limited epoch. Supposing a more definitive form of our simulations, integrating events at macroscopic and microscopic scales, were formulated, then indeed its state trajectory would not be subject to linear decomposition over any usefully prolonged epoch, as Ingber points out. Our macroscopic simulations are linear only in the sense that wave motions almost conform to an instantaneous superposition principle.

Since our treatment in the target article was non-mathematical, we did not spell out a further property of the macroscopic model's dynamic equations; this is relevant to Ingber's concern to develop methods bridging events at many scales. Their form is in fact independent of scale. If applied at the microscopic level, the sigmoid nonlinearity would become a step function, and the coupling and other parameters would be transformed by scale — but a matrix of coupled elements analogous to those in the Genesis simulations emerges at the finest level of resolution. When these equations are applied in a lumped fashion, information on individual connectivities is lost. Thus, the macroscopic simulation could never accurately describe system trajectories over prolonged durations. Conversely, the microscopic simulation is not of sufficient scale (in practice) to be considered autonomous. The importance of the near-linear superposition property of macroscopic EEG waves, for us, lay precisely in the way these state-equations could be written at different levels of spatial resolution.

Since conceptions of linearity and nonlinearity and of near- and far-from-equilibrium dynamics can be applied to our models only approximately and at the relevant scales of time and space, the defects in definition have, in effect, been swept under the carpet of interaction between scales. In our view, interaction between scales is intimately associated with stabilisation of cortical function. Hence the next question:

**R6. Is the cerebral cortex self-stabilising?**

This crucial issue is most specifically raised by Freeman, with related issues addressed by Amit and Liljenstrom. It may be useful to restate our position as strongly as possible.

(a) To be physiologically valid, our simulations must combine a low average pulse density (low firing rate) with the full range of spectral forms seen in the waking EEG. In our macroscopic simulation, these two essential features cannot be combined without imposing a limit on the stable maintenance of a low average pulse density; above this limit the simulation enters a global state of near-maximal pulse density — that is, all cells represented in the unit volumes are operating at high firing rates. The transitional pulse density, or metastable point, between these two conditions is regulated by the choice of value for the synaptic gain g; conversely, g is dictated by the choice of average pulse density at which runaway excitation will ensue.

(b) The microscopic simulations, which have been parameterised from physiological measures that provide the synaptic gain independently, reflect a similar state of affairs, that is, a metastable point between high and low firing-rate attractors, with limit-cycle behaviours at high firing rates and irregular activity at low firing rates.

We conclude that our simulations' basic principles are correct, but incomplete, that is, cerebral dynamics involve interactions across scale that depend on sharp, localised transitions across the metastable bound, with limit-cycle local attractors active during the excited states. This holds irrespective of the configuration, columnar or otherwise, chosen as the local unit network, so long as this unit local network is sufficiently large for sustained self-excitation to develop locally. If so, slower-acting processes we have not modelled must maintain longer-term cortical activity in the low firing-rate range.

If transitions about a metastable bound provide an accurate representation of short-term dynamics in the cortex, this in turn implies that cerebral dynamics are a special case of the now fashionable edge of chaos dynamics seen in class IV cellular automata (Langton et al. 1992). (It also incidentally explains why normal brains can be provoked into epilepsy by stimuli as feeble as stroboscopic visual EEG feedback.)

The alternatives to considering local dynamics metastable in the short term appear to be only that (1) the cerebral cortex requires continuous net excitatory tone to remain in the activated state — in which case it is difficult to understand why so much neuromodulator input to cortex is inhibitory — or (2) the cerebral cortex is capable of completely autonomous self-stabilising regulation, which can sustain activity in the absence of exogenous tone of any sort. This is ruled out by the effects of subcortical lesions or cerebral undercutting as mentioned by Freeman in citing Burns (1958).

The controversy is hence not an absolute one, but concerns the degree to which cortical stability depends on intrinsic versus extrinsic mechanisms, and the time scales over which these operate.

Freeman's disagreement with us begins from a position of great strength. His analyses of olfactory function are among the most original, wide-ranging, and experimentally rigorous contributions to modern neurophysiology. Our simulations lean heavily upon his findings. Freeman's earlier work using piece-wise linearisation methods is close to the body of findings and arguments we have ourselves advanced to indicate that EEG waves are almost linear waves. We have recently shown that our cellular (microscopic) simulations give rise to normalised conditional pulse-probability distributions sharing all qualitative features of Freeman's theoretical and experimental findings except the upper asymptote of the asymmetric sigmoid function — thus further emphasising both the correspondence and the difference between the models.

Our sole disagreement concerns the mechanism by which the upper asymptote of Freeman's asymmetric sigmoid function is established or, more precisely, how it is established in the cerebral neocortex. Although to the best of our knowledge the relevant physiological estimates of the asymmetric sigmoid for neocortex have not been made, it is clear that an upper bound must exist, since the neocortex always operates at a low average firing rate.
Response/Wright & Liley: Dynamics of the brain at global and microscopic scales

The mechanism limiting the average firing rate is not yet known, however, and Freeman's theoretical arguments for the form of his asymmetric function are not exhaustive. Possibilities include the cortical/subcortical interactions we hypothesise; dynamic considerations of relative excitatory and inhibitory tone generated within the cortex as Amit suggests; actions of cells such as the Martinotti cells which have not yet been specifically included in global simulations; dynamic changes in the efficacy of inhibitory and excitatory synapses; dynamic changes in dendritic responses; and doubtless others.

Considering the absence of clear experimental grounds for distinguishing among these many possible mechanisms, we know of no way to fully resolve the mechanisms of cortical stabilisation currently. Freeman's reference to Burns's (1958) findings is not a strong counterargument to our suggestions. Our simulations also go to a point attractor when isolated from input, unless prior stimulation was so strong that all cells are firing, in which case they remain firing. Our simulations, however, lack not only mechanisms of local adaptation, but also a mechanism equivalent to cellular metabolic exhaustion — either of which would bring the activity back to the electrical silence Burns observed.

Regarding the Amit and Brunel (1995) findings, cited by Amit: This proposal is very interesting and perhaps provides a formulation of some of the slower time-course effects absent from our present simulations. However, it is not known whether such a model would be capable of generating a realistic range of EEG-like behaviours while retaining a realistically low average firing rate. For us, this is the crucial issue in establishing whether the cortex is fully self-stabilising, or poised constantly on the verge of epilepsy.

Liljenström contributes his experience from modelling of the olfactory bulb. His findings pertain to ours and are similar in many respects. He gives excellent reasons why neuromodulators such as noradrenaline may indeed contribute to stabilisation from exogenous sources but still argues that intrinsic activity (in the olfactory system) is self-stabilising. Inhibitory cells appear to be more prevalent in the bulb than in neocortex, however, which may partially account for the difference in local stability demonstrable at this site.

R7. Are resonant modes of importance?

The systematic development of linearised models by Nunez and colleagues is a major conceptual source from which we have drawn. Thus we agree with the commentary of Nunez in large part. The simulations presented are for electrocorticogram and local field potential, and an extension of the simulation to scalp recordings would have to include a consideration of the signal transformation produced by the cranial structures. Our macroscopic simulation is a global electrocortical model, however, utilising a reasonably realistic scale and a closed geometry. The absence of obvious global resonant modes therefore appears to challenge Nunez's principal writings, even if it is allowed that transcranial filtering is likely to pass standing wave activity selectively, if this is present. We know of no presently available and definitive test that could determine whether global modes actually occur in the brain, although in the qualitative studies Nunez mentions the evidence for such occurrence is highly plausible.

In view of the conceptual similarity of our simulation to Nunez's linearised analytical results and the apparent ability of both approaches to account for fairly similar data (EEG spectral form, wave velocities, etc.), it is worth asking why our simulation does not exhibit global resonances. The simplest reason seems to be that in our simulation damping is simply too high to permit sustained oscillation if the system is driven from a localised source. That is, wave activity does not circle the closed geometry with sufficient residual amplitude, when meeting waves travelling in the reverse direction, to generate a distinguishable standing wave. Standing waves might arise if the simulation were driven continuously over the entire surface by diffuse input signals. We have not explored this possibility, and it cannot be ruled out as unphysiological — indeed, the driving of global wave processes by the transiently excited local network processes we hypothesise might reasonably constitute such a source of diffuse driving.

The simulation damping factors are high largely because of the simple rise-and-fall character of dendritic response to unit input, as we have modelled. While our representation of this process is but a crude approximation of dendritic dynamics as revealed with much greater precision in Freeman's findings, it is unlikely that more precise modelling would influence damping much. This is because the rise-and-fall character, without pronounced overswing beyond baseline, resembles the behaviour of an overdamped, or critically damped, second-order oscillator.

Resonant activity might emerge in our simulation with more complicated regulation of local excitatory and inhibitory tone. Some work we are pursuing in parallel with the simulation studies — utilising linearised equations and methods similar to those applied by Nunez and van Roterdam et al. (1982) but with couplings and other parameters similar to our simulations — suggests this (Liley 1995).

There is some reason to doubt, however, that global resonant modes are a prominent feature of EEG or ECoG activity much of the time. The strongest reason for this doubt is the striking similarity of ECoG spectra in animals of differing size. In our own laboratory, human EEG, sheep foetal ECoG, rat ECoG, and cat ECoG might readily be mistaken for each other on the basis of random spectral samples. This is hardly to be expected if the eigenfunctions of the resonant modes, and thus the spectral peaks, depend upon head size.

Finally, when estimates of damping coefficients are extracted from ECoG using AR techniques, these tend to agree with our simulation's properties, yielding overdamped values (Wright et al. 1990a).

R8. Are our models adequately parameterised?

Miller is concerned that we might be building models that lack fidelity. The point of our work is to find state equations for brain dynamic events in the simplest possible form, and to parameterise these equations as accurately as possible. We thank Miller for his review of physiological parameters relevant to the modification of our models, but assure him that we have not neglected such considerations (although they are not discussed in the target article). Of crucial importance here is the sensitivity of our simulation results to changes in the value of the structural parameters. We will respond to the numbered points in Miller's commentary:
Response/Wright & Liley: Dynamics of the brain at global and microscopic scales

(1) The difference in estimated mean firing rates to which Miller refers does not matter to simulation properties: the point is, that rates are very low compared with maximal rates.

(2) Wave velocity in the simulation is sensitive principally to the longest-range and fastest fibres, when fibre range and velocity parameters are manipulated independently. This has also been shown by Nunez in his general analytic work. The 9 m/sec axonal velocity estimate we have used (following Nunez’s arguments for this) thus seems a reasonable start.

(3) The simple geometry of corticocortical connections distributed in Gaussian fashion with distance and the toroidal closed form were again chosen for simplicity. The whole brain scale and range of the fibres did match real anatomy reasonably accurately. Experiments perturbing the distribution of corticocortical connection density over the full range of projection, and distortion of the same projections into elliptical rather than circular patterns about the origin, has surprisingly little effect on the reported findings. This is not to say that the details of connectivity are not functionally of vital importance, nor that they would not affect the spatiotemporal patterns of electrophysiological activity. But these details are just not very important to the aspects of the dynamics with which we are primarily concerned. An exception may be the gradient of phase-velocity with frequency referred to by Miller.

(4) We refer Miller to the basis of our earlier calculations of synaptic connectivity (Liley & Wright 1994, Wright & Liley 1995). The high corticocortical to intracortical ratio of excitatory synaptic density we used in the human simulations is less marked in our simulations of cat ECoG, yet the spectral form of results is less affected. Velocities do not depend on this aspect of connectivity.

(5) Estimating the synaptic gain is indeed problematic. We hope our discussion in section R6 above has been of some help. It would be a mistake, incidentally, to confuse the units of synaptic gain we used with those applied in other situations.

Erwin too mentions that our methods of calculation of couplings and so forth were based on very approximate assumptions. We agree, but believe these to be reasonable over large populations of cells. More detail is given in the original paper (Liley & Wright 1994).

It would be incorrect to conclude that our simulation would exhibit much the same properties whatever parameter values were applied: A much lower synaptic gain would result in a stable model with highly nonlinear waves, and the desynchronised ECoG would be associated with much higher average pulse densities than is in fact the case. Corticocortical connections of much lower velocity and shorter range would produce much slower wave motions and so on.

Thus, our claim is only that we have utilised parameter values within a physiologically and anatomically plausible domain, thus generating a close enough fit to data. We hope that much more detail and precision can be subsequently embodied within this frame. Indeed, much more precise and relevant estimates of the parameters for models of brain dynamics is urgently needed in the field.

R9. Alternative approaches

The commentaries of Zhadin and Tsuda are both very interesting. We could have begun our article from the dichotomy posed by the results from these two workers and their colleagues, rather than from those of Nunez’s and Freeman’s groups. This is because Zhadin’s model (with which we were not familiar) is one of mass action, implicitly assuming the applicability of the central limit theorem and the law of large numbers to the lumped dynamics of large cell masses. Tsuda’s work, in contrast, focuses on aspects of linked chaotic subprocesses in conditions when the input of noise does not lead to the emergence of near-equilibrium, near-linear wave states. Translating to our parlance, Tsuda is here describing conditions in which strong stochastic dependencies appear among the instantaneous values of noisy parameters describing the myriad of individual coupling events among neurons.

We note the similarity of Zhadin’s prior findings to our own. Although the models are not identical, they are close in a number of ways – both in principle and in reproducing the rightward shift of simulated ECoG power with increasing excitatory tone. The decrease in total spectral power associated with the shift – a property of Zhadin’s model – is not specifically reproduced in our own.

It is not possible to say with any certainty how our findings and those of Tsuda may be reconciled, but they do not seem to us to be necessarily in contradiction. Conclusions regarding macroscopic waves being approximately linear, deduced from our stochastic approach (which treated all instantaneous coupling parameters as stochastically independent), would remain robust in the face of the emergence of highly interdependent transient and localised events; that is, globally, waves would remain near-linear and near-equilibrium, but the transient local events would appear as an event perturbing the global system trajectory. The local events might include “noise-induced destabilisations of nonuniform chaos,” in which case Tsuda’s concepts of “chaotic itineracy” might be applied to describe local network dynamics and hence to interactions across scale.

R10. A role for rhythmic driving

We agree with Koerner that rhythmic controls ought to be important in overall control of cerebral dynamics, particularly in triggering local activity into transiently excited patterns as a wave peak passes a given locale, and perhaps later in suppressing activity within the locale as the wave trough arrives. Rather than contradicting this view of events, our macroscopic simulations may be a step toward explaining how the cortical waves arise; but, as explained below, we have not explained the details of origin of the cerebral rhythmicities. These will depend in part upon rhythmic input from subcortical structures (e.g., Steriade et al. 1990).

Pritchard asks why the simulated ECoG power spectrum is peaked at particular values and whether the noise we used to drive the simulation has to be white. To eliminate the last point: no, the noise need not be white, but it is not clear what colour it should be, other than white. Further analysis of the properties of the macroscopic simulation shows that the details of the spectral peaks at fine frequency resolution follow variation in the spectrum of the input noise, as would be expected in an approximated linear system. Variation of the level of nonspecific excitatory tone changes the system transfer function rather dramatically, so that at low levels of activation the lowest frequencies are
augmented, and, as excitatory tone increases, the distribution of power shifts to the right (as our illustrations show), until the extreme case shown in Figure R1 is reached. The "shift to the right" is not uniform with increasing non-specific activation; over a large part of the range of activation activity in the alpha band appears favoured. (There are further complications, which are not immediately relevant.) Consequently, it should be noted that our illustrations show the effects of varying only the average excitatory tone, not the particular input noise sequence. However, when the noise sequence is changed, equally realistic spectra, similar to short-epoch spectra of real ECoG, are reproduced. Ensemble averages (which we have not shown here) exhibit blurred spectra, following the envelope of the short-epoch individual spectra, thus replicating the spectral average of many separate short epochs of real ECoG.

Thus, our models do not account for subtle variations in combinations of activity at different frequencies, nor for alpha blocking in response to specifically visual input. The most obvious way to improve on this situation is to model inputs from the special sensory pathways and other subcortical systems as nonwhite, as doubtless they are in reality. This we have not attempted; a complete model would be very difficult to formulate.

R11. What range of EEG phenomena can we expect to account for? Coherence studies and the ERP

Although we are quite unable to offer much explication of EEG signal-analysis measures in situations of any degree of realistic complexity, we agree that all such phenomena must ultimately be accounted for by a really adequate model or simulation of the EEG. Bullock asks whether the distribution of coherence is a test of the model; he goes on to describe some interesting results. Coherence is a difficult measure to interpret unambiguously. We have used coherence to measure wave-velocities and to check for the near linearity of electrocortical waves in one of our experiments (Wright & Sergejev 1991). Such an approach made no allowance for confounding effects such as the development of synchrony at a distance, when several local cortical sites would be strongly activated. The local transitions about a metastable point that we hypothesise to underlie interactions across scale are bound to produce complex interacting patterns of activity, which ought to be coherent over variable distances; we are unable to be more specific.

On a similar theme, it has been rather a disappointment that frequency/wave number methods have not yielded convincing evidence of short-wavelength activity (such as might account for varying local coherences) in either real or simulated ECoG. The results shown in Figure 4 of the target article, while confirming correspondence of real and simulated results, may be accounted for by spectral leakage of long-wavelength waves, even though we took all proper steps to minimise this type of interference. Thus, we can only say that mesoscopic wave phenomena do provide the ground upon which the testing of all dynamic models must ultimately rest, but we have not gotten there yet. Needless to say, we are even more silent upon the problem posed in Bullock's last sentence.

Molnár points out that we have almost completely neglected the evoked potential. We echo Freeman (1995b), as quoted by Molnar. For what it is worth, we believe the evoked potential is an Impulse Response of the cortical system - but neither the impulse nor the system transfer function operating upon the impulse is well defined or even time-invariant. Analysis of the impulse, no matter by what means, is therefore problematic, and our present work makes little contribution to this problem.

Similarly, we can make little response to the commentary of Preissl et al., since we have not attempted to relate estimates of signal dimension to variation of state within our simulations. In view of their remarks on synchronously oscillating pools of neurons as Hebbian cell assemblies, we hope that the experiment on simulation of synchrony-at-a-distance outlined above will be of interest to them.

R12. Conclusion

A future path of development for our simulations must include their integration with each other. We envisage a simulation in which the global dynamics are represented under mass-action assumptions in part of the simulation, and a locale of individual cells, perhaps as large as a minicolumn, is represented as embedded within the global activity. The global activity might then be represented as in feedback interaction with extracortical neural structures, so that a sort of trinity of interactive processes might be concurrently simulated.

This scheme, which may be grandiose, seems nonetheless to be the necessary minimum, if events at all relevant scales in the brain are to be considered at once. Such a multilevel simulation might then be used as a testbed for the progressive introduction of increasingly realistic neurophysiological and neuroanatomical detail, and results could then be matched against experimental findings in real brains, in an iterative fashion. During this process, one hopes that answers to some of the many major questions raised in the commentaries will emerge.

References

Letters "s" and "r" appearing before author's initials refer to target article and response respectively.


250 x WORK TOWARD A THEORY OF BRAIN FUNCTION


References/Wright & Lily: Dynamics of the brain at global and microscopic scales


EEG simulation: variation of spectral envelope, pulse synchrony and \( \approx 40 \) Hz oscillation

J. J. Wright

Mental Health Research Institute of Victoria, Parkville, Victoria 3052, Australia

Received: 25 June 1996 / Accepted in revised form: 29 November 1996

Abstract. Macroscopic EEG travelling wave phenomena and cortical pulse synchronisation effects are related within a single simple simulation. Non-specific activation acts to control the transfer function of the simulated cortex, and thus determines the relative amplitude of macroscopic EEG waves generated by rhythmic inputs. When concurrent asynchronous excitatory inputs to separate, local, cortical sites are introduced, the simulation reproduces both gamma-band (40 Hz) electrocorticogram (ECoG) activity and synchronous oscillation of action potential pulse density at the separate sites. The gamma-band ECoG and pulse synchrony effects depend on different mechanisms: the former upon local excitatory/inhibitory interactions, and the latter on cortico-cortical interactions. The pattern of synchronous activity depends upon both structural and dynamic aspects of gain, and is sustained by linearised versions of the simulation’s state equations. Dynamic properties of the simulation, which are independent of scale, describe both microscopic and macroscopic phenomena, all in accord with physiological findings.

1 Introduction

1.1 ECoG elementary properties

The dendrites of cortical pyramidal neurones generate local field potentials (LFP) during processing of cognitive information (John et al. 1969; Gevins et al. 1983; Mitzdorf 1988; Picton and Hillyard 1988). Summed at a point on the cortical surface these LFP form the electrocorticogram (ECoG), and when recorded via the scalp, the electroencephalogram (EEG). The terms EEG, ECoG and LFP can be treated as interchangeable if the relative degree of spatial and temporal filtering of the underlying cortical events is kept in mind. The fluctuating surface potentials appear, in general, to be associated with, and arise from, travelling waves of local potential, which thus reflect waves of mean local dendritic potential (Thatcher et al. 1986; Burkitt 1996).

In the EEG certain spectral characteristics are commonly observed (e.g. Walter et al. 1967). These include the alpha, beta and gamma rhythms, and ‘desynchronisation’, which appears at high levels of alertness. Desynchronised ECoG typically yields a power spectrum without prominent peaks, loosely describable as a ‘\( 1/f \)’ spectrum (Freeman 1991).

The roles played by cortical and subcortical systems in generating and regulating the EEG are partially understood. Both slow-time-scale diffuse regulation of excitatory and inhibitory tone to the cortex, and the injection of rhythmic activities are involved. Inherently rhythmic neurones and resonances within the cerebrum may also play a part (Steriade et al. 1990).

1.2 LFP and action potential relations

Recently (Eckhorn et al. 1988; Gray and Singer 1989; Gray et al. 1989; Singer 1994) it has been shown that multi-unit action potentials at simultaneously stimulated sites in the cortex and elsewhere in the brain can exhibit synchronous fluctuations of pulse density. Typically these results are obtained when retinal sites are activated selectively by moving bars. Concurrently, the LFP in the neighbourhood of the synchronous pulse activity often, but not invariably, oscillates in the gamma band (around 40 Hz, varying from about 30 to 80 Hz). These findings have been recently reviewed by Singer and Gray (1995). They show that synchrony between concurrently active cortical sites is typically at zero time lag, and is manifested at several scales: intracolumnar, intercolumnar and interareal. Both structural connectivity (e.g. by cortico-cortical fibres crossing the corpus callosum) and functional dynamic state (e.g. cortical visual areas respectively specific for colour and form, responding to a particular colour/form combination) determine the regions which enter synchrony with each other. In most instances these results cannot be explained by synchronous concurrent inputs to the separate sites.
Although synchrony appears in most instances at zero lag, lag relations between stimulated sites are somewhat more complicated. For instance, cortical neurones driven by optimal stimuli appear to lead other neurones in the neighbourhood which are suboptimally stimulated, by a few milliseconds (Konig et al. 1995). It appears possible that synchrony may play a part in spatiotemporal association and segmentation of the perceptual field, and thus resolve the ‘binding’ problem (Singer and Gray 1995).

No simulation or other general explanation has been advanced to relate macroscopic EEG phenomena (which involve lagged relations between events at different cortical areas) with local, zero-lag, pulse and LFP synchrony and oscillation.

1.3 Previous simulation and explanatory findings

A number of simulations deal specifically with synchronous oscillation about 40 Hz. Whittington et al. (1995) have produced a detailed cellular model for hippocampal cells, which accounts for the oscillation as arising from activation of inhibitory cells depending upon fast (GABA\textsubscript{A}) receptors, and have further generalised this approach to the cortex (Traub et al. 1996). Alternative explanations depend on intrinsic oscillations in stellate and pyramidal cells (Linas et al. 1991) and in thalamocortical projections (Steriade et al. 1993). For the olfactory cortex at least, feedback interactions between excitatory and inhibitory cells impose an approximately 40 Hz oscillation, as evidenced in spectral and unit phase recordings (Bressler and Freeman 1980; Eeckman and Freeman 1990).

It has also been shown in simulation that local lateral inhibition might entrain synchronous oscillations (Nischwitz and Glunder 1995), that limit cycle oscillators representing single neurones can entrain to several independent synchronised clusters (Tass and Haken 1996), and that non-oscillating cells in Synfire chains can shift phase into synchrony (Arnoldi and Brauer 1996).

1.4 Basis for the present study
In Wright and Liley (1995) a simple simulation of EEG in cat and human was reported that was shown to be able to reproduce peak spectral densities about the range of the major cerebral rhythms, a ‘shift to the right’ (i.e. increasing high-frequency power) with increasing non-specific cortical excitation, travelling waves with velocities close to experimental measures (Lopes da Silva and Storm van Leeuwen 1978; Thatcher et al. 1986; Wright and Sergejew 1991; Burkitt 1994) emanating from localised inputs, and frequency/wavenumber characteristics similar to those of real ECoG. (The last result was marred by unavoidable problems of spatial aliasing.)

The same simulation methods have been applied in the present paper, firstly to clarify the spectral and travelling wave characteristics, secondly to reproduce conditions and results concerned with synchrony and oscillation analogous to experiments conducted by Singer and Gray and others, and thirdly to elucidate the dependence of the findings on particular physical mechanisms.

By these means it is aimed to show that a unified explanation of both zero-lag synchrony and travelling wave phenomena can be advanced, and that the mechanisms for these effects need not depend upon non-linear phase-locking, complicated synaptodendritic dynamics, or synchrony of cortical inputs. First steps have also been taken to show that certain properties of synchrony are scale invariant and may apply to interactions among individual cells, or interacting populations.

2 Methods
2.1 An outline of single-layered cat cortex ECoG simulation (Wright and Liley 1995)
Elements of connectivity in the simulation are sketched in Fig. 1.

Fig. 1. Schematic of connectivity in the simulation. Boxes labelled E and I represent dendritic lag processes of excitatory and inhibitory cells within each unit volume, shown shaded around a single unit volume. Straight arrows within a unit volume represent intracortical (\(\beta\)-type) axonal couplings with negligible delay, and coupling strength proportional to the density of synapses connecting the excitatory and inhibitory cells. Curved arrows represent cortico-cortical (\(\alpha\)-type) axonal couplings with lag proportional to distance, and coupling strength also proportional to synaptic density but diminishing as a Gaussian function with distance. Unit volumes are organised within a 20 \(\times\) 20 matrix, and identified by row and column, from 0 to 19. The matrix is closed edge-to-edge by \(\alpha\)-type connections to create a toroidal architecture.
2.1.1 Unit volume, pulse density and local field potential. For the $N$ cells in unit volume, pulse density, $Q$, is given by

$$Q = \frac{1}{N} \sum_{j=1}^{n} q_j$$

where $\{q_j\}$ are the states of the cells, and $q_j = 0$ or 1, depending on whether the individual cell is sub- or suprathreshold.

The local field potential, $V$, and $Q$ are related by

$$Q = (1 + e^{\alpha V - 3})^{-1}$$

which represents a sigmoidal response of pulse density with respect to $V$, consequent to the Gaussian distribution of action potential thresholds with respect to $V$. Voltage units are standard deviations (SD) of the distribution of cell thresholds, and the distribution is considered truncated at 3 SD from the mean threshold.

Parameter $\alpha$ has the value $-1.82$, rather than the conventional $-\sqrt{2}$, as this choice produces a somewhat better approximation over the usual low pulse density operating range of the simulation.

2.1.2 Synaptic density as a measure of coupling strength. Within and between unit volumes, the strength of synaptic coupling between any two masses of cells of uniform type, or the recurrent coupling of any one mass to itself, was treated as proportional to the synaptic density of the appropriate type, expressed as a fraction of the total synapses in unit volume. The nomenclature adopted reflects Braibanti and Schuiz's (1991) association of apical and basal dendritic systems with cortico-cortical (s-type) and intracortical (p-type) axo-axonic connections. Subscript $ct$, for example, indicates that the synaptic class is that of excitatory cells afferent to inhibitory cells, and similarly for other subscripts. The relevant types of synaptic density are then:

$-x_{ct}, x_{ct}$ Static potentials associated with the entirely excitatory cortico-cortical fibres. For a typical unit volume at cortical coordinates $(0,0)$, partial densities can be defined by their origins from surrounding volumes at $(x,y)$ as $x_{ct}(x,y)$. For cat cortex, $x_{ct} = 0.765$.

$-\beta_{ct}, \beta_{ct}, \beta_{ct}$ Synaptic densities of intracortical connections. $\beta_{ct} = 0.0845$, $\beta_{ct} = 0.0149$, $\beta_{ct} = 0.0228$, $\beta_{ct} = 0.004$.

$-\mu_{ct}, \mu_{ct}$ Synaptic densities of non-specific cortical afferents. A maximally conservative estimate for density of non-specific inputs to cortex was made – namely that only 1% of synapses are of this origin. $\mu_{ct} = 0.0077$, $\mu_{ct} = 0.0011$.

$-M_{ct}$, the density associated with specific cortical afferents – a value which was not required for purposes of simulation.

Calculations of synaptic density were made by methods reported in Liley and Wright (1994) and Wright and Liley (1995). We have since revised these estimates somewhat, but for ease of comparison with earlier work, the earlier values are retained here, and the same numerical procedures were used in simulation as previously. Quite small changes in these parameters and numerical approximations used in the simulation exert sensitive effects upon the range of stable performance without qualitative effects upon spectral and other properties. These influences upon stability will be discussed in later publications.

2.1.3 Dendritic delay (pulse/wave conversion), synaptic gain and axonal delay. System delays imposed by lag time of transmission through the cortical dendrites were modelled in a simplified form based on the findings of Freeman (1991). Thus where $g$ is a measure of synaptic gain, $Q$, is the afferent pulse action density, $\{w_j\}$ are discrete lag weights approximating the rise and fall of dendritic potential in response to a unit-amplitude pulse input, and $dt$ is the discrete time step (0.1 ms), the time response of local field potential is given by

$$V_i(t) = \sum_{j=1}^{n} w_j Q_j(t - jdt)$$

The $\{w_j\}$, $j = 1 \ldots n$, were in the form of an isosceles triangular function, with maximum $\frac{1}{2}$ being the 5-ms lag weight. For $j = 1$ and $n$, $w_j = 0$, and $\sum w_j = 1$.

The synaptic gain, $g$, is equal to 37 units, where each unit is a SD of threshold distribution.

Axonal delays corresponded to a uniform velocity of conduction of 9 m/s, matching physiological average estimates (Nunez 1995). Within each unit volume axonal conduction lags were treated as zero. Delays of conduction between unit volumes were calculated from $\Delta t = r_{ax}/v$, where $\Delta t$ is axonal conduction lag over the Euclidean distance $r_{ax}$ between the $p$th unit volume and the $q$th unit volume and axonal conduction velocity is $v$.

2.1.4 Coupling of unit volumes to simulate extended cortex. An extended area of cortex was simulated by unit volumes in a 20 x 20 matrix, each volume connected with its neighbours so that the coupling strengths, $x_{ct}(x,y)$, declined with $r_{ax}$ as a Gaussian function with SD 4 distance units, where a distance unit was the side of one cell of the 20 x 20 matrix. This approximates to the distribution of cortico-cortical fibres in the cat brain if the distance unit is taken as about 0.9 mm, giving an axonal delay of 0.1 ms/distance unit.

Boundary conditions were toroidal, conveniently imposing a closed condition.

2.1.5 State transition equations. Defining $V_{exc}(t)$, $V_{inp}(t)$, $Q_{exc}(t)$, $Q_{inp}(t)$, as the $p$th unit volume’s state variables, at time $t$,

$$Q_{exc}(t) = (1 + e^{\alpha V_{exc} - 3})^{-1}, \quad Q_{inp}(t) = (1 + e^{\alpha V_{inp} - 3})^{-1}$$

$$V_{exc}(t) = \sum_{j=1}^{n} w_j Q_{exc}(t - jdt), \quad V_{inp}(t) = \sum_{j=1}^{n} w_j Q_{inp}(t - jdt)$$

where $Q_{exc}(t), Q_{inp}(t)$ are afferent synaptic action densities for the $p$th unit volume, receiving cortico-cortical inputs from $q$ unit volumes, $q = 1 \ldots n$.

$$Q_{exc}(t) = \mu_{ct} Q_{exc}(t) + \mu_{ct} Q_{exc}(t) + M_{ct} Q_{exc}(t) + \mu_{ct} Q_{exc}(t) + M_{ct} Q_{exc}(t)$$

$$\sum_{j=1}^{n} \sum_{j=1}^{n} w_j (r_{ax}(t)) Q_{exc}(t - r_{ax}(t))$$

Consequent to the symmetry of long-range couplings, $\sum \sum w_j = x_{ct}$ and $\sum \sum w_j = x_{ct}$.

2.2 Inputs and control parameters

2.2.1 Definitions, and operating ranges. System inputs as defined in (5a,b) are $Q_{exc}$ and $Q_{inp}$.

$Q_{exc}$ was delivered as a zero mean Gaussian white noise (Marsaglia and Zaman 1987) added to the value of $Q_{inp}$ in selected unit volumes. The SD of the white noise was limited so that $Q_{exc}$ in the input unit volumes never became negative. This input could be used to introduce time-varying inputs to selected sites, mimicking special sensory input. $Q_{inp}$ was delivered at constant positive values to each unit volume during each simulation run, thus acting as a control parameter operating in the mean excitatory range. Applied to all unit volumes, it is analogous to non-specific cortical excitation as delivered by the reticular formation. When $Q_{inp}$ is supplied selectively to particular unit volumes in association with $Q_{exc}$ it acquires additional analogy to the mean excitatory tone associated with the specific sensory input, since the end result in the simulation is the same. The latter usage also justifies the use of values of selective $Q_{inp}$ inputs greater than 1, although the definition of pulse density limits $Q_{exc}$ to the range 0-1. Two further reasons justify this liberal range of $Q_{inp}$. Firstly, $\mu_{ct}, \mu_{ct}$ (the non-specific synaptic densities) have been very conservatively estimated in our anatomical estimates (Liley and Wright 1994). Secondly, since cortical non-specific activation depends upon relatively long-acting neuromodulators, this increases the effective non-specific pulse density.
matrix, and $Q$ was delivered to the excitatory cell compartment of the unit volumes in column zero of the matrix.

(ii) To allow the effects of conjoint excitation and driving of localised cortical sites to be studied, $Q_{ss}$ was set to 0.1 for all unit volumes except two. The two excited sites were then delivered higher $Q_{ss}$ inputs and (usually) asynchronous $Q$ inputs – i.e. noise sequences generated from different seed values.

White noise and sinusoidal forms of $Q$ were utilised as detailed with specific results.

Unless specifically stated, all simulation runs started with 0.8 s of both $Q$ and $O_S$ input, beginning with $[Q_1, Q_2]$ all zero at $t = 0$. Outputs were then analysed over the subsequent 0.8–16 s epoch.

2.3 Outputs

Outputs were the time series of $[V_{ss}, Q]$. Selected members of $[V_{ss}, Q]$ underwent spectral analysis by Fast Fourier Transform, and pairs of $[Q(t)]$ were used to obtain Pearson’s correlation coefficient as a function of lag, between selected sets. Ensemble averages for each of these measures were constructed by averaging the results obtained from the 0.8 s individual runs, usually over 100 samples, or as indicated with specific results.

2.4 Specialised forms of the simulation’s state equations

2.4.1 Linearised model form. To determine dependence of the simulation’s dynamics upon non-linear effects, it was useful to apply a linearised version of the state equations and contrast the results in runs of this type with runs generated using (4a, b) and (5a, b).

Differentiating (2),

$$\frac{dQ}{dV} = -ae^{-aV^3} + be^{-bV}$$

(6a)

and thus for any given mean value, $V_o$ of $V$

$$Q = \frac{dQ}{dV} (V_o - V)$$

(6b)

is a linearised approximation of (2), which can be generalised to (4a, b) and (5a, b). The mean values of $[V_{ss}, V]$, obtained from the standard, non-linear simulation at steady state, were thus the required $[V_o]$ to operate the simulation with equivalent gains to the standard condition, but without other non-linear signal transformations.

2.4.2 Roles of intracortical and cortico-cortical interactions. To study the roles played by different classes of synaptic coupling, special simulations were performed in which all couplings of the sets $\{\alpha_a\}, \{\beta_a\}, \{\alpha_e\}$ and $\{\beta_e\}$ were set to zero, alone or in combinations. Results were contrasted with matched standard simulation runs.

2.4.3 Effects of cortico-cortical coupling symmetry. To enable the consequences of a rich variety of asymmetrical coupling conditions to be studied by variation of a single parameter, the strength of $x$-type couplings between any two unit volumes, a function of distance of separation, $r_{m}$, was elliptically reweighted according to $r_{m} \sim \sqrt{[x^2 + (Ey)^2]}$, where $x$ and $y$ are the distances of separation along the rows and columns of the unit volumes, $E$ is equal to 1 for the usual two-dimensional Gaussian distribution of connections. For $E \neq 1$, the total input connectivities per unit volume were summed and normalised so that the total input synaptic density remained equal to $\sum_{m} \alpha_m$. Adjacent unit volumes were assigned reciprocal values of $E$ alternately along any row, column and diagonal on the matrix, thus generating asymmetry of coupling between most unit volumes. Results of simulations generated for a variety of values of $E$, were again contrasted with matched standard simulations.

2.4.4 Simulation of events at different scales. The state equations have a form which is applicable independent of scale provided that at very small scales the synaptic densities attributed to intra-cortical and cortico-cortical connections and axonal delays are appropriately corrected. This property arises since $N$, the number of cells per unit volume is arbitrary. Voltage units are defined in standard deviations of cell threshold, and the sigmoidal non-linearity can be interpreted as reflecting either the distribution of thresholds of individual cells or the probability that single cells will fire at a given level of dendritic de-polarisation. Whereas toroidal bounds are most appropriate for the simulation of events at the scale of the whole brain, for small-scale events zero boundary conditions best reflect the connectivity of a small locale of cells, considered as embedded in the larger cerebral cortex. Details are discussed further in Sect. 3.2.3.

3 Results

3.1 System properties with uniform $Q_{ss}$ inputs: travelling waves associated with low pulse densities

Figure 2 shows ensemble average spectral properties of simulated ECoG. The driving noise, $Q_{ss}$, was input synchronously (i.e. the identical time series) to each element in column zero of the matrix, and in each graph the ECoG power spectrum from the unit volume at the tenth column and row of the matrix is displayed, as representative of spectra from across the matrix. As described in Wright and Liley (1995) these spectra reflect travelling wave activity in the matrix, generated by the $Q_{ss}$ input. Within each graph, spectra are shown for progressive increments of uniform non-specific cortical excitation. At $Q_{ss} = 1.2$ the system is just short of runaway self-excitation, which occurs when the mean values of $Q_{ss}$ reach approximately 0.04.

It is seen that with progressive increase in $Q_{ss}$ total spectral power increases, and the peak spectral density moves to the right, towards a limiting case at which spectral power is distributed in the high beta to gamma range.

The top left-hand graphs show these features of ECoG spectral behaviour for the complete model of the cat brain. The top right-hand graphs were obtained with the matched linearised version. The lower two sets of graphs show the results when inhibitory to excitatory couplings, $\{\beta_{in}\}$, are set to zero, and intracortical excitatory couplings, $\{\beta_{ex}\}$, are set to zero, respectively.

It is seen that the complete (non-linear) and linearised results are little different. With the removal of excitatory/inhibitory interactions the shift of peak spectral density to the right does not take place with increasing $Q_{ss}$, and with the removal of $\beta_{ex}$ couplings spectral behaviours are similar to those of the complete model, but of diminished amplitude.

Not shown in Fig. 2 are the results of removals of $\beta_{in}$-type and $\beta_{ex}$-type couplings. The first of these cases ($\{\beta_{in}\} = 0$) gives spectral results resembling the complete model, but with a small relative increase of peak power best seen in the $Q_{ss} = 1.2$ case. The second case ($\{\beta_{ex}\} = 0$) is of course the case for the isolated unit volume, and no wave propagation in the matrix is possible.

In addition to simulation runs with synchronous noise inputs to column zero, runs were performed with sinusoidal $Q_{ss}$ inputs (single sinusoids and combinations, across the frequency spectrum of interest) and asynchronous (uncorrelated) $Q_{ss}$ noise driving to each unit.

\footnote{Here, as in earlier papers, the term zero bounds implies that axonal transmissions disappear at the boundary – equivalently, the boundary is absorbing}
volume in column zero. Asynchrony of the inputs lowered ECoG output amplitude without affecting spectral density. Sinusoidal inputs, single or in combination, yielded output ECoG in which almost all spectral power was concentrated at the input frequency or frequencies, with components several orders of magnitude lower, at integer multiples and fractions of the input frequency.

Figure 2 describes ensemble averages. Since real ECoGs are usually spectrally characterised over short epochs, Fig. 3 shows a set of examples of 0.8 s spectra. It is seen that random fluctuations in the spectra of the input noise, and the transfer function of the cortical medium under the control of \( q_{in} \), respectively control the transient fine spectral structure and the form of the spectral envelope. Both the transient examples and the ensemble averages appear similar to typical equivalent spectra from real ECoG, as seen in our laboratory and elsewhere.

3.2 Effects of non-uniform excitation

3.2.1 Zero-lag cross-correlation and ECoG gamma-band activity. Figure 4 shows results obtained from the standard form of the simulation, when equal DC inputs (\( q_{in} \)) and asynchronous noise inputs of equal RMS amplitude (\( q_{in} \)) are delivered to two separate unit volumes on the matrix – in these examples to the (6,10) and (12,10) (row, column).

In Fig. 4a the \( q_{in} \) input is of insufficient magnitude to elevate mean pulse density at the sites of input to the point where runaway self-excitation will supervene at the sites of driving – so the average pulse densities, \( \{ q_r \} \), at the driven sites remain low fractions. A surrounding field of secondarily excited unit volumes is revealed by the surface of pulse density variance. The ECoG power spectrum from the middle of this field (the 10,10 site) is representative in spectral density of ECoG from all unit volumes, and resembles the spectra produced with low uniform \( q_{in} \).

Cross-correlations of pulse densities from unit volumes of interest on the matrix show, from top down, that:

(i) Pulse density fluctuation is essentially uncorrelated at the sites of input. (The inputs themselves had a correlation of zero at all lags, and the correlations of the input sites were always low and maximal at variable lag.)

(ii) Pulse density at a driven site, and pulse density at any immediate neighbour to this site, are maximally correlated at a few milliseconds lag, i.e. the neighbouring
unit volumes are following the driven unit volume. This is shown for the (6,11) following the driven (6,10) site, but the result is general for all immediate neighbours.

(iii) Pulse densities in immediate neighbours of each of the driven sites are maximally cross-correlated at zero lag. This result, shown for the (6,11) and (12,9) unit volume sites, is general between any pair of sites neighbouring the driven sites, and extends to the sites between the driven sites – e.g. site (10,10) has maximal cross-correlation at zero lag with neighbours of either driven site, as will be shown further below.

In Fig. 4b cases are shown in which the $Q_{ns}$ inputs are similar to those in Fig. 4a but the RMS amplitude of the asynchronous driving noises is greater. The increase in noise magnitude is just such that pulse density at the driven sites fluctuates between the upper and lower bounds of $Q_\infty$. Average pulse densities remain lower than the maximum possible. In this condition pulse density cross-correlations follow the same general rules as in Fig. 4a. However, the representative ECoG power spectrum shows high amplitude at low frequency, and a strongly hyperbolic form.

Figure 4c shows effects of higher $Q_{ns}$ inputs to the driven sites. The mean pulse density at the driven sites is now essentially at the maximum. The ECoG has moved to gamma-band activity, centered around 40 Hz, and this spectrum is representative of activity throughout the matrix. Cross-correlations show the same general relations as in Fig. 4a and b.

Results obtained in the form of Fig. 4 were modified in the following ways in extended simulation runs, all using the standard model.

(i) When additional uncorrelated noise was injected to all other sites on the matrix, maximum cross-correlations declined, and ECoG spectra from remote unit volumes (e.g. 0,0) exhibited typical low $Q_{ns}$ form, rather than simply following the spectra of unit volumes nearer the highly excited sites.

(ii) When synchronous components were added to the asynchronous noises delivered to the driven sites, results were essentially similar, but maximum cross-correlations were higher.
Fig. 4a–c. Each figure shows the results when localised $Q_{in}$ input and asynchronous noise are delivered to sites $x = 6, y = 10$, and $x = 12, y = 10$, on the simulated cortical surface. Results are averaged over 100 0.8-s epochs. 

Left columns: Average pulse density, and log variance of pulse density, on the cortical surface, with the average ECoG spectrum from $x = 10, y = 10$.

Right columns: Pulse cross-correlations in excitatory cell components, calculated at 1-ms lags. Upper: Between the driven sites. Middle: Between a driven site ($x = 6, y = 10$) with an immediate neighbour at ($x = 6, y = 11$). Lower: Between sites neighbouring the driven sites, at (6,11) and (12,9). In all figures, the $Q_{in}$ input to all sites other than the driven sites was $Q_{in} = 0.1$. $Q_{in}$ input to the driven sites $= 4.84$, and noise standard deviation (SD) (in units of $Q_{in}$) $= 0.0177$.

(iii) When the driven sites were moved further apart or closer together then, in the presence of additional noise, cross-correlation maxima were reduced with increasing distance of separation of the input sites. All cross-correlations reported above were calculated at 1 ms lags. Lags as short as the step interval for the simulation ($0.1$ ms) were also used, and it was shown that sites neighbouring each of the driven sites exhibited small perturbations of maximal cross-correlation about zero lag within short simulation runs, but had maxima at zero lag on average.

An attempt was made to determine the time of onset of zero-lag cross-correlation. Precise quantification is problematic because of the effects of simulation initialisation. But zero-lag cross-correlation appeared to be present within tens of milliseconds after initialisation.

Figure 5 shows the extent of the domain of synchrony, its relation to the strength of average input to the driven sites, and the fact that synchrony is independent of the presence of gamma-band ECoG oscillation. It can be seen that in both cases a central field of synchronously oscillating pulse density is surrounded by outwardly radiating travelling waves. Calculations from these lags indicate velocities are similar to those previously calculated for traveling waves.
Fig. 5. The field of synchronous oscillation, and its surround of travelling waves. Top figures: Surfaces corresponding to those shown in Fig 4a and c. In these instances inputs have been delivered to unit volumes at (7,10) and (13,10), so that the (10,10) site is centrally positioned in the field. Left top surface: Delay of maximum positive cross-correlation of $Q_p$ between the (10,10) unit volume and all other unit volumes on the matrix. As for Fig. 4a, the $Q_p$ input to the two driven sites is 4.84. Right top surface: Corresponding result when $Q_p$ to the driven sites is 5.01, as in Fig. 4c. The small figures below each of these surfaces show the average value of $Q_p$ on the surface, and the ECoG spectrum arising from the (10,10) site. Results here are ensemble averages from 50 independent initializations of the simulation.

Fig. 6. Local excitatory/inhibitory pulse relations. Configuration of the simulation as in Fig. 5. Mean pulse densities and cross-correlations within the (10,10) site. Top left: Mean pulse density within the excitatory and inhibitory cell masses of the (10,10) unit volume, as $Q_p$ increases. Top right: Cross-correlation of pulse density in the excitatory and inhibitory components, corresponding to the shifts in the means. Bottom small graphs: Corresponding ECoG spectra at (10,10).
3.2.2 Mechanisms of synchrony and oscillation

3.2.2.1 Oscillation: local phase relations of excitatory and inhibitory cells. As described with Fig. 2, simulated gamma-band ECoG oscillation is contingent upon the presence of local excitatory/inhibitory interactions. This is true also when $Q_0$ inputs are restricted to a pair of driven sites. To indicate the mechanism of this oscillation, Fig. 6 shows the cross-correlations of the excitatory and inhibitory cells in a representative unit volume within the field of synchrony, in a simulation configuration identical to that applied to yield Fig. 5.

As the level of excitation of the two driven unit volumes is increased up to and beyond the point at which oscillation in the gamma range appears in the ECoG, excitatory and inhibitory pulse densities move towards an approximately quarter-cycle phase separation, in terms of the principal period of oscillation.

3.2.2.2 Synchrony: role of cortico-cortical ($\alpha_{ee}$) connections. Figure 7 was generated with a simulation configuration again like that used for Fig. 5, and is again plotted for every unit volume with respect to the (10,10) site. The distribution of synchrony as a function of lag, and the values of positive correlation associated with the characteristics delays, are shown.

Unlike the symmetrical $\alpha_{ee}$-type couplings between unit volumes used in the generation of Fig. 5, in these cases asymmetrical couplings ($E = 6$) were applied. It is seen that the pattern of delay is much transformed, but the central field of zero-lag synchrony, surrounded by outwardly radiating waves, is preserved.

A closely similar pattern of delays and correlations is seen in both the upper and lower sets of graphs. Yet the lower figures were generated in simulation runs in which the model was linearised (on the basis of the mean values found in the equivalent complete model: see Sect. 2.4.1.) and both $\{\beta_{ee}\}$ and $\{\beta_{re}\}$ were set to zero, thus eliminating intracortical, local excitatory/excitatory and excitatory/inhibitory interaction, and leaving only $\alpha$-type, linear interactions in an asymmetrically coupled matrix.

3.2.3 Interneuronal, versus interareal, synchrony. Results presented so far use a simulation of a cat-sized brain, resolved to about 1 mm$^2$ unit volumes. As mentioned in Sect. 2.4.4, state equations of similar form can be used to...
represent interactions of individual neurones or small populations, so long as boundary conditions, coupling coefficients and delays are appropriately reconfigured.

Results with regard to synchrony at small scales are shown in Fig. 8. These were obtained by:

(i) setting the boundary conditions of the matrix to zero (absorbing) bounds;
(ii) setting \( \{z_{\alpha \alpha}\} \) to the usual value of \( \{\beta_{\alpha \alpha}\} \) and \( \{\beta_{\alpha \alpha}\} \) to zero, so that all excitatory couplings are intracortical and there is negligible self-excitation;
(iii) setting the axonal conduction time for passage of signals across single unit volume to zero, reflecting negligible local axonal delays;
(iv) again applying a highly asymmetrical type of coupling between unit volumes, by the use of the elliptical weighting of inputs, so as to better conform to the high asymmetry of synaptic couplings at microscopic scale;
(v) raising the \( Q_{\alpha \alpha} \) inputs to both the driven sites to a higher value \( (Q_{\alpha \alpha} = 6) \), and also the \( Q_{\alpha \alpha} \) to other neurones/unit volumes to 0.5, so as to increase (for numerical reasons) the dynamic gains to a degree sufficient to offset the fall in coupling strength when only intracortical couplings are modelled.

The simulation then represents, to a first approximation, a matrix of individual cells, or small populations of cells, connected by intracortical fibres only.

It is seen in Fig. 8 that synchronous activity appears at separated sites at this microscopic scale, as at the meso/macroscopic scales.

3.2.4 Effect of total coupling gains upon synchrony. Figure 9 shows results again using a standard configuration with regard to couplings, axonal delay and boundary conditions. Here the input signals (high \( Q_{\alpha \alpha} \), asynchronous noise \( Q_{\alpha \alpha} \)) have been applied not only to two discrete sites, but to four inputs sites, arranged in a bar. In this configuration couplings are again rendered asymmetrical, by use of elliptical weighting of input couplings to unit volumes. This asymmetry has been arranged so that when the bar of inputs is applied to all sites from (7,10) to (13,10) it is acting on units which are relatively strongly coupled along the tenth row of the matrix. But when the bar is input diagonally, from sites (7,7) to (13,13), it is acting across relatively weakly coupled unit volumes. It is seen that the extent and amplitude of the field of synchronous activity are influenced by the orientation of the input.

These results depend essentially upon manipulating the gain of couplings between involved units. This can be achieved structurally, by manipulating the symmetry of type couplings as shown, or by dynamic influences, i.e. by manipulating the pattern of \( Q_{\alpha \alpha} \) inputs. The latter, equivalent result to that shown for structural manipulation, follows from the fact that the total gain between any two unit volumes is a function of the structural terms \( z \) and \( g \),
and the dynamic gain term $dQ/dV$, which is in turn a function of $Q_0$.

4 Conclusions

4.1 Summary of findings

It appears that both travelling EEG waves and zero-lag pulse synchrony and oscillation can be accounted for by a relatively simple system of state equations, even though these do not model most aspects of individual cell physiology in any detail.

Narrow bandwidth EEG activity – alpha activity in particular – is not fully accounted for, and must be taken to reflect rhythmic input to cortex from subcortical sites, rather than arising in the cortex itself. Nonetheless the regulation of the cortical wave transfer function by non-specific cortical excitation can account for much of the relative amplitude of EEG frequency components.

Synchronous oscillations in this simulation do not depend upon synchrony of inputs to separate cortical sites, and are manifest at both microscopic and meso/macroscopic scales. The extent of the field of synchrony is regulated by the pattern of gains – both structural and dynamic – within the field of coupled elements.

4.2 Mechanisms determining spectral form, synchrony and oscillation

Results described with Fig. 2 enable consideration of the ECoG spectral envelope as the sum of two essentially linear processes: one dependent on excitatory/inhibitory interactions, and one dependent on excitatory/excitatory interactions. From the results in Figs. 7 and 8, zero-lag synchrony depends upon linear excitatory couplings alone, and is robust despite asymmetry of couplings, and variation of axonal lag.

The basic mechanisms underlying the phenomena can be deduced as follows:

4.2.1 Spectral envelope. As indicated in Fig. 6, given a sufficiently excited state, excitatory/inhibitory interactions generate oscillation at a center frequency of about 40 Hz, in accord with the prediction of Freeman (1991). This process appears to give rise to one distinct component in the spectrum.

A second distinct spectral component may be accounted for as follows. Each unit volume is associated with a system of recurrent loops of cortico-cortical connections, leading back to the reference unit volume. If there are $n$ steps of one synapse in a particular recurrent loop, the mean delay in this loop is $\delta$ per synaptic step, and $\gamma$ is the mean gain at each step, then the recurrent loop’s activity is associated with resonance at $f = 1/\eta \delta$ with a gain of $\gamma^n$. If each unit volume is connected to $\eta$ other unit volumes, then the number of recurrent loops of $n$ synaptic steps which can be formed is proportional to $\eta^n$, and the recurrent gain $G(f)$ of each excitatory mass upon itself is thus proportional to $\gamma^n \eta^n$. Thus when $\eta = 1/\gamma$, $G(f)$ along pathways of any $n$ is (roughly) equal, the afferent gain, $\gamma \eta = 1$, and the system is on the verge of runaway excitation. When $\eta > 1/\gamma$, $G(f)$
is monotonically descending with $f$, and the system is stable. When $\eta < 1/\gamma$, $G(f)$ is monotonically ascending with $f$, and the system enters runaway self-excitation.\(^{2}\)

The two components together appear to account for the forms of the spectral envelope, their sum dictating the frequency of peak power transfer from inputs.

The further spectral effect which appears when levels of pulse density vary rapidly from low to maximal action potential rates, as shown in Fig. 4b, may contribute to the typical $1/f$ form seen in desynchronised real ECoG.

### 4.2.2 Zero-lag synchrony

The occurrence and spatial conformation of synchrony in the simulation depends on long-range excitatory couplings, gains and lags only, and is linear in nature. This property is not one commonly encountered in a linear wave medium. Notably there is not synchrony between the driven sites themselves, and both the neighbour sites and activity beyond the field of synchrony exhibit lag correlations with input sites consistent with the propagation of travelling waves. Thus the synchronous field must be accounted for as a property of intersecting, weakly dispersive linear waves propagating from each input site. An analytic treatment of correlation functions in an equivalent continuum wave model is in preparation.

### 4.3 Match to experimental results and other theoretical accounts

The simulation’s qualitative behaviours correspond to the experimental phenomena reported in Singer and Gray’s (1995) summary review of synchrony and oscillation results. They contrast with some of the explanations advanced by other writers. The mechanism of 40 Hz oscillation is closely akin to that proposed by Freeman (1991), but appears rather different to that modelled in hippocampus by Whittington et al. (1995). Further, there is no dependence upon subcortical rhythmic inputs (Llinas et al. 1991; Steriade et al. 1993), nor limit-cycle oscillations (Tass and Haken 1996), nor local lateral inhibition (Nischwitz and Glunder 1995), nor do the effects depend upon comparator properties of individual neurones (Konig et al. 1996). However, these various alternative mechanisms may be complementary, or applicable in different situations, in ways which are not yet clear.

Singer and Gray’s (1995) findings can be considered accounted for to the following extent:

(i) Synchrony at exactly zero lag appears over the correct time course, and is present at all neural scales, from individual neurones within a column to interareal extents. Roles for both structural connectivity gains and functional gain are provided, which concur with the interpretations given to their findings by Singer and Gray (1995).

(ii) Synchrony is not contingent upon the presence of gamma-band oscillation. Singer and Gray emphasise the separability of these phenomena in experimental situations, although experimentally they are likely to appear together when stimulus conditions are strong and non-specific activation high.

(iii) Synchrony is present not between the cells receiving input but in those surrounding the sites of input. This may explain why synchrony is best observed as a multi-unit phenomenon.

(iv) The neighbouring sites to input actually follow the input sites with a few milliseconds’ lag. If it is assumed that cortical neurones which are optimally driven by visual stimuli correspond to those more directly receiving sensory input, while those suboptimally driven by sensory input are more strongly coupled to, and following, the neurones at the sites of direct input, then an explanation of why optimally driven neurones phase-lead those suboptimally driven (Konig et al. 1995) is apparent. Yet the density of local connection would involve both types in synchronous oscillation to some extent.

(v) When visual inputs which produce synchrony in response to short bars are concurrently stimulated with a large bar which links the stimuli delivered to the two sites, synchronous oscillation is enhanced (Singer and Gray 1995). This may be accounted for by the influence of the longer bar upon functional gains in the field of synchrony. However, a detailed explanation would need to allow for the influences of overlying noise, etc., which have not been modelled in any detail.

(vi) Dependence of synchrony on subtle aspects of sensory tuning (e.g. to movement and spatial frequency) has not been accounted for in these simulations. These findings do not constitute evidence against the general validity of this model, but pose a challenge for further detailed modelling.

(vii) While reproduction of the experimental findings in simulations adds little to debates concerning the relationship of synchrony and oscillation to cortical information processing, it should help to define a test-bed for modelling of the cognitive events. The field of synchrony contains information derived from both sites of asynchronous input. Given appropriate local synaptic modification (learning) rules, local networks could subsequently identify partial inputs as belonging to particular classes of associated spatio-temporal events. Also, the extended field of synchronous oscillation suggests that this mechanism may be important in ensuring redundancy of storage.\(^{3}\)

(viii) The simulation’s properties imply that future experimentation might compare results at the intracortical, micro-level with macroscopic EEG waves in the surrounding field, and that events at

---

\(^{2}\)A more rigorous development of this argument depends upon assumptions concerning the distribution of long-range couplings, but the general conclusion regarding the spectral behaviour is robust.

\(^{3}\)In concurrent work Amit and Brunel (submitted) have shown that a neural network with many common features with the present simulation has physiologically realistic behaviour under Hebbian memory storage, and exhibits synchronous oscillation associated with low pulse densities.
these very different scales may share certain scale-invariant dynamics.

4.4 Some general modelling aspects and limitations

Very simplified state equations compared with those used in most neuronal modelling nevertheless appear to have considerable explanatory power with regard to the dynamics of the neuronal field. It remains to be emphasised in which ways the simulation is restricted, and which aspects of the parameterisation appear critical.

Firstly, without the approximation to dendritic time constants which accord with the experimental findings of Freeman (but which are shorter than those commonly accepted) oscillation in the gamma band would not appear, since the period of oscillation is four times the dendritic rise time. Similarly, wave velocities are partially dependent on the dendritic rise time, as well as on the axonal conduction velocity.

Secondly, without setting synaptic gain to a high figure, this simulation cannot combine both the range of spectral densities exhibited and the low overall average pulse rate actually present in the neocortex (Amit et al. 1990). Consequently the model readily enters runaway excitation. There is reason to believe that real neocortex may exploit such a metastability (Wright and Liley 1996), but an account of dynamic stabilising mechanisms is lacking at this stage.

Thirdly, the present model is unable to account adequately for the complex relations of ECoG RMS amplitude and spectral density commonly revealed in real ECoG—instead, amplitude rises monotonically with shift of the spectral envelope to the right. Neither does it deal adequately with pulse density variance as a function of mean pulse density.

Fourthly, the description of subcortical inputs and cortico-cortical connectivities remain crude, but also subject to reasonably simple improvement.

Acknowledgements. This work was supported by the Ross Trust, of Melbourne, Australia. The author thanks Nick Hawthorn for technical contributions.

References


Burkitt G (1996) Steady-state visually evoked potentials and travelling waves. PhD dissertation, Swinburne University of Technology, Melbourne, Australia


Propagation and stability of waves of electrical activity in the cerebral cortex

P. A. Robinson,1,* C. J. Rennie,1,2,† and J. J. Wright3,‡

1School of Physics, University of Sydney, New South Wales 2006, Australia
2Department of Medical Physics and Cognitive Neuroscience Unit, Westmead Hospital, Westmead, New South Wales 2145, Australia
3Mental Health Research Institute, Parkville, Victoria 3052, Australia

(Received 23 December 1996)

Nonlinear equations are introduced to model the behavior of the waves of cortical electrical activity that are responsible for signals observed in electroencephalography. These equations incorporate nonlinearities, axonal and dendritic lags, excitatory and inhibitory neuronal populations, and the two-dimensional nature of the cortex, while rendering nonlinear features far more tractable than previous formulations, both analytically and numerically. The model equations are first used to calculate steady-state levels of cortical activity for various levels of stimulation. Dispersion equations for linear waves are then derived analytically and an analytic expression is found for the linear stability boundary beyond which a seizure will occur. The effects of boundary conditions in determining global eigenmodes are also studied in various geometries and the corresponding eigenfrequencies are found. Numerical results confirm the analytic ones, which are also found to reproduce existing results in the relevant limits, thereby elucidating the limits of validity of previous approximations.

PACS number(s): 87.22.Jb, 87.22.As, 87.10.+e

1. INTRODUCTION

Measurement of electrical activity in the cerebral cortex by means of electrodes on the scalp or the cortical surface is a commonly used tool in neuroscience and medicine. Detailed multichannel recordings of activity resulting from neuronal firings are routinely made, showing complex spatial and temporal patterns in the cortical regions where cognitive tasks are performed. These signals, known as electroencephalograms or EEGs, display sufficient consistency that their coarse morphological and spectral features may be empirically identified and quantified. The frequency content of EEG and variations in the power spectrum with cognitive state have been well characterized [1], velocities of EEG waves have been estimated [2], and typical features of the EEG response to external stimuli (so-called event related potentials) have been measured. Unfortunately, the connection between recorded EEGs and the underlying neuronal dynamics (and a fortiori cognition) remains poorly understood. A few of the most basic properties of cortical waves appear to be established [3], but virtually everything beyond this level is the subject of considerable debate and the wealth of experimental data is largely wasted in the absence of a more solid theoretical framework within which to analyze it.

Numerous models of cortical activity have been developed at a variety of levels of description. At the most fundamental level are neural networks, which attempt to describe the interconnections between individual neurons with varying degrees of idealization [4]. We term such simulations microscopic because of their incorporation of microstructure and neglect of long-range interconnections. Most notably, Freeman has modeled the EEG arising from the olfactory bulb of animals, during the perception of odors, by uniting estimates of physiological parameters within a system of nonlinear equations [5]. However, other methods are called for when models for microscopic, highly nonlinear neuronal events are extended to the large scale required to describe the macroscopic EEG waves of the cerebral cortex. Because of the huge numbers of neurons (~10^10) in the cortex, smoothed-parameter models have been introduced to study global properties of cortical activity. Such models implicitly treat the cortex as a continuum (although they may be discretized for computation), characterized by mean densities of interconnections between neurons (which occur at synapses), mean neuronal firing rates, etc., with means taken over volumes large enough to include many neurons. Theoretical justifications for this “mass action” approximation have been given by Stevens [6] and Wright and Liley [7] and the resulting match with experimental findings has been discussed by several authors [7].

Both microscopic and continuum models typically include both excitatory and inhibitory inputs to a given neuron, which may itself be either excitatory or inhibitory in its action on other neurons. Excitatory inputs tend to increase the firing rate of a given neuron, while inhibitory ones reduce it, with both effects being nonlinear due, for example, to saturation at a maximum physiologically possible firing rate. Thus, in general, continuum models must incorporate mean densities of both populations of neurons, and of both types of interconnections, as well as the two neuronal firing rates. Delays in the propagation of signals through neurons (which are highly elongated) must also be included. These delays are of two types: dendritic lags, in which incoming signals are delayed in the dendritic fibers (see Fig. 1), and axonal delays of outgoing signals due to the finite propagation velocity along the axon.

The first continuum model [9,10] included excitatory and inhibitory populations in an infinite, linearized, one-dimensional (1D) model. With suitable adjustment of parameters, this model was able to reproduce the characteristic ~10 Hz frequency of the alpha rhythm, but omitted nonlin-
ear effects, axonal delays, and the convolutions of the cortex.

Nunez [11,12] added axonal delays in order to investigate global modes. This model permitted wave solutions and, with the imposition of boundary conditions, the excitation of global eigenmodes. Nunez solved this model analytically for a 1D loop cortex, and for two-dimensional cortices with periodic and with spheroidal boundary conditions (i.e., ignoring the more complicated convoluted form of the real cortex, and the inhomogeneity of cortical connections), interpreting observed cortical wave frequencies in terms of discrete eigenfrequencies. This model predicted global modes whose frequencies approximately match those of the major cerebral rhythms. In particular, the alpha rhythm was interpreted as being at the fundamental cortical eigenfrequency.

Wright and Liley [13–15] introduced a spatially discretized model in which the cortex is treated as 2D and divided into patches, each of which is parametrized by the mean densities of excitatory and inhibitory neurons, their mean firing rates, and their mean densities of interconnections (i.e., of synapses). Nonlinear effects and axonal and dendritic delays were all included, with a Green-function formulation describing the interconnections between patches as a function of their spatial and temporal separation. This model incorporated all relevant effects mentioned above, except convolutions and nonuniformities in cortical connectivity, while allowing for the imposition of a variety of boundary conditions. Moreover, its parameters were largely physiologically measurable, a significant advantage when comparing its predictions with measurements. However, simulations based on it have been limited to very small systems (or very coarse resolution in larger systems) due to its formulation in terms of Green functions, which are very slow to evaluate, and a numerically intensive treatment of dendritic lags.

The central purpose of this paper is to introduce a model of cortical electrical activity which includes nonlinearities, axonal and dendritic time lags, variable geometries and boundary conditions in 2D, and which permits analytic studies of wave properties and stability, while speeding computation to the point that whole-cortex simulations are possible with good resolution. This is accomplished in Sec. II by introducing a continuum wave-equation model to replace the linear parts of Wright and Liley’s [13–15] discrete Green-function one, and also by simplifying their treatment of dendritic lags. The new model is not identical to that of Wright and Liley, but incorporates the same underlying neurophysiology to a similar degree of approximation. Neither model addresses the question of filtering of cortical signals through the skull to determine the scalp EEG, a problem that can be avoided in any case by using magnetoencephalograms (MEGs) based on the magnetic signals associated with neural activity. The task of the remainder of the paper is to lay the mathematical basis for analysis of this model and obtain its basic properties. In Secs. III and IV we investigate the steady-state properties of the model and study the propagation and stability of small perturbations in the limit of an infinite medium. Periodic and spherical boundary conditions are imposed in Sec. V to investigate the properties of global eigenmodes and the eigenfrequencies are calculated for typical human parameters. An algorithm for numerical study of our model is described in Sec. VI and its output is used to verify key analytic results obtained in earlier sections.

II. CORTICAL MODEL

In this section we describe the relevant neurophysics and neurophysiology and incorporate it into a continuum model of cortical activity. The relationships of this model to those of Wright and Liley [13–15] and Nunez [11,12] are described in this section and Sec. IV, respectively.

A. Model equations

An excitatory neuron such as the one shown in Fig. 1 emits pulses (i.e., fires) at a mean rate $q_e$ that is determined by the potentials generated in the dendritic tree by the synaptic inputs of thousands of other neurons. Threshold potentials, above which high firing rates occur, are not identical for all neurons, but have a centrally peaked distribution. We can then make a continuum approximation by replacing $q_e$ with a local mean value $Q_e$, averaged over many neurons,
and introduce the mean dendritic potential $V_e$. Similar considerations apply for inhibitory neurons, denoted by the subscript $i$. Taking account of the spread of individual threshold potentials, one then finds the nonlinear relationship

$$Q_{ei} = \frac{1}{1 + e^{-c(V_{ei} - V_0)}},$$

(1)

$$\frac{dQ_{ei}}{dV_{ei}} = C e^{-c(V_{ei} - V_0)} \left[1 + e^{-c(V_{ei} - V_0)}\right]^2,$$

(2)

$$= C Q_{ei} (1 - Q_{ei}),$$

(3)

where $C$ is a positive constant and we have assumed the distribution (2) of threshold potentials relative to the mean value $V_0$ [a Gaussian distribution would be equally compatible with physiological measurements, yielding an error function in place of Eq. (1)]. In Eqs. (1) and (2), $Q_e$ is measured in units of the maximum value possible (250–1000 s$^{-1}$ per neuron), and potentials are measured in units of the characteristic standard deviation of the threshold distribution. Suitable values of the constants in Eqs. (1) and (2) are $C = 1.82$ and $V_0 = 3$ [14].

Within a particular neuron, the relationship between the rate of arrival of incoming pulses, $Q_{ae}$ or $Q_{ai}$, and the corresponding potential, $V_e$ or $V_i$, is complicated. The induced transmembrane voltage perturbation propagates along the dendrites in a way that depends on the local dendritic capacitance and resistivity [5]. However, for the situation considered here of aggregate neural masses, we adopt the empirical finding that the temporal spread and conduction delay within an individual neuron’s dendritic tree may be described by a simple impulse response. Specifically, Freeman [5] found that one can write

$$V_{ei}(r,t) = g \int_{-\infty}^{t} w(t - t') Q_{ae,ai}(r,t') dt',$$

(4)

where $w(u)$ is a non-negative weight function, with a characteristic width of $\approx 10$ ms and

$$\int_{0}^{\infty} w(u) du = 1.$$  

(5)

A suitable choice for $w(u)$ is

$$w(u) = \begin{cases} \frac{\alpha \beta}{\beta - \alpha} (e^{-\alpha u} - e^{-\beta u}), & \beta \neq \alpha \\ \alpha^2 u e^{-\alpha u}, & \alpha = \beta \end{cases}$$

(6)

(7)

for $u > 0$, where $\alpha$ and $\beta$ are positive constants. This function, shown in Fig. 2, peaks at $u_p = \ln(\beta/\alpha)/(\beta - \alpha)$ for $\alpha \neq \beta$ and at $u_p = 1/\alpha$ for $\alpha = \beta$. This peak location can be chosen to be approximately 5 ms to correspond reasonably closely to physiological parameters [5,9,10], although some authors favor somewhat larger values with $\alpha \approx \beta \approx 400$ s$^{-1}$ [16].

In general, Eq. (4) is a convolution that is difficult to handle analytically or numerically. However, the choice of

![FIG. 2. Weight function $w(u)$ given by Eq. (6) for $\alpha = 100$ s$^{-1}$ and $\beta = 350$ s$^{-1}$.](image)

Eq. (6) enables Eq. (4) to be split into two ordinary differential equations via the introduction of auxiliary potentials $U_{ei}$ and $W_{ei}$, with

$$U_{ei}(r,t) = \int_{-\infty}^{t} e^{-\alpha(t - t')} Q_{ae,ai}(r,t') dt',$$

(8)

$$W_{ei}(r,t) = \int_{-\infty}^{t} e^{-\beta(t - t')} Q_{ae,ai}(r,t') dt',$$

(9)

$$V_{ei}(r,t) = g \frac{\alpha \beta}{\beta - \alpha} [U_{ei}(r,t) - W_{ei}(r,t)].$$

(10)

We find

$$\frac{dU_{ei}(r,t)}{dt} = Q_{ae,ai}(r,t) - \alpha U_{ei}(r,t),$$

(11)

$$\frac{dW_{ei}(r,t)}{dt} = Q_{ae,ai}(r,t) - \beta W_{ei}(r,t)$$

(12)

for $\beta \neq \alpha$. For $\alpha = \beta$, one can work directly with $V_{ei}$ using the equation

$$\left(\frac{d^2}{dt^2} + 2 \alpha \frac{d}{dt} + \alpha^2\right) V_{ei}(r,t) = g \alpha^2 Q_{ae,ai}(r,t).$$

(13)

Equations (11) and (12), or Eq. (13), are much simpler to treat than the general case (4), but preserve all the essential physics. For applications in which only the characteristic time scale of the response $w(u)$ is important, one may assume $\beta \gg \alpha \approx 100$ s$^{-1}$ and omit $W$.

When a neuron fires, the pulses propagate along the axon and axonal tree to provide incoming pulses at other neurons various distances away. The strength of interaction decreases as the number of synapses decreases with increasing distance. If we assume a characteristic axonal propagation velocity $v$ and an isotropic distribution of axons in the continuum approximation, we can approximate the outward propagation of pulse density as a wave $\phi_{ei}$ generated by the source $Q_{ei}$. We thus find
where \( \gamma_{e,i} = v/r_{e,i} \) and \( r_{e,i} \) is the characteristic range of the axons (assumed to have an approximately exponentially decreasing distribution at large ranges). Appendix A discusses the connection between Eq. (14) and the axonal range distribution and explores generalizations of this equation to anisotropic media and media in which there is more than one characteristic axonal range. Typical values of the constants in Eq. (14) are \( r_c = 0.08 \text{ m} \) and \( r_i \approx 10^{-4} \text{ m} \) for humans.

The incoming potentials \( Q_{ae} \) and \( Q_{ai} \) at a particular location comprise contributions from the wave potentials \( \phi_{e,i} \) and inputs external to the cortex. These inputs are usually split into two classes: a uniform mean \( Q_{ns} \) resulting from the sum total of inputs from noncortical structures in the brain aside from those involved in a particular stimulus under study, and a specific excitation \( Q_s \) due to stimuli, which is defined here to include both noisy and coherent components which may or may not be spatially localized (e.g., in the visual cortex in response to a visual stimulus). The resulting equations are

\[
Q_{ae}(r,t) = M_e Q_s(r,t) + \mu_e Q_{ns} + a_{ee} \phi_e(r,t) - a_{ei} \phi_i(r,t),
\]

(15)

\[
Q_{ai}(r,t) = M_i Q_s(r,t) + \mu_i Q_{ns} + a_{ei} \phi_e(r,t) - a_{ii} \phi_i(r,t).
\]

(16)

The constants \( M_e \) and \( M_i \) determine the strength of coupling of specific inputs to excitatory neurons and inhibitory ones, respectively. Likewise, \( \mu_e \) and \( \mu_i \) represent the densities of synapses associated with nonspecific stimuli. The parameters \( a_{ee} \), \( a_{ei} \), \( a_{ii} \), and \( a_{ii} \) are the synaptic densities associated with excitatory and inhibitory inputs to excitatory and inhibitory neurons. Note that we have defined \( Q_{ns} \) to be constant in time and space, while \( Q_s \) may vary in time and space but is defined here to have zero spatial and temporal means.

If the range of the inhibitory axons is sufficiently short, their inhibition can be considered to be a local effect and axonal delays can be neglected. In this case, every inhibitory pulse is immediately received locally and one can replace \( \phi_j \) by \( Q_j \) in Eqs. (15) and (16) and omit the inhibitory version of Eq. (14). This local inhibition approximation limits the validity of the resulting equations to scales \( \gg r_i > 0.1 \text{ mm} \), which is not problematic in practice because the finest-scale probes currently applied to the cortical surface are arrays with separation of order 1 mm [10], while scalp electrodes typically have separations of 20–50 mm. Naturally, if one wishes to explore possible long-range inhibitory interactions, this approximation can be easily relaxed.

Our model is characterized by the system of seven equations (1), (10)–(12), and (14)–(16). Typical values of the constants in these equations are given in Table I for the cortices of mouse, cat, and human. Also quoted are values for the equivalent radius \( R_q \) of a spherical cortex with the same area as the actual convoluted one, and the linear size \( L_0 \) of a square cortex with the same property. The quantities \( M_e \) and \( M_i \) have not been measured.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mouse</th>
<th>Cat</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_{ee} )</td>
<td>0.8023</td>
<td>0.844</td>
<td>0.853</td>
</tr>
<tr>
<td>( a_{ij} )</td>
<td>0.0112</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td>( a_{ii} )</td>
<td>0.1186</td>
<td>0.122</td>
<td>0.126</td>
</tr>
<tr>
<td>( a_{ei} )</td>
<td>0.0626</td>
<td>0.022</td>
<td>0.011</td>
</tr>
<tr>
<td>( \mu_e )</td>
<td>0.0046</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td>( \mu_i )</td>
<td>0.0007</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>( r_e (\text{mm}) )</td>
<td>2</td>
<td>2.7</td>
<td>84</td>
</tr>
<tr>
<td>( v \ (\text{m s}^{-1}) )</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>( R_q \ (\text{mm}) )</td>
<td>3.8</td>
<td>5.1</td>
<td>157</td>
</tr>
<tr>
<td>( L_0 \ (\text{mm}) )</td>
<td>13</td>
<td>18</td>
<td>558</td>
</tr>
<tr>
<td>( g )</td>
<td>25</td>
<td>37</td>
<td>36</td>
</tr>
</tbody>
</table>

**B. Comparison with Wright and Liley’s model**

Wright and Liley [7,13–15] developed a model similar to the present one. Its similarities and differences are discussed here. A similar discussion for Nunzi’s linear model [11,12] is given in Sec. IV A.

The first point of similarity between the two models is that both use the form (1) for the relationship between \( Q_{re,i} \) and \( V_{re,i} \). Equation (4) is also the same in both models, but the present choice for \( w(t) \) enables the convenient forms (10)–(13) to be obtained. In contrast, Wright and Liley [7,13–15] used a less physiologically justifiable triangular function to approximate the curve shown in Fig. 2, and evaluated the convolution (4) by direct integration. Numerically, this led to large demands on processing and storage (see Sec. VI).

The main difference between the two models is in the treatment of axonal propagation. Wright and Liley made the local approximation \( \phi = Q \) (although they did not describe it in these terms) and employed a Green-function formulation in place of Eqs. (14)–(16). Their corresponding equations for \( Q_{ae} \) and \( Q_{ai} \) in terms of \( Q_s \) and \( Q_i \) were in discretized form and involved additional parameters describing the coupling of a given discrete region to itself. Discretization is an unnecessary complication here, so we give their equations in the following equivalent continuum form:

\[
Q_{ae}(r,t) = M_e Q_s(r,t) + \mu_e Q_{ns} - a_{ei} Q_i(r,t) + a_{ee} \int d^2r' \int dt' G(r,t,r',t') Q_s(r',t'),
\]

(17)

\[
Q_{ai}(r,t) = M_i Q_s(r,t) + \mu_i Q_{ns} - a_{ii} Q_i(r,t) + a_{ei} \int d^2r' \int dt' G(r,t,r',t') Q_s(r',t'),
\]

(18)

\[
G(r,t;r',t') = G(|r-r'|) \delta(t-t' - |r-r'|/v).
\]

(19)

In these equations \( \phi \) is expressed as an integral over the retarded Green function (19), which corresponds to signals
TABLE II. Relationship between symbols for quantities used here and those used by Wright and Liley in previous work. Note the reversal of the ordering of the mixed subscripts $e_i$ and $i_e$ in the present work relative to Wright and Liley’s notation.

<table>
<thead>
<tr>
<th>Symbol used here</th>
<th>Symbol used by Wright and Liley</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{ee}$</td>
<td>$a_{ee} + \beta_{ee}$</td>
</tr>
<tr>
<td>$a_{ei}$</td>
<td>$\beta_{ei}$</td>
</tr>
<tr>
<td>$a_{ie}$</td>
<td>$\alpha_{ie} + \beta_{ei}$</td>
</tr>
<tr>
<td>$a_{ij}$</td>
<td>$\beta_{ij}$</td>
</tr>
<tr>
<td>$\mu_e$</td>
<td>$\mu_{ee}$</td>
</tr>
<tr>
<td>$\mu_i$</td>
<td>$\mu_{ei}$</td>
</tr>
<tr>
<td>$M_e$</td>
<td>$M_{ee}$</td>
</tr>
<tr>
<td>$M_i$</td>
<td>$M_{ei}$</td>
</tr>
</tbody>
</table>

that propagate at a velocity $v$. This integral is slow to compute numerically and involves a large amount of storage (see Sec. VI); to date these factors have limited simulations using Eqs. (17) and (18) to relatively small grids, which do not always provide adequate resolution for the desired applications. In addition, in their discrete form, they involved additional parameters associated with the scale of the discretization. One advantage is that the form of the spatial part of the Green function can be chosen at will, with Wright and Liley using a Gaussian to approximate the decreasing synaptic density at large $|\mathbf{r} - \mathbf{r}'|$. Appendix A contains a discussion of the relationship of this Green function to the one implicit in the present work.

The notation used in the present work is somewhat different from that used previously by Wright and Liley [7,13–15]. Changes have been made partly because some of their parameters are redundant in the present formulation, and partly to conform more closely with conventional usage in physics and mathematics. The relationships are given in Table II.

### III. STEADY STATE

Understanding of the dynamics of our model begins by determining the uniform, steady-state behavior. Evaluating the integral in Eq. (4) for this special case and setting all derivatives to zero in Eqs. (14)–(16) yields

$$V_{e,i} = gQ_{ae,ai}$$

$$\phi_{e,i} = Q_{e,i}$$

$$Q_{ae} = \mu_e Q_m + a_{ee} \phi_e - a_{ei} \phi_i$$

$$Q_{aii} = \mu_i Q_m + a_{ie} \phi_e - a_{ii} \phi_i$$

where all quantities have spatially uniform steady-state values. Equation (21) can be used to eliminate $\phi_{e,i}$ from Eqs. (22) and (23). Then Eqs. (20), (22), and (23) are used to eliminate $V_{e,i}$ from Eq. (1), in favor of $Q_{e,i}$. These steps yield

$$\exp[CV_0 - gC(\mu_e Q_m + a_{ee} Q_e - a_{ei} Q_i)] = \frac{1}{Q_e} - 1.$$  (24)

Equation (24) can be used to eliminate $Q_e$ from Eq. (25) to give a single, rather cumbersome (but numerically straightforward), equation for the steady-state value of $Q_e$, whence the other steady-state quantities can be determined using Eqs. (20)–(24). Rather than treat this equation analytically here, we approximate Eqs. (24) and (25) and compare our results with numerical solutions of the exact equations.

Noting from Table I that physiological measurements imply $a_{ei} \ll a_{ee}$, Eq. (24) implies

$$Q_e \exp[CV_0 - gC(\mu_e Q_m + a_{ee} Q_e - a_{ei} Q_i)] = 1 - Q_e.$$  (26)

The left side of Eq. (26) is non-negative, with a single maximum where $gCA_{ee}Q_e = 1$; at large $Q_e$ it decreases exponentially fast. Figure 3 shows graphically that Eq. (26) can have either one solution or three, depending mainly on the values of $V_0$ and $Q_m$. One solution, always present, is located very close to $Q_e = 1$, with

$$Q_e \approx 1 - \exp[CV_0 - gC(\mu_e Q_m + a_{ee} Q_e - a_{ei} Q_i)] = 1 - Q_e.$$  (27)

The other two solutions, which only exist for small values of $Q_m$, are located at small values of $Q_e$. If we neglect $Q_e$ on the right of Eq. (26), these solutions are

$$Q_e = (gCA_{ee})^{-1} \ln \{B^{-1} \ln \{B^{-1} \ln \{\cdot \} \} \}.$$  (29)

$$Q_e = (gCA_{ee})^{-1} B \exp \{B \exp \{B \exp \{\cdot \} \} \}.$$  (30)

Equations (29) and (30) are recast in an iterative form, they become, respectively,

$$x_{n+1} = \ln(x_n/B),$$  (32)

$$x_{n+1} = B \exp(x_n).$$  (33)
FIG. 4. Steady-state values vs $Q_{ns}$ for the human parameters from Table I. The approximate solutions (29), (30), and (32) are shown as dashed curves, while exact solutions of Eqs. (24) and (25) are drawn solid. The square symbols show steady-state values obtained in the fully nonlinear simulations discussed in Sec. VI B. (a) $Q_e$. (b) $Q_i$.

respectively, where in both cases $Q_e$ is related to the limit $x_n$ by

$$Q_e = (g C a_{ee}^{-1}) x_n .$$

Equation (32) is found to converge to the larger of the two solutions, while Eq. (33) converges to the smaller solution. The corresponding values of $Q_i$ are given by

$$Q_i = Q_e \exp \left[ -g C (\mu_e - \mu_i) Q_{ns} - g C (a_{ee} - a_{ie}) Q_e \right].$$

Note that when $B = e^{-1}$, then $Q_e = (g C a_{ee})^{-1}$ according to both Eqs. (29) and (30). The solutions (29) and (30) are only valid for $B \approx e^{-1}$, which places an upper bound on $Q_{ns}$ in the low-$Q_e$ steady state:

$$Q_{ns} < \left[ CV_0 - 1 - \ln(g C a_{ee}) \right] / (g C \mu_e).$$

For non-negative $Q_{ns}$ this criterion also implies

$$g < e^{CV_0 - 1} C a_{ee}^{-1} .$$

for low-$Q_e$ steady-state solutions to exist. The solutions (27)–(35) can be substituted into Eqs. (20)–(23) to obtain steady-state values of other quantities.

Figure 4 compares the approximate solutions (29) and (30) with the numerical solution of Eqs. (24) and (25) for the human parameters from Table I. Figure 5 shows similar results for the fixed point given by Eqs. (27) and (28). In all cases the agreement is seen to be good. The approximation is least satisfactory near the rightmost point of the locus of solutions in Fig. 4. For example, Eq. (36) gives $Q_{ns} < 0.9520$ for the existence of low-$Q_e$ solutions, whereas the full equations yield $Q_{ns} < 1.0000$. The small errors present in Eqs. (29)–(35), and demonstrated in Figs. 4 and 5, arise from the neglect of $Q_e$ on the right side of Eq. (26). This neglect is justified on experimental grounds where it is found that in the normal cortex typical rates of less than 20 pulses per second per neuron are observed, compared with a maximum possible rate of 250–1000 s$^{-1}$ [4,17]. The square symbols in Figs. 4 and 5 are discussed in Sec. VI B.

In summary, we have found three fixed points of Eqs. (24) and (25). One, given by Eq. (27), corresponds to a seizure in which all neurons are firing at near their maximum possible rate. The next two, given by Eqs. (29) and (30), involve low firing rates of all neurons, similar to what is seen in the normal state of the cortex. We discuss these steady states further in Sec. IV B, once their stability characteristics have been clarified.

IV. WAVE PROPERTIES: INFINITE MEDIUM

In this section we consider the properties of small perturbations about the fixed points found in Sec. III. This yields the dispersion relations and growth or damping rates of the waves, and the linear stability boundary of the system. We do not consider nonlinear wave propagation or instabilities.
A. Wave dispersion

To determine the linear wave properties of our model, we must first linearize Eq. (1), writing

$$Q_{e,j} = Q_{e,j}^{(0)} + \rho_{e,j} [V_{e,j} - V_{e,j}^{(0)}],$$  
(38)

where $Q_{e,j}^{(0)}$ and $V_{e,j}^{(0)}$ are the relevant steady-state values from Sec. III and $\rho_{e,j} = \partial Q_{e,j}/\partial V_{e,j}$ at this point. Fourier transforming Eqs. (38) and (10)–(16), and deleting the components $(k, \omega) = (0, 0)$ (which were treated in Sec. III) then yields

$$Q_{e,j} = \rho_{e,j} V_{e,j},$$  
(39)

$$V_{e,j} = g L Q_{ee,ai},$$  
(40)

$$L = \frac{\alpha \beta}{(\alpha - i \omega)(\beta - i \omega)},$$  
(41)

$$D_{e,j} \phi_{e,j} = \gamma_{e}^{2} Q_{e,j},$$  
(42)

$$D_{e,j} = (\gamma_{e,j} - i \omega)^{2} + k^{2} v^{2},$$  
(43)

$$Q_{ee} = M_{e} Q_{e} + a_{ee} \phi_{e} - a_{ei} \phi_{i},$$  
(44)

$$Q_{ui} = M_{e} Q_{e} + a_{ie} \phi_{e} - a_{ii} \phi_{i},$$  
(45)

where the arguments $k$ and $\omega$ are implicit. It is worth noting from Eqs. (40) and (41) that $|L|$ decreases monotonically with $\omega$ and, hence, the dendrites act as low-pass filters that tend to remove frequencies $\omega > \min(\alpha, \beta)$.

We can write $Q_{e,j}$ in terms of $Q_{ee,ai}$ using Eqs. (39) and (40). Equations (44) and (45) can then be used to write $Q_{e,j}$ in terms of $\phi_{e,j}$. If the results are substituted into Eq. (43), we find

$$(D_{e} - F_{e} a_{ee}) \phi_{e} + F_{e} a_{ei} \phi_{i} = F_{e} M_{e} Q_{e},$$  
(46)

$$(D_{i} + F_{i} a_{ii}) \phi_{i} - F_{i} a_{ei} \phi_{e} = F_{i} M_{e} Q_{i},$$  
(47)

$$F_{e,j} = \gamma_{e}^{2} \phi_{e,j} + g L.$$  
(48)

A dispersion equation for $\phi_{e}$ alone results from elimination of $\phi_{i}$ from Eqs. (46) and (47). Instead of following this route, we make the local-inhibition approximation $\phi_{i} = Q_{i}$ based on the short range of the inhibitory fibers. As mentioned earlier, this approximation limits us to considering waves with wavelengths longer than a few tenths of a mm. We then find

$$(D_{e} - F_{e} a_{ee}) \phi_{e} = F_{e} (M_{e} Q_{e} - a_{ei} Q_{i}).$$  
(49)

Elimination of $Q_{i}$ in favor of $\phi_{e}$, as before, then yields

$$Q_{e} = \frac{F_{i}}{\gamma_{e}^{2} + F_{i} a_{ii}} (M_{e} Q_{e} + a_{ie} \phi_{e}).$$  
(50)

After substitution of Eq. (50) into Eq. (49) we then find the wave equation

$$(D_{e} - F_{e} a_{ee}) \phi_{e} = F_{e} M_{e} Q_{e},$$  
(51)

where some small terms have been neglected subject to the assumption that $M_{e} \gg M_{i}$ is not satisfied. The linear response of the cortex to a specific signal $Q_{e}$ is given by Eq. (51) in Fourier space. One significant point is that all the inhibitory coefficients have disappeared, having been discarded as small quantities in going from Eqs. (49) and (50) to Eq. (51). Hence it is the excitatory component that determines the long-range behavior.

For freely propagating waves, Eq. (51) with $Q_{e} = 0$ gives the dispersion equation

$$(\alpha - i \omega)(\beta - i \omega) D_{e} - \alpha \beta \gamma_{e}^{2} G = 0.$$  
(52)

The quantity $G$, given by

$$G = \rho_{e} g_{a_{ee}},$$  
(53)

is the net gain in the loop in which a low-frequency, low-$k$ perturbation of magnitude $\epsilon$ in $\phi_{e}$ gives rise to perturbations $a_{ee} \epsilon$ in $Q_{ee}, g \phi_{e} a_{ee} \epsilon$ in $V_{e}$, and $\rho_{e} g_{a_{ee}} \epsilon$ in $Q_{e}$. Hence, $G e^{i \phi_{e}} e^{-i \phi_{e}} = G e$. Equivalently, $G$ is the mean number of pulses stimulated by each pulse emitted. Equation (52) can be approximated as

$$(\alpha - i \omega) D_{e} - \alpha \gamma_{e}^{2} G = 0,$$  
(54)

$$D_{e} - \gamma_{e}^{2} G = 0,$$  
(55)

for $\beta \gg \alpha, \omega$ and for $\beta, \alpha \gg \omega$, respectively. For typical physiological parameters [10,18] one has $\alpha = 100 s^{-1}$ and $\beta = 350 s^{-1}$, and the approximations (54) and (55) are applicable for frequencies $f = \omega/2\pi$ given by $f \ll 55$ Hz, and $f \ll 15$ Hz, respectively. Only the first of these is appropriate to study the full range of normally recognized human EEG rhythms, although Eq. (55) may be semiquantitatively useful. In the single-parameter case $\alpha = \beta \approx 200 s^{-1}$, Eq. (54) does not apply but Eq. (55) is valid provided $f \ll 30$ Hz.

Equations (52) and (54) incorporate dendritic lags to generalize the corresponding linear wave equation obtained by Nunez [11,12], which omitted these lags and was of the form (55). Nunez’s equation applies in the limit in which these lags are negligible. However, the discussion in the preceding paragraph implies that this is at best semiquantitatively correct for the alpha rhythm ($f \approx 10$ Hz) and an increasingly poor approximation at higher frequencies. Examples of the solutions of Eqs. (52)–(55) are given in Sec. IV C, after we have discussed the question of stability.

B. Linear stability

Equations (52), (54), and (55) are polynomials in $u = -i \omega$ with purely real coefficients. Hence solutions for $u$ are either purely real or occur in complex-conjugate pairs. Growing solutions correspond to $\text{Im} \, \omega = \text{Re} \, u > 0$. From Eq. (55) we find

$$\omega = -i \gamma_{e} \pm i \frac{\gamma_{e}^{2} G - k^{2} v^{2}}{2 \gamma_{e}^{2}}.$$

This result immediately implies that an instability occurs for

$$G > 1 + k^{2} v^{2} / \gamma_{e}^{2} = 1 + k^{2} v_{e}^{2}.$$

WORK TOWARD A THEORY OF BRAIN FUNCTION | 273
with equality in Eq. (57) defining the instability boundary. The least stable perturbations are at \( k = 0 \). At large \( k \), the waves are damped, propagating modes, with \( \omega \approx \mp kv \pm i \gamma_r \). In the case of the quartic equation (52) at large \( k \), \( D_\alpha \) can be either large or small. If it is large, we find two purely damped modes with \( \omega \approx -i \alpha - i \beta \). If it is large, there are two propagating modes that approach the high-\( k \) solutions of Eq. (56). Since there can be only four solutions to a quartic, these four modes are the only ones. Similar reasoning can also be applied to the cubic equation (54), yielding the same modes except for the one at \( \omega \approx -i \beta \).

The quartic and cubic equations, (52) and (54), can be solved analytically in the general case, but the solutions are too unwieldy to be useful. Here we obtain the stability boundary for these equations by considering the special case where the root(s) with the largest Re \( u \) are marginally stable with Re \( u = 0 \). The stability boundary for the cubic equation can be obtained from that of the quartic in the limit \( \beta \to \infty \), so we consider only Eq. (52) here.

There are two ways in which instability can first set in and from which the instability boundary can be calculated: either a real root can reach the point \( u = 0 \) or a pair of complex-conjugate roots can reach the point where they are purely imaginary. The latter case breaks into two subcases in which the other two roots have negative real parts and are either purely real or are complex conjugates. In the case of a real root being the first to reach the instability boundary, the boundary corresponds to \( u = 0 \). Equation (52) immediately yields the criterion (57) for instability. The dendritic parameters \( \alpha \) and \( \beta \) do not enter this criterion because the frequency is zero at the point of marginal stability, whence \( L = 1 \) in Eq. (41). In Appendix B we demonstrate that the cases in which a pair of complex-conjugate roots are the first to become unstable have no consistent solution for \( G > 0 \). Hence Eq. (57) is the instability criterion for Eqs. (52), (54), and (55) in all circumstances.

For \( k = 0 \), the criterion (57) can be used to determine the stability of the fixed points found in Sec. III. Using Eqs. (3) and (53), we find stability only for

\[
Q_p(1 - Q_e)(gCa_{ee})^{-1} < 0.015.
\]

This criterion immediately implies that the high-\( Q_e \) root (27) is stable. By imposing the same approximation used in deriving (27), (29), and (30), namely that \( Q_e \ll 1 \), Eq. (58) becomes

\[
Q_p < (gCa_{ee})^{-1}.
\]

As noted above, this value of \( Q_p \) is the one at which the solutions (29) and (30) coincide, and is located at the rightmost point in Fig. 4(a). Consequently, this inequality identifies the lower of the two solutions, Eq. (30), as being stable, while Eq. (29) is unstable. Thus there are two basins of attraction in a linear approximation, corresponding to the two stable roots. 

(i) A “normal” low-activity state, corresponding to Eq. (30), in which all neurons fire at rates far below their physiological limits. This state corresponds to the stable one found previously in numerical calculations [7,13–15] and to normal levels of cortical activity in nature. 

(ii) A saturated high-activity state, corresponding to Eq. (27), where physiologically maximal firing rates are approached during a seizure. This state was also seen previously in numerical work [7,14] and presumably corresponds to a grand mal seizure in nature.

Numerical calculations show empirically that nonlinear systems initialized with \( Q_e \) below the upper of the two low-\( Q_e \) solutions tend to converge to the lower one, while systems initialized with higher \( Q_e \) converge to the solution (27). Thus Eq. (29) approximately defines the boundary between the basins of attraction even in the nonlinear case.

It has been remarked previously on many occasions that self-organizing systems such as the brain must by their very nature operate “on the border of instability” or “on the edge of chaos” [19]. Otherwise, complex behavior would not be possible because the system would either be unstable or would settle into a relatively quiescent state (or, at least, one of low complexity). We can measure the nearness of the cortex to marginal stability by taking the ratio of the damping rate at \( k = 0 \) in Eq. (57) to the corresponding rate for \( G = 0 \), since it is the cortical gain parameter \( G \) that controls stability. The resulting parameter \( (1 - G)^{1/2} \) is approximately 0.4 for the normal state of the mouse, cat, and human for the parameters in Table 1, even for \( Q_{ns} = 0 \). Thus, even without stimulation, the cortex is more than halfway to instability for the parameter values adopted here. In the more typical case in which \( Q_{ns} \) is sufficiently large that \( Q_e = 0.015 \), the stability parameter is only 0.07, implying that the cortex is very near to instability under typical conditions.

C. Numerical solutions of dispersion relations

The dispersion relations (52), (54), and (55) are straightforward to solve numerically. Several sample solutions are examined in this section to illustrate the main features of stable and unstable waves, and the similarities and differences between the waves predicted by the three equations.

Figure 6 shows the real and imaginary parts of \( \omega \) for the various modes predicted by Eqs. (52), (54), and (55), which predict four, three, and two modes, respectively. For the parameters given in the caption, the system is predicted to be stable for all \( k \) and, indeed, \( \text{Im} \omega < 0 \) throughout. The solutions labeled \( \pm 1 \) have \( \omega \approx -i \gamma_r \pm kv \) at high \( k \), in accord with the discussion in Sec. IV B. For these solutions, the group velocity \( v_g = \partial Re \omega / \partial k \) approaches \( \pm v \) as \( k \) increases, implying that axonal propagation chiefly determines the propagation of electrocortical waves in this limit. This derived velocity is in accord with cortical and scalp measurements [2,20,21]. The solutions labeled \( 0L \) and \( 0H \) have \( \omega \approx -i \alpha \) and \( \omega \approx -i \beta \), respectively, at large \( k \), also in accord with Sec. IV B. The three least damped modes are very similar in both the quartic and cubic cases, implying that the cubic approximation (54) is adequate under these circumstances. Important differences between the quadratic case and the other two are that the least damped mode is one of the pair \( \pm 1 \) in the former case, and the \( 0L \) mode in the other, and that the quadratic case has no propagating modes for \( k^2 \gamma_r^2 < G^2 \), in agreement with Eq. (56). Interestingly, the \( \pm 1 \) and \( 0H \) solutions become less damped at high \( k \), while the \( 0L \) solution becomes more heavily damped.

Figure 7 shows a case where the parameters correspond to an unstable solution to the steady-state equations at \( k = 0 \). The mode structure is the same as in Fig. 6, except that there are growing solutions where Eq. (57) is satisfied. This
boundary is the same for quadratic, cubic, and quartic equations, as the discussion in Appendix B implies. The unstable boundary is the same for quadratic, cubic, and quartic equations, whereas higher-degree dispersion relations have a different topology.

V. WAVE PROPERTIES: FINITE MEDIUM

The previous two sections have explored the properties of our model for an infinite medium. Since the cortex is finite we now examine the effects of imposing boundary conditions on our equations. Two cases are considered here: periodic boundary conditions on a square cortex (i.e., a toroidal topology, but not geometry), and a spherical cortex. We do not consider the effects of cortical convolutions or inhomogeneities in this paper, except to choose the size of our square or sphere so its area equals that of the actual, convoluted cortex seen in nature. Both convolutions (or other reductions in symmetry) and inhomogeneities will lead to splitting of degenerate eigenfrequencies found below, a point noted by Nunez in the context of his linearized analysis [11,12].

A. Cortex with periodic boundary conditions

If we impose periodic boundary conditions on a rectangular cortex with edges of length $L_x$ and $L_y$ (both of which we will denote by $L_0$ when they are equal), the wave vector $k$ is restricted to values $(2\pi n_x/L_x, 2\pi n_y/L_y)$ where $n_x$ and $n_y$ are integers. This corresponds to selecting out a series of discrete eigenmodes from dispersion curves such as those in Figs. 6 and 7 as having the only combinations of $\omega$ and $k$ allowed in the finite cortex. Note that modes with equal and opposite values of $n_x$ and/or $n_y$ are always degenerate.

One important point to note is that, for $1 < G < 1 + (2\pi r_c/\max[L_x, L_y])^2$, only the $k = 0$ mode is unstable. Since $2\pi r_c \approx L_0$ for the parameters in Table I, this can be a substantial range of parameter space. Thus, when considering the global linear stability of the brain, the problem can often be reduced to that of a single mode by virtue of the discrete nature of the eigenspectrum.

Table III lists $\omega_0$ and $\omega_1$ for solutions of Eq. (52) with human parameters from Table I, $G = 0.57$, and periodic
boundary conditions for small \( n_x \) and \( n_y \) in order of increasing \( \Re \omega \). Purely damped, nonpropagating modes are not listed. The tabulated values all lie on the branch labeled +1 in Figs. 6(a) and 6(b); the parameters of the branch labeled −1 are obtained by reversing the sign of \( \Re \omega \). Three important points to note are that (i) the imaginary part of \( \omega \) gives the characteristic width of each mode in frequency; since these widths are larger than the separation between modes for the parameters of Table III, one would not expect to see well defined resonances when these modes are excited by white noise, for example, (ii) the prominence of the resonances is reduced when, as in the present situation, the scale length of the damping, \( r_\ast \), is less that the circumference of the system, and (iii) the minimum frequency of a propagating mode is \( \sim 15 \) Hz, which lies in the typical physiological range.

### B. Spherical cortex

If we Fourier transform Eq. (52) in space, we find the form

\[
u^2 \nabla^2 \phi_e (r) = \left( \gamma_e - i \omega \right)^2 - L \gamma_e^2 G \phi_e (r), \tag{59}\]

where the temporal variation \( \sim \exp(-i \omega t) \) has been separated off. If we consider the spatial component of any wave equation on a spherical cortex, the eigenfunctions satisfy

\[
- \nabla^2 \phi_e = \frac{l(l+1)}{R_0^2} \phi_e, \tag{60}\]

where \( l \) is the principal quantum number of the eigenfunction in question, and \( R_0 \) is the radius of the sphere. Solutions are of the form \( \phi_e = e^{i \omega t} P_l^m (\cos \theta) \) where \( P_l^m \) is an associated Legendre function \([22]\), \( m \) is the azimuthal quantum number, and \( \theta \) and \( \varphi \) are standard spherical coordinates. As a result of this constraint, Eq. (59) leads to a dispersion relation very like Eq. (52) except that the quantity \( k^2 \) is replaced by \( l(l+1)/R_0^2 \). The azimuthal quantum number \( m \) does not appear in the dispersion relation, so all \( 2l+1 \) modes for fixed \( l \) are degenerate for a precisely spherical cortex.

Table IV lists \( \Re \omega \) and \( \Im \omega \) for solutions of Eq. (52) with human parameters from Table I, \( G=0.57 \), and spherical boundary conditions. Values for small \( l \) are given in order of increasing \( \Re \omega \) and purely damped, nonpropagating modes are not listed. The eigenfrequencies increase approximately as \( \sqrt{l(l+1)} \) at large \( l \) and are in the range observed for cortical frequencies. As in the periodic cortex, the mode widths exceed their separations, so we do not expect noise to excite clear resonances for these parameters.

### VI. NUMERICAL RESULTS

In order to study the dynamics of our full nonlinear model, Eqs. (1), (10)–(12), and (14)–(16) have been implemented numerically. This section outlines the methods used and numerical confirmations of the key analytic results obtained in earlier sections.

#### A. Methods

Of our seven equations, implementation of Eqs. (1), (10), (15), and (16) is trivial, since they involve no derivatives. Likewise, Eqs. (11) and (12) present no difficulties because they are ordinary differential equations in \( t \).

In solving the wave equation (14), we make the local-inhibition approximation \( \phi_i = Q_i \), which removes the need to follow \( \phi_i \) via a wave equation, but restricts us to wavelengths \( \gg 0.1 \) mm as discussed in Sec. II A. Initially, we implement Eq. (14) on a square grid with periodic boundary conditions, corresponding to a cortex with toroidal topology. By introducing the auxiliary field \( \psi = e^{-i \gamma t} \phi_e \), Eq. (14) is transformed to the form of the standard wave equation:

\[
\frac{\partial^2}{\partial t^2} - v^2 \nabla^2 \psi (r, t) = \gamma_e^2 e^{-i \gamma t} Q_e (r, t). \tag{61}\]

This enables us to use standard routines to step \( \psi \) forward in time, although we actually store \( \phi_e \) at each step to avoid
underflow at large \( t \) due to the exponential factor in Eq. (61). We solve Eq. (61) in coordinate space, rather than Fourier space, so as to make the most straightforward generalization to more complicated geometries in future, where Fourier methods are not applicable (e.g., on a convoluted cortex). Our approach will also allow inhomogeneities and anisotropies to be relatively easily incorporated in future.

On an \( N \times N \) grid, the runtime of our code scales as \( N^2 \) per time step. This represents a major improvement on the Green-function method used in previous nonlinear calculations [7,13–15], where the runtime required to evaluate the Green functions scaled as \( N^4 \) per time step. Direct comparisons of the two methods verify a speedup of order \( N^2 \), which enables us to attain adequate whole-cortex resolution. Storage is also dramatically reduced through the use of Eqs. (10)–(12) and (14), which are time local and require storage of only a few configurations of the system needed by the time-stepping routines (storage \( \sim N^2 \)). In contrast, previous methods stored of order \( 100 \) previous configurations of the system to treat the dendritic lags (storage \( \sim 100N^2 \)) and of order \( N \) previous configurations to evaluate the Green function (storage \( \sim N^3 \)).

B. Steady-state solutions

In Sec. III approximate steady-state solutions were derived, and in Figs. 4 and 5 a comparison was made with the exact solutions for \( Q_e \) and \( Q_i \). With the numerical implementation of the dynamical equations, as described above, it is possible to determine the accuracy of the analytic approximations made previously and of the resulting fixed-point estimates. This was done for several different values of \( Q_m \), including values for which there are three possible fixed points, and larger values of \( Q_m \) for which there is just one fixed point. We chose the grid ratio to be \( p = v \Delta t / \Delta x = 0.1 \). A grid of \( 100 \times 100 \) was employed, although the results are not sensitive to grid size. There is an important dependence on initial conditions though. The stable states both have their own basins of attraction characterized by low and by high firing rates, and so we set \( Q_e = Q_i = 0 \) initially for one series of simulations, and \( Q_e = Q_i = 1 \) for another. The other variables were initialized according to Eqs. (20)–(23). The results are shown as square symbols in Figs. 4 and 5 (showing low and high firing-rate steady-state solutions, respectively). As expected, the former set of simulations converged to the stable, lowest firing-rate fixed point, except for \( Q_m \approx 1.000 \), in which case no such solution exists and they converged to the stable, high firing-rate fixed point. The latter series, with \( Q_e \) and \( Q_i \) initialized to unity, always converged to the high firing-rate fixed point.

From the discussion of linear stability in Sec. IV B, we expect the upper branch of the curves in Fig. 4 to be unstable. This expectation was supported by a further set of simulations in which \( Q_m \) was arbitrarily set to 0.6 and \( Q_e \) was initialized to values in the range 0.000–0.050 at intervals of 0.005, to cover the range defined by the two low-\( Q_e \) steady-state solutions for \( Q_m = 0.6 \), namely, \( Q_e = 0.009 \) and \( Q_e = 0.032 \). For completeness, initial values \( Q_e = 0.100–1.000 \) at intervals of 0.100 were also used. With \( Q_i \) initialized to zero, it was found that simulations having \( Q_e \) initialized to 0.000–0.030 converged to the lowest of the steady-state solutions, and the remainder to the \( Q_e = 1 \) solution. With \( Q_i \) initialized to unity, almost the same result was obtained: simulations having \( Q_i \) initialized to 0.000–0.035 converged to the lowest of the steady-state solutions, and the remainder to the \( Q_e = 1 \) solution.

This series of simulations supports our theoretical linear-stability result that the upper branch of solutions in Fig. 4 is unstable, because convergence is only ever to the other two solutions. Indeed, initializing all variables as nearly as possible to those of an upper branch fixed point resulted in unstable behavior: after initial slow evolution, it eventually converged to one of the other two fixed points.

In addition to the above, we can infer the approximate form of the basin of attraction of the stable solutions. Altering the initialization of \( Q_i \) from zero to unity had only a small effect on the ultimate steady state. Whether the initial value of \( Q_e \) is larger or smaller than its value at the unstable fixed point is therefore the principal determinant of ultimate state of the system; the initial value of \( Q_i \) can affect that choice only when \( Q_e \) is close to that of the unstable fixed point.

With regard to the reproducibility of the fixed points using the full nonlinear simulation, errors in the final values of \( Q_e \) and \( Q_i \) were of the order of \( 10^{-5} \) for \( p = v \Delta t / \Delta x = 0.1 \). The accuracy of simulations worsens for larger values of \( p \), and when \( p \geq 1/\sqrt{2} \) (related to the Courant condition for the two-dimensional explicit difference method) the numerical solutions become unstable. Much smaller values of \( p \) lead to cumulative rounding errors unless more sophisticated time-stepping routines are used. The choice of \( p \) also affects the accuracy of convergence, as do the initial values of the variables. Variation of the initial value of \( Q_e \) gave rise to marginally different values of the final steady-state variables. These differences, also of order \( 10^{-5} \), provide another estimate of the accuracy of the calculations. Of course such small errors are negligible in applications because cerebral parameters are known only approximately and measurements cannot distinguish such small differences in firing rates.
FIG. 9. Spectral amplitudes for the case of periodic boundary conditions. A grid of 20×20 points was used, with \( \Delta x = \Delta y = 27.9 \text{ mm} \), driven by spatially coherent white noise \( Q_e \), along the column defined by \( x = 0 \), and the resulting activity \( Q_e \), measured at a point \( x = 10 \Delta x, y = 10 \Delta y \) for a total of 100×20.048 s. Spectral amplitudes are normalized to the spectrum of the input signal, \( Q_m = 0.7 \), \( \gamma = 100 \text{ s}^{-1} \), and \( \beta = 350 \text{ s}^{-1} \). The dashed line is for \( r_e = 0.15L_0 \), \( \gamma_c = 108 \text{ s}^{-1} \), and the solid line is for \( r_e = 1.50L_0 \), \( \gamma_c = 10.8 \text{ s}^{-1} \). All other parameters were as in Table III.

C. Wave propagation

The existence of roots of the dispersion relation having nonzero real parts means that the system supports traveling waves. This is demonstrated in Fig. 8, in which a 256×256 grid has its central 6×6 points driven by a sinusoidal signal with \( \omega = 500 \text{ s}^{-1} \). The amplitude of the driving signal was 0.01 and \( Q_m \) was set to 0.7. With reference to Fig. 4(a), this was sufficient to maintain activities (a) within the basin of attraction of the nonseizure state, and (b) within physiological limits. A concentric distribution of traveling waves is evident, and the scale length of damping (\( r_e = 83.7 \text{ mm} = 40 \) grid units) may be appreciated. This degree of damping does not allow accurate estimation of the wavelength, but the expectation, from the dispersion relation (52), of \( \lambda = 52 \) grid units is at least approximately borne out.

D. Power spectra

As another demonstration of the dynamical properties of the model, we consider the square cortex with periodic boundary conditions discussed in Sec. V A, and in Table III. Figure 9 is the result of driving this system along one column (\( x = 0, y = 0 - 19 \Delta y, \Delta y = 27.9 \text{ mm} \)) with spatially uniform white noise, and recording the values of \( Q_e \) at a site distant from the sites of stimulation. A total of 100 periods of 2.048 s were recorded, transformed to give amplitude spectra, and averaged. The nonspecific excitation \( Q_m \) was set equal to 0.7 to make this figure match the eigenfrequencies listed in Table III. With all other system parameters as in Table I (human) the dashed line in Fig. 9 was obtained. This shows no apparent resonance because for the eigenfrequencies to be clearly visible the damping length \( r_e \) must be \( \approx L_0 \). To demonstrate resonance, \( r_e \) was set to the somewhat unrealistic figure of 1.5L_0, and a resonance peak was then obtained (Fig. 9, solid line). The location of the peak is also consistent with the dispersion relation (52): with this revised value of \( r_e \), the \( n_S = n_y = 0 \) mode becomes purely damped and the first oscillatory mode is then the \( n_S = 1, n_y = 0 \) (or \( n_S = 0, n_y = 1 \)) mode having an expected frequency of 101 s\(^{-1}\). (The next mode is at 143 s\(^{-1}\), but there is so little signal at frequencies \( \approx 120 \) s\(^{-1}\) as to be beyond the precision of the analysis.)

A further point demonstrated by this example is the relationship between damping and the width of the resonance peak. Increasing \( r_e \) had the effect of reducing \( \gamma_c = \nu/r_e \) to 10.8 s\(^{-1}\), and this is reflected in the observed resonance peak width.

The main implications of this example are (i) very low values of \( \gamma_e = \nu/r_e \) are needed to see resonances, as discussed in Sec. V A, and (ii) experimentally Im\( \omega \) can be estimated from the width of the alpha resonance. The first point sheds some doubt on Nunez’s global resonance picture, unless \( \gamma_e \) is smaller than previously thought.

VII. SUMMARY AND DISCUSSION

Motivated by the need for a formulation of cerebral activity that is analytically and numerically tractable, we have formulated a set of nonlinear continuum equations that satisfies these requirements. These equations embody the nonlinear response (on average) of neurons to imposed potentials, the presence of excitatory and inhibitory populations, and axonal and dendritic lags, and provide the framework for a wide variety of analytic and numerical calculations.

Analytically, we have used our model to study the steady-state behavior of the cortex, and its stability, as well as linear waves propagating in bounded and unbounded models of the cortex. Numerically, the speed at which our nonlinear system can be simulated is of the same order as that for the corresponding linear one, thereby enabling adequate whole-brain resolution to be obtained in a 2D nonlinear model. The main results of the present study are summarized next.

(i) Dendritic lags have been treated in a way that is both simpler and closer to physiological measurements than in previous work. This enables analytic treatment and reduces numerical runtime and storage requirements.

(ii) The propagation of axonal signals, including axonal delays, has been formulated in terms of a wave equation. This bears some similarities to previous wave equations, but does not assume that the system as a whole is linear. Our wave-equation formulation is analytically tractable and dramatically faster to treat numerically than its Green-function predecessor. Numerical storage requirements are also far lower.

(iii) The results emphasize the importance of both dendritic and axonal delays in determining the dispersion relations of cortical waves and, hence, global eigenfrequencies.

(iv) The criteria for ignoring the finite range of inhibitory fibers have been made more explicit.

(v) An analytic fixed-point analysis has been done to determine the steady states of the system. Three fixed points have been found, one of which is unstable. Of the other two, one represents a “normal” state of low activity, while the other represents a saturated “seizure” state in which activity is near its maximum.

(vi) Dispersion equations have been derived for small-amplitude linear waves. These equations incorporate both axonal propagation and dendritic lags. The limit in which dendritic lags can be neglected is elucidated and it is shown that an earlier equation [12] is reproduced in this limit.
(vii) A stability boundary has been described, beyond which a seizure will set in. Under normal conditions the cortex is not far from this boundary, consistent with the view that complex, self-organizing systems must be near "the edge of stability" to function properly. This emphasizes the prospect for future work to analyze internal controls of cerebral dynamics, such as regulation of local and global inhibition, which may exploit this near-marginally stable behavior to produce much richer dynamics [7].

(viii) The effects of boundary conditions have been studied for a square cortex with periodic boundary conditions and for a spherical cortex, yielding discrete eigenfrequencies in the relevant physiological ranges.

(ix) Numerical solutions of our model equations have confirmed the existence of one steady-state solution with a high firing rate, and two steady-state solutions with a low firing rate. Of the latter pair, only the lower is stable, and both require \( Q_w \) to be less than a limiting value given approximately by Eq. (36). If \( Q_w \) is large enough the system will saturate. The system will also saturate if the state is moved to some point with \( Q_e \) greater than that of the unstable fixed point.

(x) Numerical simulations have demonstrated the existence of traveling waves.

(xi) Numerical simulations have confirmed theoretical results that sharp resonances are impossible unless the damping parameter \( \gamma_e \) is substantially smaller than previously supposed. This casts some doubt on previous suggestions that the alpha rhythm is a global resonance of the cortex. If, however, this rhythm is a global resonance, the imaginary part of its frequency can be estimated experimentally from its width.

ACKNOWLEDGMENTS

Two of the authors (P.A.R. and C.J.R.) thank the Mental Health Research Institute of Victoria and the University of Iowa for their hospitality during visits in which some of this work was undertaken. This work was supported by the Ross Trust, Melbourne.

APPENDIX A: GREEN FUNCTIONS AND GENERALIZED WAVE EQUATIONS

This appendix discusses the connection between the wave equation (14), its Green function, and the corresponding axonal range distribution. It then compares the Green function with that used by Wright and Liley [13,14], and explains how Eq. (14) can be generalized to a broader class of media.

1. Green functions

The solution of Eq. (14) can be written in terms of a Green function \( G \) as

\[
\phi\left(\mathbf{r}, t\right) = \int d^2 \mathbf{r}' \int dt' \ G\left(\mathbf{r}, t; \mathbf{r}', t'\right) Q\left(\mathbf{r}', t'\right), \quad \text{(A1)}
\]

where subscripts have been omitted for simplicity. In an isotropic medium, \( G \) depends only on the distance \( R = |\mathbf{r} - \mathbf{r}'| \) and the time difference \( \tau = t - t' \). The Fourier transform of the Green function can be immediately evaluated from the Fourier transforms of Eqs. (14) and (A1), giving

\[
G(k, \omega) = \frac{\gamma^2}{(\gamma - i\omega)^2 + k^2 v^2}. \quad \text{(A2)}
\]

The inverse Fourier transform of Eq. (A2) then yields

\[
G(R, \tau) = \int \frac{d^2 \mathbf{k}}{(2\pi)^2} e^{ikR} \int d\omega \ e^{i\omega \tau} G(k, \omega) = \int \frac{d^2 \mathbf{k}}{(2\pi)^2} e^{ikR} \int d\omega \ \frac{\gamma^2 \ e^{-\gamma \tau}}{4\pi v} \sin(kv \tau) \int_0^{2\pi} d\theta \ e^{ikR \cos\theta} \frac{1}{\sqrt{\nu^2 \tau^2 - R^2}} \Theta(v \nu - R), \quad \text{(A6)}
\]

where polar coordinates, \( R \) and \( \theta \), are used to do the integral over \( \mathbf{k} \). \( \Theta \) is a unit step function, and only the retarded part of the propagator has been retained to avoid unphysical solutions that propagate backward in time. Note that

\[
\int d^2 \mathbf{R} \int d\tau \ G(R, \tau) = \int d^2 \mathbf{R} \ \frac{\gamma^2}{2\pi v^2} K_0(R \gamma/v) \quad \text{(A7)}
\]

where \( K_0 \) is a modified Bessel function of the second kind (a Macdonald function) [22]. The result (A8) is required on physical grounds to ensure conservation of pulses. Note that the integrand in Eq. (A7) represents the time-integrated response at a distance \( R \); i.e., the total number of pulses reaching a unit area at that distance.

Figure 10 shows Eq. (A6) at various times. One point to note is that \( G(R, \tau) \) is not a \( \delta \) function at \( R = \nu \tau \), unlike in the better known three-dimensional case. Rather, it is a function that is concentrated close to the point \( R = \nu \tau \), but with a tail at smaller \( R \). Such behavior is also seen in the standard 2D wave equation without damping terms and is characteristic of wave propagation in a 2D geometry. Writing \( G(R) = \int d\tau \ G(R, \tau) \). Eq. (A6) implies \( G(R) = K_0(Rr_0)/(2\pi \gamma) \tau - (R/r_0)^{1/2} \exp(-R/r_0) \) at large \( R \), where \( r_0 = \nu / \gamma \), im-

![Figure 10. Green function (A6) for \( \gamma_e = 108 \text{ s}^{-1} \), \( r_d = 84 \text{ mm} \), and \( v = 9 \text{ m s}^{-1} \) for \( t = 2, 4, \) and 6 ms, from left to right.]
plying that axons have a characteristic range \( r_0 \). At small \( R, G(R) \sim -\ln(R/r_0) \), which may appear at first sight to be pathological. However, the total number of synapses at range \( R \) is proportional to \( 2 \pi R G(R) \), which remains finite. Also, it should be remembered that this logarithmic singularity is present intrinsically in the standard 2D wave equation, which does not yield singular responses to nonsingular inputs. If desired, the singularity can be removed by replacing \( K_0(r/r_0) \) by

\[
K_0(r/r_0) = r_0^{-2}K_0(R/r_0) - r_1^{-2}K_0(R/r_1),
\]

(A9)

with \( r_1 \ll r_0 \). This Green function has no singularity and corresponds to a pair of fields with characteristic ranges \( r_0 \) and \( r_1 \), and \( \phi \) equal to their difference. In general, the second field requires the introduction of an additional wave equation but a local approximation may be possible since it has a short range.

A Green function of the form (A6) is implicit in the work of Nunez [12], who used a wave equation of the form (14) to study linear cortical waves. Wright and Liley's, form [7,13,14] involves a \( \delta \) function, as in Eqs. (17) and (18), so the correspondence with the present work is not exact. Comparison of their distribution of axonal ranges with the present form is achieved by integrating over time, as in Eq. (A7). The result should be compared with Wright and Liley's \( G(R) \). They typically chose \( G(R) \propto \exp(-R^2/r_0^2) \), where \( r_0 \) is a constant; however, the closest large-\( R \) correspondence between the two models would be obtained by substituting the form (A7) for \( G(R) \) in Wright and Liley's model.

2. Generalized wave equations

It is straightforward to generalize the wave equation (14) in a number of ways. Most obviously, one could replace it by the form

\[
\frac{\partial^2 \phi}{\partial t^2} + 2\gamma \frac{\partial \phi}{\partial t} + \phi - v_1^2(e_1 \cdot \nabla)^2 - v_2^2(e_2 \cdot \nabla)^2 = \phi = \gamma^2 Q,
\]

(A10)

where the subscripts and arguments of \( \phi \) and \( Q \) have again been omitted for simplicity. This equation represents a case of anisotropic propagation velocity, with velocities \( v_1 \) and \( v_2 \) along orthogonal principal axes \( e_1 \) and \( e_2 \). The effective axonal ranges in these two directions are then \( v_1/\gamma \) and \( v_2/\gamma \).

A second way in which Eq. (14) can be generalized is to assume that there are a number of different axonal populations \( j \) characterized by different ranges \( r_0 \) and velocities \( v_j \). This is an extension of the procedure described in the preceding section for removing the singularity in \( G(R) \). In this more general case, one can write

\[
\phi = \sum_j \phi_j,
\]

with an equation of the form (14) or (A10) for each \( j \). This is also a suitable way to allow for a range of different propagation velocities, also parametrized by \( j \) [12]. Furthermore, it is straightforward to generalize Eq. (A10) by making the coefficients slowly varying functions of time and position.

**APPENDIX B: INSTABILITY BOUNDARY ANALYSIS**

In this appendix we outline the proof that the stability boundary for the quartic dispersion relation (52) cannot be set by a pair of complex-conjugate roots having \( \Re u = 0 \), where \( u = -i\omega \). In our proof by contradiction, we assume that there are two conjugate roots \( u_1 = iz \) and \( u_2 = -iz \) at the point of marginal stability, with \( z > 0 \). We then show that this assumption contradicts the requirement \( G \geq 0 \), with \( G \) given by Eq. (53). Hence the stability boundary is set by the condition \( u = 0 \), which leads to Eq. (56).

Equation (52) can be expanded to yield

\[
0 = u^4 + u^3(\alpha + \beta + 2\gamma) + u^2[\alpha \beta + 2\gamma(\alpha + \beta) + \gamma^2 + k^2v^2] + u[2\alpha\beta\gamma + (\alpha + \beta)(\gamma^2 + k^2v^2)] + \alpha\beta\gamma(1 - G + k^2v^2),
\]

(B1)

where the subscript on \( \gamma \) has been omitted for simplicity.

1. **Stable roots in a complex-conjugate pair**

If we suppose that, at the point of marginal stability, there are two stable roots given by \( u_3 = -x + iy \) and \( u_4 = -x - iy \), where \( x, y > 0 \), then

\[
0 = u^4 + 2u^2(x^2 + y^2 + z^2) + 2xz^2u + z^2(x^2 + y^2).
\]

(B2)

Equating the coefficients in Eqs. (B1) and (B2) and eliminating \( x \) and \( y \) yields

\[
z^2 = \frac{2\alpha\beta\gamma + (\alpha + \beta)(\gamma^2 + k^2v^2)}{\alpha + \beta + 2\gamma},
\]

(B3)

\[
G = 1 + \frac{k^2v^2}{\gamma^2} - \frac{z^2}{\alpha\beta\gamma}[(\alpha + \beta)(\gamma^2 + k^2v^2) - 2z^2].
\]

(B4)

Direct expansion of the right side of Eq. (B4), using Eq. (B3), shows that it is always negative for \( x, y, z > 0 \), contradicting the required sign of \( G \).

2. **Stable roots real**

The only remaining case is the one in which the two stable roots are real with values \( u_1 = -x \) and \( u_4 = -y \), for positive \( x \) and \( y \). If we expand the resulting equation and equate coefficients with those in Eq. (B1) we again arrive at Eqs. (B3) and (B4). Hence this case also yields a contradiction, implying that Eq. (53) is the instability criterion under all circumstances.


[2] F.H. Lopes da Silva and W. Storm van Leeuwin, in Architec-


Synchronous oscillations in the cerebral cortex

P. A. Robinson, J. J. Wright, and C. J. Rennie

1School of Physics, University of Sydney, New South Wales 2006, Australia
2Mental Health Research Institute, Parkville, Victoria 3052, Australia
3Department of Medical Physics and Cognitive Neuroscience Unit, Westmead Hospital, Westmead, New South Wales 2145, Australia

(Received 27 March 1997; revised manuscript received 19 September 1997)

The dynamics of a cortex driven by a finite number of white-noise point sources is studied using a recently developed wave-equation formulation. Green’s functions, power spectra, fluctuation levels, and two-point correlation functions are computed analytically and numerically. It is shown that a range of observed properties of so-called synchronous oscillations in the cerebral cortex can be correctly reproduced using the wave equation that involves only excitatory interactions between neurons. In particular, the observed existence of a maximal correlation at zero temporal lag between spatially separated points is reproduced and explained for a cortex driven by two white-noise sources. [S1063-651X(98)03904-X]

PACS number(s): 87.22.Jb, 87.22.As, 87.10.+e

1. INTRODUCTION

A long-standing puzzle in neurophysiology is the so-called binding problem, which may be stated as follows: Among the many concurrent patterns of neuronal activity present simultaneously in the billions of neurons in the brain, how are related aspects of a single stimulus bound together? For example, how are the disparate features of a face, each analyzed by specific cerebral areas that receive visual input and respond specifically to movement, angles, color, etc., associated and seldom confused with incidental aspects of the background, despite complex concurrent changes in the visual stimulus? Recent findings in neurophysiology indicate that the solution of this problem may lie in the brain’s use of a phenomenon termed synchronous oscillation to correlate spatially separated responses to a stimulus. The main purpose of this paper is to apply the recently developed wave-equation formulation of cortical dynamics [1] to elucidate this phenomenon.

It has been shown that clusters of neurons at simultaneously stimulated sites in the cortex and elsewhere in the brain can exhibit synchronous oscillations of neural firing rates over distances comparable to the size of the cortex and that this synchrony typically appears in circumstances when the stimulus properties are such that the features of the stimulus demand binding if a perceptual whole is to be created [2–4]. In this context synchronous oscillations are defined to be oscillations for which the temporal cross-correlation between signals at different locations exhibits a maximum at zero lag; we will also call such oscillations zero-lag oscillations on occasion.

In a recent review [5] findings were summarized that showed that synchrony appears at multiple scales, from small pools of locally connected neurons to sites on opposite sides of the brain. Both structural connectivity (e.g., by cortico-cortical axonal fibers) and functional dynamic state (e.g., by level of activation of cortical sites by inputs from subcortical sites [6]) partially determine which neuronal pools synchronize with each other. Relative lags between stimulated sites can be more complicated near the site of input: Some cells in the field of input can lead others nearby by a few milliseconds, particularly if the leading cells are particularly precisely responsive to the features of the stimulus; yet apparently precise synchrony appears between more distant sites and across most cells in the local population [7].

The mechanism(s) via which the synchronous oscillation is generated is a subject of controversy and it is now fairly widely accepted that multiple mechanisms may be involved [8]. In most instances results cannot be explained by concurrent synchronous inputs to the separated sites, although this sometimes plays a role [9,10]. A variety of other experimental and theoretical approaches have been made to the problem. These include the recognition that limit cycle oscillators representing single neurons can mutually entrain to form independent synchronized clusters [11,12]. Simulations also indicate that local lateral inhibition might entrain synchrony [13] and that nonoscillating interlocking chains of neurons (so-called synfire chains) can shift phase into synchrony [14]. Simulations of intercellular interactions in the hippocampus, which model excitatory and inhibitory neuronal interactions via specific chemical neurotransmitters in considerable detail, were found to account for both synchrony and specific patterns of firing as seen in in-vitro slices of this region of the brain [15]. However, these results also depend primarily on local interactions among inhibitory cells and encounter some difficulty in explaining the ubiquity of long-range synchrony mediated by excitatory connections.

All these attempts at elucidation and explanation have tended to concentrate upon interactions between specific cells, considered as interacting discretely with each other, while ignoring the fact of the embedding of these cells in a continuum of intercellular connections. These considerations raise the possibility that synchronous oscillations might be a continuum property of large fields of interconnected cells and might thus be best accounted for by continuum-field models of neuronal interactions. Such models have been developed primarily to account for traveling-wave properties,
and electrocortical activity more generally, which is usually recorded from the scalp as an electroencephalogram.

Prominent among models treating the properties of the extended coupled neuronal field are those of Freeman [16] and Nunez [17]. Recently we proposed a model based on similar assumptions, incorporating locally coupled excitatory and inhibitory neuronal populations, long-range excitatory connections, dendritic integration, and axonal time delays [18,19]. The resulting equations describe the spatial and temporal properties of a uniform cortex in the continuum approximation and permit simulations on any scale greater than that of the inhibitory neurons (a few tenths of a millimeter). Using these numerical simulations, it has been demonstrated recently that fields of zero-lag synchrony that reproduce general features of the experimental data can be readily generated [20]. The fields of zero-lag synchrony appear as part of a larger field of lag-correlated (traveling-wave) activity and do not depend upon synchrony of inputs, nonlinearity of the simulated neurons, or interactions between excitatory and inhibitory cells: Purely excitatory interactions are sufficient.

Recently, we proposed a nonlinear model of cortical dynamics [1], similar in physical basis to versions of the dynamical equations introduced earlier [18,19], but replacing their formulation in terms of Green’s functions by a wave-equation approach [1]. This model was not identical to the previous ones, but incorporated the same neurophysiology to a similar degree of approximation. This model allowed us to find cortical steady states and analyze their stability and to study the propagation and stability of small-amplitude cortical waves. In the present paper we use it to calculate analytically the response of the cortex to a finite number of point sources of stimulation. The results are used to show that synchronous oscillations arise naturally in the cortex and can be explained simply in terms of propagating waves.

The structure of this paper is as follows. In Sec. II we briefly review the wave-equation model and write down the linearized wave equation. We then derive the Green’s function for propagating cortical waves and use it to calculate correlation functions, levels and times of maximal correlations, and spectra of fluctuations excited by a finite number of point sources. In Sec. III we evaluate these expressions numerically for some representative cases and compare the results with direct solution of the full set of nonlinear cortical-dynamics equations. Comparisons with experimental results for synchronous oscillations are also discussed in Sec. IV.

II. THEORY

In this section we first outline the main relevant results of our recently developed wave-equation formulation of cortical dynamics [1]. These results are then applied to derive the Green’s function of propagating cortical disturbances, and the two-point correlation function and power spectrum of cortical fluctuations driven by a finite number of point sources. In all cases we restrict attention to regimes in which only stable waves exist [1]; cortical instabilities are not considered.

A. Dynamical equations

In a previous paper [1] we developed a set of nonlinear equations for cortical dynamics in the continuum limit. These equations incorporated excitatory and inhibitory neurons, dendritic integration of inputs to a given neuron, finite axonal propagation velocities, and the nonlinear relationship between inputs to a neuron and its firing rate.

The first of the central equations of our model is

$$Q_{e,i}(r,t) = \frac{1}{1 + e^{-c(V_{e,i}(r,t) - V_0)}},$$

(1)

which relates the mean firing rate $Q_{e,i}(r,t)$ of neurons (the pulse density in neurophysiological terminology) to the applied potential $V_{e,i}(r,t)$, where $e$ and $i$ denote the excitatory and inhibitory populations, and $C$ and $V_0$ are constants or order unity. Potentials are measured in units of the standard deviation of the distribution of neuronal firing thresholds.

The potential $V_{e,i}(r,t)$ at the point where conversion to neuronal pulses takes place results after inputs have been summed and filtered through the dendrites. A good approximation to $V_{e,i}(r,t)$ is given by

$$V_{e,i}(r,t) = g \frac{\alpha \beta}{\beta - \alpha} [U_{e,i}(r,t) - W_{e,i}(r,t)],$$

(2)

$$\frac{dU_{e,i}(r,t)}{dt} = Q_{ae,i}(r,t) - \alpha U_{e,i}(r,t),$$

(3)

$$\frac{dW_{e,i}(r,t)}{dt} = Q_{ae,i}(r,t) - \beta W_{e,i}(r,t),$$

(4)

where $Q_{ae,i}(r,t)$ represent arrival rates of input pulses to the dendrites, $g$ is a dendritic gain factor, and $\alpha$ and $\beta$ are constants parametrizing the dendritic response to an impulse. In effect, diffusion during dendritic propagation smears out the temporal response and the dendritic tree acts as a low-pass filter.

Outgoing pulses from each neuron propagate along its axon and axonal tree at a characteristic velocity $v$. Assuming an isotropic distribution of axons whose ranges have an approximately exponential distribution (see Ref. [1] for details), this propagation can be modeled by a wave equation for the corresponding potentials $\phi_{e,i}(r,t)$:

$$\left(\frac{\partial^2}{\partial t^2} + 2 \gamma_{e,i} \frac{\partial}{\partial t} + \gamma^2_{e,i} - u^2 \nabla^2\right) \phi_{e,i}(r,t) = \gamma^2_{e,i} Q_{e,i}(r,t),$$

(5)

where $\gamma_{e,i} = v/r_{e,i}$ and $r_{e,i}$ are the characteristic ranges of the axons [1].

The incident potentials $Q_{ae,i}(r,t)$ at a particular location comprise contributions from the wave potentials $\phi_{e,i}$ and inputs external to the cortex. These inputs are usually split into two components: a uniform mean nonspecific excitation $Q_{ns}$ that results from the total of inputs from noncortical structures in the brain and a specific excitation $Q_s(r,t)$ associated with the stimulus under investigation. Robinson et al. [1] defined $Q_{ns}$ to be constant in time and space, while
$Q_s(r, t)$ has zero spatial and temporal averages, leading to the final underlying equations of our model:

$$Q_{ns}(r, t) = M_e Q_s(r, t) + \mu_e Q_{ns} + a_{ee} \phi_e(r, t) - a_{ei} \phi_i(r, t),$$  \hspace{1cm} (6)

$$Q_{ni}(r, t) = M_i Q_s(r, t) + \mu_i Q_{ni} + a_{ie} \phi_e(r, t) - a_{ii} \phi_i(r, t).$$  \hspace{1cm} (7)

Here the constants $M_{e,i}$ determine the strength of coupling of specific inputs to excitatory and inhibitory neurons, $\mu_{e,i}$ are the coupling strengths for nonspecific impulses (i.e., the relevant fractional synaptic densities), and the coefficients $a_{nn}$ are the synaptic densities associated with excitatory and inhibitory inputs to excitatory and inhibitory neurons.  

### B. Wave equation

Robinson et al. [1] showed that Eqs. (1)–(7) have a stable low-$Q_e$ fixed point provided $Q_{ns}$ is not too large.  Linearizing the system around this fixed point, they wrote down an approximate wave equation for the excitatory wave potential $\phi_e$ alone, from which all other fields could be derived in the linear limit.  In Fourier space, this equation is

$$[D_e(\mathbf{k}, \omega) - F_e(\omega) a_{ee}] \phi_e(\mathbf{k}, \omega) = F_e(\omega) M_e Q_s(\mathbf{k}, \omega),$$  \hspace{1cm} (8)

where the cortex is driven by a specific input $Q_s(\mathbf{k}, \omega)$, $\mathbf{k}$ is the wave vector, $\omega$ is the angular frequency,

$$F_e(\omega) = gp_e \gamma_e^2 L(\omega),$$  \hspace{1cm} (9)

$$L(\omega) = \frac{\alpha \beta}{(\alpha - i \omega)(\beta - i \omega)},$$  \hspace{1cm} (10)

$$D_e(\mathbf{k}, \omega) = (\gamma_e - i \omega)^2 + k^2 v^2,$$  \hspace{1cm} (11)

$$\rho_e = C Q_e^{(0)} [1 - Q_e^{(0)}],$$  \hspace{1cm} (12)

and $Q_e^{(0)}$ is the equilibrium pulse density.

For freely propagating waves, Eq. (8) yields the dispersion equation [1]

$$(\alpha - i \omega)(\beta - i \omega) D_e(\mathbf{k}, \omega) - \alpha \beta \gamma_e^2 G = 0$$  \hspace{1cm} (13)

for $Q_s = 0$, with

$$G = gp_e a_{ee}.$$  \hspace{1cm} (14)

Equation (13) implies that only the excitatory field need be followed for low $Q_e^{(0)}$, with only excitatory quantities entering its dispersion equation.  This is reasonable given the preponderance of excitatory connections between neurons.  Robinson et al. [1] showed that this yields a good approximation to the dispersive properties of the model medium provided the wavelengths of the waves are much greater than a few tenths of a millimeter.  This is not a significant constraint in practice because typical wavelengths are a few centimeters in the human cortex.  They also showed that waves are stable for $\Gamma < 1$, which will be assumed in what follows here.  Wright [20] showed numerically that inhibitory-excitatory interactions can play a role in determining the so-called gamma-band (30–80 Hz) spectrum in a strongly stimulated cortex, but we will not consider this problem here.  We have recomputed most of the results of this paper with the inclusion of inhibition, finding only negligible changes for typical parameters, so we omit these refinements for simplicity.

### C. Green’s function

We are interested in calculating the correlation between $\phi_e(r, t)$ and $\phi_e(r’, t + \tau)$ as a function of $r$, $r’$, and $\tau$, a quantity that is experimentally measured.  Hence we first wish to calculate $G(r, \omega)$ to determine the effect of a point source of frequency $\omega$ at a range $r$.  Later we will integrate this quantity over a frequency spectrum and sum over point sources to obtain the correlation function in question.

In Fourier space, the Green’s function for the wave equation (8) is

$$G(\mathbf{k}, \omega) = \frac{L(\omega) \gamma_e^2 g p_e}{k^2 v^2 + (\gamma_e - i \omega)^2 - L(\omega) \gamma_e^2 G}.$$  \hspace{1cm} (15)

Using this relationship we can write

$$G(r, \omega) = \frac{L(\omega) \gamma_e^2 G}{a_{ee} v^2} \int \frac{d^2 k}{(2 \pi)^2} \frac{e^{ikr}}{k^2 + q(\omega)^2}$$  \hspace{1cm} (16)

$$= \frac{L(\omega) \gamma_e^2 G}{a_{ee} v^2} \int_0^\infty dk k^2 \int_0^{2\pi} d\theta \frac{e^{ikr \cos \theta}}{k^2 + q(\omega)^2}$$  \hspace{1cm} (17)

$$= \frac{L(\omega) \gamma_e^2 G}{2 \pi a_{ee} v^2} K_0[q(\omega)r]$$  \hspace{1cm} (18)

$$q(\omega) = \frac{1}{v} [(\gamma_e - i \omega)^2 - L(\omega) \gamma_e^2 G]^{1/2},$$  \hspace{1cm} (19)

where $J_0$ and $K_0$ denote Bessel and Macdonald functions, respectively.  Note that the root chosen in Eq. (20) must have $\text{Re} q(\omega) > 0$ for stable waves with $G < 1$; otherwise the Green’s function diverges with increasing $r$, which is unphysical.  For $G \approx 0$ and $\omega = 0$,

$$G(r, 0) = \frac{\gamma_e^2 G}{2 \pi a_{ee} v^2} K_0(\gamma_e r/v).$$  \hspace{1cm} (21)

Apart from the factor $G/a_{ee} = gp_e$, which represents the net gain in generating $Q_s$ from $Q_e$, this is simply the static Green’s function derived previously [1].

The coordinate-space Green’s function can be written in terms of Eq. (12) as
FIG. 1. Integrals $I(r,t)=2\pi a_v v^2 G(r,t)/\gamma_e^2 G$ from the Green’s function formula (22) vs $t$ for $r=0.1$ m, $v=9$ m s$^{-1}$, $\gamma_e=108$ s$^{-1}$, $a_v=0.853$, $Q_e^{(0)}=0.0103$, and $C=1.82$ with (a) $G=0.57$, $\alpha=\beta=2000$ s$^{-1}$; (b) $G=0.95$, $\alpha=\beta=2000$ s$^{-1}$; (c) $G=0.57$, $\alpha=\beta=400$ s$^{-1}$; and (d) $G=0.95$, $\alpha=\beta=400$ s$^{-1}$.

Unfortunately, the integral in Eq. (22) does not appear to be expressible in terms of tabulated functions, except in certain limiting cases. One important such case is the limit in which $\alpha, \beta \gg \omega$ and $G \approx 0$, where $L(\omega) \approx 1$ and

$$G(r,t)=\frac{\gamma_e^2 G}{2\pi a_v v^2} \int \frac{d\omega}{2\pi} e^{-i\omega t} L(\omega) K_0(q(\omega)r).$$

(22)

where $\Theta$ denotes a unit step function. This result is the Green’s function for $D_e(t)\delta(\omega)=0$, i.e., for the standard two-dimensional, damped wave equation.

Figure 1 shows the integral in Eq. (22), which we denote as $I(r,t)$, for fixed $r$ as a function of time for a variety of parameters. For large $\alpha$ and $\beta$ (short dendritic integration times) and small to moderate gain $G$, Fig. 1(a) shows that the Green’s function is sharply peaked just after the minimum propagation time $r/v=11$ ms to the point in question. This compares with the algebraic singularity of Eq. (23) in the limit $\alpha, \beta \to \infty$. For larger $G$, Fig. 1(b) shows that the Green’s function is increased in magnitude [even more so when $G(r,t)$ rather than the integral $I(r,t)$ is considered] and broadened in time, owing to a greater degree of “regeneration” of neural pulses at sites away from the origin when the gain is large. Corresponding results for smaller $\alpha$ and $\beta$ are shown in Figs. 1(c) and 1(d). Here the longer dendritic integration time leads to additional broadening of the response as a function of time and a consequent reduction in the peak magnitude of $I(r,t)$ relative to Figs. 1(a) and 1(c).

D. Correlation functions and spectra

The excitatory wave potential $\phi_e(r,\omega)$ at $r$ due to a monochromatic point source is given by the product of Eq. (19) with $Q_e(\omega)$. The unnormalized second-order correlation function between a potential $\phi_{em}$ due to a more general source at $R_m$ and $\phi_{en}$ due to a source at $R_n$ is given by

$$C_{mn}(r,r',\tau) = \langle \{ \phi_{em}(r,t) - \langle \phi_{em}(r,t) \rangle \} \{ \phi_{en}(r',t+\tau) - \langle \phi_{en}(r',t+\tau) \rangle \} \rangle$$

(24)

$$= \int \frac{d\omega}{2\pi} e^{-i\omega t} \phi_{en}(r,\omega) \phi_{en}(r',\omega),$$

(25)

$$= \left( \frac{\gamma_e^2 G}{2\pi a_v v^2} \right)^2 \int \frac{d\omega}{2\pi} e^{-i\omega \tau} L(\omega)^2 \times Q_e(\omega) Q_e(\omega) K_0[q(\omega)r_m] K_0[q(\omega)r_n],$$

(26)

where angle brackets denote an average over $t$ (over a time long compared to the phenomena of interest, which principally occur on time scales much less than 1 s), $r_m=|r-R_m|$, $r_n=|r-R_n|$, and we have set $M=1$ in Eq. (8) without loss of generality. Note that $Q_e(0)=0$ for all $m$ in accordance with the definition in Sec. II A, we have used the fact that $\phi_e(r,t)$ is real for all $r$ and there is assumed to be no correlation between different frequency components of any one source beyond the correlation implied by the reality of $\phi_e(r,t)$. Also, $K_0$ and $L$ both decrease at large $\omega$ for $G<1$ in Eq. (26).

In the limit with $\alpha, \beta \gg \omega$ and small $G$, one can approximate the integral in Eq. (26) for white-noise sources. Denoting this integral by $I_{mn}$, we find [21]

$$I_{mn}=\frac{\pi v}{2\sqrt{r_m r_n}} \exp[-\gamma_e(r_m+r_n)/v] \int \frac{d\omega}{2\pi} \frac{1}{(\gamma_e^2 + \omega^2)^{1/2}} \times \exp[-i\omega \tau + i\omega (r_m-r_n)/v]$$

(27)

$$=\frac{v}{2\sqrt{r_m r_n}} \exp[-\gamma_e(r_m+r_n)/v] \times K_0[\gamma_e \tau + (r_n-r_m)/v].$$

(28)

The result (28) is included for completeness, but is not used in what follows.

The total correlation function due to several sources is

$$C(r,r',\tau) = \sum_{m,n} C_{mn}(r,r',\tau).$$

(29)

The normalized correlation function can be written as

$$\tilde{C}(r,r',\tau) = \frac{C(r,r',\tau)}{[C(r,r,0)C(r',r',0)]^{1/2}},$$

(30)

which is unity if $r=r'$ and $\tau=0$. The variance of $\phi_e$ at $r$ is

$$\langle [\phi_e - \langle \phi_e \rangle]^2 \rangle = \text{var} [\phi_e(r)] = C(r,r,0).$$

(31)
For points near a particular source the correlation function and the variance are dominated by that source, owing to the singularity in $K_0(z)$ at small $z$ [22].

Useful limiting forms of the normalized correlation coefficient for two sources can be obtained when one source dominates at $r'$ or when both sources have the same amplitude there. Denoting the fluctuating part of $\phi_{en}$ by $f_n$ at $r$ and $f'_n$ at $r'$ and noting that these quantities are real, we can write

$$C(r, r', \tau) = \frac{\langle (f_1 + f_3)(f_1' + f_3') \rangle}{\langle (f_1 + f_2)^2 \rangle^{1/2} \langle (f_1' + f_2')^2 \rangle^{1/2}}.$$  (32)

If one source (say 1) dominates over the other at $r'$ and signals from the two sources have equal amplitudes at $r$, Eq. (32) then implies $\tilde{C} = \tilde{C}_{11}$ if the sources are uncorrelated and $\tilde{C} = \tilde{C}_{11}$ if they are completely correlated. (Here we use the notation $\tilde{C}_{11}$ to denote the normalized correlation due to source 1 alone.) This case applies when $r'$ is significantly closer to one source than the other. If the amplitudes of the signals from the two sources are equal at both points being correlated one always finds $\tilde{C} = \tilde{C}_{11}$.

The power spectrum $P(r, \omega)$ is the Fourier transform of $C(r, r, t)$:

$$P(r, \omega) = \sum_{m,n} \phi_{en}(r, \omega) \phi_{en}^*(r, \omega),$$  (33)

$$= \left( \frac{\gamma^2 G}{2 \pi a e \nu^2} \right)^2 \sum_{m,n} |L(\omega)|^2 Q_{mn}(\omega) Q_{mn}^*(\omega) \times K_0(q(\omega)r_m)K_0^*[q(\omega)r_n].$$  (34)

For the case of a point midway between two sources with identical power spectra

$$P(r, \omega) = A \left( \frac{\gamma^2 G}{2 \pi a e \nu^2} \right)^2 |L(\omega)Q_{1}(\omega)K_0[q(\omega)a]|^2,$$  (35)

where $2a$ is the separation of the sources, $A = 4$ if the sources are perfectly correlated, $A = 2$ if they are uncorrelated, and $A = 1$ for a single source. At points significantly closer to one source than the other, the spectrum is dominated by that of the closer source.

If we denote the argument of any Macdonald function in the previous expressions by $z$, the relationship $|\arg z| < \pi/2$ must hold for stable waves since otherwise the relevant expressions would diverge unphysically at large $r$ [22]. In this regime $K_0(z)$ can be rapidly evaluated numerically from the integral form

$$K_0(z) = \int_0^\infty e^{-z \cosh t} dt.$$  (36)

Provided its argument is not too small, the function $K_0$ can be approximated as

$$K_0(z) \approx \sqrt{\pi z} e^{-z},$$  (37)

which is very fast to evaluate and permits further analytic treatment in some limiting cases. For real $z$ at least, this approximation is semiquantitatively correct provided $\text{Re} z \approx 0.1$ and improves in accuracy for large $|z|$, with a fractional error of approximately $-1/8z$ [22].

Equations (34), (35), and (37) enable the asymptotic form of the power spectrum to be determined. For example, midway between two white noise sources one has

$$P(r, \omega) \sim \begin{cases} \omega^{-5}, & \omega \gg \alpha, \beta, \gamma_e \\ \omega^{-4}, & \gamma_e \gg \omega \gg \alpha, \beta \\ \omega^{-3}, & \beta \gg \omega \gg \alpha, \gamma_e \\ \omega^{-2}, & \beta, \gamma_e \gg \omega \gg \alpha \\ \omega^{-1}, & \alpha, \beta, \gamma_e \gg \omega \gg \alpha \\ \text{const}, & \alpha, \beta, \gamma_e \gg \omega \gg \alpha. \end{cases}$$  (38-43)

Note that $\beta > \alpha$ has been assumed without loss of generality. It is also worth noting that the strong inequalities in Eqs. (39)–(42) are seldom well satisfied in humans since $\alpha, \beta$, and $\gamma_e$ are typically of the same order [1].

Before proceeding, we stress that all the above analysis is for an infinite cortex. If a finite cortex is to be studied more accurately, we should replace Eq. (16) by a sum over all allowed wave vectors. This sum will then appear in subsequent formulas. However, if nonuniform ($k \neq 0$) modes are strongly damped (as our previous work has implied [1]), boundary conditions will not have a strong role because (i) the modes will be indistinguishable due to frequency broadening and (ii) the modes will not be much affected by boundaries if they largely dissipate before reaching them.

### III. NUMERICAL RESULTS

In this section we evaluate a range of properties of cortical oscillations driven by white-noise sources, including the correlation function, variance, and power spectrum. By exploring the effects of varying the cortical parameters and the degree of phase coherence between the sources, these results enable us to exhibit a robust candidate mechanism for what has been termed synchronous oscillation. A full parameter survey is not carried out since it is not needed for our main aim of establishing the existence of the key phenomena.

In this study we use a “canonical” set of parameters unless otherwise stated. These values are chosen for the purposes of illustration and to provide continuity with previous work. Except for $\alpha$ and $\beta$, the canonical values used are the same as in our previous work [1], namely, $v = 9 \text{ m s}^{-1}$, $\gamma_e = 108 \text{ s}^{-1}$, $g = 36$, $a_e = 0.853$, $\mathcal{Q}(0) = 0.0103$, $C = 1.82$, $G = 0.57$, and $\alpha = \beta = 400 \text{ s}^{-1}$.

We have adopted larger values of $\alpha$ and $\beta$ than previously to model more accurately dendritic integration with a mean response time of 5 ms (previously we had $\alpha = 100 \text{ s}^{-1}$ and $\beta = 350 \text{ s}^{-1}$, giving a peak response at 5 ms). The full Macdonald functions are used in evaluating analytic expressions from Sec. II, al-
though we have found that the approximation (37) is a good one for most purposes.

A. Correlated sources

Figure 2 shows results for driving by a pair of perfectly correlated white-noise sources of unit amplitude with the canonical parameters. The sources are correlated with each other, but different frequency components of a given source are mutually uncorrelated here and in all succeeding figures. Each source was constructed for the purposes of numerically evaluating the analytical expressions by choosing components of fixed amplitude but random relative phase, distributed equally in frequency from zero to a maximum frequency (≈10^5 s^-1) far higher than any inverse time scales relevant to the problem. In the present case, the phases of the two sources were chosen to be identical at each frequency to produce perfect correlation, but the relative phases of different frequency components were random.

In Fig. 2(a) we plot the time \( t_{\text{max}} \) of maximum positive correlation \( \tilde{C}(0,r',t) \) as a function of position \( r' \), with the origin at the center of the frame and the sources on the x axis a distance \( 2a = 0.1 \) m apart. The correlation function was determined by Fourier transforming the power spectrum and then \( t_{\text{max}} \) was found by searching directly for the global maximum. The uncertainty in \( t_{\text{max}} \) was less than 0.25 ms. The value of \( t_{\text{max}} \) increases toward the outside of the figure, reflecting the outward propagation of waves from the two sources (near which \( t_{\text{max}} \) is negative, with the innermost contour drawn at −10 ms). At large distances the contours approach circles, as expected for outward propagating waves from a single source (the sources cannot be distinguished at large ranges). Near each source \( t_{\text{max}} \) is negative because a given wave crest arrives there before it reaches the reference point at the origin, while the locus of zero-lag correlation is a figure-eight passing through the origin and determined by the interplay between the relative path lengths to the points in question and the relative amplitudes of the two signals.

One important point is that the correlation has the spatial dependence of self-correlations \( C_{\text{mm}} \) alone. This can be seen by substituting \( r_1 = r_2 = a \) into Eq. (26) for correlations relative to the origin, which gives

\[
C(0, r', \tau) = \frac{\sigma^2 \omega^2}{2 \pi a^2} \int \frac{d\omega}{2\pi} e^{-i\omega\tau} \left[ L(\omega) \right]^2 \times K_0[q(\omega)r] \left[ Q_{s1}(\omega) + Q_{s2}(\omega) \right] \times \left\{ Q_{s1}(\omega)K_0[q(\omega)r] + Q_{s2}(\omega)K_0[q(\omega)r] \right\}^\ast.
\]

We see from this expression that all the terms are of the same form as those in \( C_{11} \) and \( C_{22} \) and that both of these terms are important. The only role played by the relative correlation between the sources is to determine the number of factors of \( |Q_{s1}|^2 \) that will appear in the final expression. This result is in accord with the discussion following Eq. (35).

Figure 2(b) shows \( C_{\text{mm}} \), the maximum positive correlation \( C(0, r', t) \), as a function of \( r' \). A broad maximum of near-perfect correlation is seen near the center of the figure, decreasing slowly toward the edges. Naively, one might expect perfect correlation at all points for perfectly correlated sources. However, two-dimensional (2D) wave equations do not yield \( \delta \)-function propagators (see Fig. 1, for example) and this reduces the correlation below unity at large ranges. We explore this point further in Sec. III B. The remaining point visible in Fig. 2(b) is the pair of features at the sources. The lowered correlation here is due to the dominance of a single source, as explained following Eq. (32).

Figure 2(c) shows the variance \( C(r, r, 0) \) as a function of \( r \). This quantity is seen to fall off approximately exponentially with distance from each source at large distances (as seen by the nearly uniformly spaced logarithmic contours). This reflects the exponential distribution of axonal ranges. The two strong peaks are at the sources, where the response is singular.

The power spectrum (35), evaluated at the midpoint between the sources, is seen in Fig. 2(d). It exhibits a strong peak at low frequencies, falling to half maximum near 9.5 Hz. (Note that the zero-frequency component is zero in conformity with our definition.) At very low frequencies we find that the scaling (43) is approached, while the result (38) is approached at high frequencies. However, because of the relative proximity of the values of \( \alpha, \beta, \gamma \ast \), the exponents in Eqs. (39)–(42) are not clearly manifest.

Figure 3 shows results from a numerical simulation of the full nonlinear partial differential equations (1)–(7) carried out for the canonical parameters using the same methods as in our previous work [1], but omitting inhibitory effects in accord with the approximations made here. The driven points were at \( x = \pm a \), which for the 81×81 grid used here corresponds to ±7 grid units with respect to the center. The driving signal at these points \([Q_1 \text{ in Eqs. (6) and (7)}]\) was the same zero-mean Gaussian white noise, with a standard de-
viation in the time domain of 0.010 (much smaller than the steady-state driving), which was small enough to avoid nonlinear effects. Figure 3 is the result of a simulation lasting 132 s with a step size of 0.25 ms. Figure 3(a) shows that the variance (computed directly by accumulating mean and mean-square field values at each grid point) has a very similar form to that in Fig. 2(c), while Fig. 3(b) shows that the corresponding power spectrum (computed by compiling and then Fourier transforming a time series over the entire simulation) at \( r = 0 \) is also very similar to that in Fig. 2(d), with the half-power point occurring between 9 and 9.5 Hz, close to the value of 9.5 Hz in the canonical case. A fully linearized simulation also has been carried out with equivalent results, consistent with nonlinear effects not being important to the results shown in Fig. 3. The similarity between Figs. 2(c) and 2(d), on the one hand, and Figs. 3(a) and 3(b), on the other, demonstrates that boundary conditions are not important in determining the correlation and spectral properties of the wave fields in this case: Fig. 2 is for an infinite medium, while Fig. 3 was calculated using periodic boundary conditions. This insensitivity to boundary conditions arises because the waves are relatively strongly damped. Robinson et al. [1] showed that the width of modal resonances exceeded their separation for typical cortical parameters, obliterating the modal structure of spectra; this result evidently carries over to the correlation properties considered here.

**B. Single source**

Figure 4 shows a situation with the canonical parameters, except that the amplitude of the second source (at \( x = +a \)) has been reduced to zero. The \( t_{\max} \) plot in Fig. 4(a) is consistent with the two-source plot seen in Fig. 2(a) being the superposition of two single-source plots with the same phase, i.e., with self-correlations \( C_{nn} \), dominating in Eq. (29). The velocity of propagation of the point of maximum positive correlation, obtained from the ratio of the distance traveled to \( t_{\max} \), is approximately 8.5 m s\(^{-1}\). This is not the same as the velocity \( v \) of the waves because of the broadening of the Green’s function seen in Fig. 1, which weights times greater than \( r/v \) more strongly than in the limit \( \alpha, \beta \to \infty, \gamma \to 0 \).

Close analysis of the results in Fig. 4(b) shows that \( C_{\max} \) actually has a minimum at the source (the innermost contours shown correspond to values below 0.98, the value at the innermost labeled contour). This results from the imperfect correlation between points at different distances from the source owing to the non-\( \delta \)-function form of the 2D propagator. The maximum occurs on the circle \( r_1 = r_2 = a \). The variance plot in Fig. 4(c) is accordingly modified and the power spectrum in Fig. 4(d) is reduced by the expected factor of four relative to that in Fig. 2(d) [cf. Eq. (35)].

**C. Uncorrelated sources**

Figure 5 shows results for a pair of equal-amplitude uncorrelated (i.e., whose relative phase at a given frequency is a random number chosen between 0 and 2\( \pi \), in addition to all frequency components having random relative phase, as before) sources having the canonical parameters. In accordance with the remarks in the preceding section, the plot

![Figure 3](image1.png)  
**FIG. 3.** Wave quantities from numerical simulations of the full nonlinear equations (1)–(7) for the same parameters as in Fig. 2, but expressed in different arbitrary units. (a) Variance \( \log_{10} C(r, r, 0) \) vs position. (b) Power spectrum at the origin vs frequency.

![Figure 4](image2.png)  
**FIG. 4.** Wave quantities for a single source and the canonical parameters listed at the start of Sec. III. (a) Time of maximum correlation \( t_{\max} \) (ms) vs position. (b) Maximum correlation \( C_{\max} \) vs position. (c) Variance \( \log_{10} C(r, r, 0) \) vs position. (d) Power spectrum at the origin vs frequency.

![Figure 5](image3.png)  
**FIG. 5.** Wave quantities for two uncorrelated sources and the canonical parameters listed at the start of Sec. III. (a) Time of maximum correlation \( t_{\max} \) (ms) vs position. (b) Maximum correlation \( C_{\max} \) vs position. (c) Variance \( \log_{10} C(r, r, 0) \) vs position. (d) Power spectrum at the origin vs frequency.
of $t_{\text{max}}$ in Fig. 5(a) is almost identical to Fig. 2(a) in this case, consistent with our picture of outwardly propagating waves whose time of maximum correlation with the origin is determined by self-correlations rather than the degree of cross correlation between the sources.

The plot of $C_{\text{max}}$ in Fig. 5(b) is very different from the corresponding plot in Fig. 2(b). Near-perfect correlation is seen near the origin, falling off slowly along the $y$ axis at a similar rate to that in Fig. 2(b), but much more rapidly in the $x$ direction. Both features are consistent with the discussion following Eq. (32). (i) On the $y$ axis, the amplitudes of the two signals are equal and $C = \tilde{C}_{11}$. (ii) Near either source, that source’s signal will dominate the signal from the other source. Hence we expect $C \approx \tilde{C}_{11} / \sqrt{2} \approx 1 / \sqrt{2}$, consistent with the numerical value of just under 0.7. (iii) Far from both sources in the $x$ direction, the nearer source will again dominate and we expect the correlation function to decline toward $C_{11} / \sqrt{2}$, which is again consistent with Fig. 5(b).

The variance plot in Fig. 5(c) is very similar to that in Fig. 2(c), except that the variance at large distances from the sources is smaller owing to their lack of correlation. In Fig. 5(d) we see that the power spectrum fluctuates between zero and the value for perfectly correlated sources, with an average of half that value. The fluctuations occur because the relative phase between the sources at a given frequency has a single, randomly chosen value: phase is not averaged over at each frequency, although this could be done.) When smoothed with respect to frequency (or, equivalently, when averaged over many realizations of the phase distribution), the results in Fig. 5(d) correspond to the predicted value $A = 2$ in Eq. (35).

D. Effect of cortical gain

Reduction of $G$ from 0.57 in the canonical case to 0.015 ($g = 1$) scarcely changes the plot of $t_{\text{max}}$ in Fig. 6(a) from the corresponding one in Fig. 2(a), consistent with dominance of self-correlations. In Fig. 6(b) the region of near-perfect correlation is extended relative to Fig. 2(b), a consequence of reduction in magnitude of the tail of the Green’s function with decreasing $G$ that was seen in Figs. 1(c) and 1(d). Consistent with this interpretation, the velocity of the point of maximum correlation is found to be almost indistinguishable from $v = 9 \text{ m s}^{-1}$ in Fig. 2(b). Figure 6(c) shows that the variance is reduced relative to that in Fig. 2(c), principally owing to the first $G$-dependent factor in Eq. (26). However, the shapes of the contours are very similar. Likewise, the spectral power is reduced by approximately the same factor in Fig. 6(d). Here we also see that the half-power frequency is about 23 Hz, compared to 9.5 Hz in the canonical case. Robinson et al. [1] showed that these low-frequency waves become less damped as $G$ increases, with the least damping at $f = 0$. Hence one can expect a narrower spectral profile for small $G$.

In Fig. 7(a) we find that the central regions of the $t_{\text{max}}$ plot are only slightly changed from Fig. 2(a), although $g = 60$ and $G = 0.95$ in this case. In the outer regions, however, the contours of $t_{\text{max}}$ undergo a sharper changeover from near-field to far-field behavior. Consistent with the temporal broadening of the Green’s function for large $G$, seen in Fig. 1(d), Fig. 7(b) shows that the correlation function decreases more steeply at large distances than in Fig. 2(b). The variance plot in Fig. 7(c) is similar to that in Fig. 2(c), although the values are higher because of the larger value of $G$. The nearness to an instability at $f = 0$ (which sets in at $G = 1$ [1]) causes the power spectrum to be very strongly peaked at low frequencies, with Fig. 7(d) showing the half-power point at only 2 Hz.

E. Variation of relative phase of correlated sources

A series of runs has been done in which the phase of source 2 (at $x = + a$) has been advanced relative to source 1

FIG. 6. Wave quantities for two correlated sources and the canonical parameters listed at the start of Sec. III, except that the gain $G = 0.015$. (a) Time of maximum correlation $t_{\text{max}}$. (b) Maximum correlation $C_{\text{max}}$ vs position. (c) Variance $\log_{10} C(r, r, 0)$ vs position. (d) Power spectrum at the origin vs frequency.

FIG. 7. Wave quantities for two correlated sources and the canonical parameters listed at the start of Sec. III, except that the gain $G = 0.95$. (a) Time of maximum correlation $t_{\text{max}}$. (b) Maximum correlation $C_{\text{max}}$ vs position. (c) Variance $\log_{10} C(r, r, 0)$ vs position. (d) Power spectrum at the origin vs frequency.
(at $x = -a$), while maintaining their correlation. Figures 8 and 9 show results for phase advances of $\pi/4$ and $\pi/2$, respectively.

As the phase difference increases, Figs. 8(a) and 9(a) show that the zone of negative $t_{\text{max}}$ around source 2 increases, while that around source 1 decreases and eventually disappears for a phase difference of roughly $\pi/2$. Likewise, Figs. 8(b) and 9(b) show that the correlation plots become increasingly asymmetric owing to the changed relative lag introduced by the phase shift of source 2. All of these trends are consistent with our picture of the dominance of the two single-source self-correlations in determining the synchrony properties of the cortex.

F. Effect of dendritic integration

Figure 10 shows results for a case in which we have set $\alpha = \beta = 2 \times 10^3 \text{ s}^{-1}$ to minimize the effects of the dendritic integration time. (We do not argue that a value this large would be appropriate for a real cortex.) The results in Figs. 10(a) and 10(c) are very similar to those seen in the canonical case. The correlation plot in Fig. 10(b) shows a narrower zone of very high correlation relative to Fig. 2(b), owing to the narrower form of the Green’s function for high $\alpha$ and $\beta$ [see Fig. 1(a)], which requires a closer match between path lengths from the two sources for high correlations to exist. One should note that the sources are surrounded by regions of lowered correlation [cf. the discussion after Eq. (32)].

The power spectrum in Fig. 10(d) has a much longer high-frequency tail than the canonical one, with the half-power point at 12 Hz. This is consistent with Eq. (42) due to the relative unimportance of dendritic low-pass filtering in this case.

IV. DISCUSSION AND CONCLUSIONS

We have studied the dynamics of a cortex driven by a finite number of white-noise sources using our wave-equation formalism. We have computed Green’s functions, power spectra, fluctuation levels, and correlation functions analytically and numerically. The results obtained here reproduce the main features of zero-lag correlations (synchronous oscillation) observed physiologically and defined to be cross correlations that are maximal at zero time lag. They confirm numerical results reported by Wright [20] and extend these earlier results by showing that zero-lag cross correlation between two excited sites on the simulated cortical surface can be described by an expression in which only propagation time lags and wave amplitudes in the linked excitatory elements are of relevance. They show that the most important contributions are from self-correlations of each source with itself (in general both sources’ self-correlations are significant at a given point), regardless of the degree of mutual correlation between sources. This result thus carries over immediately to multiple sources. Zero-lag cross correlation of this sort is therefore expected to be a widely occurring property of neural nets, including real cortical neurons.

It should be stressed that the zero-lag cross-correlation results that we have found do not result from the existence of very weakly damped or growing global eigenmodes of the cortex: Indeed, there are no such modes under the circumstances discussed. Nor does it depend strongly on local inhibition in the regimes investigated. Qualifications to these conclusions are discussed in points (i)–(iii) below.

Since the determination of neuronal couplings to high accuracy is very difficult in most experimental situations, it
should be noted that our model makes predictions about the generation of nonzero-lag wave motions in the neural field around the locus of zero-lag synchrony. The occurrence of traveling waves surrounding the fields of synchrony should be readily demonstrable, but has not been looked for in any experiment to date, to the best of the authors’ knowledge. Such surrounding wave motion might be more readily observed in local-field potentials or electrocorticogram rather than in pulse activity, given the problems of adequate sampling in ongoing pulse trains.

The present model is not in contradiction with alternate mechanisms of synchrony and synchronous oscillation, such as nonlinear phase locking or cortical driving by synchronous inputs [9–11,13]. Both these processes may supplement the present mechanism. While zero-lag synchrony can emerge in the absence of any synchrony in the inputs, this is not a necessary condition. The essentially linear mechanism described here might also act to entrain local nonlinear phase locking among individual neurons. Likewise, there is no necessary contradiction with observations of intrinsic rhythmicity of firing in single cells, often associated with synchronous oscillation [9,10,23]. Certain models of learning depending on changes in synaptic strength indicate that cells with intrinsic rhythmicity would emerge as a consequence of induced rhythmic and synchronous firing within a population of real neurons [24].

Qualifications to our work are that real neurons have complicated properties in addition to those represented approximately in our model and that experimental observations of real neural synchrony are made in circumstances much more elaborate than can be represented by the introduction of noise inputs to two points on a plane. Nonetheless, this model appears to provide a mechanism for the occurrence of zero-lag maximum cross correlations in physiological observations, predicts a range of other observable quantities (e.g., variance and power spectra), and might be extended to cover a wider range of the relevant phenomena when appropriate allowance is made for complicating factors such as the following points.

(i) To a first approximation, our results appear adequate to explain the principal qualitative physiological finding upon which most subsequent work has been based, i.e., excitation of two points on the cortical surface by distinct, unrelated inputs causes activity on a locus in the vicinity of the inputs to attain zero-lag cross correlation [5]. The present match to experiment is qualitative only and no regard has been given to the inhomogeneity of real interneuronal coupling. The precise details of intercellular couplings, delay, and synaptic gain factors (which are not routinely measured) need to be determined in physiological experiments before our model can be fully quantitatively tested. A further physiologically important factor we have omitted is that of nonlinear dynamic influences on gain. Such effects can be nonuniform and will occur where stimuli are large enough to yield significant nonlinear responses (linear terms are included here). In physiological terms, nonuniform gain may arise from the action of the reticular activating system [6,20].

(ii) Although our results show that local inhibition is a small effect, it should be noted that the inhibitory connectivities are not precisely known. Also, the effectiveness of inhibition may be modulated via dynamical feedback via chemical neurotransmitters and neuromodulators acting on time scales ranging from milliseconds to seconds and longer. These effects, which are likely to depend strongly on brain state (e.g., attentive, relaxed, and sleeping), must be included in a full model of electrocortical wave activity and we are currently in the process of generalizing our model to incorporate them. It is possible that these effects could allow the cortex to cross the linear stability threshold \((G=1)\) into regimes of nonlinear dynamics, including limit-cycle oscillations or chaos, for example, without necessarily entering a state of near-maximum \(Q\). Weakly damped or growing waves (and hence the precise boundary conditions) would likely play a more important role in such regimes than for the parameters considered here.

(iii) No account is given in the present model for the frequent occurrence of gamma-band oscillation in association with pulse synchrony. This association has been partially accounted for by Wright [20] and will be considered further in future work.

ACKNOWLEDGMENTS

P.A.R. and C.J.R. thank the University of Iowa and the Mental Health Research Institute of Victoria for their hospitality during visits on which some of this work was undertaken. This work was supported by the Australian Academy of Science and the Ross Trust, Melbourne.

Simulation of EEG: dynamic changes in synaptic efficacy, cerebral rhythms, and dissipative and generative activity in cortex

J.J. Wright
Mental Health Research Institute of Victoria, 155 Oak Street, Parkville, Melbourne, Victoria 3052, Australia

Received: 25 May 1998 / Accepted in revised form: 1 March 1999

Abstract. A simulation of electrocortical activity based upon coupled local aggregates of excitatory and inhibitory cells was modified to include rapid dynamic variations of synaptic efficacy attributable to reversal potentials and related effects. The modified simulation reproduces the rhythmic phenomena observed in real EEG, including the theta, alpha, beta and gamma rhythms, in association with physiologically realistic pulse densities. At high levels of cortical activation, generative activity with a 40-Hz center frequency emerges, suggesting a basis for the occurrence of phase changes and “edge of chaos” dynamics. These local oscillation properties complement the dissipative traveling wave and synchronous oscillation effects attributable to longer range excitatory couplings, as previously demonstrated in related simulations. Results of variation of parameters provide a first approximation to the anticipated effects of slow physiological time variations in gains and lags, and some predictions of the model are described.

1 Introduction

1.1 Cerebral rhythms

The origins and functional significance of the cerebral rhythms revealed in the electroencephalogram (EEG) remain controversial. The EEG signals reflect the local space average of extracellular potential – local field potential (LFP) – arising from cortical dendrites (Mitzdorf 1988). Since dendritic potentials are directly involved in cortical signal processing, LFP, electrocorticogram (ECoG) and EEG reflect the processing of information at increasing scale and increasing spatial smoothing, although it remains unclear to what extent and in what ways the information processing is observable by this means. There is clearly a broad association of EEG features with level of consciousness, and suggestive evidence that EEG mirrors the cognitive state in more complex ways (John et al. 1969; Gevins et al. 1983; Picton and Hillyard 1988). This paper is concerned with both the emergence of the cerebral rhythms as collective properties of cortical neurons, and also with a putative mechanism via which complex information processing in the cortex may depend on these rhythmic activities.

The cerebral rhythms revealed by the EEG occur at restricted frequency bands (alpha, beta, gamma, theta, etc.), and are typically present singly or in combination, changing their character toward higher frequencies at increasing levels of alertness (e.g., Walter et al. 1967). Notably, the center frequencies of theta, alpha, beta, and gamma activity approximately double as the frequency spectrum is ascended (e.g., Kandel and Schwartz 1985) suggesting some sort of harmonic relationship. Their origins have been variously ascribed to neurons with intrinsic oscillation properties (e.g., Llinas et al. 1991), subcortical driving of the cortex from rhythmic systems of subcortical neurons (Steriade et al. 1990, 1993), local resonances (Freeman 1991; Traube et al. 1996; Rennie et al. 1998), the occurrence of linear standing waves of local field potential arising from travelling waves under the constraints of the closed geometry of the cortex (Nunez 1981, 1989, 1995), and to emergence of stochastic group resonances in large fields of essentially nonlinear elements (e.g., Wright 1990). They have been modelled by a variety of continuum and discrete processes, in both linear and nonlinear formulations, as is subsequently discussed.

Historically, a major problem in understanding the relationship of EEG to individual cellular events and the processing of information has been perceived to lie in the absence of correlation of action potentials with LFP (Stryker 1989). This discrepancy has been partially resolved by Freeman’s demonstration of a correlation between LFP and pulse density in the olfactory system (Freeman 1975) and by the findings of Eckhorn et al. (1988) and Gray and Singer (1989) who
have shown that when gamma rhythm (about 40 Hz LFP oscillation) is present, action potential density is phase-locked to the LFP. Freeman (1991) suggested that this 40-Hz pulse/wave association might be simply accounted for by alternating cycles of excitation and return inhibition, as excitatory cortical cells interact with local inhibitory cells. While work has been directed to stochastic neural networks with physiologically realistic low firing rates (Amit et al. 1990; Amit 1998), it has not yet been explained why low average pulse rates are associated with EEG rhythms at frequencies much lower than 40 Hz.

The picture has been further complicated by the discovery of synchronous oscillation (zero-lag positive cross-correlation, commonly of approximately 40-Hz periodicity) which develops between concurrently active cortical neuronal fields (Gray et al. 1989; Singer 1994; Singer and Gray 1995).

1.2 Self-organizing properties in neural networks

Models relating the origin of the EEG to cerebral dynamic events have usually treated the neuronal medium as a passive system for the propagation of wave-like activity – a notable exception being Freeman’s (1975) simulations of olfactory bulb activity. Network derivatives of Freeman’s model store and recall specific perceptions in limit cycle attractors (Baird 1986). It is possible that analogous processes occur more widely in the cerebral cortex. No successful model has yet emerged to relate the cerebral rhythms in general to information processing, storage and retrieval, although moves in this direction are apparent in models which treat neural dynamics in statistical mechanics terms (e.g., Wilson and Cowan 1973; Amari 1983; Peretto 1984; Grobler and Barna 1996; Arbib et al. 1998).

A proposal which may have major consequences for the dynamics of the cortex has been advanced by Langton (1986, 1990). This concept arises from von Neumann’s (1949) work on universal computation in cellular automata, and on Wolfram’s (1984) classification of artificial life simulations. Langton argues that brain-like systems capable of universal computation must exhibit a class of dynamics so balanced that the overall dynamics are associated with the onset of chaos, but do not escape into global chaos. Such systems would involve fluxes of energy and information between energy-gaining and energy-absorbing local states, and crucially in Langton’s view, the alternation of the energy state is analogous to the change of phase in thermodynamic systems. At the critical point of change of phase, Shannon entropy and mutual information between system elements reach a maximum and the potential capacity for universal computation emerges.

If Langton’s concept also applies to the brain, a distinct change of energy state analogous to a thermodynamic change of phase may be expected to occur at critical levels of excitation of cortical systems. It can be asked in what way this transition might be observable in the EEG.

1.3 Modelling issues

Attempts to address the modelling of cerebral dynamics are both enriched and afflicted by the wealth of neurophysiological detail which has been obtained, and the many deficits in quantitative data which yet remain. A major problem is to determine the necessary degree of cellular detail to describe the relevant events. Contrasts between the complex and time-varying properties of synapses and neuromodulation observed by physiologists and the fixed synaptic strength models studied in most computational models have been emphasized by Marder (1998).

Models of synchronous oscillation provide examples of the dilemmas of determining the degree of relevant detail. These range from the work of Traub, Whittington and colleagues (Whittington et al. 1995, Traub et al. 1996) who invoke details of gamma amino butyric acid (GABA) receptors and inter-inhibitory interactions, through models depending to various degrees on more abstract non-linear principles (e.g., Nischwitz and Glunder 1995; Arnoldi and Brauer 1996; Tass and Haken 1996; Wang and Buzsaki 1996) to recent findings showing by numerical and analytic means that synchrony is an inevitable property of any large network of neuron-like delay elements, and need not involve synchronous inputs, nonlinear mechanisms, or inhibitory cells at all (Wright 1990; Robinson et al. 1998). While these findings partially resolve controversies around the mechanism of synchrony (Konig et al. 1995, 1996), they do not deal with the relationship between synchrony and 40-Hz oscillation. Neither these nor other models of the cerebral rhythms capture rhythms at all relevant frequencies in a single model, nor their detailed relation to pulse densities, and it is not apparent to what degree cellular models and the various continuum models (Wilson and Cowan 1973; Freeman 1975; Zetterberg et al. 1978; Nunez 1981; van Rotterdam et al. 1982; Zhadin 1994; Hernandez et al. 1996; Jirsa and Haken 1996) are mutually compatible, each with any or all others.

The continuum models generally lack consideration of cortical-subcortical interactions and anatomical detail, as well as physiological complexity, but offer simpler insights into basic physical mechanisms, and are susceptible to later development by addition of more detailed features. Thus, a type of continuum approach is applied here.

1.4 Object

The model presented here retains features from its direct predecessors which sustain synchronous oscillation and travelling waves (Wright 1997; Robinson et al. 1998). The physiology of cortical circuitry is reduced to a system of coupled elements with gain and lag operations. Parameter values approximate physiological values where these are known, and normalizations enable some of the quantitative uncertainty to be provisionally handled. In the present work, coupling parameters were
enabled to undergo rapid time variation by an effect upon synaptic efficacy analogous to the influence of membrane reversal potentials – with the intention of subsuming within this description other rapid synaptic feedback processes which as yet remain poorly understood.

This formulation enables numerical demonstrations aimed at the following questions.

1. Can such a lumped continuum description account for all of the major cerebral rhythms as cortical resonances, using a single set of parameter values?
2. If so, then what respective roles do local feedbacks and travelling waves play in the generation of resonances and synchronous fields?
3. Are such resonances highly nonlinear, or essentially linear processes? Under what circumstances do they tend to either extreme?
4. What are the resultant relations between pulse density and LFP fluctuations at the major frequencies of resonance or oscillation, and do these approximate physiological observations?
5. In what ways may EEG features be linked to generative activity and the release of information stored in synaptic connectivity? In what conditions does dissipative activity and passive transmission of information predominate? Do both generative and dissipative processes interact via phase changes, as foreshadowed by Langton?

Earlier work summarized in Wright and Liley (1996) and Wright (1997) used similar state equations, save for the introduction of reversal potential effects. The account given of the EEG spectrum was incomplete and inaccurate, and dynamic properties were restricted to simple fixed points. As well as amending the earlier work by the introduction of local feedback effects in cortex, the opportunity is taken here to correct some other oversimplifications and numerical errors. By discrete variation of parameters, an attempt is made to ascertain both robustness of basic properties, and possible effects of slow time variation of parameters, as might arise if more complex physiological processes were introduced to the model.

2 Methods

The simulation to be presented here as the standard case has parameters chosen to accord with the following assumptions about the overall neuronal population properties.

1. Most synapses are located on the distal dendritic tree, so the mean delay from synapses to the soma is near the upper limit of physiological estimates for dendritic delay (Segev 1995; Thomson 1997).
2. Dendritic delays in excitatory and inhibitory cells, and from excitatory and inhibitory synapses, are comparable. This simplification is extreme (Thomson et al. 1996) but is here applied as a first approximation.
3. Axonal delays are range dependent, but are always small compared to dendritic delays.
4. There is a finite probability that an action potential may be emitted by a given neuron, even at membrane potentials very close to the inhibitory reversal potential.
5. Synaptic gains are in the ratio of 4:1 for inhibitory synapses versus excitatory synapses (Segev 1995) and produce a high signal amplification (Thomson 1997) so the stable operating range of cortical activity is restricted to low pulse densities.

2.1 State equations

Basic state equations are as in Wright (1997) with modification of dendritic time response as introduced in Robinson et al. (1997).

Let the $N$ cells in unit volume each have a probability of emission of an action potential $q_i$ as a function of the LFP, $V(t)$ at time $t$. Then in a mean-field approximation the pulse-probability density $Q(t)$ is given by

$$Q = \frac{1}{N} \sum_{i=1}^{N} q_i(V) .$$

By the central limit theorem, for large $N$, $Q$ will have a Gaussian distribution with respect to $V$, so $V$ and $Q$ are approximately related by

$$Q = (1 + e^{a(V-3)})^{-1} ,$$

and where $a = -\pi/\sqrt{3}$, LFP voltage units (vu) are approximate to standard deviations of the distribution of cell pulse probability over the complete range of LFP. This sigmoidal relation thus meets the requirement that $Q$ has a value close to zero where $V = 0$, and $Q$ approaches an asymptote of maximum pulse rate at $V = 6$ vu.

The time response of mean membrane potential (and by implication LFP and soma potential) is given by

$$V(t) = g \sum_{j=1}^{n} w_j Q_a(t - j\Delta t) \quad j = 1, 2, 3, \ldots, n ,$$

where $g$ is synaptic gain, $Q_a$ is affrent pulse action density, $\Delta t$ is the discrete time step, and $n\Delta t$ is large compared to the peak time response of membrane potential. In accord with Robinson et al. (1997)

$$w_j = b^2 j\Delta t e^{-b/j\Delta t} ,$$

represents the rise and fall of the membrane potential in response to input at $t = 0$, incorporating lags due to both synaptic conduction and average dendritic cable delay in a single function. Parameter $b$ regulates both the peak time and mean delay associated with this lag.

Within unit volumes, both excitatory and inhibitory cell groups are distinguished, each reciprocally and self-coupled, and each coupled at a longer range to other unit volumes by cortical-cortical fibers. Delays due to
axonal conduction between unit volumes are given by $\Delta t = r_{pq}/v$, where $\Delta t$ is axonal conduction lag over the distance $r_{pq}$ between the $p$th unit volume and the $q$th unit volume and axonal conduction velocity is $v$.

Coupling strengths are proportional to:

1. The fractional density of synaptic couplings afferent to the dendrites of excitatory and inhibitory cells respectively ($x_{ec}, \beta_{ei}, \mu_{ei}, M_{ec}$ etc., as listed in Table 2).

2. The synaptic gains of excitatory and inhibitory synapses, $g_e$ and $g_i$.

The present simulation introduces:

3. Changes in synaptic efficacy, $E'$, representing feedback effects including those of reversal potentials (Kandel and Schwartz 1985). That is, with increasing depolarization of cell membranes there is an increase in the sensitivity to inhibitory synaptic inputs and a decrease in sensitivity to excitatory synapses. These feedback relations are modelled as linear regressions of efficacy with membrane potential

$$E_{ec} = (1 - V_{e(p)}(t - \Delta t)/V_{ec}),$$

$$E_{ei} = (1 - V_{i(p)}(t - \Delta t)/V_{ec}),$$

$$E_{ie} = (1 - V_{i(p)}(t - \Delta t)/V_{ec}),$$

$$E_{ii} = (1 - V_{i(p)}(t - \Delta t)/V_{ec}),$$

where the subscripts $e$ and $i$ indicate excitatory and inhibitory potentials, and subscript $R$ a constant-valued reversal potential. Efficacies $E'$ were applied with smoothing, so that in each case $E(t) = \sum_{i=1}^{n} u_j E' (t - j\Delta t)$ where $u_j = ce^{-j/\Delta t}$, describing an exponential decay of the impact of instantaneous membrane potential upon synaptic efficacy. For high values of $c$ this decay is rapid as would be expected for reversal potentials alone. By setting $c$ lower, effects of putative slower pre- and postsynaptic feedbacks can be imitated.

State equations for the cortical system are then given for the $p$th unit volume by

$$Q_{ec(p)} = (1 + e^{(V_{ec(p)} - 3)})^{-1},$$

$$Q_{i(p)} = (1 + e^{(V_{i(p)} - 3)})^{-1},$$

$$V_{ec(p)} = \sum_{j=1}^{n} w_j Q_{ec(j)}(t - j\Delta t),$$

$$V_{i(p)} = \sum_{j=1}^{n} w_j Q_{i(j)}(t - j\Delta t),$$

where $Q_{ec(p)}, Q_{i(p)}$ are afferent synaptic action densities for excitatory and inhibitory cells in the $p$th unit volume, receiving local synaptic input at negligible axonal delay and delayed cortico-cortical inputs from $q$th unit volumes at range $r_{pq}, q = 1 \ldots n$ in accord with

$$Q_{ec(p)} = g_e \beta_{ec} E_{ec(p)} Q_{ec(p)} - g_i \beta_{ei} E_{ii(p)} Q_{i(p)} + g_e M_{ec} E_{ec(p)} Q_{i(p)} + g_{ei} \mu_{ei} E_{ec(p)} Q_{i(p)}$$

$$+ g_e \sum_{q=1}^{n} x_{ec}(r_{pq}) E_{ec(p)} Q_{ec(q)}(t - r_{pq}/v),$$

$$Q_{i(p)} = g_i \beta_{ei} E_{ii(p)} Q_{i(p)} - g_e \beta_{ei} E_{ii(p)} Q_{ec(p)} + g_i M_{ei} E_{ii(p)} Q_{e(p)}$$

$$+ g_e \sum_{q=1}^{n} x_{ii}(r_{pq}) E_{ii(p)} Q_{i(q)}(t - r_{pq}/v),$$

where $x_{ec}(r_{pq})$ and $x_{ii}(r_{pq})$ are partial input synaptic densities, such that $\sum_{q=1}^{n} x_{ec}(r_{pq}) = x_{ec}$ and $\sum_{q=1}^{n} x_{ii}(r_{pq}) = x_{ii}$. $Q_e$ and $Q_i$ are system inputs. $Q_e$ represents all time-varying components in specific cortical afferents and $Q_{ns}$ is a uniform DC input representing nonspecific cortical activation, which acts as control parameter.

2.2 Standard parameters and definition of units

State variables and parameters, their dimensions and standard values as applied in the following simulations are given in Tables 1 and 2. Where possible, physiologically accurate values have been applied. Since some crucially relevant parameters have not yet been objectively estimated with sufficient accuracy, in the following studies normalized units of LFP voltage have been used as described in Sec. 2.1, and some parameters ascribed arbitrary estimated values in these units. The justification for this approach lies in the match to established EEG properties, and the robustness of the findings under variation of the parameter values.

Synaptic densities were estimated by methods reported in Liley and Wright (1994). The nomenclature adopted reflects Breitenberg and Shuz’s (1991) association of apical and basal dendritic systems with cortico-cortical ($\alpha$ type) and intracortical ($\beta$ type) axo-synaptic connections.

Table 1. State variables and standard parameters other than synaptic densities LFP Local field potential, PPD pulse probability density

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_e$</td>
<td>Excitatory LFP</td>
</tr>
<tr>
<td>$V_i$</td>
<td>Inhibitory LFP</td>
</tr>
<tr>
<td>$Q_e$</td>
<td>Excitatory PPD</td>
</tr>
<tr>
<td>$Q_i$</td>
<td>Inhibitory PPD</td>
</tr>
<tr>
<td>$a$</td>
<td>Slope parameter</td>
</tr>
<tr>
<td>$b$</td>
<td>Dendritic time constant</td>
</tr>
<tr>
<td>$g_e$</td>
<td>Excitatory gain</td>
</tr>
<tr>
<td>$g_i$</td>
<td>Inhibitory gain</td>
</tr>
<tr>
<td>$c$</td>
<td>Decay time constant</td>
</tr>
<tr>
<td>$v$</td>
<td>Axonal velocity</td>
</tr>
<tr>
<td>$\sqrt{\tau}$</td>
<td>SD of axonal range</td>
</tr>
<tr>
<td>$V_{ec}$</td>
<td>EPSP reversal</td>
</tr>
<tr>
<td>$V_{ir}$</td>
<td>IPSP reversal</td>
</tr>
<tr>
<td>$Q_{ns}$</td>
<td>Nonspecific input</td>
</tr>
<tr>
<td>$Q_s$</td>
<td>Specific input</td>
</tr>
</tbody>
</table>
Table 2. Synaptic couplings subscripts ee, ei, etc., indicate synapses between cell types, excitatory to excitatory, excitatory to inhibitory, etc. Types of coupling are: α (cortico-cortical connections), β (intracortical connections), μ (nonspecific cortical afferents) and M (specific afferents). Synaptic density fraction is the proportion of synapses of each type in unit cortical volume. [The exact values used in the simulations are given for completeness (Liley and Wright 1994) although the precision given is greater than is justified from the anatomical data.] Afferent fraction is the proportion of synapses on the excitatory or inhibitory cell dendrites respectively, and are thus the values applied in Eq. (5).

<table>
<thead>
<tr>
<th>Synaptic coupling</th>
<th>Synaptic density fraction</th>
<th>Afferent fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>z_ee</td>
<td>0.765</td>
<td>0.8693</td>
</tr>
<tr>
<td>β_ee</td>
<td>0.0845</td>
<td>0.0960</td>
</tr>
<tr>
<td>μ_ee</td>
<td>0.0149</td>
<td>0.1242</td>
</tr>
<tr>
<td>β_α</td>
<td>0.100</td>
<td>0.8333</td>
</tr>
<tr>
<td>μ_β</td>
<td>0.0228</td>
<td>0.0259</td>
</tr>
<tr>
<td>μ_μ</td>
<td>0.004</td>
<td>0.0333</td>
</tr>
<tr>
<td>μ_M</td>
<td>0.0077</td>
<td>0.0088</td>
</tr>
<tr>
<td>M_α,β</td>
<td>0.0011</td>
<td>0.0092</td>
</tr>
<tr>
<td>M_α,β,μ</td>
<td>Not given</td>
<td></td>
</tr>
</tbody>
</table>

Axonal delays corresponded to a uniform velocity of conduction of 9 m/s, matching physiological estimates for the upper bound of velocities (Nunez 1995).

Reversal potential values qualitatively approximate to values given by Kandel and Schwartz (1985), who cite the excitation post-synaptic potential (EPSP) reversal potential as 0 mV, inhibitory post-synaptic potential (IPSP) reversal potential –60 mV, while the resting membrane potential is given as –55 mV and the threshold membrane potential as –45 mV. Since pulse probability density has not been accurately estimated as a function of LFP by any workers – a deficiency which partly motivated the choice of normalized voltage units in these studies – reversal potentials cannot be ascribed yu values with much quantitative accuracy. Notably, the inhibitory reversal potential has been chosen very close to the value of LFP for which there is a non-zero probability of emission of an action potential, thus offering inhibitory synapses high dynamic gain.

Likewise the choice of g_c, the excitatory synaptic gain has been chosen with little precision, but has been set to a high level consistent with physiological results (Thomson 1997). The ratio g_l = 4g_c was chosen to approximate the relative conductivities associated with glutamate/non-NMDA (t_peak = 0.1 – 0.3 ns) GABA/ GABA_A (t_peak = 0.4–1.0 ns) neurotransmission, respectively (Segev 1995).

The time response of the local field potential to unit impulse given by w_i = b^2 / [(c – n)/b] implies the time of peak response of the local field potential following unit synaptic input is 1/b (Robinson et al. 1997) and the mean transmission lag imposed by dendritic processes (which are directly reflected in the local field potential) is given by τ^d = b^2 / [c - b] . dt = 2/b . Thus, for a choice of b = 50 s^-1, the peak dendritic response is at 20 ms, and the mean lag imposed by the dendrites is 40 ms. This is within estimates for the response times soma for synaptic inputs mostly distributed distally on the dendritic tree (Segev 1995) assuming synaptic transmission time is largely governed by fast neurotransmission – notably by non-NMDA receptors (t_peak = 0.1–0.3 ns) and GABA_A receptors (t_peak = 0.4–1 ns) (Segev 1995).

However, the peak time of 20 ms is delayed compared to direct measurements made by some other workers, notably those estimates made from open-loop responses by Freeman (1975), although the mean lag accords with other estimates (Thomson 1997) to first approximation.

2.3 Configuration of simulation

An extended area of cortex was simulated by unit volumes in a 20 x 20 matrix, each volume connected with its neighbors so that the coupling strengths, z_ee(r_p) declined with r_p as a Gaussian function with standard deviation of 4 distance units, where a distance unit was the size of one cell of the 20 x 20 matrix. This approximates to distribution of cortico-cortical fibers in the cat brain if the distance unit is taken as about 0.9 mm. Boundary conditions were toroidal in all simulations reported. The application of absorbing boundary conditions and changes in matrix size were also studied and these changes did not affect the results to be reported.

Time step Δt was set at 0.1 ms, after trials showed that progressive decrements of the time step to 0.01 ms produced only small, asymptotically diminishing effects on spectral content of the results. Such an influence of the time step as was apparent occurred only at the lowest frequencies, including the DC offset in preliminary studies of the system impulse response. Thus, the longer time step was applied for computational convenience.

Specific inputs, Q_i, were delivered as zero-mean asynchronous white noise (Marsaglia and Zaman 1987) of small amplitude to both excitatory and inhibitory cell dendritic junctions in column zero, and outputs were recorded as V_{d}(t) from the element in the tenth row and column.

Q_{ns} inputs were delivered uniformly to all unit volumes in the matrix. The use of values of Q_{ns} greater than one, although pulse densities are more strictly defined on the range 0–1, has been justified in terms of the high efficacy of neuromodulators among other factors, in Wright (1997).

All simulation runs began with 0.8 s of both Q_e and Q_{ns} input, with \{Q_e, Q_{ns}\} all initially zero. In exceptional cases, high amplitude transient behavior immediately after initialization was apparent, so its influence was excluded by prolonging the pre-analysis duration to 4 s. Prolonged transients after initialization were apparent when temporal damping approached zero, as was the case at high Q_{ns} values as described below. Outputs were analyzed over the subsequent 0.8–1.6 s epoch. The time series \{V_{d}(t)\} underwent spectral analysis by fast Fourier transform, and ensemble averages were constructed by averaging the results obtained from the 0.8-s individual runs out of 100 samples. Power spectra were then expressed in units of \((V_{d}(Q_e))^2\), as a function of frequency.
Following completion of the analysis epoch, each simulation run was protracted for a further 0.8 s, while the $Q_e$ inputs were terminated. Time series from the second 0.4 s of this post-input epoch were automatically verified for the occurrence of persisting lightly damped activity, system saturation at $\{Q_e = 1\}$ and/or the development of undamped activity.

### 2.4 Sensitivity to parameter variation

The impact of variation of parameters was investigated for synaptic gains $g_e, g_i$ and their ratio, dendritic lag $b$, and reversal-potentials $V_{eR}, V_{IR}$.

Respective effects of multiplying $\mu_i$, multiplying $\mu_{at}$, and introducing the divisor $Z$ to the expression

$$Q_i = (1 + e^{(t'/p-3)})^{-1}$$

so that this becomes

$$Q_i = (1 + e^{(t'/p-3)})^{-1}$$

were also investigated. These manipulations influence the equations on system properties which would arise if the efficacy of IPSP on inhibitory cells differs from that on excitatory cells, if nonspecific cortical tone is directed selectively to inhibitory or excitatory cell populations, or if pulse probability as a function of LFP differs for excitatory and inhibitory populations.

Effects of variation of fiber range and axonal conduction velocity were also studied, although these are not further reported, as they were similar to earlier results.

### 2.5 Linearized and adiabatic-feedback model forms

To determine the dependence of the simulation’s dynamics upon nonlinear effects, it was useful to apply a linearized (small perturbation) version of the state equations.

Differentiating Eq. (2)

$$dQ/dV = -ae^{(V'/3)}(1 + e^{(V'/3)})^{-2}$$

and thus for any given mean value, $V_0$ of $V$

$$Q = \{dQ/dV_{i} = V_{i}\} \times (V - V_0)$$

is a linearized approximation of Eqs. (2) and (5). The mean values of $\{V_e, V_i\}$ obtained from the standard, nonlinear, simulation at steady-state, were thus the required $\{V_0\}$ to operate the simulation with equivalent gains to the standard condition, but without other nonlinear signal transformations.

The effect of changes in the time constant of local feedbacks in synaptic efficacy were studied by varying parameter $c$ so as to prolong feedback to various degrees, until this became adiabatic with relation to the state variables – that is, when $c \ll b$.

### 3 Results

#### 3.1 Performance of standard simulation

Figure 1 shows a set of simulated ECoG time series generated using the standard parameter set. (The usual steady-state epoch for spectral analysis has been prolonged to exhibit the examples more fully.) The signals increase in amplitude and dominant frequency as $Q_{ns}$ is increased. At high $Q_{ns}$ a “limit cycle” emerges. As shown, the cyclic activity diverges in amplitude to very high, unphysiological values. This divergence is a trivial consequence of the asymptotic properties of Eq. (2), and can be shown to result from a transition from a stable fixed-point condition to an unstable fixed-point condition, as $Q_{ns}$ is increased. For purpose of simplicity, the resulting undamped activity is subsequently referred to as a limit cycle.

Figure 2a shows the frequency content of the simulated ECoG as $Q_{ns}$ is increased up to the level at which limit cycles emerge, breaking the progression into two graphs to improve clarity. The lowest graph shows the concurrent changes seen in the mean excitatory and inhibitory pulse densities. Note that the pulse density remains relatively low up to the level at which limit cycles emerge. Figure 2b shows ensemble average power spectra which are cross-sections of the plots in Fig. 2a, chosen to show domains in which the spectral density is most prominently peaked.

It is seen that with progressive increase in $Q_{ns}$ prominent peaks in the power spectrum are seen first in the theta range, then sequentially in the alpha and beta ranges, with activity progressing penultimately across the gamma range, and ultimately entering limit cycles centered at a dominant frequency of 40 Hz.

Figure 3 shows similar power spectra to those in Fig. 2a, and the cross-correlation of pulse density in the excitatory and inhibitory compartments of the same element (row and column 10) for direct comparison. Leading of inhibitory compartments is greatest at low $Q_{ns}$ and tends to zero with increasing $Q_{ns}$, until limit cycles emerge, when the excitatory pulse density then leads the inhibitory.

#### 3.2 Effects of parameter variation

In Figures 4–8, the format of each graph follows that of the top frame in Fig. 2a. That is, spectra are shown up to 80% of the level of $Q_{ns}$ at which limit cycles emerge.

Figure 4 reveals that the peaks of spectral density in the EEG are sensitive to the mean dendritic lag, and with dendritic lag of 50% and 25% of the standard value, the reproduction of the lower EEG frequencies is no longer present. Activity in the beta and gamma bands is still reproduced, but the frequency content has shifted to higher bands in general, at all $Q_{ns}$.

Conversely, Fig. 5 shows that reduction of parameter $g$ (that is both $g_e$ and $g_i$) by almost 50% has little effect on the spectral content beyond a scaling effect, but that higher levels of $Q_{ns}$ are required to exert comparable effects on the spectral density to those at higher $g$. The same holds for effects on pulse densities.

Figure 6 shows that while the excitatory reversal potential exerts little effect so long as it is relatively high, there is strong spectral sensitivity to the value of the inhibitory reversal potential.
Figure 7 shows that manipulation of the strength of inter-inhibitory connections, relative excitatory tone to inhibitory cells, inhibitory synaptic gain, $g_i$, and the relative distribution of pulse-probability distribution as a function of LFP in inhibitory versus excitatory cells exert rather complicated effects. In some cases, the lower frequencies are of such low relative power that they are suppressed in the display by power in the gamma range.

As well as scaling effects, at least three effects may be discerned. Firstly, the $Q_{\text{res}}$ level required to achieve
a similar (80%) approach to the threshold of transition to limit cycles is changed. Secondly, the movement of the peak spectral density to the higher frequencies with increasing $Q_{ns}$ is altered, and thirdly, the relative power of activity in the theta, alpha, beta, and gamma bands is altered. (Similar, though relatively minor influences of this sort are also to be seen with the parameter variations in the prior Figs. 4-6). However, although the relative power is drastically altered, the progression of the dominant frequency toward the higher frequencies with increasing $Q_{ns}$ is an invariant feature. Where peaks at lower frequencies can be discerned, these remain near the dominant frequencies of the EEG.
3.3 Effects of linearization and speed of feedback

Figure 8 shows firstly, that at relatively low levels of cortical excitation, linearization of the sigmoidal pulse-wave relation has negligible effect.

Secondly, Fig. 8 shows that the oscillatory properties are sensitive to the speed of feedback. As parameter $c$ is sufficiently decreased, thus smoothing the effect of the reversal potentials, peaks in the power spectra are progressively eliminated, leaving an approximately $1/f^2$ spectral form at the extreme adiabatic case.
3.4 Effects on limit cycle dominant frequency

Figure 9 shows the spectral forms of the limit cycles generated with variation of selected parameters. For small variation of all parameters, the dominance of 4-Hz activity in the limit cycles persists, as exemplified in the \( g = 35 \) case. Other instances are shown in which variation induced a change in the dominant frequency. Bifurcations at critical values of parameters were observed with regard to this effect.
4 Discussion

4.1 Simulation properties and relation to physiological findings

The standard model results reported can be considered along with earlier simulation properties of synchronous oscillation and travelling waves (Wright 1997) which are retained in the present version, although not reported again here. Taken as a whole, the questions raised in Sec. 1.4 can be answered as follows.

1. Resonances with peak power at the theta, alpha, beta, and gamma frequencies can all be obtained in simulation under a single set of physiologically plausible parameter values. With incrementing electrocortical tone, the spectrum of the simulated EEG exhibits dominant oscillation in an ordered sequence, in the theta, alpha, beta and gamma bands, with transitional cases seen in which activity at two frequencies is prominent.

Such an explanation of the cerebral rhythms is not necessarily competitive with more conventional explanations such as the injection of bandwidth-limited signals from subcortical systems or intrinsic oscillations of neurons but selective frequencies of resonance in the cortical continuum would complement the impact of such “classical” signal sources.

2. The present results indicate that rapid local feedbacks, including the effect of reversal potentials, enhance the occurrence of resonances. Local resonance is not essential to the occurrence of either synchronous oscillation or travelling waves, as the latter properties were present in linearized versions of the state equations without local feedback of any sort (Wright 1997). Thus, the occurrence of spectral peaks is associated with rapid local feedback, while travelling waves, synchronous oscillation, and a “1/f” spectral type depend on longer range, purely excitatory couplings.

3. As shown by the effects of linearization of the sigmoid function, all the resonances other than the negatively damped 40-Hz oscillations, are associated with the transmission of frequency components in essentially linear, superposition, travelling waves. Beyond a critical level of excitation, a highly nonlinear mechanism supervenes.

4. The dissipative lower frequency waves propagate at low pulse densities, consistent with the long-standing finding that the low frequency EEG rhythms cannot be correlated with spike activity (Stryker 1989). CoCoG and EEG recordings which inevitably include volume conducted activity from a wide neuronal field, and low individual cell firing rates, offer a poor sample of the population spike rates and thus correlations would be difficult or impossible to observe. Conversely, at the highest levels of nonspecific excitation, “limit cycles” with a dominant 40-Hz center frequency appear and pulse rates increase to high levels. Correlation with LFP would become apparent in comparable physiological preparations, as is in fact the case (Eckhorn et al. 1988, etc.).

5. The transition to an undamped 40-Hz oscillation involves a change from a dissipative to a generative state, at a critical level of excitation. At the critical level a “change of phase” occurs in two senses, the reversal of local lead/lag relations in excitatory and inhibitory elements, and a change of phase in the thermodynamic sense, as alluded to by Langton (1990). (See Sect. 4.2.)

A striking discrepancy with routine EEG observations is the low power associated with low frequency signals compared with the power of high frequencies. In real
EEG observations, the reverse relation holds. However, the simulation is one of point potentials and local pulse density. No allowance has been made for the effects of volume conduction in cerebral and connective tissues, which exert spatial and temporal low-pass-filtering effects (Nunez 1981). This would act to correct the apparent discrepancy to an extent not at present established.

4.2 Relation to cellular automata theory and the “edge of chaos”

The results obtained also suggest a way in which cerebral dynamics may conform to Langton’s (1986, 1990) concepts regarding universal computation in brain-like systems. Langton’s (1990) paper lays emphasis
upon the “edge of chaos” sharing identity with the concept of thermodynamic phase transition. He gives as an example the transition between highly ordered and highly disordered dynamics, such as takes place between the solid and fluid states of matter. Such a phase transition seems to take place in this simulation at the point of emergence of 40 Hz “limit cycles”. At this transition, the energetics of the waves change from energy-dissipative to energy-generating, somewhat analogous to absorption or release of energies of fusion or evaporation in the phases of water.

The transition resembles a thermodynamic phase transition in another way. The applicability of the continuum concept depends on the notion of a disordered, stochastic independence in the large, between individual cell’s respective action potentials. When generative activity appears as the transition to active 40 Hz, oscillation takes place, the conditions for the existence of a purely stochastic medium disappear, and large-scale perturbations of pulse density appear as the result of the autonomous activity. Put in the terms used in Langton’s (1986) formulation, local 40-Hz “limit cycles” would provide the localized (energy generative limit cycles) periodic structures, while the waves of lower frequency would provide the (energy-dissipative point attractor) propagating periodic structures also required.

However, beyond the transition to generative 40-Hz activity, the simulation fails to match physiological data, since unlimited runaway does not take place in the brain. Physiological realism would require the addition of further processes to regulate the system activity level so that a complete global runaway state could not ensue.
Such a role might be played by more complicated local neurotransmitter effects as is further considered next, and perhaps also by cortical-subcortical feedback adjustment of the level of cortical activation.

### 4.3 Effects of parameter variation – implications for more detailed modelling

While the choice of standard parameters is somewhat arbitrary, all lie in the physiologically plausible range, and properties of the simulation are robust to small variations of parameters. Sensitive effects on the dominant frequencies of oscillation are exerted by the dendritic delays (parameter $b$) and the smoothing of feedback (parameter $c$) while variation of other parameters of plausible physiological ranges generally affects only the temporal damping of various resonances, the sensitivity to the mean cortical excitatory tone ($Q_m$), and the threshold of transition to 40-Hz limit cycles. Extreme variation of almost any parameter takes one or more properties of the simulation outside the domain of common EEG properties, as has been illustrated for the dominant frequency of limit cycles.

Other sets of parameters within a similar general formulation may also correspond to physiological
conditions. An example which does not exhibit the same dynamics as the standard model given here, but does have a realistic physiological parameterization and gives rise to physiologically realistic dynamics of a different type, has been recently presented (Liley et al. 1999).

More generally, the present results provide a simplified frame for the development of more complicated models, in which synaptic conductivities, channel kinetics, dendritic cable delays etc., can be introduced explicitly.

The occurrence of oscillation is dependent upon rapid feedback processes, which have been here modelled to approximate the effects of reversal potentials. The required gain of this feedback may be too great to be accounted for by reversal potentials alone, and a more physiologically realistic model may need to incorporate other and more powerful feedbacks processes to achieve a sufficient recruitment of inhibition particularly. It is not clear what physiological processes would make the best candidates for such a powerful feedback.

Since the spectral content is also sensitive to the rate of decay of feedback processes, as shown by the effects of variation of parameter $c$, this suggests that a more complete model incorporating the effects of spatial and temporal summation in synaptic actions would also effect the spectral content. The rapidly acting non-NMDA and GABA$_A$ receptors were implicitly incorporated in the present model of synapto-dendritic transmission. Inclusion of NMDA and GABA$_A$ with their much slower onsets and offsets (Segev 1995) would introduce more complex variation in $g_{e}$, $g_{i}$ and $c$ to the model, adding features of adiabatic control recently theoretically analyzed for linear models of the same class as the present one (Robinson et al. 1998; Rennie et al. 1998).

Further powerful and rapid feedback processes would be required to maintain neuronal activity near the point of transition between dissipation and generation without runaway, an aspect of dynamics completely ignored in the present model, as previously noted. A possible candidate for this role has recently been reported in the form of AMPA-mediated receptor desensitization (Tones and Westbrook 1996).

4.5 Predictions of results in further experiments

Although the model given here is both extremely simplified and general in nature, it leads to definite restrictions on the classes of dynamics predicted over a considerable range of cortical conditions.

1. Recurrence plots and complexity measures (e.g., Xu et al. 1997) applied to LFP and multiunit spike activity under experimental conditions similar to those used in studies of oscillation in the gamma band (e.g., Eckhorn et al. 1988; Gray and Singer 1989) should reveal marked transitions between background linear/ stochastic conditions, and highly nonlinear states associated with peak oscillation. Although such analysis procedures have yielded controversial outcomes from EEG signals (e.g., peer commentaries in Wright and Liley 1996), part of the problem with such measures may be that they are applied to EEG signals averaged over too great an extent of the cortex, thus sampling non-uniform dynamic types concurrently. In strictly local conditions the phase transitions ought to be much less equivocally observed during onset and offset of strong local oscillations at approximately 40 Hz, at which time marked increases in Shannon entropy and mutual information between units should be apparent. It may also be possible to observe a reversal of lead/lag relations between excitatory and inhibitory cells during such transitions.

2. According to this model, bursts of generative activity around 40 Hz will occur in critically activated domains of cortex which are also undergoing perturbation by low-frequency activity transmitted from surrounding domains in the cortex as travelling waves of LFP and pulse density. Thus, the total power in the gamma band about 40 Hz should vary systematically with cortical slow potentials.

3. Because the source of power spectral peaks is attributed to local feedback, there should be little dependence of the frequency of such an oscillation upon the wavenumber, in contrast to travelling waves conducted via longer range excitatory connections which must obey a frequency/wavenumber relation in accord with the phase velocities of the medium. This would lead to the occurrence of two classes of activity distinguishable by frequency/wavenumber analysis. A formal description of these differences in dispersion relations is given elsewhere (Rennie et al. 1998).

4. Consequent also to the local origin of cerebral resonances according to this model, boundary conditions including brain size and axonal conduction velocity should play little part in the frequency of cerebral rhythms. This prediction contrasts with the Nunez (1981, 1989) global resonance model. Studies of EEG frequencies in the maturing fetal sheep (Sergejew 1999) show no change in center frequencies of EEG during late brain maturation, in accord with the present hypothesis.

Acknowledgements. This work was supported by the Australian Research Council, and the Pratt Group of Companies. Thanks are extended to Nicholas Hawthorn for technical assistance, and to Peter Robinson and Christopher Rennie for drawing attention to a numerical error influencing the spectral properties reported in Wright (1997).

References


Mechanisms of Cortical Electrical Activity and Emergence of Gamma Rhythm

CHRISTOPHER J. RENNIE*†‡, JAMES J. WRIGHT§ and PETER A. ROBINSON*

*School of Physics, University of Sydney, New South Wales 2006, Australia, †Department of Medical Physics, Westmead Hospital, Westmead, New South Wales 2145, Australia and § Mental Health Research Institute, Parkville, Victoria 3052, Australia

(Received on 22 June 1999, Accepted in revised form on 6 March 2000)

A continuum model of the electrical activity of the cerebral cortex is described which predicts the occurrence of a resonance in the gamma range near 40 Hz. The emergence of this resonance is due to two refinements to a previous model, namely the inclusion of a modulation of synaptic strength due to finite reversal potentials, and use of parameters that better match physiological measurements. Analytical expressions for the fixed points of the system and for its linear dynamics are found in terms of average neuronal properties, and together explain the occurrence and modulation of the gamma-like resonance. The analytical results are confirmed by a numerical simulation.

© 2000 Academic Press

1. Introduction

Recordings of electrical potentials on the scalp are a common probe of brain activity, both in the clinical and cognitive research settings. However, interpretation of the recordings is largely empirical, and so an explicit analytical model of their generation would be of great value. The two major problems in modeling the large-scale activity of the brain are (i) how to deal appropriately with the anatomical complications posed by the multiplicity of neuronal interactions, and (ii) how to distinguish those aspects of neurophysiology that are significant on the large scale from those that are not. In neither case are numerical simulations of realistic neurons feasible, as the number of neurons required is too high and their functional details are too intricate. Yet recordings from the scalp (electroencephalograms or EEGs) reveal the existence of low-frequency (1–50 Hz) voltage perturbations that are coherent over distances similar to the scale of the brain, and this suggests that modeling of the collective behavior of neurons may be possible, based only on the average properties of neurons.

Regarding the first of the above problems (anatomical complexity), it was proposed first by Wilson & Cowan (1973) that cortical EEGs could be modeled with a two-dimensional continuum approximation where each point has certain crucial characteristics of neurons: integration of activity from other points, and generation of activity through a nonlinear response function. The continuum approximation also allows a simple distribution function to be used to describe the spatial extent of interconnections. This style of modeling is adopted in our work. In an early version of our model, the total input to any point in the cortex was obtained by spatial convolution...
of activity occurring elsewhere, with a kernel that approximated the typical spread of nerve fibers (Wright & Liley, 1994; Wright, 1999). A later version replaced the convolution with a wave equation, which made the continuum approximation more tractable, both analytically and (numerically Robinson et al., 1997).

The second problem stated above (that of physiological complexity) was addressed in part by Rennie et al. (1999), in which the consequences of modulation of neuronal parameters on the linear dynamics of the system were explored. The modulations can be thought of as driven by the principal system variables (firing rates or membrane potentials) which in turn affect the system variables; consequently, general formulations were proposed that represented each of the possible modulations as a form of feedback. (The feedbacks are internal to the neuron, and are distinct from interneuron feedbacks which are handled by other means in our model.) In the Appendices, the results of that study have been used to estimate the likely effect of various specific modulatory mechanisms, including: in Appendix A, the influence of recent incoming firing rate on the neuronal interconnection strength [specifically the phenomena of facilitation and depression (Magleby, 1987)]; in Appendix B, the refractoriness of neurons following firing (Johnston & Wu, 1995); and in Appendix C, modulation of synaptic strength by membrane potential due to the finite magnitude of the reversal potentials at synapses (Johnston & Wu, 1995). It is demonstrated that only the last of these feedbacks is sufficient to alter the global dynamics of the system, so it alone is added to our model in the body of the paper. Facilitation, depression, and refractoriness are generally regarded as being the major forms of modulation, but our analysis shows them to affect global dynamics only weakly. Their omission is of considerable value to any model of large-scale EEG, where tractability can be hard to achieve.

An application that suits this model is simulation of gamma rhythm (Eckhorn et al., 1989; Gray & Singer, 1989; Singer & Gray, 1995). Its characteristically high frequency of 40–60 Hz probably indicates that reciprocal corticothalamic pathways are not involved, on account of the combined axonal and dendritic delays that would be involved (Steriade et al., 1990), so a purely cortical model like this one should be adequate. The wide interest in gamma rhythm arises from the proposal by von der Malsburg (1983) that a resonant oscillation occurs when some component of an input signal is recognized, and that the spread and coupling of multiple such oscillations results in fusion of all perceptual components into a whole. Synchronous gamma rhythm is seen in strongly activated cortex and is the leading candidate for solving this so-called binding problem (Singer & Gray, 1995). Attempts to account for gamma rhythm include that of Traub et al. (1997) who obtained 40 Hz activity in a detailed numerical model, Freeman (1992a) who postulated that it arises from reciprocal activation between excitatory and inhibitory neuron pools, and Tass & Haken (1996) who invoke more abstract nonlinear principles. It has been independently argued that there are computational benefits in having the system poised on the “edge of chaos” (Langton, 1990), and that gamma rhythm in the brain may be the EEG manifestation of this computational mode (Skarda & Freeman, 1987), although the connection between gamma and brain function remains controversial (Lamme & Spekreijse, 1998).

Of the models just mentioned, those that are quantitative tend to be highly detailed and numerically intensive. In the present paper, the emphasis is on identifying the critical mechanisms underlying EEG, and this is better done with a more analytically tractable model. The methodology adopted here is that of linear analysis, since the spectral characteristics of early linearized models (e.g. Freeman, 1972; Nunez, 1974; van Rotterdam et al., 1982) showed promising similarities with recorded EEGs. Alternative modeling techniques include nonlinear analysis (e.g. Freeman, 1992b) and numerical simulation with physiologically detailed stochastic neural network models (e.g. Wilson & Bower, 1992; Lumer et al., 1997a, b). In earlier work, we have shown that continuum models similar to the model presented here can reproduce spectral powers in the 0–30 Hz range (Robinson et al., 2000; Wright, 1990), and sustain zero-lag synchrony between co-active sites in the cortex (Chapman et al., 2000; Robinson et al., 1998b). However, a supplementary mechanism must be
provided to account for physiological observations of oscillations in the gamma range.

The goal of the present paper is to infer more precisely the consequences of finite reversal potentials. We also attempt to constrain the model by using only experimentally determined parameters as far as possible, so that predictions of emergent global behavior have no element of circularity.

The relevant equations and parameter values are derived in Section 2. In Section 3, the steady-state and dynamic characteristics of the model are presented, including its sensitivity to variations in parameters and the appearance of a gamma-like resonance. The findings are summarized in Section 4.

2. Theory

Assumptions common to all continuum models of cortex are that (a) the cortex can be considered to be two-dimensional, (b) two distinct populations of neurons are considered, classified as excitatory or inhibitory according to their effect on other neurons, (c) the excitatory and inhibitory inputs to the input fibers (dendrites) and to the cell body (soma) of each neuron combine linearly to determine the transmembrane potential of the neuron, (d) the membrane potential determines the firing rate of the neuron, and (e) the emitted impulses (action potentials) are conducted away from the soma along a fiber (axon) and eventually stimulate the dendrites and somas of other neurons. Excitation received from subcortical sources is generally taken as a free parameter, and contributes to the system’s equilibrium level of activity.

An important addition to the present version of the model is the inclusion of synaptic reversal potentials. It has previously been assumed in continuum cortical models of cortical activity that the size of synaptic responses can be taken as constant, but as described in Appendix C there is a mechanism by which synaptic strengths are modulated. Following an action potential there is a perturbation in the transmembrane potential (a post-synaptic potential or PSP), the integral of which we refer to as the synaptic strength, which is a consequence of the transmembrane current induced at an activated synapse. This synaptic current is the product of the instantaneous conductance and the driving potential,

\[ I = g(V - V_{\text{rev}}). \]  

(1)

The conductance \( g \) is found to be almost independent of \( V \), at least for the majority of synapses, although it varies transiently following the arrival of an action potential. The driving potential \( V - V_{\text{rev}} \) is the difference between the transmembrane potential \( V(t) \) and the reversal potential \( V_{\text{rev}} \); the former is a function of time and the latter is the potential at which no net ionic current occurs (Johnston & Wu, 1995, p. 361). We conventionally assume that a PSP, as measured at the soma, is proportional to the synaptic current \( I \), albeit with a lag reflecting the transmission properties of the membrane. We refer to the PSP integrated over its duration as the synaptic strength and represent it by \( s_{qp} \), where \( q = e, i \) indicates the two principal classes of target neurons (excitatory and inhibitory), and \( p = e, i \), \( s \) represents the three sources of activity (excitatory and inhibitory neurons in the cortex, and subcortical sources). When temporal variation in the driving potential \( V - V_{\text{rev}} \) was neglected we could take \( s_{qp} \) as uniform and constant. Now, however, we shall assume

\[
 s_{qp}(r, t) = s_{qp}^{(0)} \left(1 - \frac{V_q(r, t) - V_q^{(0)}}{V_{p_{\text{rev}}} - V_q^{(0)}}\right) \otimes H(t)
 = s_{qp}^{(0)} R_{qp}(r, t) \otimes H(t),
\]  

(2)

which has the necessary properties that \( s_{qp} = 0 \) when the membrane potential is constant and equal to the reversal potential (i.e. \( V_q(r, t) = V_{p_{\text{rev}}} \)), and that \( s_{qp} \) is equal to the reference value \( s_{qp}^{(0)} \) when \( V_q(r, t) = V_q^{(0)} \). There is a convolution (symbolized here and subsequently by \( \otimes \)) between the modulation due to reversal potentials \( R_{qp}(t) \) and the lag function \( H(t) \), which describes the low-pass filter characteristics of the membrane. [The soma where synaptic responses are being estimated is usually remote from the synapse, and the dendritic tree is not isopotential. Hence, the lagged value of the bracketed term in eqn (2) is a better estimate of the modulation factor.] Equation (2) allows for greater generality...
than will be required below: in practice, it is possible to choose identical reference potentials for both classes of neurons, so \( V^{(0)}_e = V^{(0)}_i \), and to set \( V^{rev}_e = V^{rev}_i \), \( s^{(0)}_e = s^{(0)}_i \) and \( s^{(0)}_i = s^{(0)}_e \) by assuming that synapses from subcortical sources are physiologically identical to those from excitatory cortical neurons.

It is now possible to substitute eqn (2) into the basic dendritic equation

\[
V_q(t) = N_{q_e} S_{q_e}(t) \otimes [s_{q_e}(t) \phi_e(t, t)] \\
+ N_{q_i} S_{q_i}(t) \otimes [s_{q_i}(t) \phi_i(t, t)] \\
+ N_{q_s} S_{q_s}(t) \otimes [s_{q_s}(t) \phi_s(t, t)] \\
+ V^{rev}_q. \tag{3}
\]

This represents \( V_e(t) \) and \( V_i(t) \), the somatic membrane potential in excitatory and inhibitory neurons, respectively, as the linear combination of the potentials due to the incoming firing rates \( \phi_p \), where \( p = e, i, s \). [Linear combination of excitatory and inhibitory potentials seems to be justified even when shunting inhibition is considered, according to Koch (1999) p. 423.] Shunting inhibition is the multiplicative attenuation of all PSPs initiated at more distant synapses due to an active synapse shorting the membrane at some point in the dendritic tree.) The quantities \( N_{q_p} \) are the number of synapses of type \( p \) on neurons of type \( q \). Compared with the corresponding equations in Rennie et al. (1999) these equations have three generalizations: that modulations due to both \( V^{rev}_e \) and \( V^{rev}_i \) are included, the rest potentials \( V^{rest}_e \) are included explicitly, and the normalized time courses of synaptic responses

\[
S_{q_p}(t) = \frac{\alpha_{q_p} \beta_{q_p}}{\beta_{q_p} - \alpha_{q_p}} (\exp[- \alpha_{q_p} t] - \exp[- \beta_{q_p} t]), \tag{4}
\]

(with \( q = e, i, p = e, i, s \) and \( 0 \leq \alpha_{q_p} < \beta_{q_p} \)) can be independent, in order to reflect better the different temporal characteristics of each combination of synapse and neuron type.

As in Rennie et al. (1999) the generation and propagation of action potentials are described by

\[
Q_{e,i}(t) = \frac{Q^{max}_{e,i}}{1 + \exp[-C(V_{e,i}(t) - \theta_{e,i})/\sigma_{e,i}]}.
\]

\[
D_{e,i} \phi_{e,i}(t) = Q_{e,i}(t), \tag{5}
\]

\[
D_{e,i} = \left[ \frac{1}{\gamma_{e,i} t} + 2 \frac{\sigma_{e,i}}{\gamma_{e,i}} + 1 - r^2_{e,i} v^2 \right]. \tag{6}
\]

Equation (5) is an empirical sigmoidal relationship between mean neuronal membrane potentials and their mean firing rate, and is parameterized by an offset \( \theta_{e,i} \), a width \( \sigma_{e,i} \), and the maximal firing rate \( Q^{max}_{e,i} \). The constant \( C \) equals \( \pi/\sqrt{3} \), so that \( \gamma_{e,i} \) is the standard deviation (S.D.) of the derivative of \( Q_{e,i} \), which closely resembles a normal distribution. Equation (6) describes the propagation of activity at a finite speed throughout a two-dimensional continuum by neurons with a certain axonal range distribution (Robinson et al., 1997). The parameters are \( r_{e,i} \), the ranges of the axons of excitatory and inhibitory neurons, \( \gamma_{e,i} = v_{e,i}/r_{e,i} \) are the corresponding damping rates, and \( v_{e,i} \) are axonal conduction speeds.

The set of equations (3), (5), and (6) can be solved numerically for arbitrary geometries and initial conditions; however, we shall be concerned here only with the steady-state solutions and linear dispersion relations for an infinite medium.

### 2.1. STEADY-STATE EQUATIONS

The result of setting all derivatives equal to zero in eqn (6), together with eqns (3) and (5), are the steady-state equations

\[
\tilde{V}_q = N_{q_e} s^{(0)}_{q_e} \bar{R}_{q_e} \tilde{\phi}_e + N_{q_i} s^{(0)}_{q_i} \bar{R}_{q_i} \tilde{\phi}_i \\
+ N_{q_s} s^{(0)}_{q_s} \bar{R}_{q_s} \tilde{\phi}_s + V^{rest}_q, \tag{8}
\]

\[
\tilde{\phi}_q = \frac{Q^{max}_q}{1 + \exp[-C(V_q - \theta_q)/\sigma_q]}. \tag{9}
\]

For a given value of uniform and constant subcortical input \( \tilde{\phi}_q \), these four equations can be solved for the uniform and constant values
\( \vec{V}_e, \vec{V}_i, \vec{\phi}_e, \) and \( \vec{\phi}_i, \) which are used in the linear stability analysis below.

### 2.2. Dispersion Relations

To infer the linear dispersion relation of electrocortical waves we must linearize the equations. Each term in eqn (3) is of the general form

\[
N_{q_{p}} S_{q_{p}} \otimes (s_{q_{p}}(r, t) \phi_{q_{p}}(r, t)),
\]

so \( s_{q_{p}}(r, t) \) must be written as a constant plus a zero mean perturbation, as must \( \phi_{p}(r, t). \) Consequently, eqn (3) (with \( q = e \)) becomes

\[
V_{e}(r, t) = N_{ec} S_{ce}(t) \otimes \left[ \left( s_{ce}^{(0)} \vec{R}_{ce} - \frac{s_{ce}^{(0)}}{\vec{V}_{ce} - \vec{V}_{en}} H(t) \right) \otimes (V_{e}(r, t) - \vec{V}_{e}) \right] \otimes (\phi_{e}(r, t) - \vec{\phi}_{e}) + \ldots \tag{10}
\]

\[
\approx N_{ec} s_{ce}^{(0)} \vec{R}_{ce} \vec{\phi}_{e} - \frac{N_{ec} s_{ce}^{(0)} \vec{\phi}_{e}}{\vec{V}_{en} - \vec{V}_{e}} S_{ce}(t) \otimes H(t) \otimes (V_{e}(r, t) - \vec{V}_{e})
\]

\[
+ N_{ec} s_{ce}^{(0)} \vec{R}_{ce} S_{ce}(t) \otimes (\phi_{e}(r, t) - \vec{\phi}_{e}) + \ldots \tag{11}
\]

where we have omitted terms that are the product of two zero-mean quantities. This expression can be simplified using eqn (8) (with \( q = e \)) to give

\[
V_{e}(r, t) - \vec{V}_{e} = -\frac{N_{ec} s_{ce}^{(0)} \vec{\phi}_{e}}{\vec{V}_{en} - \vec{V}_{e}} S_{ce}(t) \otimes H(t) \otimes [V_{e}(r, t) - \vec{V}_{e}] + N_{ec} s_{ce}^{(0)} \vec{R}_{ce} S_{ce}(t) \otimes (\phi_{e}(r, t) - \vec{\phi}_{e})
\]

\[
- \frac{N_{ec} s_{ce}^{(0)} \vec{\phi}_{i}}{\vec{V}_{en} - \vec{V}_{e}} S_{ci}(t) \otimes H(t) \otimes [V_{e}(r, t) - \vec{V}_{e}] + N_{ci} s_{ci}^{(0)} \vec{R}_{ci} S_{ci}(t) \otimes (\phi_{i}(r, t) - \vec{\phi}_{i})
\]

\[
- \frac{N_{ec} s_{ce}^{(0)} \vec{\phi}_{s}}{\vec{V}_{en} - \vec{V}_{e}} S_{cs}(t) \otimes H(t) \otimes [V_{e}(r, t) - \vec{V}_{e}] + N_{cs} s_{cs}^{(0)} \vec{R}_{cs} S_{cs}(t) \otimes (\phi_{s}(r, t) - \vec{\phi}_{s}), \tag{12}
\]

which can then be Fourier transformed to give

\[
V_{e}(k, \omega) =
\]

\[
\frac{N_{ec} s_{ce}^{(0)} \vec{R}_{ce} S_{ce} \phi_{e} + N_{ci} s_{ci}^{(0)} \vec{R}_{ci} S_{ci} \phi_{i} + N_{cs} s_{cs}^{(0)} \vec{R}_{cs} S_{cs} \phi_{s}}{1 + [(N_{ec} s_{ce}^{(0)} \vec{\phi}_{e}/(V_{en} - V_{e})) S_{ce} + (N_{ci} s_{ci}^{(0)} \vec{\phi}_{i}/(V_{en} - V_{e})) S_{ci} + (N_{cs} s_{cs}^{(0)} \vec{\phi}_{s}/(V_{en} - V_{e})) S_{cs} ] H(\omega)},
\tag{13}
\]

and similarly the corresponding expression \( V_{i}(k, \omega) \) is

\[
V_{i}(k, \omega) =
\]

\[
\frac{N_{ic} s_{ic}^{(0)} \vec{R}_{ic} S_{ic} \phi_{e} + N_{ii} s_{ii}^{(0)} \vec{R}_{ii} S_{ii} \phi_{i} + N_{is} s_{is}^{(0)} \vec{R}_{is} S_{is} \phi_{s}}{1 + [(N_{ic} s_{ic}^{(0)} \vec{\phi}_{e}/(V_{en} - V_{i})) S_{ic} + (N_{ii} s_{ii}^{(0)} \vec{\phi}_{i}/(V_{en} - V_{i})) S_{ii} + (N_{is} s_{is}^{(0)} \vec{\phi}_{s}/(V_{en} - V_{i})) S_{is} ] H(\omega)},
\tag{14}
\]
The \( V_{e,i} \) can be eliminated from eqns (13)--(15) to give

\[
F_{D}D_{e}\phi_{e} = G_{ee}S_{ee}\phi_{e} + G_{ei}\phi_{i} + G_{es}\phi_{s},
\]

\[
F_{D}D_{i}\phi_{i} = G_{ie}\phi_{e} + G_{ii}\phi_{i} + G_{is}\phi_{s},
\]

which describe the dynamics of the system in terms of feedback factors \( F_q = F_q(\omega) \), wave equations \( D_q = D_q(k, \omega) \), dendritic impulse responses \( S_{qp} = S_{qp}(\omega) \), and feedforward gains \( G_{qp} \). These equations are similar to our past formulations except that \( V_p^{rev} \) was previously effectively infinite, so that \( G_{qp} = \rho qN_{qp}S_{qp}^{(0)} \) and \( F_q(\omega) = 1 \). Equations (20) and (21) have solutions

\[
\left( \frac{\phi_e(k, \omega)}{\phi_i(k, \omega)} \right) = \frac{1}{\text{Det}} \left( \frac{F_iD_i - G_{ii}S_{ii}}{G_{ie}S_{ie}} \right) \frac{G_{ee}S_{ee}}{F_eD_e - G_{ee}S_{ee}}
\]

\[
\times \left( \frac{G_{ee}S_{ee}}{G_{is}S_{is}} \right) \phi_e(k, \omega),
\]

where \( \text{Det} = (F_eD_e - G_{ee}S_{ee})(F_iD_i - G_{ii}S_{ii}) - G_{ei}S_{ei}G_{ie}S_{ie} \). In particular, the transfer function (Fourier-space Green function) \( \phi_e(k, \omega)/\phi_i(k, \omega) \) is given by

\[
\frac{\phi_e(k, \omega)}{\phi_i(k, \omega)} = \frac{(F_iD_i - G_{ii}S_{ii})G_{ee}S_{ee} + G_{ei}S_{ei}G_{is}S_{is}}{(F_eD_e - G_{ee}S_{ee})(F_iD_i - G_{ii}S_{ii}) - G_{ei}S_{ei}G_{ie}S_{ie}},
\]

and the dispersion relation is

\[
(F_eD_e - G_{ee}S_{ee})(F_iD_i - G_{ii}S_{ii}) - G_{ei}S_{ei}G_{ie}S_{ie} = 0.
\]

It is frequently preferable to write the transfer function as

\[
\frac{\phi_e(k, \omega)}{\phi_i(k, \omega)} = \frac{F_iD_iG_{ee}S_{ee} - G_{ei}S_{ei}G_{is}S_{is}}{F_eD_eF_iD_i - F_eD_eG_{ii}S_{ii} - G_{ee}S_{ee}F_iD_i + G_{ee}S_{ee}G_{is}S_{is}}M_1
\]

\[
M_1 = 1 - G_{ei}S_{ei}G_{is}S_{is}/G_{ee}S_{ee}G_{ii}S_{ii},
\]

\[
M_2 = 1 - G_{ei}S_{ei}G_{ie}S_{ie}/G_{ee}S_{ee}G_{ii}S_{ii},
\]

since under assumptions made below (see Table 2) \( G_{ei}G_{ie}/G_{ee}G_{ii} = G_{ei}G_{ie}/G_{ee}G_{ii} = 1 \), and so (assuming all \( S_{qp} \) are similar) the last terms in the numerator and denominator of eqn (25) tend to be negligible.
2.3. Parameters

Here we discuss each of the values chosen for the model’s parameters, in the order in which they appear in Table 1. Here, as elsewhere, the subscript \( q = e, i \) refers to the two classes of neurons and the subscript \( p = e, i, s \) to the three sources of dendritic inputs.

The connectivities \( N_{qp} \) are the number of synapses of type \( p \) on neurons of type \( q \). The values in Table 1 are obtained by assuming that (a) there are 5000 dendritic and somatic synapses on both excitatory and inhibitory neurons (Douglas & Martin, 1998, pp. 459, 471), (b) that 78% of neurons are excitatory (Douglas & Martin, 1998, p. 462), (c) that 84% of synapses are excitatory (Douglas & Martin, 1998, p. 471), and (d) that 2% of axons are of subcortical origin.

The synaptic strengths \( s_{qp}^{(0)} \) are the integrated response due to a single impulse of type \( p = e, i \) arriving at a neuron of type \( q = e, i, s \), as recorded at the cell body and in units of V s. The values in Table 1 are derived from Thomson et al. (1996) and Thomson (1997) by weighted averaging the subclasses identified by them, and correcting for different resting potentials. However, these values were biased in several ways due to experimental conditions that are unlike those found \( \text{in vivo} \), principally by the general suppression of spontaneous neural activity, resulting in unnatural resting conditions. In an attempt to compensate for the biases, \( s_{qp}^{(0)} \) is scaled by a factor of 1.5 as it is evident from Thomson (1997) that nearby voltage sensitive channels have a “boosting” effect on EPSPs. Also \( s_{ep}^{(0)} \) is scaled by a factor of 5 to reflect the degree of facilitation seen at levels of activity typical of the cortex (Thomson, 1997). As a result \( s_{ep}^{(0)} \approx s_{ei}^{(0)} \approx 24 \times 10^{-6} \text{ V s} \). No estimates are available for \( s_{is}^{(0)} \) so we assume \( s_{is}^{(0)} \approx s_{is}^{(0)} - 59 \times 10^{-6} \text{ V s} \), in the absence of evidence to the contrary. Furthermore, to obtain values representative of single synapses in neurons in which other synapses are active too, all values are divided by 10 as a result of two separate effects: it was demonstrated by Thomson et al. (1996) and Thomson (1997) that their measured PSPs are commonly the result of multiple synapses; and Bernander (1993) (quoted in Koch, 1999, p. 416)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_{ee} = N_{ii} )</td>
<td>4120</td>
<td>Dendritic synapses from pyramidalis</td>
</tr>
<tr>
<td>( N_{ei} = N_{ie} )</td>
<td>800</td>
<td>Dendritic synapses from interneurons</td>
</tr>
<tr>
<td>( N_{is} = N_{si} )</td>
<td>80</td>
<td>Dendritic synapses from subcortex</td>
</tr>
<tr>
<td>( s_{ee}^{(0)} = s_{ie}^{(0)} )</td>
<td>( 2.4 \times 10^{-6} \text{ V s} )</td>
<td>EPSP size due to ( \phi_e ) when ( V_{e,i} = V_{e,i}^{(0)} )</td>
</tr>
<tr>
<td>( s_{ei}^{(0)} = s_{ie}^{(0)} )</td>
<td>( -5.9 \times 10^{-6} \text{ V s} )</td>
<td>IPSP size due to ( \phi_i ) when ( V_{e,i} = V_{e,i}^{(0)} )</td>
</tr>
<tr>
<td>( s_{is}^{(0)} = s_{si}^{(0)} )</td>
<td>( 2.4 \times 10^{-6} \text{ V s} )</td>
<td>EPSP size due to ( \phi_i ) when ( V_{e,i} = V_{e,i}^{(0)} )</td>
</tr>
<tr>
<td>( V_{e,i}^{(0)} )</td>
<td>( 0.060 \text{ V} )</td>
<td>Potential at which ( s_{i\text{ep}}^{(0)} ) are estimated</td>
</tr>
<tr>
<td>( V_{e,ei}^{(0)} )</td>
<td>( -0.060 \text{ V} )</td>
<td>Rest potential (( \phi_s = \phi_i = \phi_\text{h} = 0 ))</td>
</tr>
<tr>
<td>( V_{\text{rev}} = V_{\text{rev}}^{(0)} )</td>
<td>0 V</td>
<td>Reversal potential for AMPA channels</td>
</tr>
<tr>
<td>( V_{\text{rev}}^{(0)} )</td>
<td>( -0.070 \text{ V} )</td>
<td>Reversal potential for GABA(_\text{A} ) channels</td>
</tr>
<tr>
<td>( \tau_{ae} = \tau_{ea} )</td>
<td>( 68 \text{ s}^{-1} )</td>
<td>Decay rate in pyramidalis of EPSPs</td>
</tr>
<tr>
<td>( \tau_{ai} = \tau_{ia} )</td>
<td>( 47 \text{ s}^{-1} )</td>
<td>Decay rate in pyramidalis of IPSPs</td>
</tr>
<tr>
<td>( \tau_{ei} = \tau_{ie} )</td>
<td>( 176 \text{ s}^{-1} )</td>
<td>Decay rate in interneurons of EPSPs</td>
</tr>
<tr>
<td>( \tau_{ei} = \tau_{ie} )</td>
<td>( 82 \text{ s}^{-1} )</td>
<td>Decay rate in interneurons of IPSPs</td>
</tr>
<tr>
<td>( \beta_{ae} = \beta_{ea} = \beta_{ei} = \beta_{ie} )</td>
<td>( 500 \text{ s}^{-1} )</td>
<td>Rate of rise of EPSPs in pyramidalis</td>
</tr>
<tr>
<td>( \beta_{ei} = \beta_{ie} = \beta_{is} = \beta_{si} )</td>
<td>( 500 \text{ s}^{-1} )</td>
<td>Rate of rise of IPSPs in pyramidalis</td>
</tr>
<tr>
<td>( \eta )</td>
<td>( 200 \text{ s}^{-1} )</td>
<td>Rate of rise of PSPs in interneurons</td>
</tr>
<tr>
<td>( Q_{\text{max}}^{(0)} )</td>
<td>( 100 \text{ s}^{-1} )</td>
<td>Cut-off frequency for feedback</td>
</tr>
<tr>
<td>( Q_{\text{max}}^{(0)} )</td>
<td>( 200 \text{ s}^{-1} )</td>
<td>Maximal pyramidal firing rates</td>
</tr>
<tr>
<td>( \sigma_{ae} )</td>
<td>( 0.005 \text{ V} )</td>
<td>Maximal interneuron firing rates</td>
</tr>
<tr>
<td>( \theta_{ae} )</td>
<td>( -0.052 \text{ V} )</td>
<td>Sigmoid width parameter</td>
</tr>
<tr>
<td>( \phi_s )</td>
<td>( 10 \text{ s}^{-1} )</td>
<td>Firing threshold</td>
</tr>
<tr>
<td>( \tau_e )</td>
<td>( 80 \text{ s}^{-1} )</td>
<td>Firing rate of subcortical neurons</td>
</tr>
<tr>
<td>( \tau_i )</td>
<td>( 10^5 \text{ s}^{-1} )</td>
<td>Excitatory spatial damping rate</td>
</tr>
<tr>
<td>( \tau_{\text{inh}} )</td>
<td>( 10^5 \text{ s}^{-1} )</td>
<td>Inhibitory spatial damping rate</td>
</tr>
</tbody>
</table>
has demonstrated strong attenuation of PSPs in neurons with many active dendritic synapses, as a result of shorting of the dendritic membrane by background activity.

The responses due to subcortical sources \( s_{aq} \) are excitatory and mediated by the same neurotransmitter as is released by most cortical excitatory neurons, so we set \( s_{aq}^{(0)} = s_{aq}^{(0)} \).

The reference potentials \( V_q^{(0)} \) are the transmembrane potentials corresponding to the adopted synaptic strengths \( s_{aq}^{(0)} \). Synaptic strengths are approximately linear functions of the transmembrane potential [as in eqn (2)] and those reported in Thomson et al. (1996) and Thomson (1997) were measured at various transmembrane potentials. Hence, for convenience this common reference potential is adopted, and the values \( s_{aq}^{(0)} \) include a linear correction so that they all refer to this transmembrane potential.

The rest potentials \( V_{q}^{\text{rest}} \) are the potentials assumed by neurons receiving low background firing rates. As summarized by Koch (1999, p. 417), experimental values have a wide range, either for reasons intrinsic to the neurons, or on account of differing concentrations of extracellular neuromodulators (transmitters with prolonged effects). We make the simplest assumption consistent with in vivo experiments by setting \( V_{q}^{\text{rest}} = -60 \text{ mV} \).

We assume the reversal potentials for excitatory synapses to be \( V_{e}^{\text{rest}} = V_{s}^{\text{rest}} = 0 \) in accord with the predominance of AMPA receptors in the cortex, and take \( V_{i}^{\text{rest}} = -70 \text{ mV} \) to be that of GABA \(_A\) receptors (Johnston & Wu, 1995, p. 361). These values are also consistent with experimental estimates by Thomson et al. (1996) and Thomson (1997).

We take \( \alpha_{aq} \) to be the decay rates of postsynaptic responses of type \( p = e, i, s \) in neurons of type \( q = e, i \). Average values of \( \alpha_{ee} = 68 \text{ s}^{-1} \), \( \alpha_{ei} = 47 \text{ s}^{-1} \) and \( \alpha_{ie} = 176 \text{ s}^{-1} \) are inferred from Thomson et al. (1996) and Thomson (1997). In order to obtain an estimate of \( \alpha_{ii} \) it is assumed that the more extensive dendritic geometry of excitatory neurons gives rise to longer time constants than in inhibitory neurons, and that the difference in EPSP decay rates equals the difference in IPSP decay rates: \( 1/\alpha_{ee} - 1/\alpha_{ie} = 1/\alpha_{ei} - 1/\alpha_{ii} \). (This method for obtaining a value for \( \alpha_{ii} \) is arbitrary, but does not affect results.) Also we set \( \alpha_{es} = \alpha_{ee} \) and \( \alpha_{ie} = \alpha_{ie} \) because the majority of synapses from the subcortex are thought to be mediated by glutamate acting at AMPA receptors, just like typical excitatory synapses from cortical neurons.

The rise rates of synaptic responses are \( \beta_{aq} \), and the values appearing in Table 1 are inferred from Thomson et al. (1996) and Thomson (1997).

The smoothing function \( H(t) \) is included in eqn (2) since the summed potential \( V_q \) cannot influence synapses instantaneously. The rate constant \( \eta \) is chosen to be approximately equal to the average of the membrane rate constants \( \alpha_{aq} \) and \( \beta_{aq} \).

We choose the maximal firing rates \( Q_{aq}^{\text{max}} \) in Table 1 to be near the upper extreme of observed cortical firing rates, taking into account that adaption causes that maximal firing rates to be lower than that which can be achieved transiently. The reason for setting \( Q_{aq}^{\text{max}} \) to be twice \( Q_{aq}^{\text{max}} \) is to introduce an inequality between the firing rates of pyramidal neurons and interneurons, as is suggested by experiments. There are other equally plausible ways to achieve the same end, but this is the most transparent.

The width parameters \( \sigma_q \) in eqn (5) contribute to the slope of the sigmoid relationship between \( V_q \) and \( Q_q \) and so to the gains of the system. Unfortunately, they cannot easily be determined experimentally, so will be treated as adjustable parameters in the present context. They have a powerful effect on the steady state and stability of the system, so are in practice highly constrained by the properties of actual steady states.

The parameters \( \theta_q \) are the values of \( V_q \) at which neurons (assumed to have a sigmoidal response) fire at 50% of their maximal rate. However \( Q_{aq}^{\text{max}} \) are not well defined or measurable in vivo when other nonlinearities and non-stationarities are considered. Consequently, we adopt \( \theta_{e,i} = -52 \text{ mV} \), which result in appropriate firing rates of about 5–15 s\(^{-1}\), for the values of \( Q_{e,i}^{\text{max}} \) described above.

The firing rate of subcortical neurons \( \phi_s \) is taken to be 10 s\(^{-1}\). However, subcortical activity is variable in response to events and changes of state, so values of up to 100 s\(^{-1}\) will also be considered.

Characteristic damping rates for axonal propagation are \( \gamma_q = v_q/\lambda_q \). The axonal conduction speeds \( v_e = 8 \text{ m s}^{-1} \) and \( v_i = 1 \text{ m s}^{-1} \) are
obtained from Nunez (1995, p. 85, 512). The characteristic axonal range of excitatory neuron is \( r_e = 0.1 \) m according to Nunez (1995, p. 508), and the axon range of inhibitory neurons, equal to the radius of minicolumns, is taken as \( r_i = 10^{-5} \) m (Nunez, 1995).

These values are all the result of physiological measurement, combined with inference to obtain the large-scale average values needed for the continuum model. However, many uncertainties have been noted above, to which must be added the transient, local parameter variations expected in the in vivo brain. In view of this the values in Table 1 must be considered provisional. The results of this choice of parameters are described in the next section, as well as their sensitivity to the assumptions made in this section.

3. Results

The steady state and dispersion relation solutions are described below, including an assessment of parameter sensitivities and the possibility of simplification of the model. Also there is a numerical simulation described.

3.1. STEADY-STATE EQUATIONS

A feature of any model of cortical activity is the form and level of input from subcortical structures, represented here by \( \phi_i(r, t) \). It is known to vary with attention and so, for the purpose of exploring steady-state solutions, it will be taken as a free variable \( \phi_s \) that is locally uniform and quasi-constant.

When the steady-state equations (8) and (9) are solved as a function of \( \phi_s \), with the parameters listed in Table 1, we find that the variables \( \tilde{\phi}_{e,i} \) and \( \tilde{V}_{e,i} \) have just one fixed point for most values of \( \phi_s \) (Fig. 1). In general, however, the solutions of the steady-state equations are multi-valued. It was shown by Robinson et al. (1998a) that in the absence of reversal potentials there can be 1, 3, or 5 solutions for a given \( \phi_s \), depending on the choice of parameters. With the addition of reversal potentials analysis becomes harder, but one special case serves to demonstrate that multiple fixed points are possible in principle. For example, when the parameters listed in Table 1 are adopted, then \( N_{ep} s^{(0)}_{ep} = N_{ip} s^{(0)}_{ip} \) and we see from eqns (8) and (9) that \( \tilde{V}_e = \tilde{V}_i \) and \( \tilde{\phi}_e = (Q_e^{\max}/Q_i^{\max}) \tilde{\phi}_i \). Thus, eqn (8) (with \( q = e \)) can be written as

\[
\tilde{\phi}_e = \frac{V_e^{\text{rest}} - V_e^{(0)} + N_{ee} s^{(0)}_{ee} \tilde{\phi}_e - (1 + (N_{ee} s^{(0)}_{ee} (V_e^{\text{rev}} - V_e^{(0)}) \tilde{\phi}_e)) (\tilde{V}_e - V_e^{(0)})}{-N_{ee} s^{(0)}_{ee} - N_{ee} s^{(0)}_{ee} (Q_e^{\max}/Q_i^{\max}) + ((N_{ee} s^{(0)}_{ee} (V_e^{\text{rev}} - V_e^{(0)}) + (N_{ee} s^{(0)}_{ee} (V_i^{\text{rev}} - V_i^{(0)})) (Q_i^{\max}/Q_e^{\max}) (V_e - V_e^{(0)})
\]

(26)

The solutions of eqn (26), as functions of \( \tilde{\phi}_e \), have the form of a pair of hyperbolas, while the fixed points of the system are those solutions that simultaneously satisfy eqn (9), the independent sigmoidal relation between \( \tilde{V}_e \) and \( \tilde{\phi}_e \).

The four cases depicted in Fig. 2 arise because eqn (26) can have either a negative or positive slope, and the asymptotic value can be either less than or greater than the lower asymptote of the sigmoid. In Fig. 2(a) and (b), there are two sub-cases, which demonstrate that the number of solutions can be from zero to three. (Larger numbers of solutions may also be possible.)

Figure 2 shows the theoretical possibilities arising from eqns (5) and (26), but when values from Table 1 are adopted and \( \phi_s \) is considered a free variable, we find that only those possibilities depicted in Fig. 2(a), (b) and (d) are manifest. The transition from (a) to (b) occurs when \( \phi_s = -312.5 \) s\(^{-1} \), and the transition from (b) to (d) when \( \phi_s = 2.1 \) s\(^{-1} \), and consequently Fig. 2(d) is the one of principal relevance. Multiple solutions only occur when \( \phi_s < 0 \) which is unphysical as a firing rate: for example, those of the type appearing in Fig. 2(b) are found to occur when \( \phi_s \approx -50 \) s\(^{-1} \), although in Fig. 1 they are scarcely resolved.

Regarding the transition that occurs between case (b1) and (d1) at \( \phi_s = 2.1 \) s\(^{-1} \), it is of interest to note that this occurs continuously, as is evident in Fig. 1. The topology of allowed
transitions was discussed in Robinson et al. (1998a) for the more general case where $V_e \neq V_i$, although without synaptic modulation. A new feature of the present steady-state equations is the possibility that for some choices of parameters there are no finite, uniform solutions, as in case (a0).

Once a particular steady-state solution is chosen then the various derived parameters can be evaluated. For reference, Table 2 lists values of several derived parameters for two values of $\phi_s$.

The dispersion relation is a function only of $G_{eq}$, $h_{eq}$, and the rate constants appearing in Table 1. While the rate constants in Table 1 and the gains in Table 2 are the basis for all further numerical results below, the uncertainties in the underlying parameters should be remembered. When each of several parameters is varied, the steady-state values of $\phi_e$ shown in Fig. 3 are obtained. It can be seen that those parameter variations that promote firing of inhibitory neurons generally have the effect of reducing $\phi_e$ for

---

**Fig. 1.** Steady state for the adopted parameters, as a function of $/p_{10}/p_{81}$. On the left are "ring rates of excitatory neurons ($/p_{10}/p_{67}$, solid line) and inhibitory neurons ($/p_{10}/p_{16}/p_{71}$, dashed line), and on the right are the corresponding steady-state potentials $\phi_e (V_e = V_i)$.

**Fig. 2.** Schematic plot of four relevant cases that arise when finding solutions to the simplified steady-state equations (5) and (26), showing various relations between the sigmoid curves (dashed) and hyperbolas (solid). In (a) and (b) the hyperbolas have $d/phi_e/dV_e > 0$ while the opposite is the case in (c) and (d). In (a) and (c) the asymptotic value of $\phi_e$ is $> 0$ while the opposite is the case in (b) and (d). Subcases are possible when $d/phi_e/dV_e > 0$. All possibilities are labeled with the number of solutions, the locations of which are indicated with arrows.
a given \( \phi_s \), making the relationship between \( \phi_e \) and \( \phi_s \) monotonic. Variations of the opposite kind tend to increase \( \phi_s \) and cause the locus of solutions to have an s-like form, giving rise to multiple solutions in certain ranges of \( \phi_s \) in accordance with Fig. 2(b).

It is of interest to note that all parameter variations result in essentially a single family of curves. This implies that all modulations are equivalent within limits, although the sensitivity to some modulations is greater than to others. In particular, the sensitivities to \( \sigma_q \) and \( \theta_q \) mean that their values—although not known experimentally—are strongly constrained by the requirement that \( \phi_e \) is of the order 10 s\(^{-1}\).

### Table 2

Derived parameters for two values of subcortical activation. The fixed ratio of \( \rho_e \) and \( \rho_s \), and of \( G_{ep} \) and \( G_{ip} \) arises from the assumed symmetry between the two classes of neurons, and is equal to \( Q_e^\text{max}/Q_s^\text{max} \).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( Q_e^\text{max} ) (s(^{-1}))</th>
<th>( Q_s^\text{max} ) (s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi_e )</td>
<td>6.4</td>
<td>9.9</td>
</tr>
<tr>
<td>( \phi_s )</td>
<td>12.7</td>
<td>19.8</td>
</tr>
<tr>
<td>( V_e ) (mV)</td>
<td>0.6</td>
<td>1.9</td>
</tr>
<tr>
<td>( \rho_e ) = ( \rho_s )/2(V(^{-1}) s(^{-1}))</td>
<td>2163</td>
<td>3242</td>
</tr>
<tr>
<td>( G_{ep} ) = ( G_{ip} )/2</td>
<td>21.2</td>
<td>31.0</td>
</tr>
<tr>
<td>( G_{ep} ) = ( G_{ip} )/2</td>
<td>-10.8</td>
<td>-18.2</td>
</tr>
<tr>
<td>( G_{ep} ) = ( G_{ip} )/2</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>( h_{ex} ) = ( h_{ix} )</td>
<td>-1.05</td>
<td>-1.64</td>
</tr>
<tr>
<td>( h_{ex} ) = ( h_{ix} )</td>
<td>-6.01</td>
<td>-9.37</td>
</tr>
<tr>
<td>( h_{ex} ) = ( h_{ix} )</td>
<td>-0.03</td>
<td>-0.32</td>
</tr>
</tbody>
</table>

### 3.2. Dispersion Relation

Figure 4 shows solutions of the dispersion relation (24) for two values of \( \phi_s \), and for a range of wavenumbers like that observed in EEGs. The real part of each solution represents conventional frequency, and the imaginary part is the damping rate such that solutions in the upper half-plane would grow exponentially.

The most striking feature of Fig. 4 is the pair of lightly damped solutions near \( \omega = (\pm 250 - 40i) \) s\(^{-1}\). It is notable that these roots, absent in our earliest models but anticipated in our recent general examination of local feedback (Rennie et al., 1999), have a frequency comparable to that of the gamma band (40–60 Hz) of EEG rhythms. The other pair of lightly damped roots is situated near the origin, so its effect is to give the system a strong response near \( \omega = 0 \). For the parameters in Table 1, the roots occur at \( \omega = (\pm 8 - 21i) \) s\(^{-1}\) when \( k = 0 \), which lies in the delta range (4–8 Hz) of EEG frequencies. The frequency and damping of this pair of low-frequency roots are functions of the parameters in Table 1, and moderate variations in many of these parameters (i.e. that are similar to their uncertainties) can cause the roots to have larger or smaller frequencies, or even to become purely damped. For example, if \( V_r \text{est} = V_e^{(0)} + 0.003 \) V and \( k = 0 \) then the roots occur at \( \omega = (\pm 11 - 22i) \) s\(^{-1}\), and when \( V_r \text{est} = V_e^{(0)} - 0.003 \) V they are at \( \omega = -16i \) and \( -26i \) s\(^{-1}\). Consequently, there is considerable uncertainty about the character of these roots, and in the absence of stronger or additional constraints it is difficult to comment on them further. The remainder of this section will discuss the gamma-like roots.

It is of interest to note the absence of modes with \( \omega \approx \pm k v - i/\gamma \), which appear in simpler models (see e.g. Fig. B1), and which can be unambiguously attributed to traveling waves. There is some reflection of these modes in the strong \( k \)-dependence of the lightly damped modes, but they are no longer non-dispersive in character.

### 3.3. Parameter Sensitivity

When looking at the sensitivity of the roots of the dispersion relation to parameter variations there is a natural distinction between those parameters that contribute to the gains \( G_{ep} \) and \( h_{ip} \),
parameters. In each case, the gains generalize to variations in all other gain-related which the latter instability is approached].

becomes unstable for large \( p_10 \)
ted both by the size of \( \eta \), and because we know

constant are

in eqns (17) and (18); the rate constants are \( \alpha_{qp}, \beta_{hp}, \gamma_q, \) and \( \eta \). We shall consider the two classes of parameters in turn.

Subcortical activation \( \phi_s \) is an example of a parameter that modulates the gains, as was demonstrated in Table 2. We find empirically from the dispersion relation eqn (24) that the damping of all roots varies with \( \phi_s \). In particular, the pair near the origin becomes unstable for sufficiently small \( \phi_s \), while the pair near 40 Hz becomes unstable for large \( \phi_s \) [see Fig. 4(b), in which the latter instability is approached].

As demonstrated in Table 3 these observations generalize to variations in all other gain-related parameters. In each case, the gains \( G_{hp} \) and \( h_{hp} \) must first be evaluated using the steady-state equations together with eqns (17) and (18). They are found to vary in proportion to each other. Then when each set of gains is substituted into the dispersion relation the corresponding roots collectively form a locus with increasing frequency and decreasing damping as \( G_{hp} \) and \( |h_{hp}| \) increase. This simple interdependence of all the gains imposed by the steady-state equation is a useful constraint when characterizing the dispersion relation.

Two other parameters not considered in Table 3 are \( V_e^{rev} \) and \( V_i^{rev} \). Since their values are well established, only their necessity to the model will be discussed here. Their retention is suggested both by the size of \( h_{hp} \), and because we know of no other form of modulation satisfying the criteria in Rennie et al. (1999) for a high-frequency resonance. Of the two, \( V_i^{rev} \) is of lesser significance according to Fig. 3 and comparison of \( h_{hp} \) and \( h_{hi} \) in Table 2. Also, selectively disabling \( V_i^{rev} \) has only a small impact: we find that \( G_{hp} \) are increased by about 4% and \( h_{hp} = 0 \), scarcely affecting the steady-state solutions (see Fig. 3) and leaving \( h_{hi} \) as the dominant feedback gain. In contrast, disabling \( V_i^{rev} \) increases \( G_{hp} \) by about 50%, and makes \( h_{hi} \approx 0 \) while doubling \( h_{hp} \). As a result of their observations, some computational advantages might be achieved by eliminating \( V_i^{rev} \), although there is no reason to assume that \( V_i^{rev} \) dominates in all situations.

We shall now consider the effects of the rate constants \( \alpha_{qp}, \beta_{hp}, \gamma_q, \) and \( \eta \). By perturbing each in turn we see in Table 4 and Fig. 5 that the root near 40 Hz is affected by almost all underlying rate constants. In each case, moderate (20%) parameter variations generally result in displacements of a similar magnitude, although the displacements all occur in different directions in the complex-\( \omega \) plane. Given the importance of reversal potentials in this model, the significance of \( \eta \) is not surprising. However, the significant and independent roles played by the dendritic rate constants, and the axonal damping rate \( \gamma_e \), suggests that the gamma-like mode is best considered a property of the system as a whole, rather than a simple consequence of modulation of synaptic strength.

Regarding rate constants, one special case is of particular interest. When decay rates \( \alpha_{qp} \) are
Effects of parameter variations on the principal derived parameters and on the gamma-like lightly damped roots near $\omega = (258.6 - 36.8i) \text{s}^{-1}$ when $k = 0$. Each row is the result of the indicated variations to the values listed in Table 1.

<table>
<thead>
<tr>
<th>Variation</th>
<th>$G_{nc}$</th>
<th>$G_{ni}$</th>
<th>$h_{qei}$</th>
<th>$h_{qii}$</th>
<th>Root (s$^{-1}$)</th>
<th>Displacement (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{nei}$ + 5 mV</td>
<td>24.6</td>
<td>-26.5</td>
<td>-1.2</td>
<td>-7.1</td>
<td>276 - 29i</td>
<td>17.1 + 7.7i</td>
</tr>
<tr>
<td>$\sigma_{c,i}$</td>
<td>25.4</td>
<td>-25.8</td>
<td>-1.6</td>
<td>-8.9</td>
<td>300 - 21i</td>
<td>41.0 + 15.4i</td>
</tr>
<tr>
<td>$Q_{c,i}$ * 1.5</td>
<td>31.3</td>
<td>-31.7</td>
<td>-1.5</td>
<td>-8.9</td>
<td>300 - 18i</td>
<td>40.9 + 18.9i</td>
</tr>
<tr>
<td>$s_{c,i}^{(0,1.5)}$</td>
<td>32.0</td>
<td>-32.8</td>
<td>-1.6</td>
<td>-9.1</td>
<td>303 - 16i</td>
<td>44.0 + 20.4i</td>
</tr>
<tr>
<td>$V_{nei}$ + 5 mV</td>
<td>33.1</td>
<td>-33.9</td>
<td>-1.8</td>
<td>-8.8</td>
<td>302 - 13i</td>
<td>43.2 + 23.5i</td>
</tr>
<tr>
<td>$\phi$ * 10</td>
<td>31.0</td>
<td>-36.5</td>
<td>-1.6</td>
<td>-9.4</td>
<td>309 - 14i</td>
<td>50.4 + 22.2i</td>
</tr>
<tr>
<td>$\sigma_{e}$ + 1 mV</td>
<td>41.1</td>
<td>-51.4</td>
<td>-2.9</td>
<td>-12.7</td>
<td>349 + 4i</td>
<td>90.4 + 40.9i</td>
</tr>
<tr>
<td>$Q_{e}$ * 1.5</td>
<td>84.0</td>
<td>-83.2</td>
<td>-5.3</td>
<td>-20.5</td>
<td>421 + 43i</td>
<td>162.3 + 79.8i</td>
</tr>
<tr>
<td>$s_{e}^{(0,1.5)}$</td>
<td>84.5</td>
<td>-84.0</td>
<td>-5.4</td>
<td>-20.5</td>
<td>422 + 44i</td>
<td>163.8 + 80.5i</td>
</tr>
</tbody>
</table>

TABLE 4

Effects of variations of the rate constants on the gamma-like root at $\omega = (258.6 - 36.8i) \text{s}^{-1}$ when $k = 0$. For variations of 20% to each of the rate constants, the parameter value, the displacement of the root, and the magnitude of the displacement are shown. All values are in units of $\text{s}^{-1}$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Displacement</th>
<th>Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{ei}$</td>
<td>56.4</td>
<td>13.0 + 13.1i</td>
<td>18.5</td>
</tr>
<tr>
<td>$\alpha_{ei}$</td>
<td>211</td>
<td>4.6 + 7.9i</td>
<td>9.1</td>
</tr>
<tr>
<td>$\alpha_{ei}$</td>
<td>96.0</td>
<td>6.0 + 4.4i</td>
<td>7.5</td>
</tr>
<tr>
<td>$\phi$</td>
<td>600</td>
<td>0.2 + 0.7i</td>
<td>0.7</td>
</tr>
<tr>
<td>$\beta_{p}$</td>
<td>600</td>
<td>6.2 + 6.3i</td>
<td>8.8</td>
</tr>
<tr>
<td>$\eta$</td>
<td>240</td>
<td>13.7 + 7.3i</td>
<td>15.6</td>
</tr>
<tr>
<td>$\alpha_{ei}$</td>
<td>81.6</td>
<td>6.1 + 9.2i</td>
<td>11.1</td>
</tr>
<tr>
<td>$\alpha_{ei}$</td>
<td>98.4</td>
<td>0.6 + 10.4i</td>
<td>10.4</td>
</tr>
</tbody>
</table>

identical, and since the rise rates $\beta_{qp}$ are already assumed to be identical, then $S_{qp}(\omega) = S(\omega)$ and eqn (23) simplifies to

$$\frac{\phi_{e}(\omega)}{\phi_{i}(\omega)} = \frac{D_{i}G_{ei}S}{FD_{e}D_{i} - D_{i}G_{ce}S - D_{e}G_{ei}S},$$

(27)

where $F = 1 - (h_{ge} + h_{ql} + h_{qii})S$. The values of $h_{qp}$ (see Table 2) are such that $F \approx -(h_{ge} + h_{ql} + h_{qii})S$, and thus the transfer function becomes independent of $S$, $\phi_{qp}$ and $\beta_{qp}$, and it is found that the dispersion relation is without roots at gamma-like frequencies. This particular simplification is contrary to available physiological measurements, but its interest lies in the fact that should such simplifications be made for analytical convenience, they would eliminate a phenomenon of considerable applied interest.

3.4. NUMERICAL SIMULATION

The characteristics of the gamma-like resonance can be explored further with numerical simulations in order to make realistic comparisons with experiments, and to allow for inhomogeneities in parameter values and for the nonlinearity of eqn (5). Such investigations are not described here. However, applying the numerical methods described in Robinson et al. (1997) to eqns (2)–(7) confirms the existence of
a resonant mode. In the example shown in Fig. 6, the mode near 40 Hz is brought close to instability by setting \( \phi_z \) uniformly to 100 s\(^{-1}\), with additional white noise applied along one edge of the grid. The system displays coherent activity with a frequency similar to that predicted by the dispersion relation, and similar to that observed experimentally. Related simulations are described in more detail by Wright (1999).

4. Conclusion

The use of average steady-state values to characterize neurons is the core simplification on which continuum EEG models are based. However, there are mechanisms operating at the neuronal scale that have time-scales similar to that of EEG, and so challenge this assumption. This paper investigates the consequences of modulations in neuronal characteristics, within the context of one model of the brain’s large-scale electrical activity. In the appendixes three commonly considered modulatory mechanisms are each recast as a form of feedback and the system dynamical characteristics inferred. Only one is found likely to affect the global dynamics of the system. This mechanism, which is studied more thoroughly in the main body of the paper, is the modulation of synaptic strength as a result of the finite synaptic reversal potentials. In contrast, the modulation of synaptic strength by incoming firing rate (facilitation and depression) and of firing threshold by outgoing firing rate (refractoriness) are shown to be inadequate to affect global dynamics significantly. The appendixes also serve to demonstrate that irrelevant complications such as these can be straightforwardly identified by analytical means, which is particularly relevant to achieving numerically efficient simulations of realistic neural networks. As a result of adding reversal potentials alone, the dispersion relation is only slightly more complicated than before, yet has important additional characteristics.

The derivation of each of the model’s parameters is described in some detail in Section 2.3 because there is no sufficiently accurate consensus on the correct values for the human brain. The large number of parameters, the sensitivity of results to some parameters, and the goal of a fully reductionistic model, require use of the best possible values if the emergent behavior is to be usefully compared to real EEG. Yet the difficulty in obtaining good estimates for average human in vivo parameters cannot be overemphasized, and some uncertainty remains in the values of the parameters. The average size and time course of the dendritic PSPs is thought to be modulated by the general level of background activity in ways that are only beginning to be understood (Koch, 1999, p. 416). Also the parameters of the sigmoid response function equation (5), are likely to be functions of cell type, neuromodulators, and average firing rate. Nevertheless, we have used values that are probably of the right order.

As a result of the addition of reversal potentials and a fuller and more accurate parameterization, this version of our model (a) has larger feedforward gains than previously assumed, (b) has additional feedback gains, and (c) the dendritic rate constants are no longer identical. One effect of these changes is that a resonance appears near 40 Hz. This resonance is of interest as it is broadly consistent with the characteristics of the gamma rhythm of EEG in having a frequency in the range 40–60 Hz. One of the characteristics of gamma rhythm is that it appears preferentially in highly activated cortex. In our model, increasing of the level of uniform input to the cortex—or appropriately manipulating any of the gain-related parameters in Table 1—causes the 40 Hz resonance to be enhanced, while suppressing lower frequencies in a reciprocal manner (see Fig. 4). This makes modulated 40 Hz activity a robust feature of this model.
An obvious inference from the present results is that with high levels of cortical activation, the gamma-like resonance could become undamped. The neuronal networks underlying the continuum description would thus exhibit spontaneous firing, probably in the form of limit cycle behavior. This accords with observations in physiological studies of synchronous oscillation (Eckhorn et al., 1989; Gray & Singer, 1989; Singer & Gray, 1995). More speculatively, local patches of cortex would also act as sources of cortical information and exchange, in accord with psychological models of binding and association originated by von der Malsburg (1983), although a far more detailed style of modeling is needed to explore such issues of cognitive function.

It is of particular interest to note that this resonance cannot be properly understood in terms of one physiological mechanism or structure. Rather, it is a property of the whole system, since it is sensitive to all gains (see Table 3), each of the dendritic rate constants ($\alpha_{\text{ap}}$, $\beta_{\text{ap}}$, $\eta$ in Table 4), and to the axonal damping rate $\gamma_{\text{e}}$ (see Fig. 5). This is compatible with the conclusion of Freeman, who first postulated that gamma rhythm arises from the interaction of excitatory and inhibitory neural populations (Freeman, 1975). Both experimentally and in models Eckman & Freeman (1990) found (like us) that the resonant frequency and damping were functions not just of dendritic rate constants, but also of gain, and this was considered a distinction from alternative models postulating the coupling of intrinsic oscillators. Gamma-like oscillations have also been observed in spatially discretized EEG simulations (e.g. Wilson & Bower, 1992; Lumer et al., 1997a), although the numerically intensive nature of such models prevents easy characterization of parameter sensitivities.

Compared to these models, the present work makes broadly similar assumptions about cortical anatomy and physiology. However, its tractable analytical form allows the structure and sensitivities of the steady-state solutions to be described, as can the modes of the system in the linear limit. We also emphasize the value of linear analysis in exploring particular modulatory processes, especially when assessing potential elaborations to the model (see appendixes).

The version of our model described here does not predict any resonances identifiable with $\alpha$. However, we have found separately that realistic spectral envelopes for the range 0–30 Hz can be reproduced if lower gains are used and corticothalamic feedback is added (Unpublished data). These two versions are otherwise identical, and together cover the full EEG spectrum using parameters that are obtained from physiological measurements at the neuronal scale.

This work was supported by the Australian Research Council, the National Health and Medical Research Council, and the Ross Trust.

REFERENCES


ECKHORN, R., FREEMAN, W. J. (1990). Correlations between unit firing and EEG in the rat olfactory system. Brain Res. 528, 238–244.


II. The role of neural synchrony tested through perturbations of spike timing. Cereb. Cortex 7, 228–236.


APPENDIX A

Facilitation

Facilitation refers to an increase in the amount of neurotransmitter released at an axon terminal as a result of preceding action potentials, and to the resulting enhancement in the size of the postsynaptic potential (PSP). According to Magleby (1987), the degree of facilitation depends bi-exponentially on the time $\Delta t$ since the preceding impulse, with

$$A_s(t+\Delta t) = s_0(0)(F_1 \exp[-k_1\Delta t] + F_2 \exp[-k_2\Delta t]),$$

where $s_0(0)$ is the size of the PSP in the limit as firing rate $\to 0$, and $F_1 \exp[-k_1\Delta t]$ and $F_2 \exp[-k_2\Delta t]$ are two components of the increment in synaptic response. Values quoted by Magleby are $F_1 \approx F_2 \approx 0.2$, $k_1 \approx 20 s^{-1}$, and $k_2 \approx 3.3 s^{-1}$ (Magleby, 1987). However, firing rates are typically less than 20 s$^{-1}$ in the cortex, so we shall not consider the first, more rapidly decaying component of facilitation, and instead assume that at each synapse

$$A_s(t+\Delta t) \approx s_0(0)F_2 \exp[-k_2\Delta t].$$

Considering now an ensemble of synapses at $r$, $t$, the average level of facilitation will depend on the distribution of interspike intervals (ISIs). For a Poisson process with a mean rate of $\phi_e(t)$, the distribution of ISIs is

$$I(\tau) = \phi_e \exp[-\phi_e \tau],$$

where $I(\tau)$ is defined only for $\tau \geq 0$. We also require both that $\int_0^\infty I(\tau) \, d\tau = 1$ and $\int_0^\infty \tau I(\tau) \, d\tau = 1/\phi_e$ so that $I(\tau)$ is normalized and is consistent with the specified average firing rate $\phi_e$. By combining eqns (A.1) and (A.3) the average increment to the post-synaptic response due only to the most recent action potential is $A_s(r, t) =$
\[ s_e^{(0)} F_2 \beta, \]

where

\[ \beta = \int_0^\infty I(\tau) \exp[-k_2 \tau] \, d\tau \quad (A.4) \]

\[ = (1 + k_2/\phi_e)^{-1}, \quad (A.5) \]

and more generally, when all previous action potentials are taken into account, it can be shown that the enhancement to the synaptic strength (compared with when \( \phi_e = 0 \)) is

\[ \Delta s_e = s_e^{(0)} F_2 (\beta + \beta^2 + \beta^3 + \cdots) \quad (A.6) \]

\[ = s_e^{(0)} F_2 \beta/(1 - \beta) \quad (A.7) \]

\[ = s_e^{(0)} F_2 \phi_e/k_2. \quad (A.8) \]

In order to obtain an expression that is comparable to those appearing in Rennie et al. (1999), we need to restate eqn (A.8) in terms of perturbations from the steady-state conditions,

\[ s_e(r, t) - s_e^{(0)} = \frac{s_e^{(0)} F_2}{k_2} [\phi_e(r, t) - \phi_e^{(0)}], \quad (A.9) \]

and this is of the form

\[ s_e(r, t) - s_e^{(0)} = s_e^{(1)} H(t) \otimes [\phi_e(r, t) - \phi_e^{(0)}], \quad (A.10) \]

described in Rennie et al. (1999) as Type A feedback, with \( s_e^{(1)} = s_e^{(0)} F_2/k_2 \) and \( H(t) = \delta(t) \).

We are now in a position to make use of the previously derived results for feedback of Type A, according to which the dimensionless feedback strength is

\[ h_A = [s_e^{(1)}/s_e^{(0)}] \phi_e^{(0)} = F_2 \phi_e^{(0)}/k_2, \quad (A.11) \]

\[ H(\omega) = 1, \]

and the relevant dispersion relation is

\[ D_e(D_e - G_{ee} L) - G_{ee} L(D_e - MG_{ii} L) F_A = 0. \quad (A.12) \]

In eqn (A.12), \( D_e \) and \( D_i \) are wave equations given by eqn (16), \( L \) describes the PSPs and is given by an equation similar to eqn (4), \( G_{ee} \) and \( G_{ii} \) are two components of gain of the form (17), \( M \) is usually approximately zero as explained in connection with eqn (25), and only \( F_A = 1 + h_A H(\omega) \) relates to feedback.

By substituting the values quoted above from Magleby (1987) for slow facilitation and setting \( \phi_e^{(0)} = 10 \, s^{-1} \) we find that \( h_A = 0.12 \) and \( F_A = 1.12 \), and thus facilitation in effect enhances the feedforward gain \( G_{ee} \) by a factor \( F_A \). However, \( F_A \) is too close to unity to have much effect on the roots of the dispersion relation.

A similar analysis can be carried out to investigate facilitation at excitatory synapses in inhibitory neurons, as well as at inhibitory synapses in both classes of neurons. However, such analyses are certain to predict even smaller effects, as inhibitory neurons and synapses are in the minority. In conclusion, there is little reason to believe that facilitation provides a mechanism for resonance. Furthermore, the functional description of synaptic depression found by Magleby is sufficiently similar for it also to be discounted as a source of high-frequency resonance.

\section*{APPENDIX B}

\subsection*{Relative Refractory Period}

The origin of this effect lies in a property of action potentials, where following the initial rapid positive-going phase of the action potential there is a more prolonged period of K$^+$ efflux which not only causes the membrane to repolarize but also to overshoot and hyperpolarize. This hyperpolarization is about 5 mV in magnitude and decays with a rate constant of 10–50 s$^{-1}$.

While in reality it is the membrane potential that becomes more negative following the action potential, it is more convenient here to think of the threshold \( \theta \) being raised for a period following an action potential. The effective raising of the firing threshold makes the neuron relatively less likely to fire for approximately 30 ms, and this interval is called the relative refractory period. (There is also a shorter period following an action potential during which the neuron is absolutely refractory, but that will not concern us here.)

In order to represent this effect in a way compatible with our theoretical analysis, we first take the effective threshold to vary as
\( \theta(\mathbf{r}, t) = \theta_0 + (\theta_\eta/\eta)H(t) \) in an individual neuron following the generation of an action potential, and let \( H(t) \) be exponential, to match approximately the observed recovery from hyperpolarization,

\[
H(t) = \eta e^{-\eta t}, \quad t > 0, \tag{B.1}
\]

and \( H(t) = 0 \) for \( t \leq 0 \).

In order to establish a connection with our previous work, the expression for \( \theta(\mathbf{r}, t) \) in excitatory neurons needs to be recast as a perturbation from the steady-state threshold \( \theta_e^{(0)} \), which differs from the resting (or \( Q_e = 0 \) threshold \( \theta_0 \). We note that the effective threshold for a population of neurons under steady-state conditions (and with \( \eta > Q_e \), as is the case here), is approximately proportional to the average firing rate \( Q_e^{(0)} \), with

\[
\theta_e^{(0)} = \theta_0 + (\theta_\eta/\eta)Q_e^{(0)}, \tag{B.2}
\]

so the effective threshold with respect to \( \theta_e^{(0)} \) is

\[
\theta_e(\mathbf{r}, t) = \theta_e^{(0)} + (\theta_\eta/\eta)H(t) \otimes [Q_e(\mathbf{r}, t) - Q_e^{(0)}], \tag{B.3}
\]

\[
= \theta_e^{(0)} + (\rho_e, \theta_\eta/\eta)H(t) \otimes [V_e(\mathbf{r}, t) - V_e^{(0)}], \tag{B.4}
\]

where we have made use of \( \rho_e \equiv [dQ_e/dV_e]_{ss} \) describing the relation between perturbations in \( Q_e \) and \( V_e \) at the chosen steady state.

This form of the feedback equation is directly comparable to that of Type D feedback, in the terminology of Rennie et al. (1999). Thus, it can be inferred that the corresponding dispersion relation is

\[
D_e(D_i - G_{ii}L)/F_D - G_{ee}L(D_i - I_MG_{ii}L) = 0, \tag{B.5}
\]

where \( D_e \) are given by eqn (16), \( L \) is given by an equation similar to eqn (4), \( G_{ee} \) and \( G_{ii} \) are two components of gain and are like eqn (17), \( M \) is usually approximately zero as explained in connection with eqn (25), and only \( F_D = 1 + h_BH(\omega) \) relates to feedback. The feedback gain is obtained from a comparison of eqn (B.4) with the canonical form appearing in Rennie et al. (1999): \( h_B = -\rho_e\theta_\eta/\eta \approx -0.36 \) if we use the value of \( \rho_e \) from Table 2 and assume \( \theta_\eta = 5 \text{ mV} \) and \( \eta = 30 \text{ s}^{-1} \). The frequency dependence of the feedback is \( H(\omega) = (1 - i\omega/\eta)^{-1} \). The roots of the dispersion relation are shown in Fig. B1. As expected from the weak feedback, inclusion of the relative refractory period has little effect on the dispersion relation, reinforcing the conclusion in Rennie et al. (1999) that the feedback gain \( h_B \) needs to be \( \lesssim 1 \) for resonances to appear.

**APPENDIX C**

**Reversal Potential Effect in \( s_p \)**

In this section, we look at the influence of the membrane potential on the size of the excitatory
or inhibitory impulse response $s_p$ in excitatory neurons.

Since channel conductance is largely independent of voltage for most types of ion channels, we shall assume here that $s_p$ is a linear function of perturbations in the membrane potential. There is a net ionic flux within ion channels both because of the potential gradient and because of gradients in the concentration of the various ions, but there is always some potential (the reversal potential) at which the effects of the two gradients balance and there is no net current. The reversal potential is generally non-zero and specific to the type of ion channel (Johnston & Wu, 1995).

The effects of reversal potentials can be incorporated into our model by making $s_p$, the integrated perturbation in the membrane potential, a function of the membrane potential of excitatory neurons. The following expression for the response size in excitatory neurons:

$$s_p(r, t) = s_p^{(0)} \left[ 1 - \frac{V_c(r, t) - V_c^{(0)}}{V_p^{\text{rev}} - V_c^{(0)}} \right] \otimes H(t),$$

(C.1)

has the required characteristics that $s_p = s_p^{(0)}$ at the standard reference conditions (i.e. when $V_c(r, t) \sim V_c^{(0)} = -60 \text{ mV}$), and $s_p = 0$ when $V_c(r, t) = V_p^{\text{rev}}$. The term $H(t)$ is some lag function.

This modulation of $s_p$ by $V_c$ can be identified with Type B feedback, as was described in Rennie et al. (1999), where it was assumed that the modulation of synaptic strength is described by

$$s_p(r, t) = s_p^{(0)} + s_p^{(1)}[V_c(r, t) - V_n^{(0)}] \otimes H(t).$$

Comparison of this canonical form of modulation with eqn (C.1) shows that

$$s_p^{(1)} = s_p^{(0)}/[V_p^{\text{rev}} - V_c^{(0)}].$$

We can then conclude, with reference to Rennie et al. (1999), that the dispersion relation is

$$F_B D_e (D_1 - G_{ii} L) - G_{ee} L (D_1 - MG_{ii} L) = 0,$$

(C.2)

where $F_B(\omega) = 1 - h_B H(\omega) L(\omega)$. In the present case, the dimensionless feedback gain is

$$h_B \equiv -N_e s_p^{(1)} \phi_p^{(0)} = -N_e s_p^{(0)} \phi_p^{(0)} \left[ V_p^{\text{rev}} - V_c^{(0)} \right].$$

(C.3)

With reference to Table 1, and by assuming a typical firing rate $\phi_p^{(0)} \approx 10 \text{ s}^{-1}$, we find the feedback gain has a value of $h_B = -1.6$ at excitatory synapses ($p = e$) and $h_B = -4.7$ at inhibitory synapses ($p = i$). Both satisfy the criterion $h_B \lesssim -1$, which was proposed in Rennie et al. (1999) for the emergence of a lightly damped resonant mode. Consequently, this mechanism is investigated in greater detail in the main part of this paper.
Prediction of electroencephalographic spectra from neurophysiology

P. A. Robinson,1,6 C. J. Rennie,1,2,3 J. Wright,4 H. Bahramali,3 E. Gordon,3 and D. L. Rowe1,3

1School of Physics, University of Sydney, New South Wales 2006, Australia
2Department of Medical Physics, Westmead Hospital, Westmead, New South Wales 2145, Australia
3Brain Dynamics Center, Department of Psychological Medicine, Westmead Hospital and University of Sydney, Westmead, New South Wales 2145, Australia
4Mental Health Research Institute, Parkville, Victoria 3052, Australia

(Received 22 May 2000; revised manuscript received 19 October 2000; published 18 January 2001)

A recent neurophysiological model of propagation of electrical waves in the cortex is extended to include a physiologically motivated subcortical feedback loop via the thalamus. The electroencephalographic spectrum when the system is driven by white noise is then calculated analytically in terms of physiological parameters, including the effects of filtering of signals by the cerebrospinal fluid, skull, and scalp. The spectral power at low frequencies is found to vary as $f^{-1}$ when awake and $f^{-3}$ when asleep, with a breakpoint to a steeper power-law tail at frequencies above about 20 Hz in both cases; the $f^{-1}$ range concurs with recent magnetoencephalographic observations of such a regime. Parameter sensitivities are explored, enabling a model with fewer free parameters to be proposed, and showing that spectra predicted for physiologically reasonable parameter values strongly resemble those observed in the laboratory. Alpha and beta peaks seen near 10 Hz and twice that frequency, respectively, in the relaxed wakeful state are generated via subcortical feedback in this model, thereby leading to predictions of their frequencies in terms of physiological parameters, and of correlations in their occurrence. Subcortical feedback is also predicted to be responsible for production of anticonnected peaks in deep sleep states that correspond to the occurrence of theta rhythm at around half the alpha frequency and sleep spindles at 3/2 times the alpha frequency. An additional positively correlated waking peak near three times the alpha frequency is also predicted and tentatively observed, as are two new types of sleep spindle near 5/2 and 7/2 times the alpha frequency, and anticorrelated with alpha. These results provide a theoretical basis for the conventional division of EEG spectra into frequency bands, but imply that the exact bounds of these bands depend on the individual. Three types of potential instability are found: one at zero frequency, another in the theta band at around half the alpha frequency, and a third at the alpha frequency itself.

DOI: 10.1103/PhysRevE.63.021903

PACS number(s): 87.10.+c, 87.19.La, 87.18.--h

1. INTRODUCTION

Electroencephalograms (EEGs) have diverse forms that are correlated with differing pathologies and states of arousal, as determined independently by behavioral and clinical measures [1], and are inferred to be closely connected to brain dynamics, information processing, and cognition [2,3]. These correlations are widely used diagnostically, where they have been employed to elucidate fundamental mechanisms of overall brain function and dysfunction. Even so, they are not well understood in terms of the underlying physiology, despite 125 years' work [1,4,5]. Prominent in EEGs are the alpha rhythm (narrowband at 7–11 Hz in the relaxed waking state, weakening rapidly with increased attention, or with sleep), and beta rhythm (broader band at 15–25 Hz), and sleep spindles (12–14 Hz, occurring in deeper sleep and sometimes termed sigma activity), as shown in Fig. 1. The experimental setup and procedures for acquisition of these spectra have been described in detail elsewhere, along with the methods used to ensure that artifacts do not significantly contaminate the data [6]; a brief summary is given in Appendix A. The origins of these rhythms are also obscure and have not been quantified in terms of physiology, although the thalamus is widely thought to be involved in alpha, beta, and sleep spindles [7–12], since experiments show high coherence between EEGs in the two structures [10,13–15]. Underlying the specific rhythms is a smoother background EEG spectrum whose origin and structure are even more cryptic. Recently, analogous magnetoencephalographic (MEG) measurements revealed that the low frequency component of this background was 1/f noise, whose power spectrum $P(f)$ decreased as $P(f)\sim f^{-a}$, with $a\approx 1$ [16]; at high frequencies the decrease was much steeper. In deep sleep there is a strong enhancement of the spectrum at low frequencies, in the so-called delta ($f<3.5$ Hz) and theta ($4$ Hz $<f<7.5$ Hz) ranges [1,10,17,18].

Scientific and clinical understanding of EEGs would be greatly facilitated by a theory, based on physiology, that could predict features of EEG spectra such as those mentioned above. Consequently, many attempts to model large-scale cortical electrical activity have been made, with varying degrees of physiological realism and incorporating differing aspects of the cortex. One avenue has been to average over microscopic neural structure to develop continuum cortical models on scales from millimeters up to the whole cortex [7,10,15,19–33]. Some continuum models yield wave equations for the propagation of cortical activity [10,21–24,27,31–33], which are particularly convenient for analysis and which incorporate realistic anatomical features such as separate excitatory and inhibitory neural populations.
back and filtering by overlying tissues. Analytic predictions for EEG spectra in terms of physiological parameters are then made in Sec. III, and the structure of the underlying spectrum is estimated. Parameter sensitivities are explored in Sec. IV, leading to elimination of several parameters to yield a simplified model that is well suited for comparison with data. This simplified model is used to estimate the locations, sizes, and possible instabilities of spectral peaks in Sec. V, and thereby formulate constraints on the underlying feedback physiology. These results justify the customary division of EEG spectra into distinct bands, but highlight the significance of individual variability in this subdivision. In the case of spectra predicted when the cortex is near marginal stability, a state that has been inferred from MEG observations [16], further simplifications can be obtained; these are discussed in Sec. VI.

II. THEORY

In this section we outline the main relevant results of our recent wave-equation formulation of continuum cortical dynamics [21–23], which incorporates the physiological effects discussed in Sec. I. We then generalize this model in Sec. II B to incorporate a corticothalamic feedback loop, motivated by a range of physiological considerations.

A. Basic model

The mean firing rates (or pulse densities) $Q_a$ of excitatory ($a = e$) and inhibitory ($a = i$) neurons are approximately related to the cell-body potentials $V_a$ by

$$Q_a(r,t) = \frac{Q_a^{\max}}{1 + \exp\left(-\frac{\pi}{\sigma_a^2}\left(V_a(r,t) - \theta_a\right)\right)},$$  \hspace{1cm} (1)

where $\theta_a$ is the mean firing threshold of neurons of type $a$, $\sigma_a$ is the standard deviation in this threshold in the neural population, and $Q_a^{\max} \cong 100$ s$^{-1}$ is the maximum attainable firing rate. The coordinate $r$ refers to position on the cortex, modeled as a two-dimensional sheet, which is justified by its relative thinness.

The quantity $V_a$ is the potential at the cell body after inputs have been summed and filtered through the dendrites. A good approximation to $V_a$ is [21–23]

$$V_a(r,t) = \int_{-\infty}^{\infty} L(t-t')P_a(r,t')dt',$$  \hspace{1cm} (2)

with

$$L(u) = \frac{\alpha\beta}{\beta - \alpha}(e^{-\alpha u} - e^{-\beta u})\Theta(u),$$  \hspace{1cm} (3)

where $P_a$ is the mean potential generated by action potentials arriving from other neurons at the dendrites of neurons of type $a$, $\Theta$ is the unit step function, and $\alpha$ and $\beta$ are dendritic rate constants, with $\beta \geq \alpha$ assumed without loss of generality. The Fourier transform of $L(u)$, used below, is

$$L(\omega) = (1 - i\omega/\alpha)^{-1}(1 - i\omega/\beta)^{-1},$$  \hspace{1cm} (4)
which implies that the dendrites act as a low-pass filter with cutoff frequency $\alpha$ and a steeper fall-off for $\omega > \beta$.

The potential $P_a$ at a particular location comprises contributions from fields $\phi_{a,i,j}$ that represent signals propagating from other cortical neurons, and subcortical inputs $\phi_i$ [21,23]:

$$P_a = N_{ai}s_is_i + N_{ai}s_i\phi_i + N_{ai}s_i\phi_s.$$  

(5)

Here, $N_{ab}$ is the mean number of couplings from neurons of type $b = e,i,s$ to those of type $a$, and $s_b$ is the size of the response to a unit signal from neurons of type $b$.

Outgoing pulses from each neuron propagate along its axon and axonal tree at a velocity $v \approx 10$ m s$^{-1}$. Assuming an isotropic distribution of axons with a distribution of ranges in accord with experiment (see Ref. [21] for the exact distribution), this propagation can be described by damped wave equations for the fields $\phi_a$, which represent axonal signals:

$$D_a\phi_a(r,t) = Q_a(r,t),$$

$$D_a \frac{\partial^2}{\partial t^2} + \frac{\partial}{\partial t} + \gamma_a^2 - u^2v^2,$$

(6)

(7)

where $\gamma_a = v/r_a$ and $r_a$ is the characteristic range of axons of type $a$ [21]. Strictly speaking, very short-range excitatory axons have lower $v$ than the longer range corticocortical ones, because they are not myelinated. This effect is omitted here on the grounds that the difference will not have strong effects on long-range wave propagation, but could be incorporated by splitting excitatory neurons into two populations with a corresponding division of the field $\phi_e$ into two parts.

Equations (1)–(7) determine the steady states of cortical activation, when the cortex is driven by a constant, spatially uniform stimulus $\phi_e^{(0)}$. Small perturbations relative to these steady states obey a linear wave equation which yields the transfer functions [23]

$$\phi_k(k,\omega) = \frac{G_{ee}L(\omega)}{D_a[1 - G_{ii}L(\omega)] - G_{ee}L(\omega)},$$

(8)

$$\phi_k(k,\omega) = \frac{D_eG_{ii}G_{ee}L(\omega)}{G_{es}G_{ee}L(\omega)},$$

(9)

$$D_e(k,\omega) = k^2r_e^2 + (1 - i\omega/\gamma_e)^2,$$  

(10)

in Fourier space. Here, the gain parameters $G_{ab} = \rho_{ab}N_{ab}s_b$ express the response strength in neurons $a$ due to a unit signal incident from neurons of type $b$. The parameter $\rho_{ab} = dQ_a^{(0)}/dV_a = \pi Q_a^{(0)}a_{ab}^3$ is evaluated in the steady state where $Q_a^{(0)} = 5 - 10$ s$^{-1}$ is the steady-state firing rate. In writing (8) and (9) we used the short range of inhibitory axons to set $D_e = 1$, the local inhibition approximation [21], and assumed that the numbers of interconnections between neural types are proportional to the number of available synapses, the random connectivity approximation [21–23,28,34].

### Table 1. Physiological estimates of parameters of the model, as discussed in Refs. [23] and [24]. The limits given are only approximate.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{e,i}$</td>
<td>3–8</td>
<td>mV</td>
</tr>
<tr>
<td>$\theta_{e,i}$</td>
<td>10–25</td>
<td>mV</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>50–200</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$\beta/\alpha$</td>
<td>1–5</td>
<td></td>
</tr>
<tr>
<td>$r_e$</td>
<td>70–100</td>
<td>mm</td>
</tr>
<tr>
<td>$r_i$</td>
<td>$-0.1$</td>
<td>mm</td>
</tr>
<tr>
<td>$v$</td>
<td>7–10</td>
<td>m s$^{-1}$</td>
</tr>
<tr>
<td>$t_0$</td>
<td>25–100</td>
<td>ms</td>
</tr>
<tr>
<td>$N_{ee}, N_{ie}, N_{is}$</td>
<td>2000–6000</td>
<td></td>
</tr>
<tr>
<td>$N_{ei}, N_{ei}, N_{ii}$</td>
<td>300–1000</td>
<td></td>
</tr>
<tr>
<td>$N_{es}, N_{is}$</td>
<td>30–100</td>
<td></td>
</tr>
<tr>
<td>$s_{e,i}$</td>
<td>1–3</td>
<td>$\mu$V s</td>
</tr>
<tr>
<td>$-s_i$</td>
<td>5–8</td>
<td>$\mu$V s</td>
</tr>
<tr>
<td>$\gamma_e$</td>
<td>70–150</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$\gamma_i$</td>
<td>$-10^3$</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$G_{ee}^{\text{max}}$</td>
<td>100–300</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$G_{ee}$</td>
<td>5–10</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$\rho_{es}$</td>
<td>1000–6000</td>
<td>V$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$G_{ee}$</td>
<td>2–90</td>
<td></td>
</tr>
<tr>
<td>$-G_{ii}$</td>
<td>1.3–60</td>
<td></td>
</tr>
<tr>
<td>$-G_{es}$</td>
<td>0.03–2</td>
<td></td>
</tr>
</tbody>
</table>

The parameters of the above model are physiologically measurable, although not all have yet been measured precisely. Table I lists typical values inferred from physiology [23,24], plus values of several other quantities that are discussed below.

#### B. Subcortical feedback

Several suggestions for the production of the resonances that correspond to observed cortical rhythms have been made. These include the possible occurrence of standing waves in the cortex, a potential scenario that is incorporated in the model outlined in Sec. II A, but which does not appear to yield sufficiently sharp resonances to account for observations [21–23] if physiologically realistic input parameters are used in our model. A second possibility is the presence of a thalamic “pacemaker” or “clock” [9,10], which we do not consider here, although it can be incorporated into our model. Third, it has been widely noted that the thalamus also displays alpha rhythm (and possibly beta and sleep-spindle peaks), with a high degree of correlation with the cortical one [7–10,13,14,18], which may result either from pacemaker activity or from some form of corticothalamic feedback. Hence, we are particularly motivated to include corticothalamic feedback in our model.

A key subcortical feedback loop that may underlie production of the alpha rhythm involves feedback via the diffuse thalamocortical projection system [8,9,13,14,35]. This system projects both to other thalamic nuclei and from the thalamus to the whole cortex and thus has the potential to be
involved in modulating large-scale EEG activity [8,14,18,36]; it also has the correct latency of tens of ms [8]. Other corticothalamic loops pass through the so-called specific nuclei of the thalamus (e.g., the pulvinar, lateral posterior, and ventral anterior nuclei), which themselves project preferentially back to the same cortical areas that stimulate them [1,36]. We stress that, although we assume corticothalamic feedback in the remainder of this paper, for definiteness and simplicity, the results obtained are applicable to other potential feedback loops with like characteristics. Corticothalamic feedback should be significant because the thalamus is known to strongly affect the state of attention of the cortex, implying the existence of influential projections from thalamus to cortex. Strong projections in the reverse direction also exist to complete the loop, involving an even greater number of axons [37–39].

Corticothalamic (CT) feedback is modeled approximately here by assuming that $\phi_i$ is the sum of a non-CT part $\phi_N$ and a feedback $\phi_T$, which travels along excitatory neurons that project to excitatory and inhibitory neurons in the thalamus before returning to the cortex via other excitatory neurons. This situation is illustrated in Fig. 2. The feedback signal thus passes through $n \sim 1$ additional neurons compared to direct corticocortical connections (which involve more than one neuron on average). This adds a propagation time delay $t_0$ and $n \sim 1$ extra stages of dendritic filtering with rate constants $\eta_1$ and $\eta_2$ of the same order as $\alpha$ and $\beta$. We thus approximate $\phi_T$ by

$$\phi_T = T_e(\omega) \phi_e .$$

(11)

This incorporates feedback of both excitatory and inhibitory input signals to excitatory neurons (thereby generating the field $\phi_e$) that project to the thalamus. After time delays and thalamic low-pass filtering, encapsulated in the transfer function $T_e(\omega)$, this signal returns to the cortex via other excitatory neurons. We also allow for the possibility of both direct feedback, and feedbacks that emphasize changes in cortical signals by differentiating them in the loop. Differentiation circuits are commonplace in the brain (e.g., for edge detection in the visual field), so the latter type of feedback certainly cannot be ruled out, and has been discussed previously in the context of CT feedback [8]. Limiting cases of truly specific feedback (each point in the cortex feeds back to itself) and totally diffuse feedback with uniform driving (every point feeds back to the whole cortex, making every point equivalent to every other one if the driving is at least statistically uniform, thereby reducing the system to a zero-dimensional one) are included in the form (11), but intermediate types of spatial variation would require a $k$-dependent transfer function and are not considered here.

The frequency-domain transfer function that corresponds to the above physiology can be written as

$$T_e(\omega) = G_{eT} \Psi(\omega) \tau(\omega)/G_{eT} ,$$

(12)

$$\Psi(\omega) = \psi - i \omega t_0 \psi' ,$$

(13)

$$\tau(\omega) = e^{i \omega t_0} \left(1 - i \omega/\eta_1 \right)^n \left(1 - i \omega/\eta_2 \right)^n ,$$

(14)

where $\tau(0) = 1$, and $\psi$ and $\psi'$ measure the strengths of direct and differential feedbacks. The factor $t_0$ in (13) is included to make $\psi'$ dimensionless.

The magnitudes and signs of $\psi$ and $\psi'$ in (13) are expected to vary with state of arousal as the relative responses of the excitatory and inhibitory neurons in the thalamus change. Such changes may result from the action of neurotransmitters or neuromodulators that preferentially affect one population or the other, for example. Hence, if the coupling of thalamic inhibitory neurons to thalamocortical excitatory fibers exceeds that of excitatory ones in some states, the thalamus will invert incoming signals as well as filtering, delaying, and (in part) differentiating them. In Sec. IV we will see that, for many purposes, one may approximate (14) by setting $n = 1$ and $\eta_1 = \eta_2 = \alpha$.

Combining (8)–(14), we find the transfer functions for non-CT stimuli to be given by

$$\phi_e(\mathbf{k}, \omega) = G_{eT} L(\omega) \left[D_e \left[1 - G_{eT} L(\omega) \right] \right]^{-1} ,$$

(15)

and the analog of (9).

The dispersion relation of waves in our model system is given by setting the denominator of (15) to zero, giving

$$k^2 + q^2(\omega) = 0 ,$$

(16)

$$q^2(\omega) + \frac{2}{G_{eT} \left[1 + \Psi(\omega) \tau(\omega) \right] L(\omega)} ,$$

(17)

In (11)–(17) linearization is implicitly carried out relative to steady states of the combined corticothalamic system, not the cortex alone. The steady state values in Table I refer to the combined system.
III. ELECTROENCEPHALOGRAPHIC SPECTRA

In this section we derive the form of the EEG spectrum from the transfer function (15). We then derive analytic forms for quantities such as the \( \omega = 0 \) stability boundaries of the system, the asymptotic forms of the spectrum at high and low frequencies, and breakpoints between different ranges. We also examine the effects on the spectrum of filtering through the cerebrospinal fluid, skull, and scalp.

A. Filtering by overlying structures

In calculating scalp EEG spectra (rather than intracranial ones at the cortical surface), one must consider the possibility of filtering due to shielding that results from volume conduction by the cerebrospinal fluid, skull, and the scalp itself [10,20,40]. Filtering of high spatial frequencies \( k \) can be significant, as found by Srinivasan et al. [40]. The \( k \) dependence of their low-to-moderate wave number results is reasonably well fitted by a spatial filter of the form

\[
F(k) = e^{-k^2/k_0^2},
\]

where \( F(k) \) is the square of the ratio of scalp to cortical voltage and \( k_0 \approx 30 \) m\(^{-1}\). We show in Appendix B that the precise form of this function does not significantly affect the form of the final spectra calculated, so we use the form (18) because it allows the spectrum to be evaluated in a convenient closed form.

B. Spectra

EEGs result from fields generated as charges flow in and out of neural membranes as neurons fire and the resulting signals propagate. Hence, we assume that the observed signals primarily result from the firing of excitatory neurons, since these are larger and much more numerous than inhibitory ones, and have a higher degree of spatial alignment, which implies they contribute more coherently to the total signal [10]. Further reasons for ignoring \( \phi_e \) are that its transfer function is proportional to the one for \( \phi_i \), at small \( k \) (thereby giving only a multiplicative correction), while volume conduction filters out both signals at large \( k \) [10,40]. We see below that the bulk of the EEG power is at large scales where both fields have the same spectral profile in any case.

Stimuli to the cortex have complicated temporal and spatial dependences. We approximate these here by assuming that their total \( \phi_N \) consists of white noise in space and time, which is effectively filtered via its interaction with the cortex to yield the EEG spectrum. This picture is consistent with recent results that indicate that at least the alpha frequency range is consistent with linearly filtered noise in most cases [7,41], although in 1.25% of cases possible evidence of nonlinear behavior was found.

The resulting spectrum is

\[
P(\omega) = \int |\phi_N(\mathbf{k}, \omega)|^2 F(k) d^2 k
\]

\[
= P_N \left| \frac{\phi_N^2}{r_e^2} \right| \frac{G_e L(\omega)}{1 - G_i L(\omega)} \frac{1}{q^2 r_e^2} \int \frac{F(k) d^2 k}{|k^2 + q^2|^2},
\]

where \( \phi \) is the complex argument of \( q^2 \), \( |\phi_N^2| \) is the white-noise power level in Fourier space, and \( E_1 \) is the exponential integral function [42], with

\[
E_1(z) = - \gamma - \ln z - \sum_{j=1}^{\infty} \left( \frac{(-z)^j}{j!} \right).
\]

for \( |\arg z| < \pi \), where \( \gamma = 0.5772 \ldots \) is Euler’s constant. The spectrum (21) has been calculated for the specific form (18) of \( F(k) \).

The limit \( k_0 \rightarrow \infty \) corresponds to the absence of volume conduction and, in the related limit \( |q^2| / k_0^2 \rightarrow 0 \), we find that (21) simplifies to

\[
P(\omega) = P_N \left| \frac{\phi_N^2}{r_e^2} \right| \frac{G_e L(\omega)}{1 - G_i L(\omega)} \frac{1}{q^2 r_e^2} \sin \theta.
\]

In the opposite limit, \( |q^2| / k_0^2 \gg 1 \), expansion of (21) in powers of \( k_0^2 / q^2 \) yields

\[
P(\omega) = P_{\infty}(\omega) \frac{k_0^2 \sin \theta}{q^2 r_e^2},
\]

where \( P_{\infty}(\omega) \) is the spectrum given by (24). At large frequencies, where \( |q^2| \) becomes large and \( |\theta| \approx \pi \), this implies an asymptotic frequency filtering function

\[
F(\omega) \approx \omega^{3/2} / \omega^3,
\]

with a turnover frequency of

\[
\omega_f = \gamma (2 k_0^2 r_e^2 / \pi)^{1/3}.
\]

For the parameters in Table I, the turnover frequency occurs at about 30 Hz—for too high to significantly affect the main part of the frequency spectrum. Equation (26) does not have the form \( F(\omega) = \exp(-\phi^2 \omega^2 / k_0^2) \) one might naively expect from (18) for a dispersion relation whose real part has the asymptotic form \( \omega = kv \). This difference occurs because of damping, which implies that \( \omega \) is the solution of the complex dispersion relation (16). Hence, the Fourier transform of waves satisfying this relation has a Lorentzian profile in the real part of \( \omega \), giving \( F(\omega) \propto \omega^{-2} \); the actual \( \omega^{-3} \) dependence comes about because of the detailed structure of (21), which causes the coefficient of the \( \omega^{-2} \) term in an expansion in powers of \( \omega^{-1} \) to be zero.

The shape of the spectrum depends strongly on the locus of \( q^2 \) in the complex plane, with instability occurring if this locus crosses the negative real axis, which corresponds to \( q \) itself acquiring a negative real part, consistent with earlier
work [21,22]. Stability at $\omega=0$ corresponds to $\Re[q^2(0)] > 0$, $\Im[q^2(0)]=0$, and $\theta(0)=0$, implying that we can define a stability parameter $S$

$$S = 1 - \frac{G_{ii} (1+\psi)}{1 - G_{ii}}, \quad (28)$$

that must satisfy $S \geq 0$ for the system to avoid instability at $\omega=0$ (this is a necessary, but not sufficient, condition for overall stability).

### C. Illustrative spectrum

Figure 3(a) shows an illustrative spectrum obtained from (21) for parameters typical of those in Table I. Alpha and beta peaks (plus weak, higher-frequency ones), are seen near 8 Hz and 22 Hz, respectively, superposed on a smooth underlying spectrum which has approximately $f^{-5}$ behavior at large $f$, $f^{-1}$ behavior for 0.4 Hz $\leq f \leq 3$ Hz, and levels off as $f \to 0$. Figure 3(b) shows the corresponding locus of $q^2(\omega)$ for $\omega > 0$. After starting near the origin at $f=0$, $q^2$ increases in modulus, with $q^2(\omega) \propto \omega$, where $P(f) \propto f^{-1}$. Feedback then causes the curve to loop back toward the origin several times, producing correlated alpha, beta, and higher frequency enhancements. Finally, the locus recedes along an asymptotically parabolic path in the regime with $P(f) \propto f^{-5}$ until filtering via volume conduction cuts in, and $P(f) \propto f^{-8}$ thereafter, although the latter regime is not fully attained even at 100 Hz. The onset of dendritic filtering produces the noticeable knee in the spectrum at $\omega=\alpha$, thereby restricting the potential range of the $f^{-1}$ spectrum to $\omega \leq \alpha$.

Momentarily ignoring $F(k)$, the integrand in (20) peaks where $k^2 = \max[0, -\Re q^2]$, and has a characteristic width of $|\Im q^2|$. Hence, in a cortex of characteristic linear size $l$, the mode number $M$ at which the integrand peaks is

$$M = \frac{1}{2 \pi r_e} \left[ \max[0, -\Re(q^2 r_e^2)] \right]^{1/2}, \quad (29)$$

and the characteristic width of the integrand in $M$ is of order $(l/2 \pi r_e)[|\Im(q^2 r_e^2)|]^{1/2}$. Figure 3(b) thus shows that $M$ is small up to at least the beta frequency. This implies that EEGs in this regime are spatially large scale, thereby reducing spectral distortions that would otherwise occur due to the effects of volume conduction [10,40]. This point is explored further in Sec. IV.

### D. Features of the underlying spectrum

In this section we temporarily ignore the spectral peaks (which are discussed in Sec. V) and examine the properties of the smooth, underlying spectrum. At very low frequencies (21) implies $P(\omega) \approx P(0)$, with $\theta(\omega)=0$ and, hence,

$$P(0) = \frac{P_N}{G_0^2}, \quad (30)$$

$$G_0 = 1 - G_{ii}, \quad (31)$$

provided the system is stable ($S > 0$) and $|q^2| \ll k_0^2$. This prediction is in accord with the physical expectation that a small stability parameter will be associated with high EEG power. In this regime, only $S$ and $G_0$ affect the power for constant $P_N$; none of the feedback parameters affect the power level.

If $S$ is small, $q^2(0) \approx 0$ and the behavior of $P(\omega)$ at small $\omega$ depends on the leading terms in the expansion of $q^2(\omega)$ in powers of $\omega$. If we write

$$q^2(\omega)r_e^2 = \sum_{j=0}^{\infty} A_j (-i \omega)^j, \quad (32)$$

the structure of (17) implies that all the $A_j$ are real. Explicit expressions for $A_0$ and $A_1$ are found in Appendix C.

The case $A_0=0$ corresponds to $S=0$ (see Appendix C) and, hence, to marginal stability, a condition that favors complex dynamics, such as actually seen in EEG recordings [41]. If $A_0$ is small and $A_1 \neq 0$, there is a small-$\omega$ regime in which $q^2(\omega) \approx -i \omega A_1$. In this case, one finds $\theta \approx \pm \pi/2$ for $\omega > 0$ and

$$P(\omega) \approx \frac{P_N}{G_0^2} \frac{\pi}{2 \omega |A_1|}. \quad (33)$$

Such a regime with $P(\omega) \propto \omega^{-1}$ was recently observed in magnetoencephalographic observations [16]. Significantly, all the model parameters other than $k_0$ affect the behavior in this regime, via $A_1$.

If $A_0$ and $A_1$ are both very small, the $j=2$ term in (32) will dominate in magnitude over the $j=3$ term at small $\omega$ unless $A_2=0$, while the $j=3$ term dominates the imaginary part of $q^2$. Assuming that $A_2$ is not too close to zero, two cases exist: if $A_2 > 0$, $\theta \approx \pm \pi$ and

$$P(\omega) = \frac{P_N}{G_0^2} \frac{\pi}{|A_2| \omega^2}. \quad (34)$$

If $A_2 < 0$, $\theta \approx \pi A_3/A_2$, and

$$P(\omega) = \frac{P_N}{G_0^2} \frac{1}{|A_3| \omega^2}. \quad (35)$$
Regimes exist with these behaviors provided $A_0 \approx 0$ (which implies $S=0$), and $A_1 \approx 0$. Higher order cancellations of the first few terms in (32) may occur, but such occurrences represent increasingly smaller zones of parameter space, so we do not consider them in this paper. If $A_2 > 0$ the condition $A_1 = 0$ defines a stability boundary of the system, since $q^2$ crosses the negative imaginary axis infinitesimally close to the origin at $\omega = 0$ in this case.

In applying (34) and (35), the condition $S=0$ implies an interrelationship between $G_{ii}, G_{ee}$, and $\psi$, via (28). Likewise, the condition $A_1 = 0$ at $S=0$ requires

$$
\psi\left(\frac{2}{\gamma_e} + \frac{2}{\alpha G_0} + \frac{n}{\eta_1} + \frac{n}{\eta_2}\right) - \psi' t_0 = -\left(\frac{2}{\gamma_e} + \frac{2}{\alpha G_0}\right),
$$

where Eq. (C2) of Appendix C has been used.

At high frequencies with $\omega > \alpha$ the term $(1-i\omega/\gamma_e)^2$ on the right-hand side of (17) dominates, $\theta = -\pi$ for $\omega > 0$, $|q^2 r_e^2| \sin \theta = -2 \gamma_e/\omega$, and $|L(\omega)|^2 \approx \alpha^2 \beta^2 / \omega^5$. If volume conduction is not important, these scalings give

$$
P(\omega) = P_N \sigma^2 \alpha^2 \beta^2 \gamma_e \left(\frac{2}{\omega^5}\right),
$$

in this range, provided $|G_{ii}L(\omega)| \leq 1$ [22]. In this regime, the power level is most sensitive to $\alpha$ and $\beta$, with a weaker dependence on $\gamma_e$, while the feedback parameters have no effect. If volume conduction is strong ($\omega \approx \omega_f$), the filtering factor in (25) must be included, steepening the spectrum to an $\omega^{-8}$ dependence. In addition, the existence of an $\omega^{-1}$ regime requires $|G_{ii}L(\omega)| \geq 1$. Hence,

$$
1 + \omega_{\min}^2 \beta^2 \leq |G_{ii}| \leq 1 + \omega_{\max}^2 \beta^2,
$$

where $\omega_{\min}$ is the minimum frequency of any observed $\omega^{-5}$ regime and $\omega_{\max}$ is the maximum frequency of any observed $\omega^{-1}$ regime.

IV. PARAMETER SENSITIVITIES AND SIMPLIFIED MODEL

In this section we examine how the spectrum (21) varies with each of the model parameters $\gamma_e$, $\alpha$, $\beta$, $G_{ii}$, $G_{ee}$, $\psi$, $\psi'$, $t_0$, $n$, $\eta_m$ (with $\eta_1 = \eta_2$ assumed for convenience), and $k_0$. Except for the parameter being varied at any given time, the parameters are identical to those of Fig. 3 throughout this section and $P_N$ is fixed. We term the values in Fig. 3 “nominal” for convenience. Possible correlations between changes in parameters due to their common dependence on steady state firing rates, for example, are not included here; only one parameter at a time is varied.

Toward the end of this section, we argue that several of the model’s parameters can be absorbed into one another for most purposes, leaving a simplified model with fewer parameters for fitting to the shapes of experimental spectra. We also clarify the a priori physiological constraints on these parameters. Further simplifications are made in Sec. VI, where the number of parameters is further reduced by one in the case of marginally stable spectra.

Figure 4(a) shows the effect of varying $\gamma_e$ from 70 to 150 $s^{-1}$ (nominal value: 110 $s^{-1}$). At high $f$, $P \approx \gamma_e$ is satisfied in accord with (37), whereas there is little effect at low $f$, as expected from (30).

Figure 4(b) shows the effect of varying $\alpha$ from 40 to 100 $s^{-1}$ (nominal value: 70 $s^{-1}$), with $\beta/\alpha = 4$ fixed. For $f \geq \alpha$, $P \approx \alpha^2$, in accord with the effects of dendritic filtering incorporated in (37). However, dependence on $\alpha$ (via $A_1$) is weak in the $f^{-1}$ regime, in agreement with (33). The corresponding loci of $q^2 r_e$ have loops that expand and approach the origin as $\alpha$ increases, producing the enhanced peaks seen—for a strong alpha peak to be seen at a frequency $\omega_a$, one must have $\alpha \approx \omega_a$. The effects of increasing the ratio $\beta/\alpha$ from 1 to 7 (nominal value: 4), seen in Fig. 4(c), are quite similar to those of increasing $\alpha$ itself, in the frequency ratio of interest—higher $\beta/\alpha$ corresponds to less dendritic filtering, more clearly visible spectral peaks, and a shallower (as $\omega^{-3}$ rather than $\omega^{-5}$) fall-off at frequencies $\omega \approx \omega_a$, before the onset of an $\omega^{-5}$ tail for $\omega \approx \beta$.

Figure 4(d) shows the effect of varying $G_{ii}$ from $-0.5$ to $-1.5$ (nominal value: $-1$). For $G_{ii} = 0$, $S \approx 0$ is satisfied, leading to a 1/f spectrum, with instability at $f = 0$ in the first case ($q^2$ intersects the negative real axis at $\omega = 0$), and marginal stability in the second. For $G_{ii} = -1.5$, $S > 0$ holds, and a plateau occurs at low $f$, as in (30). Below a few Hz, $P$ decreases as $G_{ii}$ increases in all cases, in accord with (30) and (33). The main effect at higher frequencies is that the alpha and beta resonances weaken with increasing $G_{ii}$, a reflection of contracted loops of $q^2$ in the complex plane. Similar effects are seen in Fig. 4(e), in which $G_{ee}$ is varied from 0.5 to 1.5 (nominal value: 1). This leads primarily to a steepening of the spectrum at low $f$ due to the onset of instability at $f = 0$ for $G_{ee} > 1$, plus some sharpening of the spectral peaks owing to expansion of the $q^2$ loops.

Figure 4(f) shows the effect of varying the direct corticothalamic feedback parameter $\psi$ from $0.5$ to $2.5$ (nominal value: 1). Larger $|\psi|$ increases the magnitudes of the spectral peaks, and the low-$f$ component of the spectrum as $f = 0$ instability is approached, while negative $\psi$ results in an inversion of the peaks and troughs, with new peaks appearing midway between the original ones, which are replaced by troughs; the new peaks are still weak for $\psi = -0.5$. Note that there is a 1/f low-frequency component for $\psi = 0.25$, whereas there is a plateau for $-0.5$ because $S > 0$ in this case. Variation of the final gain-related parameter $\psi'$ from $-0.6$ to 0.6 (nominal value: 0) is explored in Fig. 4(g). As $|\psi'|$ increases the peaks sharpen and become more pronounced. As $\psi'$ increases there is an increase in the frequencies of the alpha and beta peaks, with a corresponding reduction in their frequency ratio from just below 3 at large negative $\psi'$ to just above 2 at large positive $\psi'$, a phenomenon that will be further examined in Sec. V. Note that the second $q^2$ loop is larger than the first in this case, because the magnitude of the differential feedback term peaks for nonzero $\omega$. Finally, variations in $\psi'$ produce shifts in the normalization of the $f^{-1}$ range, as expected from (33). Although (34) and (36) imply that an $f^{-3}$ regime should be attainable for suitable $\psi'$, an instability at nonzero $f$ inter-
FIG. 4. Dependence of spectra $P(f)/P_N$ and $q^2 r^2_{r_e}$ on individual parameters, with other parameters fixed at the values used in Fig. 3. Solid, dotted, and dashed curves are used in order of increasing magnitude of the parameter being varied. In each of the rows (a)–(k) the left frame shows $P(f)/P_N$ and the right shows $q^2 r^2_{r_e}$ for $f > 0$. In each case one of the curves is the same as that in the corresponding frame of Fig. 3. (a) $\gamma_e=70,110,150$ s$^{-1}$, (b) $\alpha=40,70,110$ s$^{-1}$, (c) $\beta/\alpha=1,4,7$, (d) $G_0=-0.5,-1.0,-1.5$, (e) $G_{ee}=0.5,1.0,1.5$, (f) $\psi=-0.5,1.2,5$, (g) $\psi'=-0.6,0.6,6$, (h) $t_0=0.03,0.07,0.11$ s, (i) $n=0,1,2$, (j) $\eta_1,\eta_2=50,150,250$ s$^{-1}$, (k) $k_0=20,30,40$ m$^{-1}$.  

P. A. ROBINSON et al.  
PHYSICAL REVIEW E 63 021903
venes first for the parameters considered when $|\phi'|$ is further increased. We consider this point further in Sec. VI.

The main effects of increasing the time delay $t_0$ in the corticothalamic loop are to reduce the alpha and beta frequencies, and sharpen the peaks, as shown in Fig. 4(b), the latter effect due to there being less thalamic filtering at low frequencies. Similar effects are seen with changes in $n$, $\eta_1$, and $\eta_2$, as seen in Figs 4(i) and 4(j). There are also weak effects on the normalization of the $f^{-1}$ regime, as expected from (33). These four parameters leave the high-$f$ part of the spectrum essentially unchanged.

Figure 4(k) shows the effect of varying the characteristic wave number $k_0$ above which volume conduction filters out cortical signals. There is little effect on the shape of the spectrum below about 25 Hz, but a substantial steepening above this point for small $k_0$, in agreement with the semi-quantitative estimate (26).

In summary, Fig. 4 shows that for fixed $P_N$ only $\alpha$, $\beta$, $\gamma_c$, and $k_0$ significantly affect the high frequency spectrum, in accord with (26) and (37). The lowest frequencies are only affected by $G_{ii}$, $G_{ee}$, and $\psi$, via their effects on $S$, which occurs in (30). The $f^{-1}$ spectrum is affected moderately by all parameters, with its existence depending most strongly on the value of $S$, which replaces it with a low-$f$ plateau if it is significantly greater than zero. Large values of $\alpha$, $\beta$, $G_{ii}$, $|\phi|$, $|\phi'|$, $\eta_1$, and $\eta_2$ tend to sharpen the peaks, while $t_0$, $\eta$, $n$, $\phi$, and $\psi$ are the chief determinants of the peak frequencies.

Having examined the variation of the predicted spectrum with each of our model’s parameters. We note that when fitting experimental data, a model with fewer parameters would tend to give fits that were simpler to interpret and of higher statistical significance. Published EEG data usually span only the range 1–40 Hz or thereabout, sometimes as great a range as 0.1–100 Hz, or as little as 2–30 Hz. Figure 4 shows that these ranges are unlikely to be sufficient to distinguish all the model parameters unambiguously, particularly once the effects of noise are included. Hence, there is little point in distinguishing all these parameters theoretically. Following this argument, we note that increases in $\beta$ have effects similar to increasing $\alpha$ in the frequency range of interest, so these two parameters can be replaced by a common effective value of $\alpha$ for most purposes, giving $L(\omega) = (1-i\omega/\alpha)^{-2}$ in (4). Similarly, the effects of $n$, $\eta_1$, and $\eta_2$ are quite similar to those of $t_0$ provided $n$ is nonzero (otherwise the differential feedback becomes unphysically large in magnitude at high frequencies). Hence, we can approximate these parameters by setting $n=1$ and $\eta_1 = \eta_2 = \alpha$, which implies $\tau(\omega) = e^{i\omega \alpha}L(\omega)$ in (14). The remaining model has only seven parameters (rather than 11), aside from the overall normalization of the spectrum. In Sec. VI, the likely restriction to $S=0$ reduces this to only six parameters aside from normalization. At the cortical surface, $k_0$ is also irrelevant, leaving one fewer parameter still. There is also the possibility that the dependence of scalp spectra on $k_0$ will not be experimentally distinguishable from the effects due to $\gamma_c$ and $\alpha$.

When fitting to experimental data, we emphasize that the normalization is set by the observed root-mean-square (rms) signal, while the other parameters are not completely free, since physiology constrains them, in some cases quite strongly (see Table I). For example, the characteristic range $r_x$ of excitatory axons has been measured, as have axonal signal velocities $v$. These measurements constrain $\gamma_c$ to within roughly $\pm 40\%$ of its best estimate of 110 s$^{-1}$. Likewise, the dendritic rate constant $\alpha$ is known to be within roughly a factor of two of 100 s$^{-1}$, with a major uncertainty arising from the difficulty of assigning a single rate to describe all the different neural and synaptic types. Similarly, 500 s$^{-1} \approx 2 \beta$ is probably satisfied, and $\beta \approx \alpha$ can be assumed without loss of generality (in the simplified version of our model we set $\beta = \alpha$ and the resulting effective value of $\alpha$ then lies somewhere between 50 and 500 s$^{-1}$).

Although their signs are strictly known, the values of $G_{ee}$ and $G_{ii}$ are less well constrained by physiology, with typical magnitudes estimated to lie between roughly 1 and 100, based on the values in Table I. The ratio $|G_{ii}/G_{ee}|$ is somewhat better constrained, with a value of $1\sim 5$ times the ratio $Q^{(0)}(0)/Q^{(0)}\sim 1$ of the steady-state firing rates in the two neural populations [see (1)]. Since normal EEG and magnetoencephalographic [16] spectra peak strongly at low frequencies, $S$ must be small and non-negative to ensure a significant $f^{-1}$, $f^{-2}$, or $f^{-3}$ spectral range while avoiding instability. Likewise, if $S=0$, $A_1 > 0$ is required whenever $A_2 > 0$ to avoid a low-$f$ instability—a condition that restricts combinations of parameters via (36) and via an analogous expression for the locus on which $A_2=0$.

The magnitude of $\Psi$ must be of order unity to be consistent with physiological knowledge that the thalamus has strong, but not overwhelming, effects on the cortex and its state of attention, and that the cortex likewise projects strongly to the thalamus [37–39]. Thus, although corticothalamic connections (in both directions) are much less numerous than corticocortical ones, these connections must be more effective in stimulating the cortex than their numbers imply—e.g., as a result of larger $s_n$ values than for other fibers, or because they have more axonal synapses than average. Hence, $|\psi|$ and/or $|\psi^\prime|_{\alpha_0}$ are probably of order unity in some states, the latter being the maximal magnitude of the differential feedback term $\alpha t_0 |\phi^\prime| \tau(\omega)$ if $\alpha \sim \beta$, $\eta_1$, $\eta_2$.

The value of the delay $t_0$ (including corticothalamic dendritic delays) must be of order 25–100 ms, since the latency for signals to reach the cortex from the thalamus is $\sim 25$ ms [8] and there is presumably a similar delay in the reverse direction, plus some shorter intrathalamic delays. Although the factors $G_{ee}^2$ and $|\phi_0^\prime|$ in the normalization (22) are not separately observable, they do not constitute free parameters of the spectral shape, and variations of the magnitude of the normalization $G_{ee}^2 |\phi_0^\prime|$ can be determined experimentally. Finally, the volume conduction turnover wave number $k_0$ can be estimated from the electrical conductivities and thicknesses of the various tissues overlying the cortex [40], giving $k_0 \approx 30$ m$^{-1}$.

V. SPECTRAL PEAKS

In this section we turn our attention to the EEG spectral peaks, which are of central importance to practical electro-
encephalography, where they are often termed rhythms. We use the simplified model introduced in the preceding section to avoid an unnecessarily complicated discussion. The effects of spatial filtering are also ignored, since the preceding section implies that these are only significant for frequencies above those of the major rhythms. We estimate the locations of the peaks and the resulting frequency ratios, the sharpness of the peaks, and the conditions for onset of instabilities. In doing this, we unify all the peaks observed into a single theoretical framework in which each peak is distinguished by a unique number and a theoretical basis is obtained for the customary subdivision of the spectrum into bands.

A. Frequencies and occurrence of spectral peaks

Some numerical experimentation shows that the factor in (24) that dominates in determining the frequencies of spectral peaks is the variation of Im $q^2 r_m^2 = |q^2 r_2^2| \sin \theta$. We find that the dominant behavior can be approximated in the limits $\omega^2 \ll \alpha^2$ and $\omega^2 \gg \alpha^2$ by

$$\text{Im} \, q^2 r_m^2 = - \frac{2x}{\gamma' t_0} - B(\omega) G_{ce}[\psi \sin x - x \psi' \cos x], \quad (39)$$

$$B(\omega) \approx 1/G_0, \quad \alpha^2 \ll \omega^2, \quad \omega^2 \ll \alpha^2, \quad (40)$$

$$x = \omega t_0. \quad (42)$$

Since the first term on the right-hand side of (39) is negative (for positive $x$, which is assumed henceforth), peaks of $P(\omega)$ occur approximately where the term in square brackets has its maximum negative value of $-|\Psi|$. The factor $B(\omega)$ in (39) and the factor $\theta$ in the numerator in (24) act to downshift the frequencies from those we will shortly predict using the above argument, with this shift increasing for $\omega > \alpha$ where the fractional rate of change of $B(\omega)$ becomes large. However, the present approach is adequate to predict the main trends semiquantitatively, while numerical results easily yield more accurate peak frequencies when required.

We find that the maximum negative values of the term in square brackets in (39) are attained for

$$\sin x = -\psi/|\Psi|, \quad (43)$$

$$\cos x = \psi' x/|\Psi|. \quad (44)$$

For $\psi' = 0$, these equations yield peaks at

$$x_m = (m - \frac{1}{2}) \pi, \quad (45)$$

with peak index $m = 2, 4, \ldots$ for $\psi > 0$ and $m = 1, 3, 5, \ldots$ for $\psi < 0$. For $|\psi'/x_m| < 1$ we find $x_m$ is given by (45) with the addition of a term $\text{sign}(\psi) \pi/2$. An approximate result, valid for all $\psi'$, is

$$\omega_m t_0 = x_m + \sin^{-1}(\psi' x_m/|\Psi_m|) \text{sign}(\psi), \quad (46)$$

$$|\Psi_m| = (\psi^2 + \psi'^2 x_m^2)^{1/2}. \quad (47)$$

FIG. 5. Peak frequencies predicted from (46) as functions of $\psi'/\psi$. Solid lines represent peaks that occur for $\psi > 0$ (troughs for $\psi < 0$) while dashed lines represent peaks for $\psi < 0$ (troughs for $\psi > 0$). The approximate frequency given by (46) increases monotonically from $(m - 1) \pi$ to $m \pi$ as $\psi'$ increases from $-\infty$ to $\infty$, as was illustrated in Fig. 4(g). Note that there is a discontinuity in $\omega_m$ as a function of $\psi'$ at $\psi' = 0$ in the case $\psi = 0$.

Figure 5 shows the dependence of $\omega_m$ on $\psi'/\psi$ for various $m$, as obtained from (46). The solid and dashed lines correspond to peaks and troughs in the spectrum for $\psi > 0$, and the reverse for $\psi < 0$. On each curve $\omega_m t_0$ passes through a range of $\pi$ as $\psi'$ increases, with most of the change concentrated in the range $|\psi'| \approx 1/(\omega_m t_0)$. The frequency ratio between successive peaks decreases as $\psi'$ increases, which may enable the ratio $\psi'/\psi$ to be determined experimentally. Typical ratios $f_{\beta}/f_{\alpha}$ of the beta to alpha frequencies are very close to 2 (see Fig. 1, for example). This is consistent with the association $m = 2$ for alpha and $m = 4$ for beta, if $2 \pi \psi'/\psi \geq 1$ and $\psi > 0$ in the waking state, giving a frequency ratio just over 2. It is also consistent with $m = 3$ for alpha and $m = 5$ for beta, if $2 \pi \psi'/\psi \leq -1$ and $\psi < 0$, which leads to a frequency ratio just less than 2. Physiologically, positive feedback with $\psi > 0$ is expected in the waking state, also favoring the first scenario [12]. The sleep spindle peak appears roughly midway between the alpha and beta peaks, as illustrated in Fig. 1, qualitatively consistent with the two scenarios just mentioned, for $m = 3$ and $m = 4$, respectively. The largest qualitative difference between the two scenarios is that there is another nonzero-$f$ waking peak below the alpha peak in the case $\psi < 0$; however, in practice, this may not be distinguishable from the $1/f$ peak (present in both waking and sleeping states at marginal stability, and to which we assign the value $m = 0$ for completeness). It seems that the most promising method of distinguishing the two possibilities is detailed comparison of the ratio $f_{\beta}/f_{\alpha}$, as determined from the full spectral formula (22), with experiment—if this ratio definitely exceeds 2, one has the simplest possible correspondence between the peak index of observed and predicted peaks, as listed in Table II. In determining this ratio, one must be careful to correct for downshifts due to the decrease of the smooth underlying spectrum, an effect that becomes stronger at high frequencies.
TABLE II. Association between rhythms at $f < 40 \text{ Hz}$ and the resonances discussed in Sec. V. The first column gives the peak index $m$, the second its designation on the assumption $\psi' / \psi > 1/\pi$, the third its nominal frequency in Hz based on an alpha frequency of 10 Hz, the fourth the corresponding nominal frequency range in Hz of the associated sub-band, and the fifth whether it has a peak in sleeping or waking states. The sixth column gives the peak frequencies in Hz inferred from the data displayed in Figs. 1 and 6 (the nominal ones should be multiplied by 0.95, the ratio of the actual to nominal alpha frequencies, for comparison with these data). The frequency of the delta peak was obtained from Fig. 1. Dashes indicate where no peak was observed in this instance. The conventional bands are delta=0–3.5 Hz, theta=4–7.5 Hz, alpha=8–13 Hz, beta=14–30 Hz, and gamma=above 30 Hz.

<table>
<thead>
<tr>
<th>$m$</th>
<th>Rhythm</th>
<th>$f$</th>
<th>Band</th>
<th>State</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>delta</td>
<td>0</td>
<td>0–2.5</td>
<td>both</td>
<td>$&lt;0.2$</td>
</tr>
<tr>
<td>1</td>
<td>theta</td>
<td>5</td>
<td>2.5–7.5</td>
<td>sleep</td>
<td>$4.6 \pm 0.5$</td>
</tr>
<tr>
<td>2</td>
<td>alpha</td>
<td>10</td>
<td>7.5–12.5</td>
<td>wake</td>
<td>$9.5 \pm 0.5$</td>
</tr>
<tr>
<td>3</td>
<td>spindle (sp.)</td>
<td>15</td>
<td>12.5–17.5</td>
<td>sleep</td>
<td>$14 \pm 0.5$</td>
</tr>
<tr>
<td>4</td>
<td>beta</td>
<td>20</td>
<td>17.5–22.5</td>
<td>wake</td>
<td>$19 \pm 1$</td>
</tr>
<tr>
<td>5</td>
<td>beta-2 sp.</td>
<td>25</td>
<td>22.5–27.5</td>
<td>sleep</td>
<td>$23 \pm 1$</td>
</tr>
<tr>
<td>6</td>
<td>beta-2</td>
<td>30</td>
<td>27.5–32.5</td>
<td>wake</td>
<td>$26 \pm 2$</td>
</tr>
<tr>
<td>7</td>
<td>gamma-1 sp.</td>
<td>35</td>
<td>32.5–37.5</td>
<td>sleep</td>
<td>$33 \pm 2$</td>
</tr>
<tr>
<td>8</td>
<td>gamma-1</td>
<td>40</td>
<td>37.5–42.5</td>
<td>wake</td>
<td></td>
</tr>
</tbody>
</table>

In either scenario for the ratio $\psi' / \psi$ discussed in the previous paragraphs, the ratio $\psi' / \psi$ has the same sign in both sleeping and waking states, even when $\psi$ changes sign (i.e., $\psi$ and $\psi'$ appear to change sign at roughly the same point). The simplest feedback architecture consistent with this is one in which any thalamic signal inversion occurs at a common structure through which both the direct and differential feedback signals pass. A very likely candidate for such a structure is the thalamic reticular nucleus because there is wide agreement that, in sleep, connections from this nucleus strongly inhibit other thalamic nuclei [12,14,43]. The enhanced inhibitory influence of the thalamic reticular nucleus combined with the existence of an otherwise excitatory feedback loop from the cortex through the thalamic reticular nucleus to other thalamic nuclei and thence back to the cortex, is the reason that both $\psi$ and $\psi'$ might change sign in sleep relative to waking states. A third possible scenario is one in which the even-$m$ resonances alone determine the spectral peaks, with a shift in $\psi' / \psi$ from large positive values in waking to large negative ones in sleep, with the consequent frequency downshift allowing each resonance to account for one waking peak and a sleep peak roughly 5 Hz below it. However, this possibility requires an additional argument as to why the odd-$m$ peaks never seem to contribute. A fourth and final possibility, that suffers from the same problems as the third, is that the opposite trend in $\psi' / \psi$ may occur, allowing each even-$m$ resonance to account for a waking peak and a sleep peak roughly 5 Hz higher in frequency. We discuss the distinction between these scenarios further at the end of Sec. VI.

If we assume the association between $m$ and spectral rhythm given in Table II, we find that the alpha ($m=2$) and beta ($m=4$) peaks should be positively correlated with each other, and with a new $m=6$ peak at roughly three times the alpha frequency. We term this the beta-2 peak, a nomenclature that we will justify in the next paragraph. This peak is seen at roughly three times the alpha frequency in Fig. 1, located on the wing of the beta peak, from which it is not well resolved in this example (see the discussion of Fig. 6 below for further analysis). The predicted alpha–beta correlation has been observed in recent work [44]. Similarly, we expect that peaks at $m = 1$ (theta or delta rhythm, whose peak may blend smoothly into the $1/\pi$ peak in many practical situations), $m = 3$ (sleep spindles; which we denote by the symbol $\sigma_{\beta 1}$ since they occur in the lower part of what is commonly termed the beta band), and $m = 5$ will be positively correlated with each other, and negatively correlated with the alpha, beta, and gamma peaks. Since the latter peaks occur in the waking state, they are obviously anticorrelated with sleep spindles, which are only seen during sleep. Evidence for an anticorrelation between theta and alpha power in the waking state has also been presented recently [45], while theta and delta are well known to be most prominent in deep sleep, especially sleep stage 2 [17,18]. We term the $m=5$ rhythm, which has not previously been discussed, beta-2 sleep spindles (symbol $\sigma_{\beta 2}$, since they occupy the upper part of the conventional beta band) in analogy with normal sleep spindles. An enhancement corresponding to this rhythm is seen around 21 Hz in Fig. 1.

The above arguments unify all the rhythms discussed into a single family, with two subfamilies corresponding to waking and sleeping states. This provides a theoretical justification for dividing EEG spectra into bands, but implies that the frequency boundaries of these bands will vary somewhat from individual to individual, with an inverse dependence on $t_0$, weaker dependence on $\psi' / \psi$, and still weaker variation with the other model parameters. Such variations between individuals have been recognized in the literature [18].

The third column of Table II illustrates nominal assignments of frequency band designations based on an alpha frequency of 10 Hz. These correspond quite well with the conventional assignments of the delta, theta, and alpha bands.
The traditional beta band incorporates the spindle, beta, beta-2 spindle, and beta-2 bands of our classification scheme. As conventionally defined, the gamma band extends undifferentiated upward from 30 Hz. In principle, our analysis implies that it can be subdivided into sub-bands of order 5 Hz in width, which we label gamma-1, gamma-1 spindle (\(\sigma_1\)), gamma-2, gamma-2 spindle (\(\sigma_2\)), etc. However, in practice, the relevant peaks are unlikely to be distinguishable from experimental noise in this range (see the next subsection for details), so it is not unreasonable to label the entire range the gamma band, as is conventional. In the above terminology, the theta rhythm could be labeled the alpha spindle and denoted \(\sigma_a\), but the old nomenclature is too well established to be likely to make this renaming acceptable.

Figure 6(a) shows the ratios of the pairs of spectra in Figs. 1(a)–1(e). These ratios are remarkably similar at different electrodes, despite the relatively large differences between the actual spectra. In Fig. 6(b) the mean of the five ratios is plotted, bracketed by the interelectrode standard deviation. This ratio oscillates strongly, with an amplitude that decreases rapidly with \(f\). It has peaks at \(f\approx 5, 14, 23, \) and 33 Hz, as listed in Table II, the last of these being relatively indistinct. Troughs are also seen near 9.5, 19, and 26 Hz. We argue that these peaks and troughs correspond to the theta, spindle, beta-2 spindle, gamma-1 spindle, alpha, beta, and beta-2 rhythms, respectively. The peak at \(<0.2\) Hz, seen in Fig. 1 corresponds to delta rhythm. At high frequencies, the ratios seen in Fig. 6 approach a nearly constant plateau, as expected from (37).

Use of ratios between waking and sleeping spectra removes skull-thickness differences between electrodes, as well as the systematic frequency shifts associated with the shape of the underlying smooth spectrum, and also highlights the differences between the waking and sleeping spectra. The ratios of the frequencies obtained from Figs. 1 and 6 to the alpha frequency are close to integers and half-integers, as predicted from our theory. Figure 6 thus provides evidence of the detection of the \(m=5\) (beta-2 spindle) rhythm, with weaker evidence for the \(m=6\) (beta-2), and \(m=7\) (gamma-1 spindle) rhythms. In future work, these tentative detections will be further tested using EEGs from multiple subjects with longer acquisition times.

### B. Sharpness of spectral peaks

The prominence of spectral peaks can be quantified in terms of their \(Q\) values (which should not be confused with the firing rates \(Q_m\) and \(Q_1\) in Sec. II), which are large for strong peaks. We define an effective \(Q_m\) to be the ratio of the height of the \(m\)th peak to the trough immediately above it in frequency. This gives

\[
Q_m \approx \frac{|\text{Im} g_m^2|}{|\text{Im} g_m^c| r_m^c} \quad (48)
\]

\[
Q_m \approx \frac{|\text{Im}[2x_m + G_c B(\omega_m)]\gamma_0 t_0 \Psi_m e^{i(\omega_m t_0)}|}{|\text{Im}[2x_m + G_c B(\omega_m)]\gamma_0 t_0 \Psi_m e^{i(\omega_m t_0)}|} \quad (49)
\]

Since \(\omega_m\) and \(\omega_{m+1}\) correspond to a peak and its neighboring trough, the imaginary parts of \(\Psi_m e^{i(\omega_m t_0)}\) and \(\Psi_{m+1} e^{i(\omega_{m+1} t_0)}\) are approximately equal and opposite, with the denominator in (49) being small for a strong peak. Hence, if we approximate \(x_m\) by \(x_m\) in the first term in the numerator, and \(B(\omega_{m+1})\) by \(B(\omega_m)\), we find

\[
Q_m \approx \frac{4\omega_m}{G_c B(\omega_m) \gamma_0 |\text{Im}[\Psi_m - \Psi_{res}] e^{i\omega_m t_0}|}, \quad (50)
\]

where the exactly resonant value \(\Psi_{res}\) is complex and satisfies

\[
\text{Im}[\Psi_{res} e^{i\omega_m t_0}] = 2 \omega_m / \gamma_0 G_c B(\omega_m). \quad (51)
\]

At large and small \(\omega\) (with some deviations in between), Eq. (50) implies that the \(m\)th resonance sharpens as \(\Psi \rightarrow \Psi_m\) and that \(|\Psi|\) must certainly exceed the modulus of the right-hand side of (51) before resonance can occur. The behavior of \(B(\omega)\) given by (40) and (41) implies that high-\(Q\) resonances are only possible for \(\omega_m \approx \alpha\) and, hence, for low \(m\). The rapid fall-off in \(Q\) for high \(m\) is seen in the theoretical spectra in Figs. 3 and 4 and in the experimental ones in Figs. 1 and 6. Experimentally, it seems that only the \(m=1–4\) resonances can have large \(Q\), consistent with Fig. 6. The rapid decrease of \(Q_m\) with \(m\) underpins our comments in the preceding section that only the peaks with small \(m\) are likely to be discernible in experimental data. It is worth noting that, as was seen in Fig. 6, the \(m=1\) solution does not always produce a strong discrete peak; in some cases it blends with the \(1/f\) spectrum with no intervening minimum. This behavior is consistent with the preceding discussion, because the latter explicitly ignored the factor \(\theta\) in the numerator in (24), which cancels the \(\sin \theta\) term in the denominator at small \(\theta\). This behavior can weaken the theta peak relative to what was implied above.

### C. Instabilities

The condition \(S=0\) is necessary, but not sufficient, for cortical stability, since it guarantees the absence of instability only at \(f=0\), and then only so long as \(A_1\) does not change sign. One possibility for instability at \(S=0\) occurs if \(A_1\) changes sign, in which case the spectrum passes through a \(1/f^3\) regime, given by (34), followed by an instability that first sets in at \(f=0\) then extends to include higher frequencies. Alternatively, it is possible for instability to set in at nonzero frequency if a \(q^2\) loop cuts the negative real axis. Instability at \(\omega_m\) then sets in approximately when \(|\Psi| = |\Psi_m|\), where the spectrum can first develop a singularity. Since \(|\Psi_m|\) increases monotonically with \(m\), the first such instability usually sets a stability boundary for the cortical dynamics as a whole. Assuming the simpler association between \(m\) and peaks, this boundary is thus set by the \(m=2\) (alpha) resonance in the waking state, and the \(m=1\) (theta) resonance in sleep (or the \(m=3\) resonance if the theta resonance is weakened, as was discussed above).

If an instability sets in, the system will rapidly move to a nonlinear regime in which the linear analysis is invalid.
Hence, in the presence of instabilities, spectra computed from equations in the preceding sections will not be correct. If the ratio $\psi' / \psi$ is specified, as will be the case in Sec. VI, we can estimate the stability boundaries implied by the above discussion. When $A_1 = 0$, (36) implies that one must have

$$t' = t_0 (1 - \psi' / \psi),$$

which implies $\psi' > 1$ unless $\psi' / \psi > 1 - 2 / a t_0$. The relationship $S = 0$ then implies a boundary in the $G_{ee} - G_{ii}$ plane, with the physically allowable region satisfying

$$G_{ii} < 2 \gamma_e \alpha + t' - G_{ee} \left( \frac{2}{\gamma_e + \alpha} + t' \right).$$

(54)

This condition is a straight line with slope above $-1$ unless the denominator in (54) is negative, which occurs for

$$\frac{\psi'}{\psi} < 1 + \frac{1}{t_0 \left( \gamma_e + \frac{1}{\alpha} \right)}.$$ (55)

In the latter case, the condition

$$G_{ii} > \frac{2 \gamma_e}{\alpha} + t' - G_{ee} \left( \frac{2}{\gamma_e + \alpha} + t' \right),$$

(56)

applies and the slope of the boundary line is more negative than $-1$. At large $\psi' / \psi$, the slope of the boundary is $-1 + 2 / (\gamma_e t_0 \psi' / \psi)$.

If $|\psi' / \psi|$ is large, one has $|\Psi| = |\psi o t_0|$ and $\psi$ is forced to be small once $o t_0$ exceeds unity, which is the case at all spectral peaks with $m$ $\geq$ $1$, except at $m$ $= 1$ if one also has $\psi' / \psi < 1$. The approximate condition for the absence of nonzero-$f$ instabilities of the $m = 1$ and 2 rhythms (theta and alpha), $|\Psi| < |\Psi_m|$, then yields

$$1 - G_{ee} \left[ 1 - \frac{2}{\gamma_e t_0 |\psi' / \psi|} \right] \geq G_{ii} \geq 1 - G_{ee} \left[ 1 + \frac{2}{\gamma_e t_0 |\psi' / \psi|} \right].$$ (57)

Equation (57) defines lines with negative slopes either side of unity in magnitude. Closer analysis of the instability condition $q_c^2 = 0$ implies that the left inequality is imposed by the $m = 1$ (theta) instability (or $m = 3$ if theta is weakened, a proviso that will be assumed implicitly henceforth), and the right by the $m = 2$ (alpha) instability.

If $|\psi' / \psi|$ is sufficiently small, none of the resonances can become unstable. This can be seen by noting that, for $\psi' = 0$, Eq. (17) implies that the locus of $q_c^2$ is composed of a secular part and an oscillatory part that results from CT feedback and gives rise to loops superposed on the secular trend. These loops decrease in radius as $\tau(\omega)$ in (12) decreases in magnitude with frequency, while at zero frequency and marginal stability the radius of the loop is just sufficient for the locus to start at the origin. The only way an instability can occur is for a subsequent loop to reach the origin or the negative real axis, a possibility that is most favored if $|\tau(\omega)|$ stays as large as possible—i.e., for $\alpha, \beta, \gamma_1$, and $\gamma_2$ as large as possible. If we make these parameters infinite (the optimal case), we find that the locus $q_c^2$ of the loop centers satisfies

$$|q_c^2|^2 = \frac{\psi^2}{(1 + \psi)^2} + \frac{2\omega^2}{\gamma_e} + \frac{\omega^4}{\gamma_e^2},$$ (58)

which is monotonic increasing with $\omega$ unless $2 < \psi < -1$, a range that is forbidden by (54), which requires $\psi' > 1$ for $\psi' / \psi$ small. Hence, since the loop center moves monotonically away from the origin, and the loop radius is nonincreasing, $q_c^2 = 0$ cannot be attained and no instability starting at the origin is possible in this case. Once $q_c^2$ moves into the left half of the complex plane, the origin is no longer the closest point on the nonpositive real axis and the possibility of an instability arising due to intersection of $q_c^2$ with some other point on this axis arises. However, the imaginary part of $q_c^2$ increases monotonically in magnitude in this regime from a value which is already too large for $q_c^2 = 0$ (and, hence, $\text{Im} q_c^2 = 0$) to be attained at the point where $q_c^2$ first crosses into the left half plane. Hence, combining these arguments, no instability is possible for $\psi' = 0$.

From the above arguments, we conclude that the physically allowable region of parameter space is bounded below by the higher of the $m = 2$ boundary obtained from (57) and the $A_1 = 0$ boundary (56), if the latter applies in place of (54). It is bounded above by the lower of the $m = 1$ boundary from (57) and the $A_1 = 0$ boundary (54), unless (56) applies as a lower bound and (54) is irrelevant. For small positive $\psi' / \psi$ the $A_1 = 0$ condition sets the upper bound, but this is supplanted by the $m = 1$ boundary at larger values, particularly once (56) applies and the $A_1 = 0$ boundary becomes a lower bound. There is no lower boundary for very small $\psi' / \psi$ because the $A_1 = 0$ boundary is an upper bound and no resonances can become unstable. At larger $\psi' / \psi$ the $m = 2$ lower boundary cuts in, followed by competition between $A_1 = 0$ and $m = 2$ lower bounds, which both have the same slope at large $\psi' / \psi$. At small negative $\psi' / \psi$ the $A_1 = 0$ upper bound is again the only relevant one, with the $m = 2$ lower bound appearing at larger negative values. At large negative $\psi' / \psi$ the $m = 1$ and $A_1 = 0$ boundaries compete closely to set the upper bound, since they have the same slopes as $\psi' / \psi \rightarrow -\infty$.

We thus find that there are three possible routes to instability from the marginally stable state, apart from $S$ becoming negative: (a) instability starting at $f = 0$ due to $A_1$ changing sign, (b) $m = 1$ theta instability at $f = 5/2 \eta_0 \approx 5$ Hz when $\psi$ is negative and $\psi' / \psi$ is positive, or lower frequencies if $\psi' / \psi$ is negative, and (c) $m = 2$ alpha instability at $f = 1/\eta_0 \approx 10$ Hz when $\psi$ and $\psi' / \psi$ are positive (or at around 5 Hz if $\psi$ is positive and $\psi' / \psi$ is negative). Frequencies in the alpha range is suppressed by the $m = 2$ boundary.

021903-13
VI. MARGINALLY STABLE EEG SPECTRA

In this section we explore the dependence of spectral structure and instabilities on the gain parameters of our model and impose some restrictions on the possible locus of an arousal sequence that carries the cortex through all the states of arousal from sleep stage 4 to high-vigilance waking states (e.g., while doing complex mental arithmetic under time pressure).

Observed normal EEG and MEG spectra do not appear to exhibit any low-f plateau, except perhaps below about 0.3 Hz (see Fig. 1 and [16]). Equations (30) and (33), and the requirement to avoid \( f = 0 \) instability, then restrict \( S \) to very small, non-negative values. Hence, in this section we assume \( S = 0 \) in addition to the other simplifying assumptions made in Secs. IV and V (\( \alpha = \beta = \eta_1 = \eta_2 \) and \( n = 1 \)). If we further fix the values of \( \gamma_f = 70 \text{ s}^{-1} \), \( \gamma_\alpha = 110 \text{ s}^{-1} \), and \( t_0 = 0.07 \text{ s} \), we are left with only \( k_0 \) and the gain parameters \( G_{ee}, G_{xe}, \psi \), and \( \psi' \). The characteristic wave number \( k_0 \) can be ignored here since it does not affect the stability of the cortex. Of the remaining four parameters, \( \psi \) can be eliminated via (28) for \( S = 0 \). Physiologically, one requires \( G_{ee} > 0 \) and \( G_{ii} < 0 \), which restricts attention to a quarter of the \( G_{ee} - G_{ii} \) plane for fixed \( \psi'/\psi \). We choose the latter ratio as the third free parameter, rather than \( \psi' \) itself, since it measures the relative strengths of direct and differential feedbacks.

Figure 7 shows the regions in the \( G_{ee} - G_{ii} \) plane in which the model spectrum (21) is marginally stable (white zones) or unstable at some frequencies (gray zones) for \( \psi'/\psi = 0, \pm 1, 2 \); only the white regions represent physically allowable steady states. The locus of \( \psi = 0 \) is shown dotted in each frame and solid contours of the lowest unstable frequency are overplotted in the gray zones.

In Fig. 7(a), \( \psi'/\psi = 0 \) and we see that there is a single zone of instability at upper right, bounded by the \( A_1 = 0 \) locus (54). As expected, the instability onset frequency near this boundary is zero. The physically allowable region has no lower bound, consistent with the arguments in Sec. V that no resonant instabilities are possible in this case. Figure 8 displays a sequence of spectra and \( q^2 \) loci as the upper instability boundary is crossed, showing the development of a widening \( f^{-3} \) regime below about 5 Hz, followed by onset of instability at \( f = 0 \) initially, then moving to higher \( f \), with \( f \approx 0.8 \text{ Hz} \) in column (c). In the unstable regime the corresponding \( q^2 \) locus loops first up from the origin, then down across the negative real axis very close to the origin. At the boundary, this crossing occurs at the origin.

In Fig. 7(b), \( \psi'/\psi = 1 \) and the stable zone is now bounded above and below by the \( m = 1 \) and \( m = 2 \) instability boundaries, respectively, from (57). The crossover between the

\[ 021903-14 \]
Hence, unless low-frequency arousal sequences. Specifically, the appearance of a sign close to the locus of parameters, leaving it with the wrong sign for the discussion following [10], one would expect an arousal sequence that approximately followed a line passing through the origin in the $G_{ee}-G_{ii}$ plane. The association with $\psi>0$ in the waking state provides a more natural fit with this expectation, and also accords with the physiological expectation of positive feedback in the waking state. The results presented in Fig. 7 thus support the correspondence between $m$ values and spectral peaks made in Sec. V A.

VII. SUMMARY AND DISCUSSION

In this work we have investigated the spectral properties of a generalized version of our recent model of cortical electrical activity, including a physiologically plausible corticothalamic loop with both direct and differential feedbacks, and exploring the effects of volume conduction in tissues overlying the cortex. Differential feedback is included to allow for the possibility that temporal variations are emphasized—likely, since almost all sensory inputs to the cortex first pass through the thalamus, which plays a key role in determining the attentional state of the brain. Unlike many previous models, our model does not presuppose the existence of thalamic pacemakers or "clocks" to regulate the periods of the alpha and beta rhythms. Instead, these emerge directly from the delays in the feedback loop.

One of the main results of this work is the formula (21) for the EEG spectrum that arises from white-noise subcortical inputs, in accord with previous work [7,41], allowing for volume conduction effects. Volume conduction is found to affect the spectrum significantly only above about 25 Hz.
and contributes at most an extra factor of $f^{-3}$ relative to the unfiltered spectrum.

EEG spectra are comprised of a series of peaks superimposed on smooth underlying continuum. Equation (21) accounts for the major features of the underlying continuum observed in EEG and MEG experiments, with a low-$f$ 1/$f$ or $f^{-3}$ component occurring near marginal stability, and a steepening of the spectrum to $f^{-5}$ ($f^{-8}$ with filtering via volume conduction above about 25 Hz) at frequencies above $\alpha$, the characteristic dendritic rate constant, where dendritic filtering becomes strong. The observed knee in the spectrum occurs at around 20 Hz, consistent with $\alpha \approx 120$ s$^{-1}$, which is in turn consistent with independent physiological estimates listed in Table I.

Exploration of the parameter sensitivities of the model was carried out in Sec. IV, resulting in the development of a simplified model with fewer parameters than the initial one. This simplified version was used to analyze the frequencies, sharpness, stability, and mutual correlations of spectral peaks predicted by the model. The main results were that each peak can be denoted by a unique “harmonic” number $m$, with the zeroth peak corresponding to the 1/$f$ or 1/$f^3$ spectrum, $m = 1$ to theta, $m = 2$ to alpha, $m = 3$ to sleep spindles, and $m = 4$ to beta. The $m = 2, 4, \ldots$ peaks are predicted to be positively correlated with one another and prominent in waking. Likewise, the $m = 1, 3, 5, \ldots$ peaks should be positively correlated and prominent in sleep, whereas these two sets are anticorrelated overall and the $m = 0$ peak is always present, but stronger in sleep. Comparison with published observations confirms the predicted correlations, anticorrelations, and states in which certain peaks are expected to be most prominent. The theory is also consistent with the data shown in Figs. 1 and 6, whence we tentatively identify new peaks corresponding to $m = 5$ (beta-2 spindle), $m = 6$ (beta-2), and $m = 7$ (gamma-1 spindle) rhythms. The inferred corticothalamic feedback delay $t_0 \approx 1/\gamma_\text{ref} \approx 100$ ms is also consistent with physiological constraints on this quantity. We conclude that the simplest association between state of arousal and our parameters is for $\phi$ and $\phi'$ to be positive in waking and negative in sleep, although we cannot yet categorically rule out certain alternatives discussed in Secs. V A and VI.

Our analysis of the spectral peaks leads to the unification of all the major large-scale brain rhythms, and the prediction and tentative confirmation of two new ones. It also provides a theoretical basis for the customary subdivision of the spectrum into bands, but implies that the frequencies of these bands are not universal, but are tied to the alpha frequency of the individual subject, contrary to some conventions, but consistent with the recognition of such variations in the literature [18].

Analysis of the sharpness and stability of the spectral peaks showed that only the first few can have high $Q$, and that stability of the $m = 1$ (or $m = 3$) and $m = 2$ rhythms can impose bounds on the physically allowable states of the cortex, corresponding to instabilities near 5 Hz (or 15 Hz) and 10 Hz, respectively. Together with the $A_1 > 0$ condition required to avoid instability at $f = 0$, these conditions restrict the range of physically allowable states in parameter space. We speculate that these instabilities may be associated with the linear stages of certain generalized seizures, possibly including grand mal, petit mal, and photic epilepsies. Significantly, an $f^{-3}$ regime is predicted at low $f$ close to the $A_1 = 0$ boundary in sleep. This is consistent with the observed steepening and enhancement of the low-$f$ spectrum in sleep.

The above results indicate that our model provides a promising semi-quantitative theoretical explanation for EEG spectra. Most significantly, we explain the entire spectrum in a unified way, a major advance on previous analyses, which have generally concentrated on a single frequency band at a time. In a forthcoming paper we will compare it in detail with observations of multiple subjects in various states of arousal to determine whether it is quantitatively consistent and, in particular, whether the parameters inferred are consistent with those inferred from independent physiological measurements. We hope that these physiological links will prompt experimental physiologists to determine tighter constraints on some of the less well known parameter values.

Our model differs from some in the literature in that, rather than adding numerous parameters to describe ever finer details, it restricts attention to a set of physiologically realistic mechanisms that appear to be sufficient to capture the overall behavior of the EEG, and parametrizes these as simply as possible. This keeps the number of parameters low and their physical significance clear.

One limitation of the present work is that it does not attempt to describe spatial variations of the spectrum, although these are significant in practice (albeit less so when spectral ratios, such as those in Fig. 6 are considered). This is not a
fundamental limitation, as the model can be easily generalized to include such variations if next-order accuracy is required, but at the price of introducing additional parameters.

A second limitation is that, although the cortex appears to operate near marginal stability, we have not as yet identified a mechanism to keep it at this point. This does not affect the present discussion, but is an important point of principle. Third, although we argue that corticothalamic feedback is the most important one in determining the observed spectral rhythms, we do not exclude the possibility that other feedback loops with similar time delays may also contribute.

ACKNOWLEDGMENTS

The authors thank K. E. Crowley and I. Colrain for scoring the sleep spectra in Fig. 1. This work was supported by the Australian Research Council, the National Health and Medical Research Council, and the Ross Trust, Melbourne.

APPENDIX A: BRIEF DESCRIPTION OF THE EXPERIMENT USED TO OBTAIN FIG. 1

In this appendix we briefly outline the experiment used to obtain the data used in Fig. 1, justifying the accuracy of the overall form of the spectrum and the reality of the spectral peaks seen. A fuller description is found in Ref. [6] and more details will be published in a forthcoming work in which a large set of EEG spectra from multiple subjects will be analyzed in detail.

The data in Fig. 1 were obtained from the Cz, Pz, Fz, C3, and C4 electrodes in the International 10–20 system, with linked ears providing the reference potential. The Cz electrode is located at the crown (or vertex) of the head, furthest from the reference electrodes and from muscle groups that could generate electrical interference. The Fz and Pz electrodes are located roughly 6 cm in front of and behind this electrode, respectively, while the C3 and C4 electrodes are situated symmetrically ~6 cm either side of it. The data were low-pass filtered to remove frequencies above 50 Hz and sampled at 250 Hz for periods of 2 minutes in waking and 13 minutes in sleep. The spectra were then calculated by Fourier transforming 8 second segments of these data and averaging these transforms.

Narrow band artifacts at the mains frequency of 50 Hz and its subharmonic of 25 Hz were identified due to their presence even in the absence of a subject. As each affected only a single spectral channel, they were simply removed by deleting these channels before plotting the spectra. Ocular artifacts due to eye blinks and pendular eye movements were removed by standard algorithms [46]. When present, muscular artifacts due to jaw, neck, and other movements, were found to contaminate the high-frequency spectrum with noise having an approximately 1/f spectrum which is easily identified. Trials with a subset of subjects showed that, with appropriate instructions to minimize movement and relax jaw muscles, this class of artifact could be suppressed to the point that the high-frequency spectrum was not significantly affected. Cardiac interference at low frequencies was also found to be minimal, with no spectral peak detected at the heartbeat frequency of around 1 Hz, or its harmonics.

The waking spectra in Fig. 1 were for a normal female subject, aged 30, in a relaxed, eyes-closed state, while the sleep spectrum was for the same subject in sleep stage 2.

APPENDIX B: GENERAL EFFECTS OF FILTERING BY OVERLAYING STRUCTURES

In this appendix we show that the high frequency effects of filtering are nearly the same for any physically reasonable filter function.

If we assume that $F(k)$ only depends on the magnitude of the wave vector, and falls off rapidly for $k > k_0$, the integral in (20) can be written in the form

$$I = \pi \int_0^\infty dk^2 \frac{F(k^2/k_0^2)}{|k^2+q^2|^2}.$$  (B1)

Upon changing variables to $u = k^2/k_0^2$, we find

$$I = \frac{\pi}{2i \text{Im} q^2} \int_0^\infty du F(u) \left[ \frac{1}{u+q^2/k_0^2} - \frac{1}{u+q^2/k_0^2} \right]$$

$$\sim \frac{\pi}{\text{Im} q^2} \frac{k_0^2 \text{Im} q^2}{|q^2|},$$  (B2)

where (B2) has been expanded in powers of $k_0^2/q^2 u$ and we have used the fact that $F$ falls off rapidly for $k > k_0$. This result is of the same form as the more specialized one, Eq. (26), and shows that the asymptotic $\omega^{-3}$ filtering does not depend strongly on the structure of the filter function.

APPENDIX C: EXPLICIT EXPRESSIONS FOR $A_\alpha$ AND $A_1$

In this appendix we give explicit expressions for the coefficients $A_\alpha$ and $A_1$ in the expansion (32) for $\alpha = \beta$. Expressions for higher $A_j$ are complicated and their exact analytic form is likely to be more strongly model dependent, so we do not reproduce them here.

Direct expansion of (17) yields

$$A_0 = S, \quad (C1)$$

$$A_1 = \frac{2}{\gamma} + \frac{G_{ee}}{G_0} \left[ 2(1 + \psi) + \left( t_0 + \frac{n}{\eta_1} + \frac{n}{\eta_2} \right) \psi - t_0 \psi' \right], \quad (C2)$$

with $S$ given by (28) and $G_0$ given by (31).

[34] D.T.J. Liley and J.J. Wright, Network **5**, 175 (1994).
Abstract. Zero-lag synchronisation arises between points on the cerebral cortex receiving concurrent independent inputs; an observation generally ascribed to nonlinear mechanisms. Using simulations of cerebral cortex and Principal Component Analysis (PCA) we show patterns of zero-lag synchronisation (associated with empirically realistic spectral content) can arise from both linear and nonlinear mechanisms.

For low levels of activation, we show the synchronous field is described by the eigenmodes of the resultant damped wave activity. The first and second spatial eigenmodes (which capture most of the signal variance) arise from the even and odd components of the independent input signals. The pattern of zero-lag synchronisation can be accounted for by the relative dominance of the first mode over the second, in the near-field of the inputs. The simulated cortical surface can act as a few millisecond response coincidence detector for concurrent, but uncorrelated, inputs.

As cortical activation levels are increased, local damped oscillations in the gamma band undergo a transition to highly nonlinear undamped activity with 40 Hz dominant frequency. This is associated with “locking” between active sites and spatially segregated phase patterns.

The damped wave synchronisation and the locked nonlinear oscillations may combine to permit fast representation of multiple patterns of activity within the same field of neurons.

1. Introduction and background

Synchronisation has been suggested as a solution to the binding and segregation problems of psychology [4, 5, 13–15, 23–25, 29, 31, 40, 62]. Object features might be bound by synchronous neuronal firing and coded for by spatially separated cells or cell assemblies in the cortex [59]. The phase differences between sets of these distributed synchronous firing cells might code multiple objects in the system [61].
Empirical verification of gamma band stimulus–induced synchronisation in the brain has been found in single unit and multiunit local field potential (LFP) and electroencephalogram (EEG) recordings [4,13,15,23–25].

There is no universally accepted mechanism to account for all the experimental evidence for synchronous activity in the brain. Doubts have also been raised as to whether such synchronisation is relevant to binding or segregation [43]. Such doubts include uncertainties about the time–window for synchronisation, the degree to which synchronisation can shift time differences between neural events and, particularly, the time it takes to establish synchronisation in ongoing activity [8]. This paper attempts to account for synchronous behaviour in simple physical terms.

Distinction is made between zero–lag synchrony and synchronous oscillation. Zero–lag synchrony (synchrony for short) is defined as high positive linear cross–correlation/covariance at zero phase–lag between separate sites in cortex. Synchrony is generally associated with damped sinusoidal or nonsinusoidal cross–correlation and auto–correlation functions [13,15,16,23] but constant and slight or even negative damping profiles have been found [16,34].

Synchronous oscillation, on the other hand, is synchronisation associated with gamma band EEG oscillations and found predominantly at 40 Hz. It has often been assumed that all synchronisation in the brain is a highly nonlinear phenomenon [1] since neuron firing is nonlinear at crucial stages of signal transmission [27]. Hence, large arrays of nonlinear neuron–like oscillators have been used to model cortical tissue and the observed nonlinear phase–locking in such models has been seen as the inevitable mechanism for global synchronisation in the cortex [6,7,10–12,14,18,32,35,36,52,56–58] There is, however, conflicting evidence [61] such as observations of broad band synchronisation [4] and synchronous activity seen in continuum models of electrocortical behaviour which do not depend on essential nonlinearity [48,67,68].

The models of electrocortical activity used in this paper subsume individual properties into a collective mass action. This continuum framework produces waves almost akin to linear superposition waves which are associated with synchrony [48,67] as well as synchronous oscillation. A full discussion of these models and experimental findings can be found in a recent review [66]. The term continuum as it is used here specifically refers to the act of lumping together the activity of a group of neurons and treating them as an entity rather than referring to mathematical continuity. The numerical model utilised here is discrete rather than continuous. Properties and formulation of a partial differential version of the present discrete, and integral, model have been explored elsewhere [46,48].

In this paper we use cross–correlation methods to demonstrate synchrony in a simplified averaging model and cross–correlation and PCA methods to show synchrony, synchronous oscillation and nonlinear phase–locking within a realistic physiological model. We show the synchrony mechanism is a form of coincidence detection, or selective filtering of input signals, with very rapid onset and this essentially linear mechanism gives way to non–linear phase–locked synchronous oscillation, within the 40 Hz band in the realistic physiological model.
2. Analysis tools

The cross-correlation coefficient $\rho$ between two LFP time series, is:

$$\rho(i, j) = \frac{c(i, j)}{\sqrt{c(i, i)c(j, j)}}$$  \hspace{1cm} (1)

where $c(i, j)$ is the covariance between the $i$th and $j$th time series. If, as here, all elements are measured in the same units then covariance alone can be used in the PCA. The method involves calculation of the eigenvectors $\mathbf{e}_i$ and eigenvalues $\lambda_i$ of the covariance $C$ matrix of the $m$ state–variable system. The system matrix, $\mathbf{Z}$ for $n$ time steps is:

$$\mathbf{Z} = \begin{bmatrix}
  z(1, 1) & z(1, 2) & \ldots & z(1, m) \\
  z(2, 1) & z(2, 2) & \ldots & z(2, m) \\
  \vdots & \vdots & & \vdots \\
  z(n, 1) & z(n, 2) & \ldots & z(n, m)
\end{bmatrix}$$  \hspace{1cm} (2)

where $z(i, j)$ is the system value with temporal mean removed at spatial position $j$ at time $i$. The covariance:

$$\mathbf{C} = \frac{1}{(n-1)}\mathbf{Z}^T\mathbf{Z}$$  \hspace{1cm} (3)

satisfies the eigenvalue equation:

$$\mathbf{C}\mathbf{e}_i = \lambda_i \mathbf{e}_i$$  \hspace{1cm} (4)

with the temporal principal component vector for each eigenvector given by:

$$\mathbf{a}_j = \mathbf{Z}\mathbf{e}_j$$  \hspace{1cm} (5)

3. Models

Two simulation models are used; a simplified dendritic averaging model and a physiological model [65]. We begin with the simple model.

3.1. Simplified averaging model

The averaging model consists of positive–feedback linked linear elements and was used to specify minimum requirements for synchronisation. Connectivity strength between elements was weighted by a Gaussian function of distance. For a distance $r_{pq}$ between elements $p$ and $q$ the weighting was:

$$\omega_{pq} \propto e^{-\frac{1}{2}(\frac{r_{pq}}{\sigma})^2}$$  \hspace{1cm} (6)

A given element’s input was composed of the $V_{out}$ of other elements weighted by this connection strength, with $\sigma$ the standard deviation. The time $\delta_{tpq}$ it took voltage signals $V_{out}(t)$ to travel this distance was dependent on $r_{pq}$ and axonal transmission speed $v$ so:

$$\delta_{tpq} = \frac{r_{pq}}{v}$$  \hspace{1cm} (7)
The voltage at the \( p \)th element at time \( t_0 \) was then:

\[
V^{(p)}(t_0) = \sum_{q=1}^{m} \omega_{pq} V^{(q)}(t_0 - \delta t_{pq})
\]  

(8)

A particular lumped element’s voltage can be modelled as a summation of earlier voltage values due to the delays and fall-off characteristics associated with dendritic potentials. For \( N \) time steps (with length of time step \( \Delta t \)) this dendritic potential summation is:

\[
V^{(p)}(t_0) = \frac{\sum_{i=1}^{N} V^{(p)}(t_0 - i\Delta t)}{2N}
\]  

(9)

3.2. Physiological model

The model used here (reported in detail elsewhere [68]) is an intermediate stage in a family of models [44–48, 66, 67] which progress from the simplest possible descriptions of the cortex as a delay network. By introducing more complex aspects of cerebral dynamics and independently specified parameter values in a step-wise fashion we aim for an increasingly accurate account of cerebral dynamics [66]. The level of development used here is sufficient to reproduce the essential features of synchronous oscillation [65]. The spectral properties observed (notably the capacity for oscillation in the gamma and 40 Hz range) are dependent on rapid feedback processes operating at synaptic level. Justifications for the use of normalised units, and for the particular parameter values here applied, are given in the earlier papers [30, 33, 42, 51, 53, 68]. These values, and the match to experimental data have been improved in subsequent stages of development of the family of models, but no essential change in the class of dynamics here described is brought about by the later modifications.

3.3. State-equations

This model represents the continuum of cortical tissue as discrete cortical zones in the spirit of Wilson and Cowan [63, 64]. Transfer of afferent synaptic impulses to efferent pulses via dendritic processes is modelled by a biexponential lag function matched to physiological measurements [45, 53, 54].

The \( N \) cells in unit volume each have a probability of emission of an action potential \( q_i \) as a function of their membrane potentials. The sum of population membrane potentials is directly proportional to the LFP, \( V(t) \) at time \( t \). Then in a mean-field approximation the pulse-probability density \( Q(t) \) is given by:

\[
Q = \frac{1}{N} \sum_{i=1}^{N} q_i(V)
\]  

(10)

By the central limit theorem, for large \( N \), \( Q \) will have a Gaussian distribution with respect to \( V \), independent of individual distributions of \( q_i \), so \( V \) and \( Q \) are approximately related by:

\[
Q = (1 + e^{a(V-3)})^{-1}
\]  

(11)
Spatial eigenmodes and synchronous oscillation

Where $a = -\pi/\sqrt{3}$, LFP voltage units (vu) are approximate to standard deviations of the distribution of cell pulse probability over the complete range of LFP, with a 50% mean probability of pulse emission 3 standard deviations from complete polarisation of the neural population.

The time response of mean membrane potential (and by implication LFP and soma potential) is given by

$$V(t) = g \sum_{j=1}^{n} w_j Q_a(t - j \Delta t) \quad j = 1, 2, 3...n$$ (12)

where $g$ is synaptic gain, $Q_a$ is afferent pulse action density, $\Delta t$ is the discrete time-step, and $n \Delta t$ is large compared to the peak time response of membrane potential. In accord with [46]

$$w_j = b^2 j \Delta t e^{-bj \Delta t}$$ (13)

models the rise and fall of membrane potential in response to input at $t = 0$; incorporating lags due to both synaptic conduction and average dendritic cable delay in a single function. Parameter $b$ regulates both the peak time and mean delay associated with this lag. Time step $\Delta t$ was set at 0.1 ms, after trials showed progressive decrements of time-step to 0.01 ms produced only small, asymptotically diminishing effects on spectral content of the results.

Within unit volumes both excitatory and inhibitory cell groups are distinguished, each reciprocally and self-coupled, and each coupled at longer range to other unit volumes by cortico-cortical fibres. Delays due to axonal conduction between unit volumes are given by $\Delta \tau = r_{pq} / \nu$, where $\Delta \tau$ is axonal conduction lag over the distance $r_{pq}$ between the $p$th unit volume and the $q$th unit volume and axonal conduction velocity is $\nu$.

Coupling strengths are proportional to:

- The fractional density of synaptic couplings afferent to dendrites of excitatory and inhibitory cells respectively ($\alpha_{ee}$, $\beta_{ei}$, $\mu_{ei}$, $M_{ee}$ etc., as listed in Table 2).
- The synaptic gains of excitatory and inhibitory synapses, $g_e$ and $g_i$.
- Changes in synaptic efficacy, $E^\prime$, representing feedback effects including those of reversal potentials [30]. These feedback relations are modelled as linear regressions of efficacy with membrane potential

$$E^\prime_{ee}(t) = (1 - V_{e(p)}(t - \Delta t)/V_{eR})$$
$$E^\prime_{ei}(t) = (1 - V_{e(p)}(t - \Delta t)/V_{eR})$$
$$E^\prime_{ie}(t) = (1 - V_{i(p)}(t - \Delta t)/V_{iR})$$
$$E^\prime_{ii}(t) = (1 - V_{i(p)}(t - \Delta t)/V_{iR})$$ (14)

Subscripts $e$ and $i$ indicate excitatory and inhibitory potentials; subscript $R$ a constant-valued reversal potential. Smoothed efficacies $\{E^\prime\}$ were applied so $E(t) = \sum_{j=1}^{n} u_j E^\prime(t - j \Delta t)$ where $u_j = ce^{-cj \Delta t}$. For large $c$, decay is rapid; analogous to reversal potentials alone.
State equations for the $p$th unit volume of the cortical system are then:

$$Q_{e}(p) = \left(1 + e^{a(V_{e}(p) - 3)}\right)^{-1}$$

$$Q_{i}(p) = \left(1 + e^{a(V_{i}(p) - 3)}\right)^{-1}$$

$$V_{e}(p) = \sum_{j=1}^{n} w_{j} Q_{ae}(p)(t - j\Delta t)$$

$$V_{i}(p) = \sum_{j=1}^{n} w_{j} Q_{a}(p)(t - j\Delta t)$$

(15)

$Q_{ae}(p), Q_{ai}(p)$ are $p$th unit volume afferent synaptic action densities for excitatory and inhibitory cells which receive local synaptic input (at negligible axonal delay) and delayed cortico-cortical inputs from $q$th unit volumes at range $r_{pq}$, $q = 1...u$. in accord with:

$$Q_{ae}(p) = g_{e} \beta_{ee} E_{ee}(p) Q_{e}(p) - g_{i} \beta_{ie} E_{ie}(p) Q_{i}(p)$$

$$+ g_{e} M_{ee} E_{ee}(p) Q_{s}(p) + g_{e} \mu_{ee} E_{ee}(p) Q_{ns}(p)$$

$$+ g_{e} \sum_{u} \alpha_{ee}(r_{pq}) E_{ee}(p) Q_{e}(q)(t - r_{pq} / \nu)$$

$$Q_{ai}(p) = g_{e} \beta_{ei} E_{ei}(p) Q_{e}(p) - g_{i} \beta_{ii} E_{ii}(p) Q_{i}(p)$$

$$+ g_{e} M_{ei} E_{ei}(p) Q_{s}(p) + g_{e} \mu_{ei} E_{ei}(p) Q_{ns}(p)$$

$$+ g_{e} \sum_{u} \alpha_{ei}(r_{pq}) E_{ei}(p) Q_{e}(q)(t - r_{pq} / \nu)$$

(16)

$\alpha_{ee}(r_{pq})$ and $\alpha_{ei}(r_{pq})$ are partial input synaptic densities. $\alpha_{ee}(r_{pq}) = \alpha_{ee}$ and $\sum_{i} \alpha_{ei}(r_{pq}) = \alpha_{ei}$. $Q_{s}$ and $Q_{ns}$ are system inputs; $Q_{s}$ represents all time-varying components in specific cortical afferents and $Q_{ns}$, acting as a control parameter, is a uniform DC input modelling nonspecific cortical activation. See [44] for analysis on the physics of this class of models.

3.4. Configuration of simulation

In both the simplified averaging and physiological models studied here an extended area of cortex was simulated by unit volumes in a 20 x 20 or 20 x 40 matrix, each volume connected with its neighbors so the coupling strengths, $\alpha_{ee}(x, y)$ declined with $r_{pq}$ as a Gaussian function with standard deviation of 4 distance units. A distance unit was the side of one cell of the given matrix and $m = 400$ or 800 in equation (2). This approximates the distribution of cortico-cortical fibres in cat brain if the distance unit is taken as about 0.9 mm. Boundary conditions were toroidal in all simulations. Absorbing boundary conditions and matrix size changes did not qualitatively affect the reported results.

3.5. Parameter values

Model parameters are given in Tables 1 and 2.

4. Methods and results

Methods and results are presented together, since the latter methods are contingent on the earlier results.

A time step of $\Delta t = 0.1$ milliseconds allowed a 10000 Hz maximum sampling frequency; far more than needed to capture model frequencies – mainly below 100 Hz. Non-decimated data was retained for analysis to facilitate study of temporal precision in the full model.
Table 1. State-variables and standard parameters other than synaptic densities. LFP = local field potential, PPD = pulse probability density.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_e )</td>
<td>Excitatory LFP</td>
<td>vu</td>
</tr>
<tr>
<td>( V_i )</td>
<td>Inhibitory LFP</td>
<td>vu</td>
</tr>
<tr>
<td>( Q_e )</td>
<td>Excitatory PPD</td>
<td>Dimensionless</td>
</tr>
<tr>
<td>( Q_i )</td>
<td>Inhibitory PPD</td>
<td>Dimensionless</td>
</tr>
<tr>
<td>( a )</td>
<td>Slope parameter</td>
<td>(-\pi/\sqrt{3}) (vu(^{-1}))</td>
</tr>
<tr>
<td>( b )</td>
<td>Dendritic time-constant</td>
<td>50 s(^{-1})</td>
</tr>
<tr>
<td>( g_e )</td>
<td>Excitatory gain</td>
<td>65 vu</td>
</tr>
<tr>
<td>( g_i )</td>
<td>Inhibitory gain</td>
<td>260 vu</td>
</tr>
<tr>
<td>( c )</td>
<td>Decay time-constant</td>
<td>1000 s(^{-1})</td>
</tr>
<tr>
<td>( \nu )</td>
<td>Axonal velocity</td>
<td>9 m s(^{-1})</td>
</tr>
<tr>
<td>( \sqrt{r_q} )</td>
<td>SD of axonal range</td>
<td>4 mm</td>
</tr>
<tr>
<td>( V_{ep} )</td>
<td>EPSP reversal</td>
<td>12 vu</td>
</tr>
<tr>
<td>( V_{ip} )</td>
<td>IPSP reversal</td>
<td>(-0.02) vu</td>
</tr>
<tr>
<td>( Q_{ns} )</td>
<td>Nonspecific input</td>
<td>Dimensionless</td>
</tr>
<tr>
<td>( Q_s )</td>
<td>Specific input</td>
<td>Dimensionless</td>
</tr>
</tbody>
</table>

4.1. Simplified averaging model

Each lattice element started with a pulse density of zero. Two linearly uncorrelated driving noise inputs with 0.0 mean [37, 41] and standard deviation 10.0 were given to the lattice at row 11 in columns 8 and 14. The two noise signals are referred to as an asynchronous noise source.

The system was allowed to attain stationary temporal evolution about a steady state mean. Then multichannel data sets of length 20000 were generated for a range of summations in equation (9). Using a reference channel located at column 11 on

Table 2. Synaptic couplings subscripts \( ee, ei, \) etc., indicate synapses between cell types, excitatory to excitatory, excitatory to inhibitory, etc. Types of coupling are: \( \alpha \) (cortico-cortical connections), \( \beta \) (intracortical connections), \( \mu \) (nonspecific cortical afferents) and \( M \) (specific afferents). Synaptic density fraction is the proportion of synapses of each type in unit cortical volume. (The exact values used in the simulations are given for completeness although the precision given is greater than is justified from the anatomical data.) Afferent fraction is the proportion of synapses on the excitatory or inhibitory cell dendrites respectively.

<table>
<thead>
<tr>
<th>Synaptic coupling</th>
<th>Synaptic density fraction</th>
<th>Afferent fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_{ee} )</td>
<td>0.765</td>
<td>0.8693</td>
</tr>
<tr>
<td>( \beta_{ee} )</td>
<td>0.0845</td>
<td>0.0960</td>
</tr>
<tr>
<td>( \beta_{ei} )</td>
<td>0.0149</td>
<td>0.1242</td>
</tr>
<tr>
<td>( \alpha_{ii} )</td>
<td>0.100</td>
<td>0.8333</td>
</tr>
<tr>
<td>( \beta_{ie} )</td>
<td>0.0228</td>
<td>0.0259</td>
</tr>
<tr>
<td>( \beta_{ii} )</td>
<td>0.004</td>
<td>0.0333</td>
</tr>
<tr>
<td>( \mu_{ee} )</td>
<td>0.0077</td>
<td>0.0088</td>
</tr>
<tr>
<td>( \mu_{ei} )</td>
<td>0.0011</td>
<td>0.0092</td>
</tr>
<tr>
<td>( M_{ee,i} )</td>
<td>Not given</td>
<td></td>
</tr>
</tbody>
</table>
the 13th row of the lattice, calculations of maximum cross–correlation for 100 timesteps into the future and past were made for each channel in the lattice. The delay at which the maximums occurred was also noted. The process was repeated for a selection of noise seeds and a typical case is plotted in Figure 1.

As summation length increased a pattern of maximum cross–correlation in elements around the driving sites developed. Results are calculated from the analogue of LFP but, equivalently, could be obtained using pulse–density as the observed state variable. A comparison of all cases showed the synchronous field was most highly developed for \( N = 100 \), where a maximum correlation close to 1.0 was

<table>
<thead>
<tr>
<th>Summation Length</th>
<th>Maximum Cross–Correlation</th>
<th>Delay to Maximum Cross–Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
<tr>
<td>10</td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>100</td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
</tbody>
</table>

(a) ![Image](image7.png) (b) ![Image](image8.png)

**Fig. 1.** Cross–correlation results for 20 x 20 version of Simplified Averaging model with two driving sites (see text) input asynchronous Gaussian noise. (a) Grey scale spatial plots of maximum cross–correlation, of the site marked ‘x’, with every other site for lags of \( \pm 100 \) timesteps, (b) Delay associated with the cross–correlation maximum at each site bar indicating \( \pm 2 \) milliseconds.
Spatial eigenmodes and synchronous oscillation

seen for time delays less than 5 timesteps or 0.5 milliseconds. For a 2000 Hz or less sampling system this would appear as zero–lag cross–correlation in agreement with Eckhorn and others [13].

The findings of zero–lag synchrony in the simplified averaging model, show nonlinear phase–locked oscillations are not needed for production of positive cross–correlations; only a system of coupled linear elements with delay summation on previous state values and connectivity with delay according to distance is necessary. In opposition to the conjectures of some authors [17,18,58] inhibitory elements are not required.

4.2. The physiological model

An examination of the physiological model was carried out to study the impact of local feedback dynamics upon synchronisation. In all cases correlations were obtained from LFP, and pulse–density results are equivalent.

In the standard two–input case of the complete model, nonspecific input $Q_{ns}$ DC of 20.0 was input to all elements of the lattice while the driving sites on row eleven of the lattice in columns 8 and 14 each received $Q_s$ Gaussian white noise signals of zero mean and standard deviation 0.005 to excitatory and inhibitory cell dendritic junctions. Multichannel time series of length 20000 points were acquired after an initialisation, and cross–correlation analysis carried out as before.

As for the Simplified Averaging case a pattern of maximum cross–correlation (in the range 0.7–1.0) for delays of less than 5 timesteps (0.5 milliseconds) exists as a field surrounding the driving sites. See Figure 2.

4.2.1. Spatial dependencies of eigenvectors

PCA was employed to study global cooperative modes in both Simplified Averaging and Physiological models with the results in the synchrony ranges being illustrative of the same phenomena. Therefore only the physiological model PCA is presented here.

Driving sites were removed prior to analysis to facilitate viewing of the spatial field patterns. A 20000 length PCA was chosen from a study of the asymptotic properties of the eigenvalues and is discussed elsewhere [8]. Ensemble averages were taken to obtain standard errors on the variance associated with each mode. For a selection of noise seeds, the ensemble average of results over 25 runs are in Figure 3.

Two dominant modes, containing over 99% of the variance of the original multichannel signal, were found. The first mode was similar to the pattern of zero–lag cross–correlation in Figure 1; the second mode consisted of two lobes with opposite sign loadings. The first eigenvector had associated variance about four times the second’s. Third and higher modes had significantly smaller eigenvalues and will not be examined.

4.2.2. Odd and even components of input signal

To discover what caused the dominant modes, two input paradigms were investigated. Firstly, two identical noise signals were fed to the driving sites – the
Fig. 2. Cross–correlation results. Physiological model 20 x 20 lattice with toroidal boundary conditions. Two specifically driven input sites (see text) received Gaussian mean zero white noise. Non–specific cortical activation ($Q_{ns}$) set at 20.0. (a) Grey scale spatial plots of maximum cross–correlation, of the site marked ‘x’, with every other site for lags of ±100 timesteps, (b) Delay associated with the cross–correlation maximum at each site bar indicating ±2 milliseconds.

$synchronous$ noise case. Secondly, the same noise signals were input to both sites with the input to one channel multiplied by minus one – the $anti–synchronous$ noise case. This amounted to inputting the even components of the two–dimensional noise signal for the former and the odd components of the same signal for the latter.

With purely synchronous or antisynchronous inputs, the ensemble eigenvectors (eigenmodes) are explained by odd and even components in the twin driving inputs. See Figure 3. In the synchronous noise case the first eigenmode occupies 99% of the variance and has the spatial form of the first mode in the asynchronous case. In the antisynchronous case the first eigenmode occupies 99% of the variance and has the spatial form of the second mode in the asynchronous case. The asynchronous case can then be explained by dominance of a synchronous eigenmode over an antisynchronous so the system acts as a coincidence detector; enhancing the transiently correlated inputs in the two noise signals, while reducing the effect of the negatively correlated ones. This process is made possible by dendritic summation in the elements.

The temporal characteristics of the synchronous field were examined using the principal mode time evolutions. Ensemble averaged spectra of these evolutions in Figure 4, indicate movement to the right through the different designated EEG frequency bands as cortical activation ($Q_{ns}$) was increased. A similar shift in power
spatial eigenmodes and synchronous oscillation

4.2.3. Effect of separation of sites on synchrony

Since coupling connectivity was a function of distance the effect of different separation of driving sites on synchrony was examined.

The physiological model was simulated on a 20 x 40 lattice to allow separation to an order of 19 intervening sites. Two sites situated symmetrically about the midline, and separated by 3 cells in row 11 of the lattice, were submitted to asynchronous Gaussian mean zero standard deviation 0.005 noise. The driven sites were then progressively moved apart and the multichannel output analysed with PCA.

Spatial eigenmodes for different driving site separations are in Figure 5. As separation increased the synchronous field broke down as first and second eigenmodes became two separate synchronous fields, one around each of the driving sites.

Additional coupling was employed to see if the fall–off of synchrony with distance was due to decreasing connection strength or the size of axonal time delays. The additional couplings were imposed between elements which were equal horizontal distances from the midline, and also declined in coupling strength with Gaussian distance. Standard axonal time delays were retained.

The dominant synchronous mode in this case retained the same concurrent fields of synchrony around both driving sites, and the first and second eigenmodes still partitioned responses to even and odd parts of the input signals. The first mode occupied a greater percentage of the variance than in the system with standard Gaussian...
Fig. 4. Power spectra of first two temporal principal component vectors associated with the eigenmodes in Figure 3. Ordered from top to bottom, $Q_{ns} = 0$, $Q_{ns} = 20$, $Q_{ns} = 40$ and $Q_{ns} = 50$. Ensemble averages over 10 runs. Simulation timestep is 0.1 ms.

coupling; evidence for the role of relative connection strength in synchrony. See Figure 6.

4.2.4. Time of onset of synchrony in physiological model

If synchrony is in fact an important mediator of perceptual binding, then physiologically and in any plausible candidate model, the field of synchrony must be capable of flexible adaptation on a time–scale of milliseconds. Since perceptual coding is on
Fig. 5. First two eigenvectors and associated variances for PCA done over a 10000 timestep (1 second) interval of a 20 x 40 lattice simulation of the Physiological model. Two specifically driven input sites were located at separations of 5, 11 and 19 intervening elements with $Q_{ns} = 20.0$ and $Q_{s} = 0.005$ standard deviation Gaussian mean zero white noise. Standard Gaussian coupling between sites was employed.

The order of a few milliseconds [2, 26, 29, 31, 38, 49, 55, 61] the temporal evolution of the system eigenmodes was studied.

A two–dimensional asynchronous noise signal was input as before, with $Q_{ns}$ of 20.0 to all elements. A repetition of the simulation was carried out for a matching signal, except for an aberrant 100 step (10 millisecond) noise signal at 15000 steps (1.5 seconds). The aberrant period had the same first order properties of mean and autocorrelation but was phase–randomised using the method of surrogate data [50].

After removal of a 5000 timestep initialisation, PCA was applied over 20000 timesteps and the principal component temporal vectors calculated for both ‘normal’ and ‘aberrant’ input cases using equation (5). The temporal evolution associated with the first eigenvector was studied since it contained the dominant spatial synchronous effect (similar results were found for higher order modes).
Fig. 6. First and second eigenmodes and associated variances for PCA done over a 10000 timestep (1 second) interval of a 20 x 40 lattice simulation of the Physiological model. Two specifically driven input sites were located at separations of 5 and 19 intervening elements with $Q_{ns} = 20.0$ and $Q_s = 0.005$ standard deviation Gaussian mean zero white noise as in Figure 5. Here both standard Gaussian coupling and extra long-range Gaussian coupling between sites was employed.

The magnitude of the difference between the normal and aberrant time series was calculated. This procedure was repeated twenty-five times for a selection of pairs of noise signals. An ensemble average of the magnitude of the difference between the two series was then found. See Figure 7.

The aberrant noise input was taken to have been registered by the spatial field of the lattice when the magnitude of the difference between the two time series became large. Large, in this context, meant the change in magnitude of the difference at a given time point was larger than the maximum deviation from the mean observed at any time point in the 10000 magnitude of the difference time series before the aberrant noise onset.

A separation in the two temporal evolutions occurred at about 25–30 timesteps. See Figure 7. The synchronous field of the entire system will therefore respond to a different stimulus within a few milliseconds. Much smaller sections of the lattice respond even faster because of their closer connection proximity to the source.
4.2.5. Nonlinear phase-locking results

If the level of non-specific input \( Q_{ns} \) was increased sufficiently to the two driving sites then each element in the system attained limit cycle oscillations. In particular, a state was obtained in which the elements of the lattice started to oscillate at a frequency close to 40 Hz.

As a first step, the phase relationships between the individual oscillating elements in the lattice were examined. The driving sites were completely phase-locked and locking was also found between other elements at a lag to the driving sites. A graph of the phase difference between the left-most driving site and every other element in the lattice is in Figure 8. Segments of the lattice were found to be phase-locked as expected for a time-delay nonlinear oscillator network. The operation of the wave mechanism described earlier allows phase-locking of outer array elements at a large phase lag to the driving sites. A similar phenomena of segregated phases has been found in a global system of nonlinear oscillators [60].

These oscillations were linked in zero-lag phase with other elements to form clusters of zero-lag phase and required local inhibition for their occurrence – in agreement with abstract nonlinear models of synchronous oscillation proposed by others [17,18,58].
Fig. 8. Phase relationships between the left hand driving site and every other element of the lattice at a stage in which limit cycle oscillations are present. $Q_{ns} = 50$.

As $Q_{ns}$ was increased the temporal evolution of individual lattice elements underwent transitions from damped or stable focus states through limit cycles to unstable foci. See Figure 9 for a selection of such transitions for one of the elements.

Complex periodic behaviour associated with a synchronous spatial field, similar to in the simple periodic case, was observed for lower level input driving and high values of $Q_{ns}$.

A qualitative check was made on how sensitive the limit cycle synchronous modes were to the input of noise. It was found the system remained stable even for relatively large noise inputs (standard deviation of 0.1) when $Q_{ns}$ was 40 and $Q_s$ was 0.6. The noise perturbed the system in and out of exact phase-locking but did not lead to completely uncontrolled excitation. The time course of two elements is in Figure 10.

The results are reminiscent of stable limit cycles with stochastic or nonlinear bursting seen in some physiological conditions [19, 20] and thought to be associated with specific perceptual events. The very limited use of complex neurotransmitter regulation in the present model precludes close comparison.
5. Discussion and conclusions

The results obtained in these simulations indicate synchronisation, between separated sites in the brain, may arise via fairly distinct mechanisms dependent on the level of cortical activation.

Firstly, synchrony is an inevitable property of delay elements similar to dendrites joined by long-range couplings with relatively rapid transmission and can be explained by simple linear models operating on uncorrelated inputs such as the Simplified Averaging model employed here. There is no need to assume special co-occurrence detection properties of dendritic membranes or complicated local dynamics to explain the way synchrony can lead to the elimination of asynchronous components in the activity of two concurrently firing sites in cortex. Instead, this property emerges directly from an eigenfunction decomposition of the travelling waves which radiate from the active sites. As a corollary, because of the almost linear superposition properties of travelling waves in this media, the basic principles of decomposition of eigenmodes illustrated above still hold for experiments with multiple uncorrelated inputs as reported in a following paper [9]. Synchrony can be associated with rhythmic broad-band electrocortical activity, including that in the gamma band. It is also seen as damped autocorrelation and cross-correlation.
functions. The onset of synchrony via this mechanism can co–ordinate events in an extended neuronal field very rapidly, and is therefore well suited to the role of mediation of binding in cognitive and perceptual processes. This synchrony is quite distinct from oscillation – the oscillation arising locally and requiring local excitatory/inhibitory interactions – while synchrony requires only excitatory processes. Synchrony can arise over both long and short ranges, but the form of the synchronous field is sensitive to the specific coupling within the field. Specific corticocortical connections might provide a histological counterpart to the additional long–range connections employed in these situations [3,21,22]. Such connections have been associated in the visual cortex with similar receptive fields and orientation preference properties in spatially separated sites [28] that give rise to synchrony at distances up to 7 mm [23].

Secondly, at sufficiently high levels of activation of the physiological model, a nonlinear mechanism with phase–locking emerges. This mechanism is potentially capable of mediating much more complicated dynamic interactions between cortical sites. Oscillation and both excitatory and inhibitory synaptic transmission are essential to the mediation of this type of synchronisation. This may be equated

Fig. 10. Transitions from limit cycle to stochastic bursting with input of noise. $Q_{ns}$ input of 40 and a $Q_{s}$ DC of 0.6 with zero mean standard deviation 0.1 Gaussian white noise sent to driving sites 208 and 214 in a 20 x 20 lattice. Plots on the left–hand side are with respect to driving site 208 over different length time scales. Plots on the right–hand side are with respect to site 10 on the first row of the lattice, over different length time scales.
with the role in image segmentation ascribed to synchronous mechanisms in some models [10]. Notably, this synchronous mechanism could be partially distinguished from the former type by the cross–correlation and autocorrelation functions which it gives rise to. These functions would be virtually undamped for at least a transient period, and bandwidth limited (the 40 Hz band in this model), rather than broad–band.

Of the two mechanisms, the former linear synchrony would be much more readily observable physiologically, wherever large fields of cells interact regardless of scale. This may explain why broad–band synchrony is observed widely over many cortical areas [4]. The nonlinear mechanism would be seen only episodically and strictly locally, and may arise only in circumstances in which new information is emerging from autonomous local activity in the brain, rather than with the binding of simple sensory and perceptual information. These distinctions, although plain enough in simulation, are unlikely to be readily apparent physiologically. For instance, small changes in model parameters (which quantify more complex processes likely to be slowly time–varying in physiological reality) might shift the frequency of the nonlinear oscillation somewhat. Also, a distinction between noise–perturbed limit cycles and lightly damped linear oscillations in the gamma range cannot readily be made either on auto/cross–correlation profiles or with any other time series method – particularly since these separate dynamic processes may both occur transiently in the same recordings. First steps toward empirically distinguishing the two types of synchronisation using rotated PCA are reported in a following paper [9].

References

60. Wang, D.: Object Selection Based on Oscillatory Correlation. Department of Computer and Information Science and Center for Cognitive Science, The Ohio State University, Columbus, Ohio 43210, USA, 12, 1996
Synchronous oscillation in the cerebral cortex and object coherence: simulation of basic electrophysiological findings

J. J. Wright, P. D. Bourke, C. L. Chapman
Mental Health Research Institute of Victoria, 155 Oak Street, Parkville, Melbourne, Victoria 3052, Australia

Received: 22 December 1998 / Accepted in revised form: 16 March 2000

Abstract. A lumped continuum model for electrocortical activity was used to simulate several established experimental findings of synchronous oscillation which have not all been previously embodied in a single explanatory model. Moving-bar visual stimuli of different extension, stimuli moving in different directions, the impact of non-specific cortical activation upon synchronous oscillation, and the frequency content of EEG associated with synchrony were considered. The magnitude of zero lag synchrony was primarily accounted for by the properties of the eigenmodes of the travelling local field potential superposition waves generated by inputs to the cortex, largely independent of the oscillation properties and associated spectral content. Approximation of the differences in cross-correlation observed with differently moving bar stimuli, and of the impact of cortical activation, required added assumptions on (a) spatial coherence of afferent volleys arising from parts of a single stimulus object and (b) the presence of low-amplitude diffuse field noise, with enhancement of cortical signal/noise ratio with respect to the spatially coherent inputs, at higher levels of cortical activation. Synchrony appears to be a ubiquitous property of cortex-like delay networks. Precision in the modelling of synchronous oscillation findings will require detailed description of input pathways, cortical connectivity, cortical stability, and aspects of cortical/subcortical interactions.

1 Introduction

Discovery of the phenomenon of synchronous oscillation in the cerebral cortex (e.g. Eckhorn et al. 1988; Gray et al. 1989; Singer 1994; Singer and Gray 1995) was partly motivated by a proposal by von der Malsburg (1983) that patterns of synchronous activity in the brain would offer a mechanism whereby the many sensory input features contributing to the perception of a unified object might be linked to form a whole. This offered a solution to the binding problem posed by the need to register unique combinations of all possible stimulus features (Singer and Gray 1995; Livingstone 1996).

As has been pointed out by Palm and Wennekers (1997), interest in this solution to the binding problem, rather than the simpler requirements of direct modelling of experimental results, has driven much of the attempts to find an explanatory mechanism for synchronous oscillation. A strong presumption that linked non-linear oscillations underlie the phenomenon (e.g. Abarbanel et al. 1996) has motivated most attempts at modelling the phenomenon. Yet there is little evidence that phase-locking of non-linear oscillators is necessarily involved (Wennekers and Palm 1997), considering the broad spectral band over which synchrony has been found (Bressler et al. 1993).

The significance of the experimental data now published on synchronous oscillation is very substantial, but remains controversial. Theoretical opinions vary from the view that synchrony may be essentially irrelevant to synaptic interactions (Amit 1998) to the possibility that synchrony is vital to the co-ordination of synaptic modifications in the brain (Phillips and Singer 1997). Experimental conditions leave a number of variables (connectivities, signal/noise ratios, etc.) unspecified. Meantime, attempts to model the process must be selective. Some experimental aspects of synchronous oscillation appear reasonably clearly defined.

Firstly, it is known that synchrony is most often observed in association with oscillation in the gamma band (e.g. Eckhorn 1988) but not uniquely so (Bressler et al. 1993).

Secondly, it appears that synchrony is strongest between cortical sites in which neurones seem to be involved in co-processing of sensory input, e.g. between columns of cells with similar orientation preference (Gray and Singer 1989; Livingstone 1996).
Thirdly, it appears that increased cortical activation, such as is produced by driving the mesencephalic reticular formation, enhances synchronous oscillation (Munk et al. 1996).

Fourthly, properties of the stimulus object(s) play a part. One large bar moving across the visual field produces stronger synchrony than two small bars concurrently stimulating each of the relevant points in the retina, and bars moving in separate directions generally produce less synchrony than the same bars moving in the same direction across the visual field (Livingstone 1996; Neuenschwander and Singer 1996).

To attempt to provide a concise account for these experimental properties, we used a lumped continuum model of the cortex which provides an account of the general spectral content of EEG (Wright 1999). This approach is intended to complement other simulations which approximate physiological realism using feedforward networks with inhibitory surrounds, or single and multiple orientation domains (Schillen and Konig 1994; Fuentes et al. 1996; Xing and Gerstein 1996; Juergens and Eckhorn 1997). Our object was to ascertain the minimal assumptions needed to reproduce the experimental data. The model’s parameterisation is as of yet approximate, and physiological detail is incomplete, but these limitations are not relevant to the demonstrations we will report. Models closely related to that applied here have been analysed with regard to cross-correlation properties under the condition that two points on the cortical surface are driven by independent white noise (Wright 1997; Chapman et al., in press; Robinson et al. 1998), yielding results which have guided the experimental design used here. In this model, synchrony depends wholly upon relatively long-range excitatory connections in a continuum field – dendritic lag-summations and relatively rapid axonal transmission being the essential ingredients – while short-range excitatory/inhibitory interactions appear crucial only to the occurrence of oscillation. These characteristics are very similar to those observed in neural network models with intracortical couplings only, observed by Wilson and Bower (1991). The continuum formulation supplements the neural network approach, by enabling a different insight into the physical nature of synchrony, since the essential non-linearity of individual elements is avoided, and the stochastic and essentially linear properties of the neuronal mass is emphasized. Likewise, these intracortical two-dimensional models contrast with the work of Lumer et al. (1997a,b), who studied conditions for synchrony and oscillation in a neural-network model of the thalamocortical system. They observed synchrony at many levels, which they attributed largely to re-entrant activity at multiple levels in the pathway.

In the present model, reproduction of the physiological experiments required additional simple assumptions about the signal-to-noise ratio of the cortex, and cross-correlation in afferent volleys associated with individual stimulus objects, as will be described with Sect. 2.

2 Methods

The simulation used here has been reported in Wright (1999). Parameters were chosen to accord with the following assumptions about the overall neuronal population properties:

1. Most synapses are located on the distal dendritic tree, so mean delay from synapses to soma is near the upper limit of physiological estimates for dendritic delay (Segev 1995; Thomson 1996, 1997).
2. Dendritic delays in excitatory and inhibitory cells, and from excitatory and inhibitory synapses, are comparable to first approximation.
3. Axonal delays are range dependent, but are always small compared to dendritic delays.
4. There is a finite probability that an action potential may be emitted by a given neuron, even at membrane potentials very close to the inhibitory reversal potential.
5. Synaptic gains are in the ratio of 4:1 for inhibitory synapses versus excitatory synapses (Segev 1995) and produce a high signal amplification (Thomson 1997) so that the stable operating range of cortical activity is restricted to low pulse densities.

2.1 State equations

The $N$ cells in a unit volume each have a probability of emission of an action potential $q$, as a function of their membrane potentials. The sum of population membrane potentials is taken to be directly proportional to the local field potential (LFP), $V(t)$ at time $t$. Then, in a mean-field approximation, the pulse-probability density $Q(t)$ is given by

$$Q(t) = \frac{1}{N} \sum_{i=1}^{N} q_i(V) . \quad (1)$$

By the central limit theorem, for large $N$, $Q$ will have a Gaussian distribution with respect to $V$, whatever the individual distributions of $q_i$, so $V$ and $Q$ are approximately related by

$$Q = (1 + e^{(V-3)})^{-1} . \quad (2)$$

This sigmoidal relation is used to approximate the sigmoidal error function implied by the Gaussian population pulse probability distribution. Where $a = -\pi/\sqrt{3}$, LFP voltage units (vu) are approximate to standard deviations of the distribution of cell pulse probability over the complete range of LFP, with a 50% mean probability of pulse emission 3SD from complete polarisation of the neural population. Thus, $Q$ has a value close to zero when $V = 0$, and approaching an asymptote of maximum pulse rate at $V = 6$vu.

The time response of mean membrane potential (and by implication LFP and some potential) is given by

$$V(t) = g \sum_{j=1}^{n} w_j Q_j(t - j\Delta t), \quad j = 1, 2, 3 \ldots n , \quad (3)$$
where \( g \) is synaptic gain, \( Q_a \) is afferent pulse action density, \( \Delta t \) is the discrete time-step, and \( n \Delta t \) is large compared to the peak time response of membrane potential. In accord with Robinson et al. (1997),

\[
w_j = b^2 j \Delta t e^{-b j \Delta t} \tag{3a}
\]

represents the rise and fall of membrane potential in response to input at \( t = 0 \), incorporating lags due to both synaptic conduction and average dendritic cable delay in a single function. Parameter \( b \) regulates both the peak time and mean delay associated with this lag. Time step \( \Delta t \) was set at 0.1 ms, after trials showed that progressive decrements of the time step to 0.01 ms produced only small, asymptotically diminishing effects on the spectral content of the results.

Within unit volumes, both excitatory and inhibitory cell groups are distinguished, each reciprocally and self-coupled, and each coupled at longer range to other unit volumes by cortico-cortical fibres. Delays due to axonal conduction between unit volumes are given by \( \Delta t = r_{pq} / v \), where \( \Delta t \) is axonal conduction lag over the distance \( r_{pq} \) between the \( p \)th unit volume and the \( q \)th unit volume and axonal conduction velocity is \( v \).

Coupling strengths are proportional to:

1. The fractional density of synaptic couplings afferent to the dendrites of excitatory and inhibitory cells respectively (\( z_{ee}, \beta_{ei}, \beta_{ai}, M_{ee}, \) etc as listed in Table 2).
2. The synaptic gains of excitatory and inhibitory synapses, \( g_e \) and \( g_i \).
3. Changes in synaptic efficacy, \( E' \), representing feedback effects including those of reversal potentials (Kandel and Schwartz 1985). That is, with increasing depolarisation of cell membranes there is an increase in sensitivity to inhibitory synaptic inputs and a decrease in sensitivity to excitatory synapses. These feedback relations are modelled as linear regressions of efficacy with membrane potential

\[
E_{ee}(t) = \left[ 1 - V_e(t) (t - \Delta t) / V_{er} \right],
\]

\[
E_{ei}(t) = \left[ 1 - V_e(t) (t - \Delta t) / V_{ir} \right],
\]

\[
E_{ie}(t) = \left[ 1 - V_i(t) (t - \Delta t) / V_{ir} \right],
\]

\[
E_{ii}(t) = \left[ 1 - V_i(t) (t - \Delta t) / V_{ir} \right],
\]

where the subscripts e and i indicate excitatory and inhibitory potentials, and subscript R a constant-valued reversal potential. Efficacies \( \{ E' \} \) were applied with smoothing, so that in each case \( E(t) = \sum_{j=1}^{n} w_j E_j(t - j \Delta t) \), where \( w_j = c e^{-c j \Delta t} \), describing an exponential decay of the impact of instantaneous membrane potential upon synaptic efficacy. For high values of \( c \), this decay is rapid, as would be expected for reversal potentials alone.

State equations for the cortical system are then given for the \( p \)th unit volume by

\[
V_e(t) = \frac{\sum_{j=1}^{n} w_j Q_{ee}(t - j \Delta t)}{\sum_{j=1}^{n} w_j},
\]

\[
V_i(t) = \frac{\sum_{j=1}^{n} w_j Q_{ii}(t - j \Delta t)}{\sum_{j=1}^{n} w_j},
\]

where \( Q_{ee}, Q_{ai} \) are afferent synaptic action densities for excitatory and inhibitory cells respectively, in the \( pq \)th unit volume, receiving local synaptic input at negligible axonal delay and delayed cortico-cortical inputs from \( q \)th unit volumes at range \( r_{pq} \), \( q = 1 \ldots u \), in accord with

\[
Q_{ee}(t) = g_e \beta_{ee} E_{ee}(t) Q_{ee}(t) - g_i \beta_{ei} E_{ei}(t) Q_{ei}(t) + g_e M_{ee} E_{ee}(t) Q_{ns}(t) + g_e \sum_{l=1}^{n} z_{ee}(r_{pq}) E_{ee}(t - r_{pq} / v) \tag{5a}
\]

\[
Q_{ai}(t) = g_e \beta_{ei} E_{ei}(t) Q_{ei}(t) - g_i \beta_{ai} E_{ai}(t) Q_{ai}(t) + g_e M_{ei} E_{ei}(t) Q_{ns}(t) + g_e \sum_{l=1}^{n} z_{ai}(r_{pq}) E_{ei}(t - r_{pq} / v) \tag{5c}
\]

where \( z_{ee}(r_{pq}) \) and \( z_{ai}(r_{pq}) \) are partial input synaptic densities, such that \( \sum_{l=1}^{n} z_{ee}(r_{pq}) = \sigma_{ee} \) and \( \sum_{l=1}^{n} z_{ai}(r_{pq}) = \sigma_{ai} \). \( Q_e \) and \( Q_{ns} \) are system inputs. \( Q_s \) represents all time-varying components in specific cortical afferents and \( Q_{ns} \) is a uniform DC input representing nonspecific cortical activation, which acts as control parameter.

### 2.2 Standard parameters and definition of units

State variables and parameters, their dimensions and standard values are given in Tables 1 and 2. Where possible, physiologically accurate values have been applied, and certain difficulties of parameterisation have been avoided by the use of normalised units. For details, see Wright (1999).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_r )</td>
<td>Excitatory LFP</td>
</tr>
<tr>
<td>( V_i )</td>
<td>Inhibitory LFP</td>
</tr>
<tr>
<td>( Q_e )</td>
<td>Excitatory PPD</td>
</tr>
<tr>
<td>( Q_i )</td>
<td>Inhibitory PPD</td>
</tr>
<tr>
<td>( a )</td>
<td>Slope parameter</td>
</tr>
<tr>
<td>( b )</td>
<td>Dendritic time constant</td>
</tr>
<tr>
<td>( \theta_e )</td>
<td>Excitatory gain</td>
</tr>
<tr>
<td>( \theta_i )</td>
<td>Inhibitory gain</td>
</tr>
<tr>
<td>( c )</td>
<td>Decay time constant</td>
</tr>
<tr>
<td>( v )</td>
<td>Axonal velocity</td>
</tr>
<tr>
<td>( \sqrt{r_{pq}} )</td>
<td>SD of axonal range</td>
</tr>
<tr>
<td>( V_{ce} )</td>
<td>EPSP reversal</td>
</tr>
<tr>
<td>( V_{ci} )</td>
<td>IPSP reversal</td>
</tr>
<tr>
<td>( Q_{ns} )</td>
<td>Non-specific input</td>
</tr>
<tr>
<td>( Q_s )</td>
<td>Specific input</td>
</tr>
</tbody>
</table>

### Table 1. State variables and standard parameters other than synaptic densities. LFP Local field potential, PPD pulse probability density, EPSP excitatory post-synaptic potential, IPSP inhibitory post-synaptic potential
Table 2. Synaptic couplings subscripts ee, ei, etc. indicate synapses between cell types, excitatory to excitatory, excitatory to inhibitory, etc. Types of coupling are: e (cortico-cortical connections), \( \theta \) (intracortical connections), \( \mu \) (nonspecific cortical afferents) and \( M \) (specific afferents). Synaptic density fraction is the proportion of synapses of each type in unit cortical volume. [The exact values used in the simulations are given for completeness (Liley and Wright 1994) although the precision given is greater than is justified from the anatomical data.] Afferent fraction is the proportion of synapses on the excitatory or inhibitory cell dendrites respectively, and are thus the values applied in Eq. (5)

<table>
<thead>
<tr>
<th>Synaptic coupling</th>
<th>Synaptic density fraction</th>
<th>Afferent fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_{ee} )</td>
<td>0.765</td>
<td>0.8693</td>
</tr>
<tr>
<td>( \rho_{ee} )</td>
<td>0.0845</td>
<td>0.0960</td>
</tr>
<tr>
<td>( \rho_{ei} )</td>
<td>0.0149</td>
<td>0.1242</td>
</tr>
<tr>
<td>( \rho_{i} )</td>
<td>0.100</td>
<td>0.8333</td>
</tr>
<tr>
<td>( \rho_{e} \rho_{i} )</td>
<td>0.0228</td>
<td>0.0259</td>
</tr>
<tr>
<td>( \rho_{i} )</td>
<td>0.004</td>
<td>0.0333</td>
</tr>
<tr>
<td>( \rho_{ei} )</td>
<td>0.0077</td>
<td>0.0088</td>
</tr>
<tr>
<td>( \rho_{ei} )</td>
<td>0.0011</td>
<td>0.0092</td>
</tr>
<tr>
<td>( M_{ei,i} )</td>
<td>Not given</td>
<td></td>
</tr>
</tbody>
</table>

2.3 Configuration of simulation

An extended area of the cortex was simulated by unit volumes in a \( 20 \times 20 \) matrix, each volume connected with its neighbours so that the coupling strengths, \( x_{ee} \) \( (r_pq) \) declined with \( r_{pq} \) as a Gaussian function with standard deviation of four distance units, where a distance unit is the side of one cell of the \( 20 \times 20 \) matrix. This approximates to distribution of cortico-cortical fibres in the cat brain if the distance unit is taken as about 0.9 mm. Boundary conditions were toroidal in all simulations reported. The application of absorbing boundary conditions and changes in matrix size were also studied and these changes did not affect the results to be reported.

2.4 Added assumptions

For the results which follow, two further assumptions proved necessary. Discussion of these assumptions is reserved to Sect. 4:

1. A single object moving in the sensory field stimulates feature detectors in primary sensory neurones so that, over any short epoch, afferent volleys in the sensory pathway are correlated at zero lag – i.e. inputs from a single stimulus object give rise to spatially coherent input at the cortical level. Distinct stimulus objects are thus uncorrelated with each other by definition.

2. Cortical signal-to-noise ratio rises with cortical activation – i.e. the spatially coherent component of inputs from stimulus objects is greater relative to background noise at higher levels of cortical activation.

2.5 Simulation inputs and outputs

Two configurations were used in these studies.

2.5.1 Stationary driving inputs. These simulations follow Wright (1997), Robinson et al. (1998), and Chapman et al. (in press) and are here used to succinctly demonstrate the basic physical mechanism by which synchronous oscillation appears in the simulation, as a prelude to the moving-bar studies.

Specific inputs, \( Q_s \), imitated time-variation of a complex localised input to selective sites in the cortex. These were delivered as two independent (asynchronous) time series of zero-mean white noise (Marsaglia and Zaman 1987) of small amplitude, input to the unit volumes situated at sites in one row of the matrix of elements, and separated by five intervening elements. Input was received at both excitatory and inhibitory cell dendritic junctions of the driven sites.

Non-specific inputs, \( Q_{ns} \), imitated the action of reticular, catecholaminergic and other diffuse inputs to cortex, and were delivered uniformly to all unit volumes in the matrix as constant non-zero inputs, throughout the duration of each simulation run.

Although both types of input are extremely simple compared to real cortical input, results are sufficiently general for interpretation of results from inputs of greater complexity, as will be discussed.

Outputs were recorded as \( V_c(t) \) and \( Q_e(t) \) from all other elements in the field, and were used to calculate lagged cross-correlations, and to calculate the major principle components (spatial eigenmodes) of activity in the simulated cortical field, after prior removal of the signals from the \( Q_s \)-driven sites. Spatial eigenmodes were calculated using spatial principal component analysis (PCA) (Preisendorfer 1988) to resolve the wave activity generated in the simulation into equivalent modes of spatially stationary synchronous activity.

Simulation runs began with \( \{ Q_s, Q_i \} \) all initially zero, and ran for 1 s, which ensured steady-state (signal mean stationary) conditions were achieved. Outputs were analysed over the subsequent second, to obtain data for cross-correlations, and 2 s to obtain data for principal component analysis – the results of both these measures differing negligibly with further prolongation of the run time.

2.5.2 Moving bars. The simulation of moving bars implicitly defines a single moving object as a set of cross-correlated inputs, more or less spatially contiguous.

The simulated bars, which are \( Q_s \) inputs, moved across the simulated cortical surface, while non-specific activation, \( Q_{ns} \), was delivered uniformly as before. In some experiments the \( Q_{ns} \) inputs included spatiotemporal white noise in addition to the usual DC component, as will be reported. This diffuse noise was provided by input of white noise time series initiated from different seed values to each element in the cortical matrix, allowing a more realistic representation of cortical activation. The rms amplitude of specific inputs was 0.01 units, and the point amplitude of diffuse field noise is reported in Sect. 3 in the same units.
Moving bars were orientated along the rows of the matrix, and moved up or down the columns. The standard apparent velocity of movement of the simulated cortical surface was 24.4 mm/s, so that a complete single sweep of the bars occurred in 0.8192 s. Results from these runs required ensemble averaging to yield sufficient confidence on magnitude, although robust synchrony was apparent in individual sweeps. Bar speed was varied in control experiments.

Short bars, of 3 mm length in cortical projection, and a long bar of 20 mm (the width of the matrix) were simulated, each with an apparent width of 1 mm (one cortical element). The long bar thus reached completely across the simulated cortical surface, while the short bars were aligned so as to pass close to the two reference sites subsequently used for cross-correlation. Usually bar movements were such that the reference sites were crossed simultaneously by the two short bars, or by the length of the long bar, whatever the direction of movement. This occurred when the reference sites were in the tenth row of the matrix. A control condition was the case in which the reference sites were located in the fifth row of the matrix—that is, as far as possible from the zeroth and tenth rows of the matrix, the two positions where oppositely moving bars pass each other. This exceptional condition we called the “remote passing” condition, to contrast it to the more general “proximal passing” condition.

In the experiments to be reported, the two short bars moved so as to straddle both reference sites, without directly driving either. This is not critical to the results obtained, and it is relevant only that the short bars pass close to the reference sites.

To prevent stepping discontinuities, the forward movement of the bars was smoothed in imitation of a 1-mm-wide bar moving continuously.

In both long and short bars, only every second element along the length of the bar was driven, with a 50% probability of receiving input at each time step. In no case did the reference sites themselves receive direct specific (Q_e) inputs, as to do so confounds cross-correlations to be made with regard to field effects in the simulation with the direct content of the inputs. Both zero-mean and non-zero mean Q_e inputs were applied in different simulation runs.

Finally, the white noise time series used to drive any single bar were arranged according to either of two conditions. In the “synchronous bar” condition, all inputs to a single bar were synchronous (identical), while they were asynchronous (uncorrelated) between separate bars. This condition meets the definition of a single moving object, as mentioned above. In the “asynchronous bar” condition, all inputs within, as well as between bars, were uncorrelated.

3 Results

3.1 Stationary driving inputs

Figure 1 shows the essential mechanism whereby synchronous oscillation arises in these simulations, as reported in Wright (1997), Robinson et al. (1998), and Chapman et al. (in press). We wish to emphasise that zero-lag synchrony is a universal attribute of continuum fields or neural networks which have summing junctions with delay (dendrites) andcouplings (axons) with transmission delay which is small compared to dendritic delay. Synchrony then occurs by summation of even (in phase) components in the separated input signals, and cancellation of odd (out of phase) components. This effect does not require non-linearity.

When twin uncorrelated signals are delivered to the cortical surface, a pattern of zero-lag synchrony develops around the driving sites, as is shown in the upper diagrams in Fig. 1. Similar results can be obtained for cross-correlations of any reference point close to either of the sites of white noise input. The pattern of zero-lag synchrony has been shown to be sensitive to inhomogeneities in the strength of connections between elements of the matrix, whether these reflect structural or dynamic coupling gains. Conversely, uniform multiplications of the connection strength throughout the field have no effect on the pattern of synchrony. Similarly, uniform increases in cortical activation (Q_{ms}) have little effect, so long as the level of activation remains below a critical level at which limit cycles develop. A highly non-linear dynamic emerges at very high Q_{ms} and this condition was avoided throughout the current studies.

The middle diagrams in Fig. 1 show the first and second eigenmodes revealed by PCA. The first eigenmode is similar in form to the field of zero-lag maximum cross-correlation, and consumes the majority of the variance in the field. It can be shown (Chapman et al., in press) that the first eigenmode follows even (in-phase) transients in the driving inputs and the second eigenmode follows odd (anti-phase) transients in the same inputs. The lower diagrams of Fig. 1 show schematically the way in which the summation of even components in the surrounding field creates the first eigenmode, which predominate over the second eigenmode—the form of the second mode being attributable in part to cancellation of the odd signals spreading into the field from the inputs. Predominance of the first mode is greatest in the field near the driving sites—hence the field of zero-lag synchrony.

3.2 Impact of concurrent noise and cortical activation—static case

Figure 2 shows the cross-correlations found between two sites within the field of zero-lag synchrony generated by the two stationary sites of driving with asynchronous noise, while the rms amplitude of input to the driving sites was held constant and diffuse noise throughout the field was increased.

With increasing field noise, the cross-correlation decreases as expected, but there is little effect on the cross-correlations obtained when the degree of cortical activation is changed fourfold. The lack of sensitivity of cross-correlation to Q_{ms} alone is accounted for by the fact that the first and second eigenmodes of field activity increase concurrently in amplitude in response to the
Fig. 1. Essential properties of synchronous oscillation. **Top figures** A representation of the simulated cortical surface. **Open squares** represent the sites of input of uncorrelated white noise. The **filled square** is the reference point from which cross-correlations are calculated with respect to the rest of the field. **Top left** Maximum positive cross-correlations. **Top right** Delay associated with maximum cross-correlation. **Middle figures** The first and second principal eigenmodes of activity on the simulated cortical surface, obtained from the same data used to generate the top figures. **Bottom figures** Schematic “freeze frame” images of potentials (or pulse densities) on the simulated cortical surface seen when the twin inputs are sine waves, of any single frequency. **Bottom left** Potentials in the field when the sine waves are in phase. **Bottom right** Potentials in the field when the sine waves are of reverse phase.
uniformly increasing dynamic gain (dQ/dV), produced by increase in Q_m. Such small increase in cross-correlation as occur with increasing Q_m, is manifest at higher "noise-to-signal" ratios.

3.3 Moving bars

Figure 3a shows the simulation's reproduction of the findings of Gray et al. (1989) and Eckhorn et al. (1988) for moving bars. It can be seen that LFP time series, power spectra, and cross-correlations closely approximate experimental findings. Only a single-sweep LFP time series is shown here, but superposition of LFP from repeated sweeps shows that they are not phase-locked to the presentation of the stimulus, in the same manner as noted experimentally by Eckhorn et al. (1988). The particular frequency content of the LFP time series, which matches the predominance of power in the gamma range typically seen in experiments, is a consequence of the choice of Q_m applied in these particular simulations. At much lower values of Q_m, the spectral content resembles the EEG at lower levels of activation — yet the cross-correlation values are little affected.

Notably, the highest zero-lag cross-correlation is seen for the single large bar, next for two short bars moving in the same direction, and lowest for the two short bars moving in opposite directions. The relative magnitude of these cross-correlations is essentially independent of Q_m.

Figure 3b shows data from simulations similar to those shown in Fig. 3a, except that the "asynchronous bar" condition has been applied. Notably, the long bar now yields no higher cross-correlation than two bars moving in the same direction.

Figure 3c uses conditions matched to those in Fig. 3b, and cross-correlations are shown for both the standard "proximate passing" and the "remote passing" conditions. Variation in the situation of the reference sites relative to the position of passing has little effect on the degree of cross-correlation, so long as the sites remain in a symmetric relation to the tracks of passage of the bars.

In all permutations of the various conditions applied, we found that when the diffuse field noise was zero, no differences could be demonstrated in cross-correlation for two bars moving in the same, versus opposite, directions. This result may appear paradoxical, since the two bars moving together are always closer to each other, therefore interact more strongly, and create larger amplitude travelling waves in the medium. However, this greater wave amplitude does not lead to any significant difference in the partition of energy between the first and second eigenmodes of the waves. Thus, cross-correlations (which normalise absolute magnitude) are not sensitive to the relative direction of bar movement so long as the paths travelled are otherwise identical.

The introduction of diffuse field noise produces sensitivity to the relative direction of movement of the bars, as shown in Fig. 3d. This result is accounted for as follows: when the two bars move in the same direction, the absolute magnitude of both the first and second eigenmodes is increased, although their relative magnitudes are unchanged. With equivalent levels of additive field noise, the signal/noise ratio, which is measured by cross-correlation, is then higher for the case in which the bars move together, and lower when they move in opposite directions.

3.4 Effect of cortical activation with concurrent increase in cortical signal-to-noise ratio – two moving bars

Figure 4 shows that, so long as constant amplitude diffuse spatio-temporal noise is applied to the field, then the zero-lag synchrony increases as cortical activation is increased. These results are closely similar to the experimental results reported in Munk et al. (1996).

3.5 Consequences of other alterations of signal properties

Qualitatively, the results in Figs. 3 and 4 do not depend upon the input signals associated with the moving bar having a non-zero mean. The results shown are for zero-mean input signals. Non-zero mean of the moving-bar inputs inevitably adds to the cross-correlation observed, unless high-pass filtering is applied to the output signals before cross-correlation is performed. (This offset or envelope effect is often removed experimentally by the use of high pass filters to allow for envelope of the input signal.) No input signal correlations are required for synchrony to emerge, no matter what the size or direction of movement of the bars.

The results shown are also independent of the choice of bar speed, which was doubled and halved from the standard speed without significant consequence on the
results. This is not congruent with the finding of Eckhorn et al. (1988) that increasing stimulus speed was associated with gamma-band oscillation at increasing frequency, and implies that such a relation would require stimulus speed to be associated with the level of cortical activation.

With regard to Fig. 4, if the amplitude of the specific stimuli does not increase with cortical activation, or if there is no diffuse field noise, then the maximum cross-correlation does not alter significantly.

4 Conclusions

The above results appear to capture concurrently the physiologically observed cross-correlation results of moving-bar experiments and influences of reticular formation stimulation, along with realistic representation of LFP spectral power, and absence of phase-locking of response to the input signal.

The simulated cortical medium has the capacity to selectively eliminate odd components in the asynchronous input signals. This basic property carries over into more complex moving stimuli, is little affected by the level of cortical activation and, correspondingly, does not require oscillation in the gamma range for synchrony to occur.

The ubiquity of synchrony without dependence upon any correlation in the inputs does not provide a sufficient explanation for either the experimental differences dependent on the direction of bar movement, or the effects of cortical activation upon synchrony. Reproduction of these physiological results depends upon two additional assumptions – firstly, that a single stimulus object can be defined as a spatially coherent set of inputs, and secondly that the signal-to-noise ratio of inputs increases concurrently with cortical activation.

In the results reported above, the first assumption proved necessary only to account for the enhanced magnitude of cross-correlations found when the stimulus is a long single bar, rather than two short bars moving in the same direction. This is a robust finding, both physiologically and within our simulations.
Fig. 3b. Results shown are derived as for 3a except that uncorrelated white noise was delivered to all parts of each bar, as well as being uncorrelated between bars. Diffuse field noise amplitude 0.000025

Fig. 3c. Cross-correlations for the asynchronous bar case, with diffuse field noise 0.000025, for bars moving in the same and opposite directions. Reference points at the matrix row at which the oppositely moving bars pass each other (proximate pass) or in the same columns but remote from the passing row (remote pass). See text for details. Differences are between the remote and proximate conditions. The small differences remain statistically significant

The first assumption may be usefully regarded as providing a definition of a stimulus object. Extension of this definition leads to the notion that any spatially synchronous pattern of activity in the brain is an “object” – either the representation of a physical object via sensory input, or a coherent pattern partly internally generated, and thus a representation of a mental object. This is equivalent to von der Malsburg’s (1983) original proposition that synchronous oscillation is the means by which binding is mediated. Concurrent stimulation of receptor cells in the cortex or elsewhere must produce synchrony in the afferent bombardment over at least some frequencies, as has been remarked in physiological experiments (e.g. Kreiter and Singer 1996; Neuenfuss-wander and Singer 1996; Steriade et al. 1996) and demonstrated in models of the visual pathways (Ghose
Fig. 3d. Cross-correlations from proximate-pass reference sites with differing levels of diffuse field noise, for both asynchronous and synchronous bar conditions. Differences are with respect to the directions of bar movement shown in the graphs to the left of each difference graph. The difference produced by direction of movement is zero at zero lag in the absence of diffuse field noise.
and Freeman 1997). It is not necessary that coherence in the afferent volleys be complete at all frequencies to give rise to the effects reported.

The second assumption we have shown to account for experimentally observed increases in cross-correlation with increasing cortical activation. More generally, a degree of diffuse field noise seems essential to account for the sensitivity of cross-correlation to relative direction of movement of multiple bars. There are implications here for the coherent informax hypothesis, regarding synaptic modification and learning in relation to synchronous oscillation (Phillips and Singer 1997). The information-theoretical derivation of coherent informax leaves unexplained how information from different stimulus configurations becomes selectively distributed over synapses. Our findings suggest that noise may be required to enhance contrasts of relative synaptic activation when stimulus objects move in relation to each other in different ways.

A relation between cortical activation and signal/noise ratio is without general proof from physiological data. It is inherently plausible, since cortical activation is associated with enhanced cortical information processing. It is loosely supported by the close association of pulse and LFP in cortical activity in the gamma band, as opposed to the notoriously low associations of EEG and pulse activity at lower EEG frequencies and cortical activation (Stryker 1989). A part of the enhancement of signal versus noise may be attributed to a further property of synchronous oscillation which is only weakly apparent in these simulations. As shown in Fig. 2, increasing cortical activation does somewhat enhance cross-correlation at medium ranges of signal-to-noise. This effect would be further enhanced were the effects of cortical depth or repeated passage of signals through layers of relay cells such as the cell layers of the lateral geniculate body included.

We were forced to make these assumptions because of the simplicity of representation of the input pathways in our model. The model of Lumer et al. (1997a,b), in contrast, dealt in detail with the input pathways, and reproduced synchronous oscillation as a general property of jitter stimuli introduced to their neural networks. Similar considerations hold for the earlier cited works on feedforward modelling (Schillen and Konig 1994; Fuentes et al. 1996; Xing and Gerstein 1996; Juergens and Eckhorn 1997). It would appear that these approaches and ours may be complementary.

One property of experimental synchronous oscillation which has not been reproduced is the relation of bar speed to the LFP spectral content (Eckhorn et al. 1988). Within the present model, this could be explained by the further assumption that faster moving stimuli contribute more strongly to cortical activation, whether directly or via collaterals in the reticular-activating system. This would have the effect of shifting the LFP spectrum further to the right, in accord with the experimental observation. In a study of somatic sensory synchrony by Ruiz et al. (1995), faster stimuli were associated with increased cell-firing rates, consistent with an association between stimulus velocity and total activation.

The present results do not include two further factors which may be crucial to quantitative reproduction of synchronous oscillation, whereas we have been able to obtain only the correct relative changes in cross-correlation magnitude according to different experimental conditions.

Firstly, we have here considered only symmetrical cortical couplings and uniform cortical activation whereas, in reality, both these factors are inhomogeneous, and systematically so. This may explain why our simulations did not capture the extremes of cross-correlation seen experimentally, in which cross-correlation may be sometimes reduced to zero for bars moving in opposite directions, while quite high correlations are seen for movements in the same direction. As earlier remarked, there is other evidence in similar simulations...
that non-uniform structural and dynamic coupling strengths change the patterns of synchronisation for static inputs. Introduction of non-uniform couplings would open the possibility of detailed simulation of synchrony – for example, between cortical sites of similar orientation preference – since such sites are more strongly coupled than sites of dissimilar orientation preference (Malach et al. 1993; Yoshioka et al. 1996). Like-to-like connectivity might also hold for velocity characteristics of stimuli, as is predicted by certain Hebbian models for the development of intracortical intrinsic connections (Alexander et al., in preparation).

Secondly, these results do not allow for the possible occurrence of non-linear phase locking as a mechanism of synchronous oscillation. At higher levels of cortical activation, non-linear phase locking occurs in these simulations, as reported in Chapman et al. (in press). This additional mechanism opens the door to modelling of the putative role of synchrony in image segmentation, as shown in abstract simulations by Wang (1996). The form of synchronisation modelled here acts as a means of rapid synchronisation of cortical areas active for whatever reason, i.e. the present mechanism would appear sufficient to mediate binding.

Beyond these considerations of intracortical dynamics and input path properties, still more complete accounts of synchronous oscillation will likely require an account of cortical/subcortical interactions, processes which appear required to organise the binding and uncoupling of cortical assemblies during ongoing perception and cognition (Miltner et al. 1999; Rodriguez et al. 1999).

Thus, extensions of these experiments may be expected to enable testing of the mechanisms proposed against increasingly complicated physiological experiments – at the price that experimentally, concurrent levels of cortical activation and the detailed local connectivity of cells must be explicitly considered.

References


Lumer ED, Edelman GM, Tononi G (1997b) Neural dynamics in a model of the thalamocortical system. II. The role of neural synchrony tested through perturbations of spike timing. Cerebral Cortex 7: 228–236


Unified neurophysical model of EEG spectra and evoked potentials

C. J. Rennie¹,², P. A. Robinson¹, J. J. Wright³

¹ School of Physics, University of Sydney, New South Wales 2006, Australia
² Department of Medical Physics, Westmead Hospital, Westmead, New South Wales 2145, Australia
³ Mental Health Research Institute, Parkville, Victoria 3052, Australia

Received: 22 May 2001 / Accepted in revised form: 8 January 2002

Abstract. Evoked potentials – the brain’s transient electrical responses to discrete stimuli – are modeled as impulse responses using a continuum model of brain electrical activity. Previous models of ongoing brain activity are refined by adding an improved model of thalamic connectivity and modulation, and by allowing for two populations of excitatory cortical neurons distinguished by their axonal ranges. Evoked potentials are shown to be modelable as an impulse response that is a sum of component responses. The component occurring about 100 ms poststimulus is attributed to sensory activation, and this, together with positive and negative feedback pathways between the cortex and thalamus, results in subsequent peaks and troughs that semiquantitatively reproduce those of observed evoked potentials. Modulation of the strengths of positive and negative feedback, in ways consistent with psychological theories of attentional focus, results in distinct responses resembling those seen in experiments involving attentional changes. The modeled impulse responses reproduce key features of typical experimental evoked response potentials: timing, relative amplitude, and number of peaks. The same model, with further modulation of feedback, also reproduces experimental spectra. Together, these results mean that a broad range of ongoing and transient electrocortical activity can be understood within a common framework, which is parameterized by values that are directly related to physiological and anatomical quantities.

1 Introduction

Evoked response potentials are transient electrical responses to sensory stimuli that can be recorded from the scalp or from within the brain itself. The biophysical mechanism for their generation is thought to be the same as that of the ongoing electroencephalogram (EEG), that is, caused by the extracellular currents induced by neuronal activity. Evoked potentials consist of ‘early’ components (20–60 ms after the stimulus) that can be ascribed to initial processing of the stimulus, but their most prominent components are the so-called late evoked response potentials (ERPs), which begin about 100 ms after the onset of any stimulus, have the appearance of a damped oscillatory waveform lasting ≈200 ms, and resemble EEGs in amplitude (≈10 μV) and frequency content (≈1–20 Hz). Moreover, ERPs are systematically modulated by the tonic level of attention, arousal, cognitive strategy, and by certain mental disorders, and so are a useful objective tool in psychology and psychiatry (Regan 1989).

Figure 1a is the spectrum of an ongoing EEG, and shows a prominent alpha (≈10 Hz) peak, typical of nonfocal cortical activation, such as when the subject is relaxed and not attending to external stimuli. Figure 1b illustrates typical ERPs and the extent of their modulation when experimental conditions are manipulated. In this case a pseudorandom sequence of frequent low tones and infrequent high tones are presented, and the average ERPs corresponding to each are seen to be strong functions of pitch probability. This differentiation requires attention being paid to the pitch differences, and identification of the infrequent and task-relevant tones. Experimental details of the data in Fig. 1 are described elsewhere (Bahramali et al. 1999).

Correlations like this between ERPs and behavioral measures of brain function are easily demonstrated and robust, but tend to lack sensitivity and specificity so that diagnosis and appropriate treatment of patients cannot be inferred from ERPs alone. Yet if their physiological origin were better understood, inverse modeling could be used to infer values for physiological parameters, and thereby enhance the diagnostic value of EEGs and ERPs. Recently we have used this approach to help understand ongoing EEG activity (Rennie et al. 1999; Robinson et al. 1997, 1998; Wright 1999; Wright and Liley 1994), using a continuum approximation of the cortex that included both excitatory and inhibitory
neural populations, range-dependent connectivities, dendritic delays, and a nonlinear response function. Subsequently this was generalized by the addition of feedback between cortex and thalamus (Robinson et al. 2001b), and was able to relate typical neuronal parameters to overall spectral shape (distinct power-law dependences in different frequency ranges) and to peaks in EEG spectra (forming an approximately harmonic series). The match between the model and observations was demonstrated both for states dominated by alpha rhythms (~8–12 Hz) and for sleep spectra characterized by delta and theta rhythms (1–8 Hz) and by sleep spindles (~14 Hz) (Robinson et al. 2001b). We use this model as the starting point for the present model for ERPs.

Of other models of large-scale collective neuronal behavior, one of the earliest that mentions ERPs was by Wilson and Cowan (1973), who first introduced the two-dimensional (2-D) continuum approximation for the cortex, reasoning that at scales greater than a few millimeters it should be possible to base a model exclusively on average neuronal properties. At the same time Lopes da Silva et al. (1974) proposed a lumped-parameter thalamocortical model for alpha rhythm. A 2-D continuum model by Nunez (1974, 1995) emphasized the role of long-range excitatory corticocortical connections in large-scale EEG phenomena, perhaps leading to standing waves at the alpha frequency. Elements of these models were combined by Jansen and coworkers in order to model the alpha component of EEGs and visual ERPs (Jansen and Rit 1995; Jansen et al. 1993). There is broad compatibility between all these models regarding the general manner in which to represent neural masses: the differences lie in the chosen anatomical simplifications and the specific phenomena being modeled.

However, a model that addresses the broad range of EEG and ERP phenomenons would be of value. As a step towards that end, the current paper describes the extension of our EEG spectral model to the simulation of ERPs. The modifications include more faithful representation of the thalamus and cortex, and the introduction of impulse-like events as precursors of ERPs, as suggested by Wright et al. (1990) based on inverse filtering of experimental data.

Our physiologically based approach contrasts – particularly regarding the choice of parameters – with phenomenological models that are often used for quantifying ERPs; e.g., using band-pass filtering (Robinson 1999), factor decomposition (Makeig et al. 1999), superpositions of damped sine waves (Shibasaki et al. 1987), and wavelets (Samar et al. 1995).

The main characteristics of ERPs that we address are the relative timing and amplitudes of several peaks in ERPs, particularly those evoked in response to trains of discrete stimuli (of any sensory modality) where there are two randomly interspersed variations of the stimulus – the so-called oddball paradigm – as shown in Fig. 1b. The characteristics that we aim to reproduce are (i) response onsets that are appropriately delayed with respect to the stimuli, (ii) a damped oscillatory response, (iii) an approximately 4-Hz dominant frequency, and (iv) the occurrence (typically observed following simple stimuli) of negative and positive extrema of similar magnitude, that are not followed by additional extrema. Moreover this should be done in a way that retains the ability to reproduce the ongoing EEG spectrum without any alteration to the model – apart from the form of the driving signal, and parameter modulations that are physiologically plausible.

Section 2 describes the extended model, including expressions for the frequency- and time-domain characteristics of the system. The general characteristics of the modeled spectra and evoked potentials are demonstrated in Sect. 3. The spectral equations, which are generalized in several ways from our previous work, are shown to remain compatible with observations. Various predictions of the ERP equations are then explored, using parameters that are consistent with the EEG spectra. The results are discussed in Sect. 4.

2 Model

This section summarizes the basic model for collective neuronal dynamics developed in our previous work, and extends it to represent the cortex and thalamus in more detail. The intention is to derive a model of the large-scale dynamics of the EEG; and so we employ simplifications that are extreme when compared with some other models, but are expected to be valid in the linear limit and for a large-enough ensemble of neurons.

In the linear limit and for randomly interconnected excitatory and inhibitory neurons, a model for an unbounded homogeneous cortex can be derived with the transfer function (Robinson et al. 2001b)

$$\phi_e(k,\omega) = D_e(k,\omega) \frac{G_{ee}L(\omega)}{1 - G_{ii}L(\omega)} \frac{G_{ee}L(\omega)}{1 - G_{ii}L(\omega)}$$ (1)

The function $\phi_e(r,t)$ is the rate of action potentials arriving at a point $r$ from excitatory neurons elsewhere.

**Fig. 1a,b.** Typical example of EEG spectrum and evoked response potentials (ERPs) obtained from a group of 40 subjects. a EEG spectral power density for subjects who were relaxed with eyes closed: experimental spectrum shown as solid lines, fitter spectrum (described below) shown as dotted lines. b Auditory ERPs in response to frequent, irrelevant (‘background’) tones, and to rare, higher-pitched (‘target’) tones, which are interspersed among the frequent tones. Times are with respect to the stimulus onset, and the peaks are labeled according to the convention for auditory ERPs. Typically ‘N1’, ‘P2’, ‘N2’, and ‘P3’ extrema occur respectively at 100 ms, 160 ms, 210 ms, and 320 ms after the stimulus
in the cortex, and \( \phi_s(r, t) \) is the activity from subcortical neurons.

The remaining functions appearing in (1) are

\[
L(\omega) = (1 - \text{i} \omega / \alpha)^{-1}(1 - \text{i} \omega / \beta)^{-1},
\]

which approximates the low-pass response of the dendritic portion of each neuron [in the time domain the impulse response is \( z(\beta(e^{-zt} - e^{-\beta t})/(\beta - z)) \) for \( \tau = 0 \)], and

\[
D_s(r, \omega) = k^2 r_s^2 + (1 - \text{i} \omega / \gamma_s)^2,
\]

which is obtained from the 2-D damped wave equation that we use to approximate the spread of activity through the cortex (\( r_s \) is the characteristic length for excitatory axons and \( \gamma_s = v / r_s \) is the characteristic damping rate assuming a uniform speed \( v \) of action potentials along axons). In deriving (1) it was assumed that inhibitory axon lengths were short compared with the scale of EEGs, with the result that the inhibitory analog of (3) is \( D_s(r, \omega) = 1 \). The transfer function describing the firing rate at inhibitory synapses is closely related to (1):

\[
\frac{\phi_s(r, \omega)}{\phi_s(r, \omega)} = \frac{G_{es} D_e \phi_e}{G_{es} \phi_e}.
\]

Equations (1) and (4) can be written in terms of more fundamental neuronal parameters, including the average number \( N_{\text{syn}} \) of synapses from neurons of type \( b = \text{e,i,s} \) on neurons of type \( a = \text{e,i,s} \); the sizes \( s_b \) of postsynaptic potentials; and the shape of the sigmoidal relationship between membrane depolarization \( V_a \) and firing rate \( Q_a \).

[The subscripts refer to excitatory (e), inhibitory (i), and subcortical (s) neural populations.] However, these quantities are more conveniently combined into the dimensionless gains \( G_{ab} = \rho_b N_{\text{syn}ab} \rho_a / dQ_a / dV_a \) is the derivative of the assumed sigmoidal response function,

\[
Q_a = \frac{Q_a^{\text{max}}}{1 + \exp[-(\pi / \sqrt{3})(V_a - \theta_a) / \sigma_a]},
\]

evaluated at the steady state of the system (Robinson et al. 1997). Note that the common factors in the expressions for the six gains \( G_{ab} \) mean that they are not independent. Moreover, we assume that the number of interconnections between neuronal types is proportional to the number of available synapses, so the six values \( N_{ab} \) are interdependent. The existence of these implicit interrelations among the gains is the reason that only three appear explicitly in (1).

To make use of (1), some form must be given to the quantity \( \phi_s(r, \omega) \). The signal reaching the cortex from subcortical gray matter is known to be a mixture of sensory signals and feedback activity. The latter involves pathways from the cortex, via the basal ganglia, thalamus, or hippocampal formation, and back to the cortex. Rather than attempt to represent the considerable complexity of these pathways in a functional form, we shall adopt a simple form for \( \phi_s \), consisting of a single sensory source and one feedback loop between thalamus and cortex, and generalize to multiple loops later, if necessary. The details of this simplification follow.

2.1 Thalamic model

The thalamus consists of roughly 15 relay nuclei, within which are excitatory neurons that either relay sensory signals to specific areas of the cortex (primary relay circuits) or feed cortical activity back to the cortex (secondary relay circuits). The latter project with various specificity to the cortex, and can be loosely divided into ‘specific’ and ‘nonspecific’ nuclei. In addition there is the reticular nucleus of the thalamus and the perigeniculate nucleus, which envelop and project inhibitory connections to the relay nuclei, as well as receiving excitatory collaterals from the interconnections between relay nuclei and the cortex. [For brevity in what follows we shall refer only to the thalamic reticular nucleus (TRN) and to thalamic relay nuclei collectively.] Considerable interest surrounds the TRN, as its structure and connections gives it the potential to influence activity throughout the thalamus, and so is believed to have an important role in the gating of cortical inputs during sleep and arousal, and in effecting focal attention. The former relies on general inhibition of the relay nuclei to gate sensory activity reaching the cortex (Yingling and Skinner 1977), while the latter proposal – known as the ‘searchlight hypothesis’ – presumes there to be some mechanism mediated by the TRN by which one portion of the thalamocortical connections is selected, while the remaining parts are inhibited (Crick 1984). Subsequent investigations have identified possible physiological mechanisms postulated by the searchlight hypothesis, but difficulties in interpreting the evidence have prevented identification of which, if any, result in attentional focus. Others have emphasized the role of the feedback loop between cortex and thalamus, where ‘...the cortical influence can either promote or antagonize the relay of information, depending on the state of the dendrites of the reticular neurons’ (Destexhe 2000).

This paper will adopt the general idea that information processing is facilitated by transient modulations in the operating characteristics of the thalamus, which can be represented by varying thalamic gains. We shall consider later the origin of the presumed gain modulations, although we note now the existence of modulatory inputs to the thalamus mediated by acetylcholine and biogenic amines, and that there is evidence that the reticular nucleus is modulated in this way during the generation of ERPs (Steriade et al. 1990; Swick et al. 1994).

Figure 2 depicts the relationship between cortex, reticular nucleus, and a typical specific or secondary relay nucleus. Also shown are the transmission delays inherent in corticothalamic fibers and various gains \( G_{ab} \) defined similarly to cortical gains as the ratio of firing rate changes in target and source neuronal populations, where the changes are with respect to the steady state values. The reticular nucleus is unlike all other thalamic nuclei in that (i) it receives input from all relay nuclei (different
relay nuclei are not directly interconnected, and (ii) its neurons are inhibitory in function (whereas the majority of relay neurons are excitatory). Note that on the basis of known physiology all gains shown are expected to be positive except for \( G_{st} \), which is negative due to the inhibition of specific and secondary relay nuclei by the reticular nucleus.

We assume here that feedback between cortex and thalamus approximately preserves spatial relationships. This is indeed the case for specific relay nuclei (e.g., those reciprocally connected with primary sensory cortex), but less true for secondary relay nuclei (e.g., those connected with association cortex), which can receive convergent activity from the cortex and project divergently to the cortex, thereby filtering out larger wave numbers. In order to model this blurring in the thalamocortical pathways conveniently, we shall permit the driving signal \( \phi_n \) to be spatially nonwhite, although we shall assume it to be temporally white, in common with previous work. Accordingly we find that \( \phi_s(k, \omega) \) is given by

\[
\phi_s(k, \omega) = \mathcal{P} \phi_n + \mathcal{I} \phi_e ,
\]

where \( \mathcal{P} \) is the thalamocortical transfer function and \( \mathcal{I} \) is the corticothalamocortical transfer function, and

\[
\mathcal{P}(\omega) = \frac{L_e G_{tn}}{1 - L_e G_{tn} L_c G_{ra} e^{i \omega t_0 / 2}} ,
\]

\[
\mathcal{I}(\omega) = \frac{L_e G_{tc} + L_c G_{ta} L_e G_{rc} e^{i \omega t_0}}{1 - L_e G_{ta} L_c G_{rc}} .
\]

Equation (6) assumes the noise to arise at or below the level of the thalamus. However, alternative sources might be postulated for the driving signal \( \phi_e \); e.g., the noise might be intrinsic to the cortex, or else it might arise in some cortical region and reach other cortical regions via the thalamus. These options can be accommodated in (6) by replacing the thalamocortical transfer function \( \mathcal{P} \) by 1 or by \( \mathcal{I} \), respectively.

### 2.2 Cortical model

The linear damped wave equation used by us to represent the spread of activity through the cortex corresponds to a distribution of axon ranges proportional to the modified Bessel function of the second kind, \( K_0(|r|/r_s) \) (Robinson et al. 1997). This function is monotonic decreasing with an integrable singularity at the origin, and has been used to approximate the axonal range distribution of excitatory neurons. However, given that the majority of excitatory neurons are thought to have connections limited to the minicolumn in which they reside, it can be argued that the proportion of short-range connections may be even greater than implied by the Bessel function distribution. To accommodate this we should include a second excitatory population to our model, which can be described in the same way as the first, but with axonal ranges \( \sim 0.1 \) mm; i.e., much less than both \( r_s \) and the typical spacing of electrodes. This generalization is important when later modeling ERPs, and we discuss the likely relative contributions of the two populations in Sect. 3.

If we indicate the additional population with the subscript \( l \), then the field equations are

\[
D_e \phi_e = L_e G_{tec} \phi_e + L_c G_{tc} \phi_t + L_c G_{te} \phi_i + L_e G_{te} \phi_s ,
\]

(9)

\[
D_l \phi_l = L_l G_{tli} \phi_e + L_l G_{tc} \phi_t + L_l G_{te} \phi_i + L_l G_{ts} \phi_s ,
\]

(10)

\[
D_l \phi_i = L_l G_{tical} \phi_e + L_l G_{tc} \phi_t + L_l G_{te} \phi_i + L_l G_{si} \phi_s ,
\]

(11)

which describe the dendritic summation and the cortical propagation of the resultant activity for each of the three classes of neurons.

With the help of (6), (9)–(11) can be written as

\[
\begin{bmatrix}
D_e \phi_e - L_e G_{tec} \phi_e - L_c G_{tc} \phi_t - L_c G_{te} \phi_i - L_e G_{te} \phi_s \\
- L_l G_{tli} \phi_e - L_l G_{tc} \phi_t - L_l G_{te} \phi_i - L_l G_{ts} \phi_s \\
- L_l G_{tical} \phi_e - L_l G_{tc} \phi_t - L_l G_{te} \phi_i - L_l G_{si} \phi_s
\end{bmatrix}
= \begin{bmatrix}
\phi_e \\
\phi_l \\
\phi_i
\end{bmatrix}
= \begin{bmatrix}
D_l G_{ls} \\
L_l G_{ls}
\end{bmatrix}
\mathcal{P} \phi_n .
\]

(12)

After solving for \( \phi_e, \phi_l, \) and \( \phi_i \) we then find

\[
\begin{bmatrix}
\phi_e \\
\phi_l \\
\phi_i
\end{bmatrix} = \frac{1}{\det
\begin{bmatrix}
D_l G_{ls} + D_l G_{li} & D_l G_{ls} + D_l G_{li} & D_l G_{ls} + D_l G_{li} \\
D_l G_{ls} + D_l G_{li} & D_l G_{ls} + D_l G_{li} & D_l G_{ls} + D_l G_{li} \\
D_l G_{ls} + D_l G_{li} & D_l G_{ls} + D_l G_{li} & D_l G_{ls} + D_l G_{li}
\end{bmatrix}
- \text{diag}
\begin{bmatrix}
D_l L_l G_{lic} + D_l G_{lic} \phi_t + D_l L_l G_{ls} \\
D_l L_l G_{lic} + D_l G_{lic} \phi_t + D_l L_l G_{ls} \\
D_l L_l G_{lic} + D_l G_{lic} \phi_t + D_l L_l G_{ls}
\end{bmatrix}
\bigg(\begin{bmatrix}
D_l L_l G_{lic} + D_l G_{lic} \phi_t + D_l L_l G_{ls} \\
D_l L_l G_{lic} + D_l G_{lic} \phi_t + D_l L_l G_{ls} \\
D_l L_l G_{lic} + D_l G_{lic} \phi_t + D_l L_l G_{ls}
\end{bmatrix}
\bigg)
\times
\begin{bmatrix}
L_c G_{re} \\
L_l G_{ts}
\end{bmatrix}
\mathcal{P} \phi_n ,
\]

(13)
where
\[ \det = D_c D_1 - (L_c G_{cc} + L_c G_{ca} \mathcal{P}) D_i D_l \]
\[ - D_l L_i G_{il} D_i - D_i D_l L_i G_{ii}, \]
and diag denotes a diagonal matrix with the elements shown. Hence,
\[ \frac{\phi_l}{\phi_n} = \frac{D_i D_l L_i G_{cc} \mathcal{P}}{\text{det}}, \quad (15) \]
\[ \frac{\phi_i}{\phi_n} = \frac{D_i D_l L_i G_{cc} \mathcal{P}}{\text{det}}, \quad (16) \]
\[ \frac{\phi_l}{\phi_n} = \frac{D_i D_l L_i G_{cc} \mathcal{P}}{\text{det}}. \quad (17) \]

These results use simplifications that follow from the random connectivity assumption for cortical neurons: specifically the symmetries \( G_{ab} G_{cd} = G_{ac} G_{bd} \), where \( a, c = e, l, i \) and \( b, d = e, l, i, s \) (Robinson et al. 1997).

From the above equations it can be shown that
\[ \frac{\phi_e}{\phi_n} = \frac{\lambda L_c G_{cc} \mathcal{P}}{k^2 r^2_e + q^2 r^2_e} M, \quad (18) \]
where
\[ \lambda = \frac{1}{1 - L_i G_{ii} - L_i G_{il}}, \quad (19) \]
\[ q^2 r^2_e = (1 - \text{i} \omega / \gamma_a)^2 - \lambda L_c G_{cc} - \lambda L_c G_{ca} \mathcal{P}, \quad (20) \]
\[ M = 1 + \frac{D_i L_i G_{il} D_i (1 - D_i) + D_i D_l L_i G_{ii} (1 - D_i)}{\text{det}}. \quad (21) \]

We note also that the axon lengths \( n \) and \( r_e \) are sufficiently small that \( D_n = k^2 r^2_e + (1 - \text{i} \omega / \gamma_a)^2 \approx 1 \) for \( a = e, l \). The reason is that wavenumbers are limited to \( k_{\text{max}} < 2 \pi / r_e \) by the wide reach of population \( e \) axons, and frequencies are limited to \( \omega_{\text{max}} < \gamma_a \) by dendritic filtering, so that \( k_{\text{max}} \gamma_a \approx 2 \pi r_e / \gamma_a \approx 1 \), and similarly \( \omega_{\text{max}} / \gamma_a \approx 2 \pi r_e / v \approx 1 \). As a result of \( D_i \approx D_i \approx 1, M \approx 1 \) and (18) becomes identical to the corresponding expression used in earlier work, other than replacement throughout of \( L_i G_{ii} \) by \( L_i G_{ii} + L_i G_{il} \). The newly introduced excitatory population (with subscript \( i \)) affects \( \phi_e / \phi_n \) only as far as predictions now depend on this lumped measure of the two short-axon populations. If the dendritic time constants of the two populations are similar, so that \( L_i \approx L_i \), then inverse modeling from real data will not be able to estimate \( G_{ii} \) and \( G_{il} \) separately.

The form of the response in short-axon excitatory neurons \( \phi_i(k, \omega) \) is best viewed in relation to \( \phi_e(k, \omega) \). If we take (15) and (16), we find
\[ \frac{\phi_i}{\phi_n} = \frac{D_i L_i G_{il} \phi_e}{D_i L_i G_{cc} \phi_n}. \quad (22) \]

Then
\[ \frac{\phi_i}{\phi_n} = \frac{\lambda L_i G_{il} \mathcal{P} M \left( k^2 r^2_e + (1 - \text{i} \omega / \gamma_a)^2 \right)}{k^2 r^2_e + q^2 r^2_e}, \quad (23) \]
\[ = \frac{\lambda L_i G_{il} \mathcal{P} M}{D_i} \left[ 1 + \frac{\lambda L_i G_{cc} + \lambda L_i G_{ca} \mathcal{P}}{k^2 r^2_e + q^2 r^2_e} \right], \quad (24) \]
so,
\[ \frac{\phi_i}{\phi_n} = \frac{\lambda L_i G_{il} \mathcal{P} M}{D_i} \phi_n + \frac{\lambda L_i G_{il} \mathcal{P} M}{D_i} \phi_e + \frac{\lambda L_i G_{il} \mathcal{P} M}{D_i} \phi_e \phi_n. \quad (25) \]

Equation (26) is obtained from the preceding equation by relations implied by the random connectivity assumption (Robinson et al. 1997). The corresponding result for \( \phi_i(k, \omega) \) can simply be obtained by exchanging the subscripts \( i \) and \( e \). As a result, the remainder of this section will deal only with the \( e \) and \( l \) populations, and the latter can be thought of as referring either to local excitatory or local inhibitory neurons.

### 2.3 Spectra

In deriving the expected form of EEG spectra we assume that the scalp potential is proportional to some linear combination of the synaptic firing rates, as discussed below in Sect. 3.1. Hence \( |V(\omega)|^2 = |w_e \phi_e + w_l \phi_l + w_i \phi_i|^2 \), where the weights \( w_i \) depend on many (constant) neuronal and volume conduction factors. Here we shall consider just the two functions, \( |\phi_e|^2 \) and \( |\phi_l|^2 \), since \( |\phi_i|^2 \) can be obtained trivially from \( |\phi_i|^2 \); the various cross terms are intermediate in character between the two principal cases, and absolute estimates are not needed in the present context.

From (18) and (22),
\[ |\phi_e(k, \omega)|^2 = \left| \frac{\lambda L_i G_{cc} \mathcal{P}}{k^2 r^2_e + q^2 r^2_e} M |\phi_n|^2 \right|^2, \quad (27) \]
\[ |\phi_l(k, \omega)|^2 = \left| \frac{D_i L_i G_{il}}{D_i L_i G_{cc}} |\phi_e|^2 \right|^2. \quad (28) \]

To obtain expressions for the conventional spectral measures \( |\phi_e(\omega)|^2 \) and \( |\phi_l(\omega)|^2 \) we commonly assume that \( \phi_n \) is spatially white. However, thalamocortical axons individually and collectively have nonzero spread within the cortex, and so it is more realistic to assume a limit on the range of wavenumbers reaching the cortex. We shall assume
\[ |\phi_n(k, \omega)|^2 = (r_e^2 / 2 \pi) \exp(-k^2 r_e^2 / 2)|\phi_n(\omega)|^2, \quad (29) \]
which limits wavenumbers to \( k \leq 2 \pi / r_e \). The spatial component of \( |\phi_n(k, \omega)| \) is the Fourier transform of a 2-D Gaussian distribution with standard deviation \( r_e \), and \( |\phi_n(k, \omega)|^2 \) satisfies the condition \( \int |\phi_n(k, \omega)|^2 d^2 k = |\phi_n(\omega)|^2 \). Then
\[ |\phi_1(\omega)|^2 = \frac{1}{2\pi} \int \frac{|M|^2 |\phi_1(k, \omega)|^2}{|k^2 r_c^2 + q^2 r_c^2|} \, dk , \]

\[ \approx |\phi_n(\omega)|^2 \frac{1}{2\pi} \int \frac{|M|^2 e^{-k^2 r_c^2/2}}{|k^2 r_c^2 + q^2 r_c^2|} \, dk , \]

\[ = |\phi_n(\omega)|^2 \frac{1}{2\pi} \int \frac{-\kappa}{\text{Im} q^2 r_c^2} \times \text{Im} \left( e^{q^2 r_c^2 E_1(q^2 r_c^2 \kappa)} \right) , \]

where we have introduced \( \kappa = r_c^2 / 2 \) as the measure of the wavenumber content of the driving signal, and \( E_1 \) is an exponential integral (Abramowitz and Stegun 1965). Equation (32) follows from the integrand's effective wave number cutoff \( \sim 2\pi/r_c \) and frequency cutoff \( \sim x \), for which \( |M|^2 \approx 1 \). If \( r_c \ll r_e \) (the driving signal is spatially nearly white) then (32) simplifies to

\[ |\phi_1(\omega)|^2 = \frac{|\phi_n(\omega)|^2}{\frac{1}{2\pi} \int \frac{|M|^2 |\phi_1(k, \omega)|^2}{|k^2 r_c^2 + q^2 r_c^2|} \, dk} \]

\[ \approx |\phi_n(\omega)|^2 \frac{1}{2\pi} \int \frac{-\kappa}{\text{Im} q^2 r_c^2} \times \text{Im} \left( e^{q^2 r_c^2 E_1(q^2 r_c^2 \kappa)} \right) . \]

The equivalent expression for \( \phi_1 \) is

\[ |\phi_1(\omega)|^2 = \frac{|\phi_n(\omega)|^2}{\frac{1}{2\pi} \int \frac{|M|^2 |\phi_1(k, \omega)|^2}{|k^2 r_c^2 + q^2 r_c^2|} \, dk} \]

\[ \times \int \frac{d^2 k}{2\pi} \frac{|M|^2 |\phi_1(k, \omega)|^2}{|k^2 r_c^2 + q^2 r_c^2|} \, dk \]

\[ \approx |\phi_n(\omega)|^2 \frac{1}{2\pi} \int \frac{-\kappa}{\text{Im} q^2 r_c^2} \times \text{Im} \left( e^{q^2 r_c^2 E_1(q^2 r_c^2 \kappa)} \right) . \]

In this case the presence of \( D_e = k^2 r_c^2 + (1 - i\omega/\gamma_e)^2 \) in the numerator means that the integrand is not low-pass limited to \( k \lesssim 2\pi/r_e \) as in (32). Instead the integrand has an effective high-frequency cutoff in wave number due to the Gaussian envelope of the noise function, and in frequency due to the dendritic filter functions \( L_a \).

Provided the spatial and temporal frequencies are limited to \( \ll 2\pi/r_1 \) and \( \ll \omega/\gamma_1 \), respectively, then both \( D_t \approx 1 \) and \( |M|^2 \approx 1 \), and thus

\[ |\phi_1(\omega)|^2 \approx |\phi_n(\omega)|^2 \frac{1}{2\pi} \int \frac{|M|^2 |\phi_1(k, \omega)|^2}{|k^2 r_c^2 + q^2 r_c^2|} \, dk \]

\[ \times \left[ 1 - \frac{\kappa}{\text{Im} q^2 r_c^2} \text{Im} \left( q^2 r_c^2 e^{q^2 r_c^2 E_1(q^2 r_c^2 \kappa)} \right) \right] \]

\[ + \frac{2(1 - \kappa \omega^2/\gamma_e^2)}{\text{Im} q^2 r_c^2} \text{Im} \left( q^2 r_c^2 e^{q^2 r_c^2 E_1(q^2 r_c^2 \kappa)} \right) \]

\[ - \frac{(1 + \kappa \omega^2/\gamma_e^2)^2}{\text{Im} q^2 r_c^2} \text{Im} \left( e^{q^2 r_c^2 E_1(q^2 r_c^2 \kappa)} \right) \right] . \]

The corresponding expression for \( |\phi_1(\omega)|^2 \) can be obtained by exchanging the subscripts \( l \) and \( i \).

### 2.4 Evoked response potentials

For the purposes of modeling ERPs we shall take the same model as was discussed in the context of spectra, and replace incoming white noise with an impulsive stimulus. The driving signal will be taken as a pulse of limited spatiotemporal extent,

\[ \phi_e(k, \omega) = e^{-k^2 r_c^2/4} e^{-\omega^2 r_c^2/2} , \]

which is normalized in coordinate space to have unit integral and standard deviations of \( r_e \) spatially and \( t_e \) temporally. This expression has the advantage of being more general and realistic than \( \delta \)-function alternatives. The spatial component of \( \phi_e \) means that the impulse produces a cortical activation that has the form of a 2-D Gaussian, the extent of which \( (r_e) \) is really a property of the divergent sensory thalamocortical pathways, but it is convenient to lump it with the temporal distribution of the driving signal. In the derivations that follow, \( t_e \) will be assumed to be \( \sim 10 \) ms and of the same order as \( 1/\gamma_e \), \( 1/\beta_e \), but much greater than \( 1/\gamma_{Le} \) or \( 1/\gamma_{Se} \). The appropriate spatial extent \( r_e \) of the stimulus is less definite, but is plausibly \( > 10 \) mm (observed to be the smallest extent of cortex that can be activated by a stimulus) and certainly \( \ll 500 \) mm (the circumference of the brain). These bounds on allowed wavenumbers and frequencies permit the approximations \( M \approx D_t \approx 1 \) to be made below.

With the above assumptions, the response \( \phi_e(R, t) \) at some distance \( R \) from the center of the impulse is

\[ \phi_e(R, t) = \int_0^\infty \frac{dk}{2\pi} J_0(kR) \times \text{F}^{-1} \left\{ \frac{\phi_e}{\phi_n} e^{-k^2 r_c^2/4} e^{-\omega^2 r_c^2/2} \right\} , \]

\[ \approx \int_0^\infty \frac{dk}{2\pi} J_0(kR) \times \text{F}^{-1} \left\{ \lambda L_e \phi_n \frac{e^{-k^2 r_c^2/4} e^{-\omega^2 r_c^2/2}}{k^2 r_c^2 + q^2 r_c^2} \right\} , \]

where \( \text{F}^{-1} \) represents the inverse temporal Fourier transform, and \( M \approx D_t \approx 1 \) is assumed.

For the same stimulus, \( \phi_1(k, \omega) \) is

\[ \phi_1(k, \omega) = \frac{D_t L_e \phi_n}{D_t L_e \phi_n} \frac{e^{-k^2 r_c^2/4} e^{-\omega^2 r_c^2/2}}{\text{F}^{-1}} \]

and like (26) this can be written as the sum of three terms, the latter two of which are closely related to \( \phi_e \):

\[ \phi_1(k, \omega) = \lambda L_e \phi_n \text{F}^{-1} \left\{ \frac{e^{-k^2 r_c^2/4} e^{-\omega^2 r_c^2/2}}{k^2 r_c^2 + q^2 r_c^2} \right\} \]

\[ + \frac{\lambda L_e \phi_n}{D_t} \phi_e + \frac{\lambda L_e \phi_n}{D_t} \text{F}^{-1} \phi_e . \]

The three terms in (40) have different spatial and temporal characteristics, which we shall consider in
turn. Assuming, for reasons given above, that $M = D_h = 1$, the first term can be spatially inverse transformed giving
\[
\frac{1}{\pi^2 r_s^2} e^{-r^2/r_s^2} \mathcal{F}^{-1} \left\{ J L G_0 \mathcal{F} e^{-a^2 r_s^2/2} \right\} .
\] (41)

This term is limited spatially to $R \leq r_s$, and has no terms relating to long-axon neurons. Physically, it is the initial response of short-axon neurons to subcortical activation.

The second term in (40) is the component of $\phi_t$ that is due to direct driving by $\phi_c$. The factor $L_t G_{0c}$ describes the filtering and gain experienced by a signal passing through a local neuron; the factor $\lambda = 1/(1 - L_t G_{0i} - L_t G_i)$ describes the modulation of $\phi_c$ by the short-range neural populations, and is of a form that is commonly encountered in systems where there is additive filtered feedback. Only corticocortical connections are involved in this term. The presence of the term $L_t$ means that a delay of order $1/\lambda + 1/\beta$ with respect to $\phi_c$ can be expected.

The last term in (40) can also be considered to be due to $\phi_c$, but in this case includes $s$, equal to the corticothalamic transfer function $\phi_c/\phi_s$. Consequently this term describes feedback of $\phi_c$ from the cortex, through the corticothalamic loop, to short-axon neurons in the cortex. There is a delay of order $1/\lambda + 1/\beta$, as with the preceding term, but there is an additional delay $\delta_0$ due to the finite conduction speed in the corticothalamic loop (see Eq. 8), as well as the delay within the thalamus.

### 3 Characteristics of the model

This section compares the theory of Sect. 2 with experimental data. In order to obtain representative parameter values for numerical examples in this section, (33) was first fitted to the experimental spectrum shown in Fig. 1a, which yielded the values listed in Table 1. Note that in general the equations allow only certain compound gains to be inferred: $G_{0c} G_{ec}$ is the net gain of the direct corticothalamic feedback; $G_{0c} G_{ec} G_{ec}$ is the net corticothalamic feedback via the TRN; and $G_{0c} G_{rs}$ is the net gain between the relay nuclei and TRN.

#### 3.1 Spectra

We have discussed fitting of an earlier version of the model to experimental spectra elsewhere (Robinson et al. 2001b), so here we shall cover only the variations due to the additional features of the current version of the model. They include (i) the site of the noise generator, (ii) the effect of spatial filtering of subcortical feedback, (iii) resonances internal to the thalamus, and (iv) the relative contributions of $\phi_c$, $\phi_t$, and $\phi_s$ to spectra recorded from the scalp.

It was noted in connection with (6) that the generator of noise might be hypothesized to lie in various parts of the brain, and that the principal alternatives to the version in Fig. 2 could both be accommodated by a suitable redefinition of $\mathcal{F}$. Anatomically the choice is unclear, since the reticular activating system provides a continuous driving signal to both the thalamus and cortex; but at the same time there is spontaneous firing within the cortex which could also be seen as a driving signal. Mathematically, the consequences of these alternatives are simply understood since the spectral expressions (32) and (35) are proportional to $|\mathcal{F}(\omega)|^2$. When $\mathcal{F}$ is defined by (7), corresponding to noise entering at the thalamus, its magnitude is asymptotically proportional to $\omega^{-2}$; as is the case when $\mathcal{F} \rightarrow \mathcal{F}$ (white noise arriving at the cortex via the corticothalamic pathway from elsewhere in the cortex), and in contrast to the case when $\mathcal{F} \rightarrow 1$ (noise arriving unfiltered at the cortex). Given sufficiently noise-free data it should be possible to identify which of the three possibilities best resembles experimental data through their different cut-off characteristics and asymptotic behavior. However, the parameters $\lambda$, $\beta$, $\gamma_c$, $\eta_1$, and $\eta_2$ all also influence the spectral shape, and we find in practice that the noisy experimental spectrum in Fig. 1a can be equally well fitted with all three source alternatives, when suitable adjustments are made to the rate constants. Furthermore, independent estimates of the five rate constants have uncertainty factors of $\sim 2$, and the critical high-frequency asymptote is obscured by biological and instrumental noise, and so none of the three alternatives can be confidently eliminated at present. Table 1 was derived by assuming $\mathcal{F}$ to be given by (7), corresponding to a noise source projecting to the thalamus.

The second of the issues listed above – the effect of spatial filtering – arose in Sect. 2.3 from the desire to employ a thalamocortical signal with a realistic spatial spectrum. Limiting of the spatial frequency content of this signal could be intrinsic to the noise source or be imposed by convergence and divergence within the corticothalamic feedback loops discussed in Sect. 2.1. The contribution of the different wave numbers has been

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_c$</td>
<td>142</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>36</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>730</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$\eta_1$</td>
<td>51</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$\eta_2$</td>
<td>1000</td>
<td>s$^{-1}$</td>
</tr>
<tr>
<td>$G_0 + G_i$</td>
<td>-5.9</td>
<td></td>
</tr>
<tr>
<td>$G_{ec}$</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>$G_{0c} G_{ec}$</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>$G_{0c} G_{ec} G_{ec}$</td>
<td>-0.98</td>
<td></td>
</tr>
<tr>
<td>$G_{0c} G_{rs}$</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>$\delta_0$</td>
<td>0.08</td>
<td>s</td>
</tr>
<tr>
<td>$\nu$</td>
<td>10</td>
<td>m s$^{-1}$</td>
</tr>
</tbody>
</table>
examined elsewhere in the context of modal analysis, where it was found that boundary effects are negligible except when strong resonances occur and the waves are very weakly damped (Robinson et al. 2001a). Here we shall look instead at varying the wave number cutoff parameter, \( r_s \). We take the circumference of the brain as an upper limit to the range of wavelengths to consider. As a lower limit, we again use \( \sim 10 \text{ mm} \), the smallest cortical extent that can be activated by a stimulus. Taking these values, those in Table 1, and (32) and (35), we can calculate numerical estimates of the expected spectra, as seen in Fig. 3.

The asymptotic behavior and characteristics of the spectral peaks of \(|\phi_s|^2\) were discussed by Robinson et al. (2001b) for the spatially white noise case, \( r_s \rightarrow 0 \). In the same limit the constant term in (35) tends to dominate, resulting in a featureless spectrum for \(|\phi_s|^2\) described mainly by the term in front of the integral. In the other limit \( r_s \rightarrow \infty \), corresponding to spatially uniform noise, \(|\phi_s|^2\) and \(|\phi_s|^2\) differ only by a factor \(|1 - io/\gamma_s|^2\), so are similar at frequencies \(< \gamma_s \approx 20 \text{ Hz} \). Other than near peaks, both \(|\phi_s|^2\) and \(|\phi_s|^2\) are steeper when \( r_s \) is large than when \( r_s \rightarrow 0 \). In practice we find that experimental spectra generally resemble \(|\phi_s|^2\) for \( r_s \approx 0 \), but the same spectra might also be due to \(|\phi_s|^2\) with an intermediate value of \( r_s \). As such it is difficult to infer the value of \( r_s \) from spectral fits alone.

Concerning the third of the issues listed above, it is of interest to note that the thalamocortical transfer function \( \varphi \), which appears in both (32) and (35), has a dispersion relation \( 1 - L_s G_s L_r G_r = 0 \). This yields a fourth-order polynomial that can be solved for \( \omega \) in terms of the thalamic rate constants and gains \( G_s \) (a negative quantity) and \( G_r \). We find

\[
\omega = -i \left( \frac{\eta_1 + \eta_2}{2} \right) \pm \sqrt{ \left( \frac{\eta_1 - \eta_2}{2} \right)^2 + \eta_1 \eta_2 |G_s/G_r| },
\]

and the two roots that are least damped become linearly unstable, at a frequency \( \sqrt{\eta_1 \eta_2} \) when \( G_s/G_r = (\eta_1 + \eta_2)/(\eta_1 \eta_2) \sim 10 \). The instability is due to undamped feedback between the two elements of the thalamus, and so is unrelated to the corticothalamic resonance responsible for the peaks in Fig. 3. The resonant frequency \( \sqrt{\eta_1 \eta_2} \approx 135 \text{ s}^{-1} \approx 20 \text{ Hz} \) is in the range of practical interest, so is potentially relevant if the intrathalamic gains are sufficiently strong.

A final issue is the modeled variables \( \phi_e \), \( \phi_i \), and \( \phi \), relate to the potentials measured on the scalp. Each impulse arriving at a synapse triggers a depolarization or hyperpolarization of the dendritic membrane, which largely attenuates with a time constant \( \sim 2 \text{ ms} \) (Koch 1999, p. 75). The variables \( \phi_e \), \( \phi_i \), and \( \phi \) are therefore good proxies for the extracellular currents induced by synaptic action. Also, we take these signals to be approximately proportional to their corresponding scalp potentials, under the assumption that many of complications diminish on large, slow scales typical of EEGs. However, this simplification is contingent upon further investigations in the area of anatomy and modeling of extracellular potentials.

This problem of how to express the EEG potential explicitly in terms of the synaptic firing rates arises (i) from the difficulty of accurately representing the volume conduction of extracellular currents, and (ii) from the fact that the contribution of each \( \phi \) needs to be weighted by a morphological factor, which depends on the number and geometric distribution of synapses of each type on the many different morphological categories of neurons. Cell morphology matters, since spherically symmetric activation of a neuron will produce less current at distant (e.g., scalp) sites than asymmetric activations. Pyramidal neurons are elongated, aligned, and numerous, and so their activation by \( \phi_e \), \( \phi_i \), and \( \phi \) is likely to be the main contribution to EEGs. Ideally we would like to calculate the power spectrum for some linear combination of these signals, and fit that to experimental data. Unfortunately current physiological knowledge of neuronal interconnections does not allow estimation of the relative contributions of \( \phi_e \) and \( \phi_i \) [\( \phi_e \] is likely to be less important since the proportion of inhibitory connections is known to be small (Liley and Wright 1994)]. In the absence of firm evidence to the contrary, we shall tentatively adopt the values in Table 1, obtained by fitting (33) to the spectrum in Fig. 1a. In particular, they will be used in the following discussion of ERPs.

### 3.2 Evoked response potentials

The basic equations describing impulse responses of the system depicted in Fig. 2 are (38) and (40). Introducing an impulse into the system, assuming parameters from Table 1, results in the responses in Fig. 4, which are unlike observed ERPs: both \( \phi_e(R,t) \) and \( \phi_i(R,t) \) show evidence of a damped resonance at the alpha frequency that is superimposed on an upward baseline shift: cf. Fig. 1b.

However, as mentioned in Sect. 2.1 and as described by other authors (Destexhe 2000; Steriade et al. 1990; Swick et al. 1994), there are physiological reasons for expecting thalamic gain modulations after stimuli.

---

**Fig. 3a,b.** Modeled spectral power density functions (arbitrary units). Spatial and temporal white noise is assumed, except that the spatial wave numbers of the noise are limited by a Gaussian function \((r_s^2/2\pi) \exp(-k^2r_s^2/2)\). Spectra are shown for three values of \( r_s \), and other parameters are from Table 1. a \(|\phi_s(\omega)|^2\) given by (32); b \(|\phi_s(\omega)|^2\) given by (35)

This problem of how to express the EEG potential explicitly in terms of the synaptic firing rates arises (i) from the difficulty of accurately representing the volume conduction of extracellular currents, and (ii) from the fact that the contribution of each \( \phi \) needs to be weighted by a morphological factor, which depends on the number and geometric distribution of synapses of each type on the many different morphological categories of neurons. Cell morphology matters, since spherically symmetric activation of a neuron will produce less current at distant (e.g., scalp) sites than asymmetric activations. Pyramidal neurons are elongated, aligned, and numerous, and so their activation by \( \phi_e \), \( \phi_i \), and \( \phi \) is likely to be the main contribution to EEGs. Ideally we would like to calculate the power spectrum for some linear combination of these signals, and fit that to experimental data. Unfortunately current physiological knowledge of neuronal interconnections does not allow estimation of the relative contributions of \( \phi_e \) and \( \phi_i \) [\( \phi_e \] is likely to be less important since the proportion of inhibitory connections is known to be small (Liley and Wright 1994)]. In the absence of firm evidence to the contrary, we shall tentatively adopt the values in Table 1, obtained by fitting (33) to the spectrum in Fig. 1a. In particular, they will be used in the following discussion of ERPs.
Following the implications of the searchlight hypothesis (Crick 1984), where focal enhancement and general inhibition of cortex is postulated to underlie attention, we shall hypothesize that after a stimulus, relay gains are altered to induce a positive corticothalamic feedback loop of a more focal kind. (The pathways and physiological mechanisms supporting this hypothesis are discussed in Sect. 4.) Specifically, if the positive feedback is so focal that throughout most of the cortex direct corticothalamic (i.e., positive) feedback is suppressed \( G_{cf}G_{tf} \rightarrow 0 \), then we find that the impulse responses are as shown in Fig. 5a and b. (It is also assumed that the thalamic modulations take effect at any time prior to the arrival at the thalamus of the impulse response from the cortex; i.e., \( \leq 100 \) ms poststimulus.) By comparing Fig. 5a and b with Fig. 4a and b it is apparent that both \( \phi_{t}(R, t) \) and \( \phi_{e}(R, t) \) are affected by these modulations of thalamic gain. Most notable are the elimination of the \( \sim 100 \) ms periodicity and the prolonged upward offset.

The appearance in Fig. 5b of a minimum at \( \sim 0.2 \) s is of particular relevance to ERP modeling. It can be understood by considering the three components of \( \phi_{t}(R, t) \) plotted in Fig. 5c-e, and corresponding to the three terms in (40). Component 1 of \( \phi_{t}(R, t) \) contributes little to the total because it is limited in extent to the stimulated area; component 2 is more widespread and is delayed by cortical propagation; and component 3 is like component 2 except that it is further delayed by the corticothalamic feedback delay and inverted by the negative feedback character of the corticothalamic loop \( G_{ce}G_{te}G_{ec} \approx -0.98 \). The contribution of component 1 is further limited by the fact that thalamocortical connections are largely to spiny stellate neurons in the sensory cortex, which are not expected to give rise to scalp potentials due to their isotropic dendritic structure. As a result of initial propagating excitation and the delayed negative feedback, \( \phi_{t}(R, t) \) [and to a much lesser extent \( \phi_{e}(R, t) \)] have a biphasic appearance like that of a basic ERP, such as in Fig. 1b. Another similarity between the modeled response in Fig. 1b and experimental ERPs is that both typically consist of extrema of similar size and opposite sign. This is not the case for the resonant system implied by Table 1 and demonstrated in Fig. 4, but occurs robustly in this model once thalamic modulation is introduced, as above.

Figure 5 demonstrates the basic character of impulse responses expected from our model. However, there are several additional details that will now be considered: (i) the location of the impulse source, (ii) the extent of the area stimulated by the driving signal, and (iii) size of the relative contributions to the observed response from each of the three neuronal populations.

We note that sensory stimuli affect the thalamus broadly: by direct input, via collaterals, and via the brainstem reticular system. Also the N1 ERP component is generated almost simultaneously in the central and secondary auditory cortices (Näätänen and Picton 1987), suggesting a subcortical source triggered at about the time the sensory stimulus reaches the thalamus. Accordingly, in Fig. 5 and in all subsequent figures, we assume an impulse to be produced within the thalamic

---

**Fig. 4a-b.** Firing rates – a \( \phi_{e}(R, t) \), and b \( \phi_{i}(R, t) \) – for parameters from Table 1, expressed as deviations from the steady state values. The stimulus occurs at \( t_{onset} = 0.03 \) s and is Gaussian in form, with a temporal standard deviation of \( 0.02 \) s and a spatial standard deviation of \( r_{s} = 0.12r_{c} = 0.01 \) m at the cortex. Responses are plotted for four distances from the focus of the stimulus: \( R = 0.5r_{c} \) (thick line), \( r_{c} \), \( 1.5r_{c} \), and \( 2r_{c} \). Vertical units are arbitrary. Increases in \( \phi \) are shown as negative values, in accordance with the expected changes in scalp potential when \( \phi_{e} \) acts on the apical dendrites of pyramidal cells, or when \( \phi_{t} \) acts on the basal dendrites or the cell body.

**Fig. 5a-e.** Responses \( \phi_{e}(R, t) \) a and \( \phi_{i}(R, t) \) b. The three components of \( \phi_{t}(R, t) \) in (40) are plotted separately as c-e. Vertical units are arbitrary, and each figure shows responses at four distances from the focus of the stimulus: \( R = 0.5r_{c} \) (thick line), \( r_{c} \), \( 1.5r_{c} \), and \( 2r_{c} \). As in Fig. 4, Parameters are as in Table 1 and Fig. 4. Note the diminishing size and increasing delay of the peaks as the distance \( R \) from the stimulus increases, as expected from the finite axonal conduction speed, \( v \approx 10 \) m s\(^{-1} \). In e all responses are essentially flat except for the case \( R = 0.5r_{c} \).
relay nuclei at ~30 ms poststimulus. Regarding the size of the cortical area stimulated by the driving signal, more anatomically detailed modeling of the pathways may help to resolve this uncertainty, but is arguably small due to the fine functional structure of the primary sensory cortex (specifically, $r_c = 10$ mm is assumed). The third matter mentioned above relates to microscopic anatomy, as $\phi_e$, $\phi_i$, and $\phi_t$ all contribute to the scalp potential, but with unknown proportions that depend on the number, morphology, and connectivity of the various classes of neurons. Dominance of $\phi_t$ or $\phi_t$ is favored by the present evidence since $\phi_t(R,t)$ does not have the required biphasic character of real ERPs. Regarding the relative significance of $\phi_e$ and $\phi_e$, we note that if inhibitory synapses are concentrated near the base of pyramidal cells then they are not as effective at generating current dipoles. Accordingly, in the following figures only $\phi_e(t)$ will be shown.

Although Fig. 5b is a fair approximation to the background ERP shown in Fig. 1b, the N2 and P3 extrem a observed in target ERPs call for additional hypotheses. In generating Fig. 5 we simply assumed that the majority of the cortex is purely under the influence of negative corticothalamic feedback. In the case of target stimuli, however, we propose that the area of cortex subject to thalamic inhibition will not be so extensive since the stimulus is more significant. Such stimuli will require more perceptual and memory resources compared with background stimuli, and so the cortical area needing to be activated (the breadth of the spotlight) will be correspondingly larger. Activation of a greater area of cortex could be achieved by partitioning the cortex and thalamus into two parts: one for which the corticothalamic loop gain is positive (activated), and the remainder for which the loop gain is negative. This is facilitated by the compartmentalization of the thalamus and corticothalamic pathways (Alexander et al. 1990).

In our model, we can approximate background and target responses by supposing that $G_{eG_{cG_e}} \rightarrow 0$ following a background stimulus (i.e., little positive feedback), but that this parameter increases according to stimulus significance. Figure 6 shows the result of fits to experimental data when most parameters are held constant, and Table 2 lists the values of the remaining parameters. The inferred values of $G_{eG_{cG_e}}$ support this principles just described. It is notable that the remaining model parameters are largely identical to those obtained separately from a fit of our model to the subjects’ average EEG spectra. Modulation of $G_{eG_{cG_e}}$ is also expected, since the extent and degree of inhibition is also likely to be a function of stimulus significance. The difference in the value of $t_0$ is not implied by the above proposals, but nor is it surprising given that several corticothalamic loops have been distinguished on anatomical and functional grounds (Alexander et al. 1990), and that slower loops might be activated in the case of significant stimuli. Note that overall normalization is arbitrary since we have no way of knowing the relationship between firing rates and scalp potential, however it was constrained to be the same for both the ERP fits since we expect the relationship to be constant.

The additional upward-going component of the target response, Fig. 6b, which would be labeled ‘N2’ in ERP literature, can be attributed to the initial cortical signal passing through the positive feedback loop from cortex to secondary relay nuclei and back to cortex, with a gain parameterized by $G_{eG_{cG_e}}$. The time difference between the N1 and N2 peaks is closely related to the delay in the corticothalamic loop, $t_0$. As a result, background ERP is reproduced simply through thalamic gain modulations, and the N1–P2–N2 sequence of peaks in the target ERP is similarly reproduced by our model, although the final P3 peak is distorted. The characteristics of the model will be discussed further in Sect. 4, as will possible reasons for the remaining discrepancies in Fig. 6a and b between the model and experimental ERPs.

### Table 2. Thalamic parameters corresponding to the spectral fit in Fig. 1a and ERP fits in Fig. 6. The three fits are distinguished only by the parameters listed here: the remainder are as in Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>EEG</th>
<th>Background ERP</th>
<th>Target ERP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_{eG_{cG_e}}$</td>
<td>6.9</td>
<td>0.0</td>
<td>2.9</td>
</tr>
<tr>
<td>$G_{eG_{cG_e}}$</td>
<td>$-1.0$</td>
<td>$-1.0$</td>
<td>$-3.7$</td>
</tr>
<tr>
<td>$t_0$ (s)</td>
<td>0.08</td>
<td>0.08</td>
<td>0.10</td>
</tr>
</tbody>
</table>

The additional upward-going component of the target response, Fig. 6b, which would be labeled ‘N2’ in ERP literature, can be attributed to the initial cortical signal passing through the positive feedback loop from cortex to secondary relay nuclei and back to cortex, with a gain parameterized by $G_{eG_{cG_e}}$. The time difference between the N1 and N2 peaks is closely related to the delay in the corticothalamic loop, $t_0$. As a result, background ERP is reproduced simply through thalamic gain modulations, and the N1–P2–N2 sequence of peaks in the target ERP is similarly reproduced by our model, although the final P3 peak is distorted. The characteristics of the model will be discussed further in Sect. 4, as will possible reasons for the remaining discrepancies in Fig. 6a and b between the model and experimental ERPs.

### 4 Discussion

In this paper we have introduced two new structural elements to our previous model – a more realistic representation of the thalamus and an additional population of cortical neurons – and raised the question of the location of the driving signal. Each of these elaborations contributes to the results in Sect. 3.

Looking first at spectral estimates of EEG, many of the consequences of delayed corticothalamic feedback have been explored in a simpler thalamic model by Robinson et al. (2001b). The more detailed effects of the present thalamic representation, as well as questions of the location of the noise source and the relative contri-
butions of the different cortical populations, are found to have relatively minor consequences. These refinements are the subject of current investigation and are not critical to the present ERP results, so they will not be discussed further here.

The effects of neural population (i.e., axon range) and various thalamic gains manifest clearly in impulse responses, which we identify here with ERPs. We see from Fig. 5 that $\phi(R,t)$ has a rapidly damped response that is little affected by the corticothalamic feedback. The characteristic time constant $r_c/v$ for cortical propagation is small ($\approx 10$ ms), so the response is heavily damped, although this also depends on the gain of the feedback loop. This is in contrast to the case of $\phi_1(R,t)$, for which analytical expressions are naturally represented as multiple components. Thus the first two components of $\phi_1(R,t)$ in (40) are closely related to the initial impulse reaching the cortex, and are nearly coincident. The third component includes the factor $\sqrt{f}$ which is the corticothalamocortical transfer function defined by (8), and allows both for positive and negative corticothalamic feedback with the parameters $G_{se}G_{re}$ and $G_{se}G_{re}$, respectively. These parameters are critical to reproducing ERPs.

Figure 7 demonstrates many of the possible responses of the system in terms of $G_{sa}G_{se}$ and $G_{sa}G_{re}$. We see that the main effect of $G_{se}G_{re}$, the strength of the positive corticothalamic feedback via secondary thalamic relay nuclei, is to modulate the amplitude of the peak at $\approx 0.2$ s. Since the neurons in this path are excitatory, the $\approx 0.2$-s peak has the same sign as the $\approx 0.1$-s peak, and the separation of these two peaks is approximately $t_0$ plus the delay $1/\tau_1 + 1/\tau_2$, occurring in the thalamus. The value $G_{se}G_{re}$ quantifies the strength of the alternative corticothalamic pathway. This pathway involves an additional nucleus (the reticular nucleus) consisting of inhibitory neurons, and so results in a trough at a time that is slightly delayed with respect to the $\approx 0.2$-s peak.

On the basis of the pattern of results in Fig. 7 we expect the values $G_{se}G_{re} = 6.9$ and $G_{se}G_{re} = -1.0$ from Table 1 to result in multiple periodic maxima in $\phi_1$, which is indeed similar to the waveforms in Fig. 4. Based on the searchlight hypothesis, it is proposed that both $G_{sa}G_{se}$ and $G_{sa}G_{re}$ may be modulated following a stimulus. We find that responses resembling backgrounds occur near the left of Fig. 7 and have P2-like components that can largely be attributed to the negative feedback loop through the reticular nucleus, as shown explicitly in Fig. 5e. In target-like responses – roughly the middle of Fig. 7 – the same minimum is present, but the increased positive feedback causes an N2-like component at about $0.2$ s and the emergence of two separate minima, which are suggestive of the P2 and P3 components of target responses. The responses at the bottom left of Fig. 7 hint at the emergence of a theta-like rhythm for sufficiently strong negative feedback, while a combination of a baseline shift and alpha rhythm occurs for strong positive feedback (upper right). Robinson et al. (2001b) showed, in a delayed corticothalamic feedback model resembling the present one, that spectra are dominated by $\approx 5$ Hz when feedback is strongly negative, and by frequencies near zero and $\approx 10$ Hz when feedback is strongly positive, in agreement with limiting cases of the present model. They also considered the linear stability of such systems.

In summary, according to this model all ERP components are the result of a single impulse occurring in a system whose spectral characteristics can also be matched with an identical model and largely identical parameters. The main types of response (alpha and theta rhythms, and background and target ERPs) are seen as resulting from modulations of the gains of the relay and reticular nuclei of the thalamus.

Figure 8 shows the effects of variations to individual parameters other than $G_{sa}G_{se}$ and $G_{sa}G_{re}$. The rate constants $G_{sa}$, $G_{se}$, and $G_{re}$ generally have little effect on the impulse responses. The cortical gains $G_{sa} + G_{sa}$ and $G_{se}$ simply scale the response magnitude, and variations in the corticothalamic loop delay $t_0$ affect the separation of the components. Variations in the stimulus parameters $\alpha_{in}$, $L_{in}$, and $r_{in}$ offset the response or alter the shape of the response in ways that do not affect its overall character.

As mentioned in Sect. 3.1, $G_{sa}G_{re}$ becomes relevant when $G_{sa}G_{re}$ approaches a critical value. Near this limit, a new unstable resonance occurs between the secondary relay and reticular nuclei, whose properties have been examined in detail by Lopes da Silva and coworkers (e.g., Lopes da Silva 1991). The resonance has a characteristic frequency close to $\sqrt{\eta_1\eta_2} \approx 130$ s$^{-1}$ $\approx 20$ Hz, which is distinct in mechanism and frequency from the alpha and theta rhythms in Fig. 7. The gain $G_{sa}G_{re}$ is constrained by physiology to
be negative, so the other limit is a value of zero. At this extreme (dotted line in Fig. 8l) the response has a deepened and prolonged trough, which may used to account for the discrepancy observed in Fig. 6b.

Another possibility for improving the quality of the fits in Fig. 6a and b is suggested by Fig. 9. The parameters of the partially successful fit in Fig. 6b fail to reproduce the depth of the P3 trough, however if the thalamic gains are increased by about 40%, then this failure can be overcome, leaving only the latency difference to overcome. (A model in which target stimuli activate longer and slower corticothalamic feedback loops would have the required effect.) The same strengthening of thalamic gains also results in low amplitude ~10 Hz oscillations near 0.4 s, which may help in fitting the later part of the background response, where a similar oscillation is observed. However these refinements are hard to investigate with our present analytical methods, and call for a numerical approach.

The model deals only with one effective corticothalamo-cortical pathway, and identifies separately modulated positive and negative feedback loops as critical to ERP generation. However it is of interest to note the similar proposal by Houk (1995) of a positive feedback loop linking the cortex, basal ganglia, and thalamus, but which assumes a negative feedback character in response to novel situations (cf. our ‘target’ stimuli). Likewise Kropotov and Etlinger (1999) have independently proposed a similar circuit based on extensive intracerebral recordings. Moreover, the existence of multiple corticosubcortical loops has been pointed out by Alexander et al. (1990). Consequently we see the present results as having relevance beyond the particular case shown in Fig. 2, since they are easily generalized to include loops of these alternative types.

It is difficult to provide a definitive argument for a single specific physiological mechanism for thalamic gain modulation, since there are several modulatory influences acting on the thalamus and cortex. One candidate is noradrenaline from the locus ceruleus, which has long been known to be associated with tonic arousal (Moruzzi and Magoun 1949) and with transient attention shifts (Pineda et al. 1987). Another is acetylcholine, which is released from the parabrachial region of the brainstem, and constitutes the principal input to the thalamus from the brainstem (Sherman and Koch 1998). Acetylcholine has been shown experimentally and computationally to be the likely cause of the radical changes in the character of the thalamus, where there is modulation of the efficacy of the TRN leading to switching between slow wave and spindle sleep states (Destexhe 2000). The effect of both these metabotropic modulators is to alter the effective gains connecting the thalamic elements, and consequently alter the resonant properties of the system. The phenomenon of alpha blocking might be reflecting these shifts in system parameters. Additional factors contributing to the effective gain...
values are the spatial coherence (Destexhe 2000) and temporal synchrony (Sillitto et al. 1994) of action potentials.

It is interesting to note that gain modulation has long been invoked in the neuropsychological literature. The searchlight hypothesis is one example; another is the proposal by Pribram and McGuinness (1975) for corticothalamic loops to mediate arousal and activation.

Modulation of gain could also occur as the result of the intensity of firing rate arriving at thalamic nuclei. The present results were obtained by linearizing the neuronal response function (5) so that \( G_{ab} \) are all constant over time, but for sufficiently large impulses the static nonlinearity of this function will cause each \( G_{ab} \) to become (to a first-order approximation) proportional to the instantaneous average depolarization within the nucleus. These temporal gain modulations are in addition to those due to neuromodulators. It is of interest to note that Jansen and coworkers used this static sigmoidal nonlinearity to model a relationship between pre-stimulus alpha phase and ERP amplitude (Jansen and Rit 1995; Jansen et al. 1993), and so this mechanism may be an alternative explanation for event-related gain modulations.

## 5 Conclusion

We have described a new version of our continuum model of the EEG that has a more realistic representation of thalamic anatomy, and reflects the broad classes of cortical neurons more accurately. The continuum approximation permits large-scale dynamics to be modeled, and complements the network style of modeling. As with previous versions of this model, analytical expressions for the expected EEG spectrum could be obtained, which are suitable for parameter estimation through inverse modeling. The spectrum of the long-range excitatory activation field \( \phi_e \) is identical to that from our earlier model, while that of the newly introduced short-range activation field \( \phi_i \) is qualitatively similar, although it is a weaker function of frequency. When model parameters were obtained from experimental spectral data, and combined with thalamic gain modulations consistent with psychological theories relating to the orienting response, the model was shown to result in impulse responses \( \phi_i(t) \) that resemble ERPs. The critical parameters that differed from those obtained from the same subjects’ EEGs were gains within the TRN and the thalamic relay nuclei. ERPs resembling typical ‘background’ responses were obtained whenever corticothalamic feedback became predominantly negative, while ‘target’ tones could be modeled by retaining some degree of positive feedback. These were argued to be the same changes as those expected from the principles contained in the ‘searchlight hypothesis’ (Crick 1984). The modeled ERPs satisfy the criteria that the response onsets are correct, the responses are transient rather than being sinusoidal, they have the appropriate sequence of positive- and negative-going sequence of extrema, and the majority of parameters are identical to those obtained from fitting the corresponding model of EEG spectra to experimental data.

The proposed sequence of events following a stimulus is as follows (see Fig. 10):

1. An impulse, consisting of the sensory signal and associated activity due to collaterals, arrives at the thalamus 20–30 ms after a stimulus.

2. The thalamic response is relayed to the appropriate part of the cortex, resulting in the first negative (N1) peak. Activation of pyramidal neurons with long axons is propagated through the cortex; activation of locally connected pyramidal neurons is not propagated beyond the area of the initial cortical activation. The latter population reinforces the initial cortical excitation, helping to create a synchronized cortical activation, which is in part directed at the thalamus. This second volley to the thalamus reaches it ~100 ms after the first.

3a. By a mechanism probably involving neuromodulators or thalamocortical collaterals, the initial volley to the thalamus also widely suppresses the normal feedback route between the cortex and the thalamic relay nuclei. This alters the transfer characteristics of the thalamus such that the positive feedback is largely removed, while the negative feedback route (via the TRN) remains. The second cortical volley is thus inverted with respect to the N1 component, resulting in the P2 ERP component, and is delayed by ~100 ms.

![Fig. 10a–d. Evoked potentials (arbitrary units) corresponding to four combinations of positive (\( G_{es}, G_{je} \)) and negative (\( G_{es}, G_{je} \)) corticothalamic feedback strengths, illustrating the four distinct states described in the text: a \( G_{es}, G_{je} \to 0 \), and \( G_{es}, G_{je} \) unchanged from resting value; b \( G_{es}, G_{je} \approx |G_{es}, G_{je}|; c \( G_{es}, G_{je} \approx |G_{es}, G_{je}|; d \ G_{es}, G_{je} < |G_{es}, G_{je}| > \text{resting value. In each case the initial thalamic and cortical responses (solid lines) are plotted separately from the responses resulting from corticothalamic feedback (dotted lines), and are labeled as in the text. The thalamic responses are plotted below the cortical responses, and so precede the latter by an amount approximately equal to the thalamocortical conduction delay.}
3b. If the significance of the stimulus requires a greater portion of the relay nuclei to be disinhibited, then the second volley will be a mixture of a positive and negative feedback. The negative feedback component is delayed and broader due to having passed through the TRN, and so the mixed response is asymmetric, resulting in the N2 and P3 ERP components.
3c. States of nonfocal attention have positive feedback characteristics \((G_{\alpha}G_{\theta} \gg G_{\alpha}G_{\beta}G_{\theta})\), which are associated with an alpha-like resonance.
3d. States of focal attention may have isolated areas with positive corticothalamic feedback, but generally the system will have negative feedback characteristics \((G_{\alpha}G_{\theta} \ll G_{\alpha}G_{\beta}G_{\theta})\), which are associated with a theta-like resonance.

This scheme accounts for ERPs in terms of thalamic gain changes having durations at least as long as that of ERPs, and shows that these gain changes result in impulse responses similar to those observed. It also is compatible with general EEG distinctions between alpha and theta states. Nevertheless, there is a large body of work on the phenomenology of ERPs which has scarcely been addressed in the present work: there are distinctions that can be drawn between different kinds of P3, prolonged negativities, ERPs obtained during sleep, ERP topography, and responses to missing (but expected) stimuli (Naatanen and Picton 1987). To improve the general applicability of this model we need to make explicit how the limbic and neuromodulator systems influence the critical thalamocortical gains identified above. It would also be desirable to incorporate more information about cell morphology and connectivity in order to clarify the extent to which \(\phi_{s}, \phi_{t}, \) and \(\phi_{l}\) contribute to scalp potentials.

Further refinements to the model will improve the quantitative predictions of both EEG spectra and ERP waveforms, and will be reinforced by performing fits to a range of normal, clinical, and experimental data. But even with the current version we have demonstrated qualitative matches for both EEGs and ERPs using an identical underlying model, and parameters that are physiologically plausible. This encourages us to believe that much of the phenomenology of ERPs might be explained by a simple linear model of wave propagation through the cortex and subcortical structures.

References


Moruzzi G, Magoun HW (1949) Brain stem reticular formation and activation of the EEG. Electroencephalogr Clin Neurophysiol 1: 455–473
Pribram KH, McGuinness D (1975) Arousal, activation, and effort in the control of attention. Psychol Rev 82: 116–149

WORK TOWARD A THEORY OF BRAIN FUNCTION  I  395


Toward an integrated continuum model of cerebral dynamics: the cerebral rhythms, synchronous oscillation and cortical stability


Abstract

Continuum models of cerebral cortex with parameters derived from physiological data, provide explanations of the cerebral rhythms, synchronous oscillation, and autonomous cortical activity in the gamma frequency range, and suggest possible mechanisms for dynamic self-organization in the brain. Dispersion relations and derivations of power spectral response for the models, show that a low frequency resonant mode and associated travelling wave solutions of the models’ equations of state can account for the predominant $1/f$ spectral content of the electroencephalogram (EEG). Large scale activity in the alpha, beta, and gamma bands, is accounted for by thalamocortical interaction, under regulation by diffuse cortical excitation. System impulse responses can be used to model Event-Related Potentials. Further classes of local resonance may be generated by rapid negative feedbacks at active synapses. Activity in the gamma band around 40 Hz, associated with large amplitude oscillations of pulse density, appears at higher levels of cortical activation, and is unstable unless compensated by synaptic feedbacks. Control of cortical stability by synaptic feedbacks offers a partial account of the regulation of autonomous activity within the cortex. Synchronous oscillation occurs between concurrently excited cortical sites, and can be explained by analysis of wave motion radiating from each of the co-active sites. These models are suitable for the introduction of learning rules—most notably the coherent infomax rule. © 2001 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Brain dynamics; Cortical continuum models; Spectral content of electroencephalogram; Synchronous oscillation; Coherent infomax; Cortical information processing

1. Introduction

The operation of the brain requires the coordinated interplay of billions of neurones via their synapo-dendritic couplings. The development of a concise mathematical description of this inter-
play is a major goal of neuroscience, but attempts to attain this goal encounter problems of a fundamental nature.

What are the essential cellular properties needed to account for the local cell pulse characteristics, and the macroscopic fields (notably the electroencephalogram, EEG) emitted by the working brain? How are the observable pulses and fields related to information processing in the brain? Which are the essential features of the brain’s gross anatomy which need to be taken into consideration? How do the dynamics of neurons interact with growth and other plastic modifications, to enable adaptive learning?

Accounts of interactions within populations of neurons have generally utilized neural network methods, in which the elementary units considered are the neurons (e.g. Amit et al., 1990; Arbib et al., 1998; Traub et al., 1996; Whittington et al., 1995; Lumer et al., 1997a,b; Wilson and Bower, 1991). Despite their many virtues, such models are limited by the rapid increase in their numerical complexity as the scale and detail of simulation is increased. This complexity also generally makes them unsuited to mathematical analysis.

This paper describes the current attempts being made by our group to account for the dynamics of the brain in as simple a way as possible. We utilize continuum modelling methods pioneered in the works of Freeman (1975), Wilson and Cowan (1973), Nunez (1981, 1995), van Rotterdam et al., 1982 and subsequent), Haken (e.g. Jirsa and Haken, 1996), Zhadin (1994) and others. Our work has been guided also by the following principles:

- We have attempted to fit all simulations and analyses within a simplified conception of the brain’s overall organization, with a view to achieving logically compatible models of brain dynamics at all spatial scales.
- We assume that at microscopic scale neuronal interactions are highly nonlinear and discrete, yet at macroscopic scale observable fields such as the EEG emanate from a stochastic and essentially linear continuum. (Wright, 1990; Wright and Liley, 1996). This means that well established linear methods can be used to aid analysis of properties often initially demonstrated by numerical simulations.
- Parameters (dendritic time constants, synaptic densities, etc.) compatible with independent physiological estimates (e.g. Braitenberg and Schuz, 1991; Thomson et al., 1996; Thomson, 1997; Liley and Wright, 1994; Rennie et al., 1999b) are used to constrain model fitting to experimental data.

2. A simplified concept of brain organization

Fig. 1 shows the basic aspects of the brain’s operation we seek to analyze.

Three principal scales of organization need to be considered.

Firstly, at the microscopic scale in any locale of cortex, local interactions between a mixed population of excitatory and inhibitory cells take place. Interactions are largely mediated by fast neurotransmitters. The existence of synaptic and dendritic delays, and synapto-dendritic feedbacks, means that local oscillations of mean dendritic potential and pulse density are likely in response

Fig. 1. The top left panel of Fig. 1 shows two representative cells within the cerebral cortex—an excitatory and an inhibitory cell. Populations of these cells are linked together densely in the cerebral cortex. Top right panel shows the gamma band local oscillation which emerges when the locale of cells becomes sufficiently excited. The middle panels of Fig. 1 show how at a larger scale, these foci of excited cortex generate waves of cortical electrical activity spreading into the less excited surrounding cortical tissue (middle left panel). The resulting wave activity can be analyzed by cross-correlation, as shown in the middle right panel. Here the lag time for maximally correlated activity (with reference to the recording site shown in the left middle panel) is displayed for the extended field. It is seen that the foci of activity have entered ‘synchronous oscillation’. The lower panels in Fig. 1 show the overall brain, and EEG activity as generated in simulations—from low frequency ‘theta’ activity, through the alpha, beta and gamma ranges, to 40 Hz. These progressive changes in frequency content reflect the overall level of cortical excitation. At the highest levels of excitation, the self-excited cortical state described in the top panels has been reached.
Fig. 1.
to any perturbation. The properties of such local oscillation determine the signal transfer characteristics and attractor dynamics exhibited by local neural networks.

Secondly, at a scale from fractions of a millimeter to many centimeters of cortex, patches of active cells have been experimentally observed to enter into synchronous oscillation. That is, cross-correlations of pulse density, or of mean local field potential at the separated locii are maximal at zero lag. This phenomena has been widely, although controversially, considered to act as a substrate for association processes in the cortex. (e.g. Eckhorn et al., 1988; Singer, 1994; Singer and Gray, 1995; Stryker, 1989; Bressler et al., 1993; Livingstone, 1996; Mitmier et al., 1999; Neuenschwander and Singer, 1996; Palm and Wennekers, 1997; Steriade et al., 1996; Gray and Singer, 1989; Gray et al., 1989). Such interactions depend particularly upon excitatory cortico-cortical fibres of medium to large scale.

Thirdly, at the scale of the whole brain, the total cortical system interacts with the subcortical systems. These interactions include reciprocal interactions of thalamus and cortex, and limbic and cortical projections to the basal ganglia and elsewhere. These descending pathways ultimately exert return controls via the reticular arousal system and related fibres, back upon the cortex (e.g. Steriade et al., 1990; Alexander et al., 1990; Posner and Petersen, 1990). By these means the cortex is capable of self-control of both its mean level of activity (cortical arousal) and spatial pattern of activity (an aspect of attention), as indirectly observed experimentally (e.g. Gevins et al., 1983; Munk et al., 1996; John et al., 1969; Walter et al., 1967; Xu et al., 1997).

3. State equations

Fig. 2 summarizes the elementary physiological properties considered basic in our recent models, and their mathematical expression (Robinson et al., 1997, 1998a,b; Rennie et al., 1999a,b; Wright, 1999; Wright et al., 2000). A field of mixed excitatory and inhibitory neurones are considered as a continuum approximating the cortex in two dimensions. The control parameter for cortical activity is the diffuse excitation delivered by the reticular formation of the brain stem.

The response of dendritic membranes to synaptic inputs is nonlinear, because of the influences of reversal potentials, Rqp. Thus the gross postsynaptic potential (PSP size) is a function of afferent synaptic inputs, modified by reversal potentials and further synaptic feedback processes, which may be included in the term H(t). The time-course of average dendritic potentials (the PSP shape, as measured at the neuronal soma) is modelled by a biexponential function with fixed time constants, alpha and beta. Mean soma potential, V, (considered linearly proportional to the population local field potential) is related to mean efferent pulse rate, Q, by a sigmoidal function which takes its form from the distribution of cell action potential thresholds. The propagation of action potentials in the continuum of surrounding neurones is expressed as a linear wave equation, which relates afferent and efferent pulses in time and space, and a spatial damping term, gamma. Gamma is determined by the velocity of axonal
conduction and the geometry of axonal projections.

Inhibitory cells we have for the most part considered as wholly local in their action, with long range cortical interactions occurring by excitatory connections.

To these basic state equations we have subsequently added consideration of co-resonance of cortex with the thalamic system (Robinson et al., 2000), and more complex feedback processes operating at synaptic level, as will be discussed in later sections.

The basic state equations enable direct calculation of steady state fixed points for the system, dispersion relations for the medium, and the power spectrum of local field potential (EEG) in response to white noise inputs. They also permit simulation of an extended field of cortex, enabling inspection of time series and spatial properties of activity in the medium.

4. Parameters

Table 1 lists the basic parameters of the state equations, with current values applied in our work. Values for these parameters have been obtained largely from the work of Braitenberg and Schuz, Thomson et al., Liley and Wright, and Rennie et al., cited in Section 1.

There is some sensitivity of emergent system properties, even with parameter variation within the approximate physiological range. To aid in melding cellular physiological parameters with experimental observations on EEG, we have partly relied upon initial numerical simulations with a non-dimensional model (e.g. Wright, 1999) and partly upon analytical explorations of the basic state equations (Robinson et al., 1998a,b; Rennie et al., 1999a). Parameters have been subsequently set with some ‘tuning’ within the physiological

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_e = N_p$</td>
<td>Dendritic synapses from pyramids</td>
<td>4120</td>
</tr>
<tr>
<td>$N_i = N_p$</td>
<td>Dendritic synapses from interneurons</td>
<td>800</td>
</tr>
<tr>
<td>$S_{ep}^{00} = S_{ep}^{00}$</td>
<td>EPSP size due to $\phi_e$ when $V_e = V_e^{00}$</td>
<td>$2.4 \times 10^{-6}$ V s</td>
</tr>
<tr>
<td>$S_{ip}^{00} = S_{ip}^{00}$</td>
<td>IPSP size due to $\phi_i$ when $V_i = V_i^{00}$</td>
<td>$-5.9 \times 10^{-6}$ V s</td>
</tr>
<tr>
<td>$V_c^{00} = V_c^{00}$</td>
<td>Potential at which $S_{ep}^{00}$ is estimated</td>
<td>$2.4 \times 10^{-6}$ V s</td>
</tr>
<tr>
<td>$V_c^{00}$</td>
<td>Rest potential ($\phi_e = \phi_i = 0$)</td>
<td>$-0.060$ V</td>
</tr>
<tr>
<td>$V_c^{00}$</td>
<td>Reversal pot. for AMPA channels</td>
<td>0 V</td>
</tr>
<tr>
<td>$V_c^{00}$</td>
<td>Reversal pot. for GABA A channels</td>
<td>$-0.070$ V</td>
</tr>
<tr>
<td>$\alpha_e$</td>
<td>EPSP decay rate in pyramids</td>
<td>68 s$^{-1}$</td>
</tr>
<tr>
<td>$\alpha_i$</td>
<td>IPSP decay rate in pyramids</td>
<td>47 s$^{-1}$</td>
</tr>
<tr>
<td>$\alpha_o$</td>
<td>EPSP decay rate in interneurons</td>
<td>176 s$^{-1}$</td>
</tr>
<tr>
<td>$\alpha_i$</td>
<td>IPSP decay rate in interneurons</td>
<td>82 s$^{-1}$</td>
</tr>
<tr>
<td>$\beta_{eq}$</td>
<td>Rate of rise of PSPs in pyramids</td>
<td>500 s$^{-1}$</td>
</tr>
<tr>
<td>$\beta_{eq}$</td>
<td>Rate of rise of PSPs in interneurons</td>
<td>500 s$^{-1}$</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Cut-off frequency for feedback</td>
<td>200 s$^{-1}$</td>
</tr>
<tr>
<td>$Q_{max}$</td>
<td>Maximal pyramidal firing rates</td>
<td>100 s$^{-1}$</td>
</tr>
<tr>
<td>$Q_{max}$</td>
<td>Maximal interneuron firing rates</td>
<td>200 s$^{-1}$</td>
</tr>
<tr>
<td>$\sigma_x$</td>
<td>Sigmoid width parameter</td>
<td>0.005 V</td>
</tr>
<tr>
<td>$\theta_{x}$</td>
<td>Firing threshold</td>
<td>$-0.052$ V</td>
</tr>
<tr>
<td>$\phi_x$</td>
<td>Firing rate of subcortical neurons</td>
<td>10 s$^{-1}$</td>
</tr>
<tr>
<td>$\gamma_e$</td>
<td>Excitatory spatial damping rate</td>
<td>80 s$^{-1}$</td>
</tr>
<tr>
<td>$\gamma_i$</td>
<td>Inhibitory spatial damping rate</td>
<td>10$^5$ s$^{-1}$</td>
</tr>
</tbody>
</table>
Fig. 3. Roots of dispersion relations for the cortical model with parameters in Table 1. See text for further description. Left hand graph—roots with relatively low cortical activation. Right hand graph—system roots with high cortical activation.

These two roots give rise to spectral content of $1/f$ character, and activity associated with a sharp spectral peak in the gamma range (around 40 Hz $\pm$ ), respectively.

The occurrence of a potentially unstable root in the gamma range immediately suggests a basis for autonomous cortical activity at this frequency range. This is in keeping with the many experiments on synchronous oscillation cited in Section 2, which reveal powerful local pulse and local field potential oscillations around this frequency, during processing of inputs.

Obviously, sustained unstable activity as illustrated in Fig. 3, is unphysiological. For this reason we are presently modifying the basic state equations to include negative feedbacks at synaptic junctions. Table 2 lists some of the more prominent of the factors known to modify synaptic neurotransmission. Of particular importance in this context are the fast feedbacks, operating in the 1–5 ms time range. These complement the fast negative feedback actions induced by membrane reversal potentials. Probably the dominant fast effect of this type arises from the post-synaptic activation and deactivation of glutamate receptors (e.g. Tones and Westbrook, 1996).

When fast negative feedbacks are introduced to the basic state equations, two major influences on system dynamics are apparent.

Firstly, synaptic feedbacks can introduce local resonances. Fig. 4 shows time series and power spectra associated with increasing cortical activation. Similar patterns of resonance also arise when strong negative feedback is present at excitatory synapses. It can be seen that as cortical activation increases (down Fig. 4) EEG-like activity appears which exhibits multiple peaks in the power spectrum, progressively moving to higher frequencies, approximately imitating activity in the theta, alpha, beta, and gamma bands. At sufficiently high cortical activation runaway unstable activity, in the 40 Hz range, appears. These results were obtained from our non-dimensional model (Wright, 1999). The values of the parameters in this version have been tuned to optimize the resemblance to EEG. Although preliminary, these results serve to show that rapid synaptic range, so as to better fit observed EEG frequencies, etc (Rennie et al., 1999b).

5. Microscopic dynamics

Fig. 3 shows the roots of dispersion relations calculated from our basic state equations, with the parameter values listed in Table 1 (Rennie et al., 1999b). Dispersion relations encapsulate the system’s favoured modes of resonance when driven by white noise inputs, and thus act as an analytic adjunct to numerical simulations. Resonance occurs at specific combinations of frequency and wavenumber. Frequency of resonance (omega) is measured by the horizontal distance of a root from the origin. Temporal damping is estimated by the value on the imaginary axis. Shown here are only those roots which are relatively lightly damped. For each root the squares show the resonance associated with a wavenumber ($k$) of zero, and the tails indicate the root values as wavenumber increases. The two cases shown are those obtained for relatively low (10/s) and relatively high (350/s) input pulse rates of cortical activation delivered by the non-specific reticular activation system.

Two relatively lightly damped roots are apparent — a low frequency resonance at $k = 0$, which is also associated with a travelling wave system, and a high frequency resonance which becomes less damped as cortical activation is increased, until it becomes unstable, as indicated by the root’s placement above the horizontal (real) axis.
Fig. 4. Time series and power spectra of a simulation with parameters approximating (but not identical to) those represented in Fig. 3, and with the addition of strong synaptic feedbacks. Cortical activation (associated with the control parameter Qns) increases as the page is descended.
feedbacks enhance local resonances not apparent as lightly damped roots in the basic state equations. Thus resonances in the EEG range, including resonances near the frequencies of the great cerebral rhythms, may arise locally within the cortex, as a consequence of synaptic feedbacks. As will be discussed in Section 7 the major cerebral rhythms can be quantitatively well accounted for at macroscopic scale, by cortico-thalamic mechanisms. There is thus no need to appeal to local cortical low-frequency resonances, to account for macroscopic EEG in the sub-gamma range. But co-resonance of microscopic and macroscopic neuronal fields is an implied mechanism for interaction between scales, in our models.

Secondly, sufficiently strong negative feedback operating at excitatory synapses, or positive feedback at inhibitory synapses, or various combinations, can act as a time-varying control upon otherwise unstable resonance in the gamma range. With appropriate choice of the time constants of these feedbacks, autonomous generation of activity develops, yet remains locally controlled at pulse rates within the physiological range.

Table 2
An incomplete list of synaptic feedback mechanisms

<table>
<thead>
<tr>
<th>Fast effects (1–5 ms)</th>
<th>Medium effects (10–100 ms)</th>
<th>Slow effects (300+ ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Excitatory Neurotransmission</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>· Activation of AMPA receptors</td>
<td>· Recovery of AMPA receptors from desensitization</td>
<td>· Desensitization of NMDA receptors</td>
</tr>
<tr>
<td>· Removal of glutamate from synapse</td>
<td>· Allosteric modulation of AMPA receptors by intracellular polamines</td>
<td>· Recovery of kainate and NMDA receptors from desensitization</td>
</tr>
<tr>
<td>· Dissociation of glutamate from AMPA receptors and closure of ion channels (deactivation)</td>
<td>· Activation of NMDA receptors (requiring co-activation by glycine and removal of magnesium, from ion channel)</td>
<td>· Activation of pre- and post-synaptic metabotropic glutamate receptors</td>
</tr>
<tr>
<td>· Desensitization of AMPA receptors</td>
<td>· Deactivation of NMDA receptors</td>
<td>· Activation of most monamine receptors</td>
</tr>
<tr>
<td>· Activation, deactivation and desensitization of kainate receptors</td>
<td>· Allosteric modulation of NMDA receptors by polyamine, zinc, nitric oxide, pH, redox state, neurosteroids, and possible endogenous ligand acting at the PCP binding site</td>
<td>· Enzymic modification of receptors (e.g. phosphorylation)</td>
</tr>
<tr>
<td>· Allosteric modulation of AMPA receptors by endogenous ligands (not yet identified) acting on ampakine, benzothiadiazine and 2,3-benzodiazepine binding sites</td>
<td>· Dissociation of GABA from receptor and deactivation</td>
<td>· Declustering of receptors at synapses</td>
</tr>
<tr>
<td>· Allosteric modulation of GABA&lt;sub&gt;A&lt;/sub&gt; receptors by endogenous inverse agonists (not yet identified)</td>
<td>· Desensitization of GABA&lt;sub&gt;A&lt;/sub&gt; receptors</td>
<td>· Activation of pre- and post-synaptic metabotropic GABA&lt;sub&gt;B&lt;/sub&gt; receptors</td>
</tr>
<tr>
<td><strong>Inhibitory Neurotransmission</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>· Activation of GABA&lt;sub&gt;A&lt;/sub&gt; receptors</td>
<td>· Recovery of GABA&lt;sub&gt;A&lt;/sub&gt; receptors from desensitization</td>
<td>· Activation of most monamine receptors</td>
</tr>
<tr>
<td>· Removal of GABA from synapse</td>
<td>· Allosteric modulation of GABA&lt;sub&gt;A&lt;/sub&gt; receptors by endogenous ligands (not yet identified) acting on neurosteroid, 1,4-benzodiazepine and barbiturate binding sites</td>
<td></td>
</tr>
</tbody>
</table>

404 | WORK TOWARD A THEORY OF BRAIN FUNCTION
Fig. 5. Freeze-frame snapshots of surface electrocortical potential from a simulation operating at high cortical activation, under the constraint of powerful, fast, synaptic feedbacks. Frames are separated by approximately 10 ms, and scales are normalized units. The upper, folded, surfaces show the spatial conformation of local field potentials, with increasing membrane depolarization associated with the upward direction. The shading on this surface is present merely to enhance the visual form of the contour. The underlying flat surfaces show concurrent pulse density, and here the darker shading indicates higher pulse density, associated with greater mean membrane depolarization. Other than an initiating impulse of spatio-temporal white noise, there is no random driving. The activity is self-sustaining.

To illustrate the impact of this feedback control, Fig. 5 shows snapshots from a small scale cortical simulation which includes first approximations to these strong synaptic feedbacks. The continuous emergent activity invites comparison to the dynamics within the gamma range of spon-
taneously firing pools of neurones, in which bursting patterns of action potentials are observed. Chaotic and limit cycle types of activity may be present within various basins of attraction, but as of yet little work on their classification has been undertaken. In contrast, at lower levels of cortical activation, only point attractors are present.

6. Synchronous oscillation — mesoscopic interactions at longer range

During the development of these simulations it became apparent that synchronous oscillation, with close analogy to that observed physiologically appeared as a property of the simulations without any additional assumptions (Wright, 1997a). We have since developed analytical and numerical treatments of zero-lag cross correlation (Robinson et al., 1998a; Chapman et al., 2001), and matched the phenomenon we observe to some physiological results (Wright et al., 2000). The following account of synchronous oscillation applies to all frequencies, but, as is the case in physiological experiments, large oscillations in the gamma band, at about 40 Hz, are associated with the largest amplitudes, and occur when the cortex is excited by higher levels of input.

If two points on the simulated cortical surface are driven by separate, completely uncorrelated, inputs (or if uncorrelated autonomous local cortical activity emerges at two separated points) then, within a few milliseconds cross-correlated activity, maximal at zero lag, appears in the neighbourhood of both active sites. For this effect to occur, the two points must be relatively strongly coupled, and axonal delay must be small compared to the rise and fall time of the dendritic response. The speed of onset depends mainly on the axonal delay. This effect does not depend upon system non-linearity, as is frequently assumed by those familiar with stochastic resonance in other situations. It is instead associated with linear wave transmission in the simulated cortex.

Fig. 6 sums up our current understanding of the mechanism of synchronous oscillation. In the top frames of Fig. 6 it can be seen how a large field of zero-lag synchrony surrounds both sites of uncorrelated input. In the middle frames eigenmode decomposition shows that the first eigenmode of the wave activity radiating from both driven sites is predominant, and this eigenmode defines the field of synchrony. The bottom frames attempt to convey the essence of the physical process, and can be explained as follows:

The uncorrelated inputs can be decomposed into their even and odd components — roughly, the parts of each driving signal which are in phase with the other, and those parts of reversed phase. Wave activity radiates from both sites as outwardly propagating, dissipative, linear waves, creating a small local field of ’self-synchrony’ about each site, by the summation of signals transmitted by pathways of similar lag. The travelling waves obey simple superposition rules subject to dissipation — but interactions of afferent signals at summing junctions (dendrites) in the surrounding field dissipate even and odd components in their inputs selectively. Even components reinforce one another, while odd components tend to cancel about the signal mean. Thus, in the field of both driven sites, the activity induced by the even components in the driving signals dominates, and a field of zero-lag correlated activity emerges as the first eigenmode of the wave activity. It can be shown that the magnitude of this zero-lag correlated field is sensitive to dendritic delay time, axonal conduction lag between sites, and the relative strengths of couplings in the field.

Fig. 7 gives an example of how the results produced in simulation can be made to qualitatively reproduce those seen in physiological experiments on synchronous oscillation (Wright et al., 2000). Cross-correlation of pulses from a pair of recording sites is greatest when the stimulus is a single large bar, next for two smaller bars moving in the same direction, and least for smaller bars moving in opposite directions. We believe the match to experimental data might be rendered quantitatively precise if relative coupling strengths, coherence within signals in the input pathways, and strengths of overlying noise, were more closely matched between simulation and experiment.
Fig. 6. A simulated cortical field driven by uncorrelated white noise inputs at two sites on the cortical surface. Top: Sites of input shown by white squares. Cross-correlations and delays, with respect to the reference site at the black square. Middle: first and second eigenmodes of the travelling waves radiating out from the sites of input. Bottom: A schematic representation of the way in which the first and second eigenmodes arise from addition of even components, and cancellation of odd components, in the extended cortical field.
7. The macroscopic EEG field. A thalamo-cortical resonance model

When spectral responses of the simulations applied in Figs. 4 and 5 are compared to the spectral content of the macroscopic EEG, the simulations prove to be inadequate with regard to both relative amplitude and width of spectral peaks. A purely cortical account ignores the role long ascribed to thalamic sources in gen-

Fig. 7. Simulation of local field potential time series, power spectral content, and cross-correlations of two sites in the cortical receptive field, stimulated by combinations of moving bars in the visual field. Input signals within a bar are spatially coherent, while the signals input to separate bars are uncorrelated. Signal/noise ratio (point amplitude of bar driving signals / point amplitude of incoherent background noise) = 300.
Fig. 8. Left hand figure: EEG power spectral average for 50 subjects. Upper set of dotted curves for eyes closed, lower set for eyes open (scaled for clarity). Fits to model shown as solid lines. Parameter values shown on graph: $t_0$ is the conduction delay in the thalamocortical loop; $n$ is the number of neurons in this loop (here set to zero to lump corticothalamic delays into the single parameter $t_0$). Gamma and alpha are delay parameters related to the equivalent terms in Table 1. Fits were achieved by varying only the fraction of thalamocortical feedback and the general level of cortical gain, including cortical activation. Right-hand figure: Complex loci of the wavenumber-independent part of the spectral content attributable to the thalamo-cortical resonance alone. Solid line eyes closed—dashed, eyes open.

It turns out that the addition of a simple feedback loop to the basic state equations, while retaining parameters close to those listed in Table 1, gives a good account of the entire macroscopic EEG spectrum over low to high levels of cortical activation (Robinson et al., 2000). The loop time required is consistent with an interaction between cortex and a subcortical system (assumed to be the thalamus) with a synaptic-dendritic delay appropriate for 1–2 synapses. The background $1/f$ spectral content of EEG at low frequencies, with a ‘knee’ around 15 Hz above which power decays as $1/f^5$, is well accounted for by the low frequency root system, similar to that seen in Fig. 3. The thalamocortical resonance then accounts for the occurrence of spectral peaks at the alpha, beta, and gamma ranges, as shown in Fig. 8. These resonances are associated with low wavenumbers and can be thought of as global resonances, as opposed to the local resonances associated with fast synaptic feedback described in Section 5.

We have found that the spectral progression with increasing cortical activation is well fitted by this model, for all experimental data obtained over a wide variety of states of cortical activation, from sleep to high alerting.

Finally it can be observed that the thalamocortical spectral model also permits the simulation of the cortical event related potential (ERP), by treating the ERP as the cortical impulse response to brief sensory stimuli, with subsequent modulations imposed by the cortex upon the cortico-thalamic resonance (Rennie et al., 2001). Examples
8. Some wider implications

The classes of dynamics seen in our models suggest the prospect of a future unification of...
8.1. ‘Edge of chaos’ and universal computation, versus locally autonomous firing and synchronous oscillations in the gamma range

In several controversial papers Langton (1986, 1990) has drawn upon the work of Wolfram (1984) and von Neumann (1949) to propose that physical systems undergoing local second-order phase transitions may include universal computation as a potential property. Such systems exhibit prolonged transitional states, and have been termed ‘edge of chaos’. Langton has suggested the brain may exploit the ‘edge of chaos’ to achieve cognition, in ways he has not stated in physiological terms. In conformity with this suggestion, local phase transitions in the gamma band are implied by our findings. With these local transitions stabilized in the large by rapid synaptic feedbacks and/or adiabatic actions of slower neurotransmitters and polysynaptic cortical–subcortical interactions, the regulation of cerebral phase transitions on all spatio-temporal scales is implicit.

8.2. Cortical dynamics, coherent infomax and storage of information in cortical networks

Kay and Phillips (1997) and Phillips and Singer (1997) have described a learning rule related to the Hebb principle, which they believe is physiologically realistic and which exploits the occurrence in the brain of synchronous oscillation. Synchrony between separate sites of input in an abstract passive-filter neural network is imposed by assuming the existence of hypothetical ‘Contextual Field’ (CF) fibres. The learning rule operates to maximize the storage in synaptic connections, of relations among the information streams introduced at separate ‘Receptor Field’ (RF) sites of input. This information storage principle they term ‘coherent infomax’. The physiological validity of the learning rule seems further confirmed by its application to learning of connectivity in models of cortical anatomy (Alexander et al., 2000) where we have found this rule more efficient than the simpler Hebb formulations in generating realistic visual cortex connectivities. In these simulations of cortical connectivity we have shown that the learning rule results in a mapping of the visual field onto cortex such that the visual field becomes tiled with small homotypic maps of the visual field — each tile about the size of a macrocolumn. This property may explain the way that contextual information is introduced into local cortical information processing (Phillips and Singer, 1997; Phillips and Pfieger, 2000).

Our account of the mechanism of synchronous oscillation appears to obviate the need for the special assumption of CF connections giving rise to synchrony. It can be shown that our model of synchrony can account for the transfer of information about stimulus properties to synapses distributed widely in the cortical field (Wright et al., 2000). Further, as indicated in another paper in the Agora Symposium, the 1/f background activity of the less excited cortex may enable optimum information transfer among the active cortical sites, in accord with the coherent infomax requirement that passive-filter information transfer take place between active RF sites.

There appears to be no reason why the coherent infomax principle might not include the extraction of informational relations among activity patterns in autonomously active patches of cortex, as well as among discrete external sources of input. (Wright, 1997b).

That is, locally autonomous, ‘edge of chaos’, attractor dynamics would give rise to an internal type of RF input. These internal RF inputs, and RF inputs from sensory pathways, linked into fields of coherent oscillation by the cortical wave medium would act to store informational relations between RFs of both internal and external origins. Applying these considerations to an imaginary organism interacting with an environment, the coherent infomax rule would thus allow association of environmental stimuli with the organism’s internally generated cognitive and motor activity. As the organism’s motor activity, sensory and cognitive activity would all bring about re-
warding and punishing reinforcement influences mediated by hard-wired, survival selected, subcortical systems (e.g. Olds and Milner, 1954), we may conjecture that adaptive learning could result. This conjecture depends upon the further assumption that the reward and punishment systems act to preferentially select the synaptic storage of information over appropriate time scales.

9. Conclusion

It appears that a relatively simple set of state equations with parameter values close to those determined by independent physiological measurements, accounts for observed dynamics of cortical neurones at microscopic, mesoscopic and macroscopic scales. At microscopic level, fast synaptic feedbacks appear to be crucial to the stability of the excited cortex. At mesoscopic scale interactions via dissipative wave transmission between excited cortical areas create fields of synchrony, and at macroscopic scale low-frequency resonances and travelling waves generate the $1/f$ background spectrum of the cortex. When this intracortical dynamic is supplemented by resonance between cortex and thalamus, much of the spectral content of the EEG is accounted for.

The emerging picture of cerebral dynamic organization is different to that envisaged in alternate models — notably those of Nunez (1981) — and much may be gained by future experimental comparison of competing formulations. Our models are testable in a number of ways — for example, we predict dependence upon brain size, axonal conduction velocity and dendritic delays, of the ‘knee’ frequency between $1/f$ and $1/f^3$ limbs of the background spectrum.

The present work offers the prospect that dynamical brain events taking place across a wide range of scales might be unified within a single account. It also offers the more remote prospect of unification of dynamics with specific learning rules. Whether such a unified account is truly possible should be revealed as progressively greater anatomical and physiological detail, and more exact specifications of parameters are included. As indicated in the text, at the time of writing we are attempting several lines of development of this kind. These include simulations with more realistic anisotropic cortical connectivity at macroscopic and microscopic scales, the simulation of cortical evoked potentials and the improved specification of synaptic feedback mechanisms.

References


Tones, M.V., Westbrook, G.L., 1996. The impact of receptor


Electroencephalographic baselines in astronaut candidates
estimated by computation and pattern recognition tech-

chronised oscillations in interneuron networks driven by
metabotropic glutamate receptor activation. Nature 373,
612–615.

oscillatory behaviour in primary visual cortex. Neural

Wilson, H.R., Cowan, J.D., 1973. A mathematical theory of
the functional dynamics of cortical and thalamic nervous

Wolfram, S., 1984. Universality and complexity in cellular

Wright, J.J., 1990. Reticular activation and the dynamics of

global and microscopic scales. Neural networks and the

Wright, J.J., 1997a. EEG simulation: variation of spectral
envelope, pulse synchrony and approx. 40Hz oscillation.

Wright, J.J., 1997b. Local attractor dynamics will introduce
further information to synchronous neuronal fields. Behav.
Brain Sci. 20, 701–702.

Wright, J.J., 1999. Simulation of EEG: Dynamic changes in
synaptic efficacy, cerebral rhythms, and dissipative and

Wright, J.J., Bourke, P.D., Chapman, C.L. 2000. Synchronous
oscillation in the cerebral cortex and object coherence:
simulation of basic electrophysiological findings. Biol. Cy-

transmission in human cerebral cortex. Physica D 106,
363–374.

bioelectrical activity of the cerebral cortex. Biophysics 39,
133–150.
Simulated Electrocortical Activity at Microscopic, Mesoscopic, and Global Scales

JJ Wright*,1,2,3,6, CJ Rennie1,4,5, GJ Lees2, PA Robinson1,4, PD Bourke3, CL Chapman3, E Gordon1,7 and DL Rowe1,4

1Brain Dynamics Centre, Westmead Hospital and University of Sydney, Westmead, Australia; 2Department of Psychiatry, University of Auckland School of Medicine, Auckland, New Zealand; 3Mental Health Research Institute, Parkville, Victoria, Australia; 4Theoretical Physics Group, School of Physics, University of Sydney, Australia; 5Department of Medical Physics, Westmead Hospital, Westmead, Australia; 6Liggins Institute, University of Auckland, Auckland, New Zealand; 7Department of Psychological Medicine, Westmead Hospital and University of Sydney, Westmead, Australia

Simulation of electrocortical activity requires (a) determination of the most crucial features to be modelled, (b) specification of state equations with parameters that can be determined against independent measurements, and (c) explanation of electrical events in the brain at several scales. We report our attempts to address these problems, and show that mutually consistent explanations, and simulation of experimental data can be achieved for cortical gamma activity, synchronous oscillation, and the main features of the EEG power spectrum including the cerebral rhythms and evoked potentials. These simulations include consideration of dendritic and synaptic dynamics, AMPA, NMDA, and GABA receptors, and intracortical and cortical/subcortical interactions. We speculate on the way in which Hebbian learning and intrinsic reinforcement processes might complement the brain dynamics thus explained, to produce elementary cognitive operations.


Keywords: gamma rhythm; cortical synchrony; evoked potentials; NMDA; AMPA; GABA

INTRODUCTION

Neuropsychopharmacology is a discipline attempting the unification of events at different scales in the brain. This is part of the larger task of relating brain events to conscious experience and learning processes of normal people and people suffering from mental disorders. The problems involved in this unification are legion, as is well known to readers of this journal. It is commonly found that different types of events in the brain are difficult to relate to each other, except as purely empirical findings. For example, it is seldom clear how findings made by PET and fMRI scanning are functionally related to most EEG measures. Nor how EEG measures might reflect specific changes in synaptic physiology. Nor how neurochemical effects can be functionally related to psychological events. These are all difficult questions, despite the huge growth in detailed knowledge about specific processes and components in the brain. No simple one-to-one ways exist, to relate these very complex phenomena, one to the others. Yet unifying principles must be sought, if the discipline is to advance in a coherent fashion. One way of moving towards this unification is through the discipline of Brain Dynamics.

The aim of the subject of Brain Dynamics is to give a simplified, but sufficient, mathematical description of the operation of the brain, in terms of the brain’s observable electrical activity (Freeman, 1975; Basar, 1976; Nunez, 1981, 1995).

Obtaining a ‘simplified, but sufficient, mathematical description’ can be broken down into a number of subtasks. These are:

(a) Abstraction from physiological data of the most important properties of the neurones of the brain, to be used in the mathematical model.
(b) Determination of the form of the state equations, and the values of parameters.
(c) Comparison of the equations’ properties to observable aspects of brain function, at as many temporal and spatial scales as possible.

The observable aspects of brain function to be accounted for include the global EEG, the electrocorticogram and local field potentials, and the statistics of action potentials. If a wide match to experimental data can be achieved, then the model that has emerged should exhibit dynamic properties akin to a real brain, and it might then be asked what consequences these dynamic properties have for adaptive
Does the introduction of plausible learning rules then lead to an explanation of adaptive behavior?

Attempts to approach a theory of the brain along these lines stem from the work of McCulloch and Pitts (1943), and have led to extensive work on the properties of artificial neural networks (e.g., Amit, 1989). Lines of theoretical development more explicitly concerned with physiological dynamics, and especially the dynamics underlying EEG, flow from Wilson and Cowan (1973), Freeman (1975), Nunez (1981), Lopes da Silva (van Rotterdam et al., 1982), Haken (Haken et al., 1985), and others (e.g., Arbib et al., 1998). Work by Freeman (1975 and subsequently), and of Singer and Gray and others (e.g., Gray and Singer, 1989; Gray et al., 1989; Eckhorn et al., 1988) has been of particular importance in revealing the dynamics of the brain from an experimental perspective.

The work to be described from our own group has evolved from early attempts to define circumstances in which linear methods of analysis could be applied (Wright, 1990). Simple numerical simulations followed (Wright and Liley, 1996; Wright, 1997, 1999), leading to more advanced methods, including the development of wave equations (Robinson et al., 1997, 1998a, b, 2001). We have progressively introduced more refined physiological parameters, and descriptions of anatomical organization (Liley and Wright, 1994; Rennie et al., 1999, 2000, 2002), with the object of developing a single model to account for events in the brain at a number of different scales.

OVERALL BRAIN ORGANIZATION

Figure 1 highlights elementary features of gross brain organization. The cortical mantle is the terminating area for major sensory pathways, and the source of much of the signals organizing motor activity. The cortex never acts alone, but via continuous interaction with subcortical systems, notably the thalamus, the limbic system including hippocampus, the basal ganglia, etc (Alexander et al., 1990; Posner and Petersen, 1990). Any model complete to first approximation must consider interactions within the cortex, and also between cortex and subcortical structures, including events at microscopic, mesoscopic, and macroscopic (global or whole-brain) scales.

CORTICAL DYNAMICS AT MICROSCOPIC SCALE

Figure 2 shows the two elementary components of the cortex—excitatory (pyramidal) cells which make up about 90% of cortical cells, and which send axons to remote cortical locations, as well as interacting with near neighbors via intracortical axons—and inhibitory cells, which give only local intracortical axons. The excitatory cells utilize glutamate as a neurotransmitter, and act principally on fast-acting AMPA receptors, and slower-acting NMDA receptors. The NMDA receptors are voltage dependent—that is, they influence the postsynaptic dendrites when the receiving cell is depolarized, and thereby emitting action potentials. The inhibitory cells use GABA as their neurotransmitter. These are the most important fast forms of neuronal interaction.

A mathematical account of interactions in the cortex of these restricted types is given in the appendix, and the values we have obtained for all the parameters required in this mathematical model are given in Table 1. The values of the parameters are taken largely from the findings of Braitenberg and Schuz (1991), Thomson et al (1996), and Thomson (1997), either directly, or from further calculations based on these (Liley and Wright, 1994; Rennie et al., 2000). Recent additions regarding synaptic physiology...

Our account treats the cortex as a continuum, rather than as a set of separate cells, although the properties of the continuum are closely related to the physiology of individual neurones. It is thus an account of population dynamics, rather than of cell-by-cell interactions, and is suitable for numerical solution of the events over a large extent of the cortical surface, as well as in small domains of

Neuropsychopharmacology
macrocoluminar, or lesser, dimension. In addition to numerical solutions, the equations can be solved for some purposes in linear approximation, to yield wave equations and dispersion relations, etc (Robinson et al, 1997, 1998a,b; Rennie et al, 2000). Comparable treatments have been given in Jirsa and Haken (1996), Liley et al (1999, 2002), and Jirsa et al (2001).

The way the mathematics are related to the cortical anatomy and physiology is contained in three main equations, next described, with explanation of the minor terms and parameters left to the appendix and Table 1.

In the following formulae, subscript \( p \) represents association with a presynaptic neurone, and subscript \( q \) association with a postsynaptic neurone. Thus, \( p \) or \( q \) may each be replaced by either \( e \) (excitatory) or \( i \) (inhibitory), and subscript \( pq \) indicates a property arising from the action of \( p \) upon \( q \).

The principal observable property, to which the EEG can be directly related, is the average dendritic membrane potential, or local field potential, arising from the pyramidal firing rate of cells of type \( r \) later at elsewhere in the cortex. This synaptic flux, \( r \) activity.

This random perturbation of instantaneous pulse density posed on the mass-action influences within the continuum. An individual neurone reflects random processes superimposed by a Poisson distribution of action potentials emitted from \( r \). With the introduction of \( \pi \) and \( \sqrt{3} \), the parameters \( \theta \) and \( \sigma \) correspond, respectively, to the average membrane potential at which pulse density reaches 50% of maximum, and the standard deviation of pulse density as a function of average membrane potential. The probability that a randomly chosen cell within the continuum is emitting an action potential at a randomly chosen time, \( t \), is given by \( Q_q = Q_q^\text{max} / 1 + e^{-\pi(V_e - \theta)/\sqrt{3}\sigma} \).

Within the continuum formulation, \( Q_q(\mathbf{r}, t) \) can be further treated as stochastic variables, by assuming that the Poisson distribution of action potentials emitted from individual neurones reflects random processes superimposed on the mass-action influences within the continuum. This random perturbation of instantaneous pulse density thus contributes one type of driving to the net cortical activity.

The second equation describes the conversion of average dendritic membrane potentials within a small locale at a position \( \mathbf{r} \), into a density of action potentials within the locale, as a sigmoidal function. Thus, \( V_q(\mathbf{r}, t) \), the average membrane potential of excitatory or inhibitory cells becomes converted into \( Q_q(\mathbf{r}, t) \), the instantaneous average firing rate of cells of type \( q \), within the locale

\[
Q_q = \frac{Q_q^\text{max}}{1 + e^{-\pi(V_e - \theta)/\sqrt{3}\sigma}}.
\]

With the introduction of \( \pi \) and \( \sqrt{3} \), the parameters \( \theta \) and \( \sigma \) correspond, respectively, to the average membrane potential at which pulse density reaches 50% of maximum, and the standard deviation of pulse density as a function of average membrane potential. The probability that a randomly chosen cell within the continuum is emitting an action potential at a randomly chosen time, \( t \), is given by \( Q_q = Q_q^\text{max} / 1 + e^{-\pi(V_e - \theta)/\sqrt{3}\sigma} \).

Within the continuum formulation, \( Q_q(\mathbf{r}, t) \) can be further treated as stochastic variables, by assuming that the Poisson distribution of action potentials emitted from individual neurones reflects random processes superimposed on the mass-action influences within the continuum. This random perturbation of instantaneous pulse density thus contributes one type of driving to the net cortical activity.

The second equation describes the way that action potentials reach the synapses at \( \mathbf{r} \) from all the neurones elsewhere in the cortex. This synaptic flux, \( \varphi_p \), arises from \( Q_p \), which arose at many positions, \( \mathbf{r}' \), and were conveyed at the velocity of axonal conduction, \( v_p \), to arrive later at \( \mathbf{r} \):

\[
\varphi_p = \int f(|\mathbf{r} - \mathbf{r}'|, r_p) Q_p(\mathbf{r}', t - |\mathbf{r} - \mathbf{r}'|/v_p) \, d^3 \mathbf{r}'.
\]

The final equation is more complicated, and describes the dynamics of the synapses, and the dendritic membranes, in converting the synaptic flux into the average dendritic potential, which then gives rise to further action potentials. This is

\[
V_q = V_q^{(0)} + \sum_p N_{qp} H_{qp}^{(2)} \otimes (s_{qp} \Gamma_p \varphi_p)
\]

including the densities of synapses of different types (\( N_{qp} \)), the time characteristics of EPSP and IPSP (\( H_{qp}^{(2)} \)), the effects of reversal potentials and backpropagation of action potentials into the dendritic tree (\( s_{qp} \)), and the modulation of AMPA, GABA, and NMDA receptor actions (\( \Gamma_p \)). The last term, \( \Gamma_p \), has parameters derived from recent models of the kinetics of channel opening and closing in receptors, consequent to varying equilibria among various tertiary molecular configurations of the receptor-channel complex (Lester and Jahr, 1992; Dominguez-Perrot et al, 1996; Hauser and Roth, 1997; Partin et al, 1996). The derivation of these receptor parameters will be reported elsewhere.

When compared to related and ancestral continuum models, the novel features of the present treatment are contained in the terms \( \Gamma_p \) and \( s_{qp} \).

The features describing the adaptations of AMPA, NMDA, and GABA receptors confer two properties. Firstly, because of the relatively rapid decrease in efficacy of the excitatory receptors compared to that of the inhibitory receptors as both types of synaptic flux increase, there is a tendency to stabilize locally the rate of generation of action potentials. Secondly, the voltage-dependent operation of NMDA receptors confers a capacity for the transient amplification of fields of synchronous oscillation, as will be shown below.

The features encapsulated in \( s_{pq} \) have consequences for the frequency spectrum of local cortical activity. They confer a switch-like property, such that for relatively low pulse densities the 1/f background spectrum typical of EEG is the predominant feature. At a critical point as firing rate increases, a switch to oscillation in the gamma band occurs, of greater local amplitude than the 1/f background. This switch depends crucially on a convolution, \( H_{qp}^{(2)} \otimes V_q \), contained in \( s_{pq} \) (see appendix). Physically, \( H_{qp}^{(2)} \) acts as a sharply adjusted filtering process conferring the switching effect, which is attributed to the impact of backpropagation of action potentials into the proximal dendritic tree (Stuart and Sakmann, 1994).

The combination of these properties enables simulation of a two-dimensional wave medium, which can be compared to experimental data, as is shown in Figures 3 and 4.

(a) At low levels of nonspecific cortical activation, when stimulated by weak afferent volleys (approximated as a white noise) to any part of the cortical surface, or in response to the intrinsic perturbations of \( Q_q(\mathbf{r}, t) \) alone, wave motion propagates outward from the site of stimulation. The wave motion propagates at realistic speed for electrocortical activity and has a 1/f' frequency content, identical to that seen in background EEG activity. At very low levels of cortical activation, a small peak of activity is also seen in the theta range, but this cannot account fully for the power seen in the theta band in matched experimental data (Figure 3). The failure to account fully for the theta activity is one example of why the cortical simulation needs to be supplemented with consideration of cortical/subcortical interactions—see further below.
(b) At slightly higher levels of nonspecific cortical activation, broadband activity in the gamma range appears against the 1/f cortical background activity, and can be matched to experiments (Figure 3).

(c) At still higher levels of nonspecific cortical activation, with or without introduction of specific sensory input, oscillation appears in the simulated cortical field, and this has close similarity to high levels of gamma or ‘40 Hz’ observed in many cortical experiments. Distant from the site of this autonomous oscillation, it is found that the gamma activity has acted as a driving source for the background 1/f activity, by contributing a further increase in the amplitude of random fluctuations of pulse density, (Figure 4). At the site of the autonomous gamma oscillation, the signal amplitude is very high compared to the background 1/f activity at remote sites, and is closely correlated with local pulse activity. Pulse densities are comparable to those observed physiologically (Steriade et al, 2001). At shorter distances of transmission, or with high coupling strength between locales, the direct transmission of activity in the gamma band is apparent.

Thus, depending upon local input conditions, transitions can take place between a condition of autonomous oscillation in the gamma band, and a dissipative, point-attractor condition, in which the cortex is a passive medium of wave transmission. Transition between these states occurs sharply, depending on a very small change in local cell firing rate, and near this transition, both the 1/f background activity and the gamma activity have a stability factor close to zero. This capacity for sharp transition of state, analogous to a thermodynamic change of phase, bears comparison to the ‘edge of chaos’ type of dynamics described by Langton (1986, 1990), without necessarily implying that the cellular dynamics are chaotic in either state. The occurrence of fluctuating changes of state at a critical level of activation accords with physiological observations (Freeman and Rogers, 2002; Phillips and Pfieger, 1999). There are, in turn, possible implications for information processing, storage and recall from memory (Kay and Phillips, 1997; Phillips and Singer, 1997; Wright, 1997b; Arhem and Liljenstrom, 2001; Liljenstrom, 2002). Also, patches of active cortex can enter into synchronous oscillation, as is next described.

CORTICAL INTERACTIONS AT MESOSCOPIC SCALE

At a scale from fractions of a millimeter to many centimeters of cortex, patches of active cells have been experimentally observed to enter into synchronous oscillation—that is, cross-correlations of pulse density, or of mean local field potential at the separated loci are maximal at zero lag. Synchronous oscillation has been widely, although controversially, considered to act as a substrate for association processes in the cortex. (e.g. Eckhorn et al, 1988; Singer, 1994; Singer and Gray, 1995; Stryker, 1989; Bressler et al, 1993; Livingstone, 1996; Miltner et al, 1999; Neuenschwander and Singer, 1996; Palm and Wennekers, 1997; Steriade et al, 1996; Gray and Singer, 1989; Gray et al, 1989).

Results very similar to classic findings of synchronous oscillation readily appear when suitable stimuli are introduced into the simulated cortex, as is shown in Figure 5. This property was first observed in numerical simulations (Traub et al, 1996; Wright, 1997a; Wright et al, 2000), and then explained analytically (Robinson et al, 1998a; Chap-
Cross-correlation at zero lag can be shown to occur at all frequencies, but, as is the case in physiological experiments, large oscillations in the gamma band, at about 40 Hz, are associated with the largest amplitudes.

If two points on the simulated cortical surface are driven by separate, uncorrelated, inputs (or in response to uncorrelated autonomous local cortical activity at two separated points) then, within a few milliseconds cross-correlated activity, maximal at zero lag, appears in the neighborhood of both active sites. For this effect to occur axonal delay must be small compared to the rise and fall time of the dendritic response, and the speed of onset of synchrony depends mainly on the axonal delay.

Figure 6 sums up the mechanism. The results shown in the top two rows of the figure are derived from a simulated field of electrophotential activity, which is not shown explicitly. The relevant state-variable in the simulated field was \( V_e(t, x) \), the mean dendritic potential of excitatory cells, and the field activity was generated by delivering two independent and uncorrelated time series of white noise to two points on the simulated cortical field. Wave activity radiated outward from both sites of input. The relevance for the generation of synchronous oscillation in the brain lies in the way these two independent fields of travelling waves can be shown to interact.

The top frames of Figure 6 show how a large field of zero-lag synchrony surrounds both sites of uncorrelated input. The middle frames show eigenmode decomposition of the same field of wave activity, which breaks the field of activity into components of spatial activity, each independent of the others. The first (dominant) eigenmode defines the field of synchrony. The bottom frames convey the essence of the physical process, which can be explained as follows. The uncorrelated inputs can be decomposed into their even and odd components—roughly, the parts of each driving signal which are coincidentally in phase with the other, and those parts of coincidentally reversed phase—the coincidences...
occurring randomly over time. Wave activity radiates from both sites as travelling waves, but since dendrites act as summing junctions, they dissipate even and odd components in their inputs selectively. Even components arriving at each junction reinforce one another, while odd components tend to cancel about the signal mean. Dendrites also perform a running time average of coinciding signals, over a period longer than the axonal conduction lags. Thus, the activity induced by the even components in the driving signals at the two sites dominates throughout the field, and emerges as the first eigenmode of the wave activity.

It can be shown that the magnitude of the zero-lag correlated field is sensitive to dendritic delay time, axonal conduction lag between sites, and the relative strengths of couplings in the field. Figure 7 shows the importance of the strength of coupling between the two active sites. Notably, when two sites on the cortex are concurrently active, the strength of synchronous oscillation between the sites will be a function of both their pre-existing degree of anatomical connectivity, and of the activation of voltage-dependent excitatory receptors. Thus, in these simulations, NMDA acts as a dynamic coupling, enhancing synchrony, and can contribute to the generation of fluctuating fields of synchrony, in appropriate conditions of cortical activation.

INTERACTIONS WITH SUBCORTICAL SYSTEMS

Figure 8 expands the notions of cortical and subcortical interaction sketched in Figure 1, with emphasis now on the first-order loops of interaction of cortical and thalamic structures. This architecture has been introduced (Robinson et al., 2001) to a mathematical description of the cortex similar to that described above, but omitting the properties encapsulated in $\mathcal{T}_p$ and $s_{ip}$, which confer the capacity for autonomous gamma-band oscillation. Since corticothalamic interactions take place at large scale, between generally quiescent masses of neurones, it has been considered sufficient at present to treat the cortex at large scale as if driven by spatio-temporal white noise, of extrinsic or intrinsic origin. In future, the detailed interaction with locally active cortex must be considered. With that proviso, introduction of coresonance with the thalamus can account for all the additional frequency bands of activity regularly seen in the EEG, in sleep and waking (Walter et al., 1967; Steriade et al., 1990).

Emphasis has been placed on the contribution made to global resonance by the shortest paths of interaction of cortex and subcortex—that is, in loops which are a few neurones long, modulated by interactions between the cortex and the excitatory and inhibitory components within the thalamus, or related systems. Obviously, longer pathways of interaction are of great functional importance, but within the parameters applied in our models, longer pathways cannot generate such relatively 'sharp' resonances, as dendritic processes act as a low-pass filter, progressively blurring the response in loops of many neurones. Multineurone loops appear functionally suited to adiabatic regulation of cortical tone, and thus may be reflected in slow cortical potentials, rather than the major EEG rhythms. Again, this adiabatic regulation is a feature yet to be introduced into any of our models.

Figure 9 shows examples of the way the corticothalamic model can be fitted to EEG power spectra over a wide range of states of cortical activation, from sleep to waking. Here, the process of curve fitting is the inverse of the way that the simulation is used to generate simulated EEG data. In each case, a theoretical curve has been fitted to the experimental data by adjusting a number of free parameters, which are not important. The impact upon synchrony of concurrent transition into autonomous gamma activity at separate cortical locations is currently under investigation.

It can be shown that the magnitude of the zero-lag correlated field is sensitive to dendritic delay time, axonal conduction lag between sites, and the relative strengths of couplings in the field. Figure 7 shows the importance of the strength of coupling between the two active sites. Notably, when two sites on the cortex are concurrently active, the strength of synchronous oscillation between the sites will be a function of both their pre-existing degree of anatomical connectivity, and of the activation of voltage-dependent excitatory receptors. Thus, in these simulations, NMDA acts as a dynamic coupling, enhancing synchrony, and can contribute to the generation of fluctuating fields of synchrony, in appropriate conditions of cortical activation.
mathematically related to groups of parameters used in the cortical model. As noted above, it is assumed that the system is driven by diffuse white noise, which may be equated with specific and nonspecific inputs to cortex, or with the effect of random perturbations of pulse density partly associated with small patches of autonomous gamma activity, or both. At optimal fit, the free parameters obtained have values in the range expected from the independently estimated parameters used in the cortical model.

A further point of interest is that this account of the origin of the gross EEG spectrum and the major cerebral rhythms does not depend on specific details of the cerebral boundary conditions, as has been proposed by Nunez (1981). In our work it can be readily shown that the spatial damping is very high, and thus global resonant modes play no significant part in the generation of wave activity (Wright, 2000; Robinson et al, 2001). Nunez’s (1995) more recent work deals with the interaction of local and global scales of cortical activity, and our present work is in closer harmony with this modification to his earlier work.

Figure 10 extends the explanatory power of the model to include auditory-evoked potentials (ERP). This figure shows real auditory-evoked potentials, obtained by averaging over time-locked epochs of EEG, obtained as the individual was responding to ‘target’ tones, and ignoring ‘nontarget’ tones, in a conventional ‘oddball’ experimental paradigm. Also shown are simulated ERP produced in accord with methods described in Rennie et al (2002). The prestimulus ‘resting’ EEG of the subject recorded in epochs just before the delivery of each stimulus was characterized by fitting the corticothalamic model in the same way as in Figure 9, to obtain a set of descriptive parameters. These parameters define the transfer function of the corticothalamic system, assuming that the sources driving cortical activity over the prestimulus periods are equivalent to white noise.
transfer function thus obtained can then be used to predict the average time-locked response, or impulse response, of the cortex, to any additional, but time-locked, input. As shown, this enables the ERP to be accounted for as time-locked responses to short bursts of activity generated by sensory input, and from within the brain. Unfortunately, there is no means currently available to us to generate the second input from within the simulation, as this must depend upon a specific cognitive event—the subjects’ decision that the particular tone heard could be (correctly) identified as a target, requiring a response. In principle at least, we expect that this cognitive process might be reduced to associative processes mediated by mechanisms as described in the earlier sections on cortical gamma activity and synchronous oscillation. This step toward closing the paths of causality within brain modelling, is, of course, a daunting task—and one that might best be initially addressed in approximately steady-state conditions, in accord with the methods described by Jirsa et al (2001).

The corticothalamic model so far developed has some interesting implications, which may influence future modelling. Firstly, the role of the lower-frequency cerebral rhythms on biasing transitions of state in the cortex is largely unexplored. Secondly, the occurrence of long-range inhibitory interactions in the thalamus, mediated by the reticular nucleus, contrasts with the predominance of excitation at long range in the cortex. This implies a reversal of phase relations generated by synchronous oscillations in thalamus, to those in cortex. Since thalamic association nuclei and cortex interact with a high degree of one-to-one mapping, the interactions in the entire system might further modulate the onset and offset of fields of synchrony in a dynamic way. Effects on alpha phase suggestive of the operation of such a mechanism have been observed, both experimentally, and in a related type of thalamo-cortical simulation (Suffczynski et al, 2001).

Thirdly, the interactions of cortex and thalamus with other subcortical systems over longer pathways may act to modulate the fields of cortical activation, creating analogues...
of attention and selective arousal. Such modulations would be relatively slow compared to the cerebral rhythms, as interactions between cortex and thalamus of order greater than two neurones are generally smoothed by dendritic filtering. Subcortical pathways might thus regulate ‘adiabatic’ shifts in cortical activation, akin to the slow potentials of the cerebral cortex. Fourthly, there is some evidence that some theta activity might be related to pulsed interactions between the cortex and the thalamus (Rennie et al., 2002).

Increased understanding of this, and of other sources for theta rhythm may lead to explanation of the failure of our cortical model to account for the theta content observed in experiments, as noted in relation to Figure 3.

Finally, it may be noted that by negative feedback mechanisms, corticothalamic and other cortical/subcortical interactions may contribute to maintaining cortical activation near the level of transition between gamma and 1/f conditions.

DISCUSSION

The results described above indicate that representation of brain activity in mathematical terms is practicable, albeit in a preliminary manner. Further work is needed to overcome the numerous points upon which our treatment encounters limitations, as pointed out in the preceding text. As such problems are addressed, further physiological and anatomical detail might thus be introduced into an organized framework, always subject to tests against experimentally observed properties of the brain’s global properties.

The properties observed in our simulations suggest approximation toward an account of the unified operation of the brain may be possible. Within the simulations there occur properties considered, in related contexts, to confer universal computation, information storage, association and recall, and self-organization (Freeman and Rogers, 2002; Arhem and Liljenstrom, 2001; Liljenstrom, 2002; Langton, 1986, 1990; Singer, 1994; von Neumann, 1949; Wolfram, 1984; Phillips and Singer, 1997; Kording and Konig, 2000). Work showing that realistic anatomical connectivity can appear on the basis of Hebbian learning, in certain simplified neural networks (von der Marlsburg, 1973; Swindale, 1996; Alexander et al., 2000) is also of interest, as this property appears likely to be transferable to the more dynamically realistic continuum models described here.

A further step would see, as well as the introduction of Hebbian learning, the incorporation of innate positive and negative reinforcement systems analogous to those known to exist in the brain (Olds and Milner, 1954). It may be presumed that such reinforcement systems act, via neuro-
modulation, to regulate synaptic consolidation. Some survival behaviors related to activity in the reinforcement pathways appear to be hard-wired, thanks to species evolution. Such a priori reinforcement systems might also act to supervise adaptive learning, if the reinforcement system was initially activated by only a small subset of the total environmental stimuli, but could, by associative learning, come to be activated by more complex inputs, and by internally generated brain states. Internally generated states that did not acquire association with basic survival behaviors and the operation of the reinforcement mechanisms would be extinguished if they did not activate the survival behaviors and the synaptic consolidation mechanisms over some critical time period. Thus, survival-consistent behaviors would remain, and develop in complexity.

Another goal is the fitting of models of this type to a much larger range of data, particularly the large normative and standardized data set under construction by Gordon (2000, 2002), which includes EEG and fMRI data from both normal people and people with a variety of psychopathologies. It is our hope that this will help us to contribute to the unification of findings in neuropsychopharmacology, as discussed at the beginning of this paper.

ACKNOWLEDGEMENTS

We thank Professor Walter Freeman for access to his library of electrocortical recordings, Mr Nicholas Hawthorn for technical assistance, and the University of Sydney, the University of Auckland, the Swinburne University, Melbourne, and the Mental Health Research Institute, Melbourne, for the use of computer facilities. We thank also our hosts at the Agora for BioSystems, Stockholm, and the University of Potsdam, at whose conferences this paper was delivered.

REFERENCES


Neuropsychopharmacology

APPENDIX: A MORE DETAILED DESCRIPTION OF THE MATHEMATICS OF A CORTICAL MODEL AND OF METHODS OF SIMULATION

In the following formulae, subscript \( p \) represents association with a presynaptic neurone, and subscript \( q \) association with a postsynaptic neurone. Thus \( p \) or \( q \) may each be replaced by either \( e \) (excitatory) or \( i \) (inhibitory), and subscript \( qp \) indicates a shared property. The extended subscripts \( ic \) and \( cc \) indicate that either intracortical or corticocortical synapses are specifically referred to.

The principal state variables are:

\[
Q_q(r, t), \quad \text{the point pulse densities of neurones of type } q, \quad \text{where } t \text{ is the time, and } r \text{ is the position on the cortical surface (s⁻¹)}.
\]

\[
V_q(r, t), \quad \text{the corresponding point local potentials (V).}
\]

\[
\phi_p(r, t), \quad \text{the synaptic flux densities (s⁻¹)}.
\]

The rate at which pulses are generated is given by

\[
Q_q = \frac{Q^\text{max}_q}{1 + e^{-(V_q - V_{th,q})/\sqrt{2}a_q}},
\]

where \( Q^\text{max}_q \) is the maximum pulse rate sustainable by neurones of type \( q \), and \( \theta_q \) and \( a_q \) are parameters defining pulse density distribution as a function of \( V_q \). Perturbation of the value of \( Q_q(r, t) \) by random local factors in individual neurones provides an internal source of white noise driving to the system.

The synaptic flux densities are given by equations of the form

\[
\phi_p = \int f(|r - r'|, \; r_p)Q_q(r', \; t - |r - r'|/v_p) \; d^r r',
\]

where \( r_p \) is the axonal range and \( v_p \) is the average velocity of axonal conduction. The axonal distribution functions \( f(|r - r'|, \; r_p) \) may have any form, but are usually treated as two-dimensional Gaussian distributions. Intracortical connections are also usually approximated as having no significant extension in space, unlike the corticocortical connections, and thus

\[
\phi_e = \phi_{e,ic} + \phi_{e,cc},
\]

\[
\phi_{e,ic} \approx Q_x,
\]

\[
\phi_{i,ic} \approx Q_i,
\]

\( \phi_s \) and \( \phi_{ns} \) are specific and nonspecific inputs to the cortex from all sources, and are not explicitly described here, but act as an external source of driving inputs to the cortical system.

Local field potential is considered directly proportional to the pyramidal cell point membrane potential, which is given by

\[
V_q = V_q^0 + \sum_p N_{qp} H_q^{[D]}(\phi_p),
\]

where \( V_q^0 \) is the resting membrane potential, \( N_{qp} \) is the number of synapses of type \( p \) per dendritic tree, and \( \otimes \) indicates convolution over time.

Normalized average EPSP/IPSP at the soma are given by

\[
H_q^{[D]}(t) = \frac{a_{qp}b_{qp}}{b_{qp} - a_{qp}}(e^{-a_{qp}t} - e^{-b_{qp}t}), \quad t \geq 0,
\]

where \( 1/b_{qp} \) and \( 1/a_{qp} \) are the rise and fall time constants.

The standardized synaptic strength is represented by \( g_q^{[e]} \). It is the integral over time of the PSP, as recorded at the soma, when the neurone is at its resting membrane potential \( V_q^0 \). The actual synaptic strength \( s_{qp} \) is related to this, through the ion channel reversal potentials, \( V^{rev}_q \), and a measure of the membrane potential, thus:

\[
s_{qp} = \frac{g_q^{[e]} V^{rev}_q - V_q^{[M]} \otimes V_q}{V^{rev}_q - V_q^0}.
\]

The convolution by \( H_q^{[M]} \) is introduced to account for an effect of retrograde propagation of action potentials into the dendritic tree. We assume that at low synaptic flux and firing rates, cells are principally sensitive to synapses proximate to the soma. During active firing, the site of synaptic summation and further action potential generation is shifted more distally, into the dendritic tree, and now more remote synapses contribute greater weight to determination of whether or not firing continues. The contribution of remote synapses depends upon the arrival of EPSP, IPSP generated at delay times in the distal dendrites, and at earlier values of the membrane potential.

For the population of cells, the delay because of this effect increases as average membrane potential decreases, so to first approximation,

\[
H_q^{[M]}(t) = c_q e^{-c_q t}, \quad t \geq 0,
\]

where \( c_q \) is determined by a linear regression,

\[
c_q = a(V_q - V_q^0) + b.
\]

(The linear regression can be replaced by a more physiologically realistic nonlinear function, which will be reported in detail elsewhere.)

The impact of this membrane-voltage-dependent switching of a mean delay process in the dendrites is to regulate partially the transition of local electrocortical activity into, and out of, an active state in which gamma activity appears, and generates strong fluctuations in the local pulse-density and local field potential. This contributes a third type of driving to wave propagation in the system, in addition to that provided from external sources and stochastic perturbations of pulse density.

Neurotransmitter receptor adaptation to continuing input, as is included in \( \Gamma_p \), for which parameters \( \lambda_p, \mu_p, k_p, \alpha, \beta \) have been obtained from modeling the behavior of receptors, from physiological measurements of transformations among receptor tertiary molecular configurations (references are given in main text).

\[
\Gamma_p = H_p^{[R]}(\phi_p) + H_p^{[R]}(\phi_p) \otimes \Gamma_p^{[R]}.
\]

The superscript \([R] = [R_e], [R_i] \) indicates whether or not the receptor is voltage dependent. Thus if \( p = e \), \( R_e \) indicates characteristics of an AMPA receptor, and \( R_i \) an NMDA receptor, while if \( p = i \), then \( R_i \) indicates a GABAa receptor.
and no voltage-dependent channels are considered.

\[ G_p^{[R]} = k_p^{[R]} e^{-(R_p) \phi_p}, \]
\[ G_p^{[V]} = k_p^{[V]} e^{-(R_v) \phi_p} \times Q_q / Q_q^{max}. \]

The coefficients \( k_p^{[R]} \) describe the relative amplitude of the two components of \( G_p \), and the multiplication by \( Q_q / Q_q^{max} \) is introduced in the second of the above equations because voltage-dependent receptors are active only in that fraction of neurones, which are currently firing.

The onset and offset of this modulation of synaptic gain are described by normalized time functions, analogous to \( H_p^{[R]}(t) \):

\[ H_p^{[R]}(t) = \left[ B_1^{[R]} / \beta_1^{[R]} + B_2^{[R]} / \beta_2^{[R]} + \ldots - A_1^{[R]} / \alpha_1^{[R]} \right]^{-1} \times [B_1^{[R]} e^{-\beta_1^{[R]} t} + B_2^{[R]} e^{-\beta_2^{[R]} t} + \ldots - A_1^{[R]} e^{-\alpha_1^{[R]} t} - A_2^{[R]} e^{-\alpha_2^{[R]} t} - \ldots]. \]

These modulations of synaptic gain help to maintain cortical stability close to the boundary between 1/f activity, and active gamma, since AMPA activity is more strongly downregulated with decreasing membrane voltage than is GABA activity. The voltage dependence of NMDA activity contributes a type of dynamic gain, enhancing synchronous oscillation.

The simulation results shown in Figures 3 and 4 applied random perturbations of the pulse densities \( Q_q(r, t) \) in accord with the assumed stochastic nature of individual neurone firing. In Figures 5, 6, and 7, this perturbation was not applied, although the moving bars or static inputs were treated as zero-mean signals of white-noise type. The theoretical functions applied to derive the results in Figures 9 and 10 assumed driving of the model system by diffuse white noise. These differences represent, in part, stages in the evolution of the family of models rather than differences of principle. All relevant properties are retained in simulations that include the more complex treatments of the sources of driving.

The cortical simulations were applied with the cortical surface 'lumped' into a 20 \( \times \) 20 matrix of elements, with toroidal boundary conditions. Variation of boundary conditions is without significant effect.
Neurophysical Modeling of Brain Dynamics

A recent neurophysical model of brain electrical activity is outlined and applied to EEG phenomena. It incorporates single-neuron physiology and the large-scale anatomy of corticocortical and corticothalamic pathways, including synaptic strengths, dendritic propagation, nonlinear firing responses, and axonal conduction. Small perturbations from steady states account for observed EEGs as functions of arousal. Evoked response potentials (ERPs), correlation, and coherence functions are also reproduced. Feedback via thalamic nuclei is critical in determining the forms of these quantities, the transition between sleep and waking, and stability against seizures. Many disorders correspond to significant changes in EEGs, which can potentially be quantified in terms of the underlying physiology using this theory. In the nonlinear regime, limit cycles are often seen, including a regime in which they have the characteristic petit mal 3 Hz spike-and-wave form.


Keywords: brain dynamics; EEG; biophysics; modeling

INTRODUCTION

Correlations of EEGs with brain function are widely used diagnostically, and close connections to dynamics, cognition, and mental disorders are inferred (Niedermeyer and Lopes da Silva, 1999). Yet the detailed link between EEGs and the underlying physiology is not well understood, despite over 125 years’ work (Niedermeyer and Lopes da Silva, 1999). Similar remarks apply to evoked response potentials (ERPs). Still more cryptic are EEGs seen in disorders, including epilepsy, whose relation to normal EEGs is not understood. As a result EEG studies are not integrated within any overall framework, nor with other branches of neuroscience.

EEGs result from cortical electrical activity aggregated over scales much larger than individual neurons or that can be modeled using neural networks. Hence, in one class of models averages are taken over microscopic neural structure to obtain continuum descriptions on scales of millimeters to the whole brain, incorporating realistic anatomy of separate excitatory and inhibitory neural populations (pyramidal cells and interneurons), nonlinear neural responses, multiscale interconnections, dendritic, cell-body and axonal dynamics, and corticothalamic feedback (Wilson and Cowan, 1973; Lopes da Silva et al, 1974; Nunez, 1974, 1995; Freeman, 1975; Steriade et al, 1990; Wright and Liley, 1996; Jirsa and Haken, 1996; Robinson et al, 1997, 2001, 2002; Rennie et al, 2002).

We have developed a physiologically based continuum model of corticothalamic dynamics that reproduces and unifies many features of EEGs, including the discrete spectral peaks in the slow wave, ‘delta’, ‘theta’, ‘alpha’, and ‘beta’ bands, seen in waking and sleeping states (Robinson et al, 2001, 2002), ERPs (Rennie et al, 2002), measures of coherence (Robinson, 2003), generalized epilepsies, EEG entrainment and seizure activation by stimuli (Robinson et al, 2002), and low-dimensional seizure dynamics (Robinson et al, 2002). Many behaviors were derived from moderate parameter changes of a few mechanisms in a single model, thereby enabling classification of different states using these parameters. Our approach averages over microstructure to yield mean-field equations in a way that complements cellular-level and neural-network analyses: these other approaches can be employed to elucidate the connections between microstructure and mean-field quantities, while the large-scale fields provide the background against which microscopic neural activity takes place. In the following sections we outline our model and its main results to date.

METHOD

Corticothalamic Model

The details of the model have been published elsewhere (Robinson et al, 2002); owing to space limitations, we restrict ourselves here to a brief outline.
The first feature incorporated is the neural response to the cell-body potential. Mean firing rates $Q_a$ of excitatory $(a = e)$ and inhibitory $(a = i)$ neurons are nonlinearly related to mean potentials $V_a$ by $Q_a(r, t) = \Sigma [V_a(r, t)]$, where $\Sigma$ is a sigmoidal function that increases from 0 to $Q$ as $V_a$ increases from $-\infty$ to $+\infty$. We use

$$\Sigma [V_a(r, t)] = Q[1 + \exp\{-[V_a(r, t) - \theta]/\sigma']^{-1},$$

(1)

where $\theta$ is the mean neural firing threshold and $\sigma'/\sqrt{3}$ is its standard deviation.

The potential $V_a$ results after dendritic inputs have been filtered and smeared out in time while passing through the dendritic tree, then summed. It obeys a differential equation (Robinson et al., 1997, 2001)

$$D_s V_a(r, t) = v_a e \phi_e(r, t) + v_a i \phi_i(r, t)$$

$$+ v_a a \phi_a(r, t - t_0/2),$$

(2)

where $\phi_e$ is the strength of the response to a unit signal from other cortical neurons, and inputs $\phi_i$ from thalamic relay nuclei, delayed by a time $t_0/2$ required for signals to propagate from thalamus to cortex. In (2) $V_a = N_{ab} s_b$, where $N_{ab}$ is the mean number of synapses from neurons of type $b = e, i, s$ to type $a = e, i$ and $s_b$ is the strength of the response to a unit signal from neurons of type $b$.

Each part of the corticothalamic system produces a field $\phi_a$ of pulses that travels at $v = 10 \text{ m s}^{-1}$ and obeys a damped wave equation (Robinson et al., 1997, 2001):

$$\left(\frac{1}{\gamma_a} \frac{d^2}{dt^2} + \frac{2}{\gamma_a} \frac{d}{dt} + 1 - r_a^2\gamma_a^2\right) \phi_a(r, t) = \Sigma [V_a(r, t)],$$

(4)

where $\gamma_a = \sqrt{v r_a}$ and $r_a$ is the mean range of axons $a$. If intracortical connectivities are proportional to the numbers of synapses involved, $V_i = V_e$, and $Q_i = Q$, (Wright and Liley, 1996; Robinson et al., 2001), which let us concentrate on excitatory quantities. The smallness of $r_i$ also lets us set $\gamma_i \approx \infty$ (Robinson et al., 1997).

The model incorporates corticothalamic connectivities and thalamic nonlinearities. Figure 1 shows the connectivities considered, including the thalamic reticular nucleus that inhibits relay nuclei. The latter relay external stimuli $\phi_a$ to the cortex as well as corticothalamic feedback. The cell-body potentials then satisfy

$$D_s V_e(r, t) = v_e e \phi_e(r, t - t_0/2) + v_e i \phi_i(r, t)$$

$$+ v_e a \phi_a(r, t) + v_e \phi_e(r, t),$$

(5)

where there is a delay $t_0/2$ for signals to travel from cortex to thalamus, $c = r, s$, $V_e = V_m = 0$, and $\phi_e(r, t) = \Sigma [V_e(r, t)]$ (Robinson et al., 2001) applies because the small size of the thalamic nuclei enables us to assume $\gamma_e \approx \infty$ and $r_e \approx 0$ in (4).

Our model has 15 parameters: $Q, \theta, \sigma', z, \beta, \gamma, \delta, r_e, s_e, v_e, \phi_e, \phi_i, \phi_a, v_a$, and $k$ enough to allow realistic representation of the anatomy and physiology, but few enough to yield useful interpretations. The parameters are approximately known from the experiment (Robinson et al., 2001, 2002), leading to the nominal values in Table 1, which are indicative only—some vary severalfold between individuals, arousal states, and disorders. We use only values compatible with physiology.

**RESULTS**

**Steady States, Linear Waves, and Stability**

Setting all derivatives to zero in (3) and (4) yields steady states when the system is driven by a constant, uniform mean stimulus level $\phi_0$. The equations are easily solved numerically, yielding a single low-$\phi_e$ solution, which corresponds to a normal state.

Small perturbations of steady states allow use of linear analysis. A stimulus $\phi_a(k, \omega)$ of angular frequency $\omega$ and wave vector $k$ has the transfer function to $\phi_a(k, \omega)$:

$$\phi_a(k, \omega) = \frac{G_{ee} L}{1 - G_{ee} L} \frac{G_{ai} L e^{-i\omega t_0/2}}{q^2 r_e^2 + k^2 r_e^2},$$

(6)

$$q^2 r_e^2 = (1 - \omega^2/\gamma_e^2)^2$$

$$- \frac{L}{1 - G_{ee} L} \left[\frac{(S_a + S_i) L}{1 - S_i^2 L} e^{-i\omega t_0}\right],$$

(7)

$$G_{ab} = (\phi_a/\sigma') (1 - \phi_a/Q) v_{ab},$$

(8)

Neuropsychopharmacology

PA Robinson et al
where $L = (1 - i\omega/\alpha)^{-1}(1 - i\omega/\beta)^{-1}$ and $\phi_r$ is the steady-state value of $\phi_r$. This function is the cortical excitatory response per unit external stimulus, and encapsulates the relative phase via its complex value (Robinson et al, 2001; Rennie et al, 2002); it is the key to linear properties of the system. The gain $G_{ab}$ is the differential output produced by neurons $a$ per unit input from neurons $b$, and the static gains for loops in Figure 1 are $S_d = G_{ed}G_{ee}$ for feedback via relay nuclei only, $S_r = G_{es}G_{re}$ for the loop through reticular and relay nuclei, and $S_i = G_{ei}G_{ri}$ for the intrathalamic loop.

Waves obey the dispersion relation $q^2(\omega) + k^2 = 0$, with instability boundaries where this equation is satisfied for real $\omega$ (Robinson et al, 1997, 2001). In most circumstances, waves with $k = 0$ (ie spatially uniform) are the most unstable (Robinson et al, 1997), and it is found that only the first few (ie lowest frequency) spectral resonances can become unstable. Analysis for realistic parameter ranges finds just four $k = 0$ instabilities, leading to global nonlinear dynamics (Robinson et al, 2002): (a) Slow-wave instability ($f = 0$) which leads to a low frequency spike-wave limit cycle. (b) Theta instability, which saturates in a nonlinear limit cycle near 3 Hz (see Figure 2(a), where $t_0 = 0.2$ s for clarity, giving a frequency nearer 2 Hz (see below)), with a spike-wave form unless its parameters are close to the instability boundary. (c) Spindle instability at $\omega \approx (\beta/\alpha)^{1/2}$ (see Figure 2, in the alpha band for physiological $\alpha$ and $\beta$, leading to a limit cycle near 10 Hz. (d) Alpha instability giving a limit cycle near 10 Hz, with a waveform like in Figure 2(b).

The occurrence of only a few instabilities, at low frequencies, enables the state and physical stability of the brain to be parameterized in a 3D space with axes

\begin{align}
x &= G_{ee}/(1 - G_{ei}), \\
y &= (S_d + S_r)/[(1 - S_r)(1 - G_{ei})], \\
z &= -S_r\alpha\beta/(\alpha + \beta)^2.
\end{align}

Figure 2 Sample time series from the model in regimes corresponding to (a) theta instability and (b) spindle instability.

Figure 3 Stability zone for nominal parameters in Table 1, except $x = 60 s^{-1}$. The surface is shaded according to instability, as labeled (dark gray = spindle, light gray at right = alpha, light gray at left = theta), with the front right face left transparent as it corresponds to a slow-wave instability. Approximate locations are shown of EO, EC, S1, S2, S4, REM (R), anesthesia (A), and alpha coma (C) states, petit mal onset (P), and the parameters in Table 1 (T), with each state located at the top of its bar, whose $x-y$ coordinates can be read from the grid.

Figure 4 shows excellent agreement with observed spectra if is white noise in space and time, including occurrence of alpha and beta rhythms at frequencies $f \approx 1/t_0$, $2/t_0$, and the asymptotic low- and high-frequency behaviors; key differences between waking and sleep spectra can also be reproduced (see below and Robinson et al, 2001). The low-frequency 1/f behavior is a signature of edge-of-stability dynamics, which allow complex behavior (Robinson et al, 1997, 2001). To test our model and estimate some of its parameters, we fitted its linear spectrum to 103 normal adults’ eyes-closed (EC) and eyes-open (EO) spectra (Robinson et al, 2002). This yielded mean parameters near those in Table 1.

A 1D wave-number spectrum results if one integrates $|\phi_r(k, \omega)|^2$ over one component of $k$:

\begin{equation}
P(k, \omega) = |\phi_r(k_x, k_y, \omega)|^2dk_y.
\end{equation}

which parameterize cortical, corticothalamic, and thalamic stability, respectively (Robinson et al, 2002). In terms of these quantities, the brain occupies a stability zone illustrated in Figure 3. The back is at $x = 0$ and the base at $z = 0$. A pure spindle instability occurs at $z = 1$, which couples to the alpha instability on the upper boundaries, with spindle dominating at top and left, and alpha at right. At small $z$ the left surface is defined by a theta instability. The front right surface corresponds to slow-wave instability and follows the plane $x + y = 1$ to $y = y_c \approx -0.2$. The boundaries correspond to onsets of generalized seizures (Robinson et al, 2002).

Spectra, Evoked Potentials, and Coherence

The EEG frequency spectrum is obtained by squaring the modulus of $\phi_r$ and integrating over

\begin{equation}
P(\omega) = \int |\phi_r(k, \omega)|^2dk.
\end{equation}

Figure 4 shows excellent agreement with observed spectra if is white noise in space and time, including occurrence of alpha and beta rhythms at frequencies $f \approx 1/t_0$, $2/t_0$, and the asymptotic low- and high-frequency behaviors; key differences between waking and sleep spectra can also be reproduced (see below and Robinson et al, 2001). The low-frequency 1/f behavior is a signature of edge-of-stability dynamics, which allow complex behavior (Robinson et al, 1997, 2001). To test our model and estimate some of its parameters, we fitted its linear spectrum to 103 normal adults’ eyes-closed (EC) and eyes-open (EO) spectra (Robinson et al, 2002). This yielded mean parameters near those in Table 1.
We find that this spectrum is flat at small \( k_x \), then approximates a power law, with \( P(k_x, \omega) \sim k_x^{-g(\omega)} \).

Figure 5 compares the exponent with EC data from Shaw (1991), showing excellent agreement for physiologically realistic parameters, except at \( f \sim 5 \text{ Hz} \), where the data are affected by phase distortions (Shaw 1991). In particular, spectral steepenings at resonances are reproduced.

The inverse Fourier transform of (6) gives the ERP that results from an impulse stimulus. Initial work shows that the result agrees well with the experiment (Rennie et al., 2002), as illustrated by the example of a background (as opposed to target, in an oddball paradigm) ERP in Figure 6. Significantly, the model parameters used are essentially the same as those that reproduce the same subject’s prestimulus EEG spectrum.

The cross spectrum \( P(r, r', \omega) \) is the phase average of \( \phi_{r, \omega}(r, r') \). The coherence function is then

\[
\gamma^2(r, r', \omega) = \frac{|P(r, r', \omega)|^2}{P(r, r, \omega)P(r', r', \omega)}. \tag{14}
\]

Figure 7 shows that this gives good agreement with observations for model parameters close to those used in obtaining the other plots in this work (Robinson, 2002). A rise seen in \( \gamma^2 \) at large \( R = |r - r'| \) (Nunez et al., 1999) is also reproduced (Robinson, 2002).

States of Arousal

Nonseizure states lie within the stability zone in Figure 3. Detailed arguments regarding the arousal sequence, from alert to deep sleep, and including REM sleep and sleep stages 1–4 (S1–S4), constrain the relevant regions of parameter space and place this sequence as shown in

---

Neurophysiological modeling of brain dynamics

PA Robinson et al
States such as anesthesia can also be represented.

Figure 8 shows model time series for parameters illustrating EO, EC, S2, and S4 states, holding $Q_e$, $\gamma_e$, $t_0$, $b/a$, $n_{ei}$, and $n_{sn}$ at their nominal values, and varying $\alpha$ and the other $v_{ab}$ only moderately. The features seen strongly resemble those of corresponding experimental data (Niedermeyer and Lopes da Silva, 1999). Known differences between EEG spectra for subjects with differing disorders enable classification of these conditions into different parts of the stability zone, while seizures correspond to departures from this zone, as discussed next.

Petit Mal Seizures

Petit mal is a common generalized epilepsy. Seizures are mostly seen at ages 4–20, last 5–20 s, cause loss of consciousness, and show a spike-wave cycle which starts and stops abruptly across the whole scalp (Niedermeyer and Lopes da Silva, 1999). The frequency falls from around 4 Hz to under 3 Hz in most cases, and non-REM sleep is a powerful seizure activator. Experiments show that the loops in Figure 1 are essential for petit mal, with the cortex synchronizing thalamic activity (Niedermeyer and Lopes da Silva, 1999; Steriade et al., 1990). GABA antagonists such as penicillin can start spike-wave oscillations, in some cases converting spindles to spike-wave complexes, similar to those also seen in some partial seizures (Niedermeyer and Lopes da Silva, 1999).

Our model gives an $\approx 3$ Hz spike-wave cycle as the nonlinear stage of theta instability (Figure 2) and we conclude that this corresponds to petit mal (Robinson et al., 2002). Analysis shows that this cycle consists of a flip-flop (it alternates between two states) in the limit $\gamma_e, \alpha \to \infty$, a residue of which is seen in Figure 2(a). The high-$\phi_e$ part corresponds to large $\phi_e$ incident on the cortex as a result of low $\phi_i$ a time $t_0/2$ earlier and low $\phi_e$ a time $t_0$ earlier; the low-$\phi_e$ part corresponds to the converse, with near silence in relay nuclei, as seen experimentally. Signals make two circuits of the system before it returns to its original state, giving a period $2t_0$. At finite $\alpha$, signals traveling via the reticular nuclei are delayed by $\approx 1/\alpha$ more than those that only pass through relay nuclei. Hence, when $\phi_e$ flips to its upper state, there is a short period $t_0/2$ later when $S_{\phi} > |S_i|$, resulting another $t_0/2$ later in a spike of duration $\approx 2^{-1}$. Finite $\alpha$ and $\gamma_e$ also round off the other side of each square wave and finite $\alpha$ leads to damped spindle oscillations at $\omega = (\alpha \beta)^{1/2}$. These mechanisms accord with the experimental inferences above, and observed EEGs often show a residual flip-flop plateau in each cycle (Niedermeyer and Lopes da Silva, 1999; Robinson et al., 2002).

Estimation of the petit mal period gives $T \approx 2t_0 + 6/\alpha + 6/\beta + 4/\gamma_e$ (Robinson et al., 2002), consistent with observations and the insensitivity of $T$ to most parameters. The main features of the waveform, apart from spindles, are found in cases with $\beta, \gamma_e = \infty$ but finite $\alpha$, which implies that the 3D system resulting in that limit contains the essential dynamics. This accords with findings that dimensions of time series of petit mal and related seizures are low (Babloyantz and Destexhe, 1986).

A typical onset point for a petit mal seizure is shown in Figure 3. Transformation of spindles into petit mal is inferred to occur by moving from the vicinity of S2 to the theta instability zone, with a rapid switch of activity from roughly 10 Hz to 3 Hz. Large values of $\gamma_e$ favor instability,
which may explain the onset of petit mal at around age 4, since \( \gamma_e \) rises in children because of myelination.

**DISCUSSION**

We have developed a model that incorporates the main features of corticothalamic physiology and anatomy using only 15 parameters. Its predictions provide a unified description of a wide range of phenomena, with six parameters fixed across all states, and the others only varying moderately. Of key importance is the xyz parameter space in which the stability zone of the brain is easily visualized, and in which disorders, states of arousal, etc, can be classified. Within this zone, linear analysis gives good approximations to EEG spectra, ERPs, coherence and correlation functions, and related measures. Zone boundaries are identified with onsets of generalized seizures, consistent with known features of their time series and patterns of occurrence. The model also explains low-dimensional dynamics in petit mal seizures and other nonlinear behaviors such as alpha entrainment and seizure activation are also reproduced (Robinson et al, 2002). A key feature of our approach is that we extract a broad range of behavior from modest changes in the parameters of a single model, without postulating extra mechanisms.

Fitting the model’s predictions to data provides a noninvasive probe of large-scale physiology that yields parameter values consistent with independent measures. This enables states of arousal, seizure onsets, and pathologies to be assigned to distinct regions of parameter space. We have found that the normal arousal sequence has a simple form in the xyz space, that clinically observed waking states lie at inferred locations, and that seizure onsets lie close to the most commonly seen precursor states (see also Robinson et al, 2002). This space thus provides a physiologically based organizing framework for a wide variety of phenomena. Its topography may indicate new connections among phenomena in neighboring regions, and enable the significance of the parameters that distinguish the various cases, and cause transitions between them, to be studied systematically.

The approach discussed here provides a powerful framework for further studies: It remains to investigate what factors control progression along inferred arousal sequences, or onset of seizures, for example. This will require inclusion of neuromodulator dynamics and brainstem reticular activation, possibly using additional feedback loops. Spatial variations are also under investigation, with successful initial application to understanding split alpha peaks seen in a few percent of the population (Robinson et al, 2003).

**ACKNOWLEDGEMENTS**

This work was supported by the University of Sydney’s Sesqui Grant Scheme and the Denison Bequest.

**REFERENCES**


Generation and control of cortical gamma: findings from simulation at two scales

J.J. Wright *

Liggins Institute, and Department of Psychological Medicine, University of Auckland, Auckland, New Zealand
Brain Dynamics Centre, University of Sydney, Sydney, Australia

A R T I C L E   I N F O
Article history:
Received 8 October 2007
Received in revised form 15 April 2008
Accepted 6 November 2008

Keywords:
Gamma activity
Synchronous oscillation
Cortical self-regulation
EEG

A B S T R A C T
A continuum model of electrocortical activity was applied separately at centimetric and macrocolumnar scales, permitting analysis of interaction between scales. State equations included effects of retrograde action potential propagation in dendritic trees, and kinetics of AMPA, GABA and NMDA receptors. Parameter values were provided from independent physiological and anatomical estimates. Realistic field potentials and pulse rates were obtained, including resonances in the alpha/theta and gamma ranges, 1/f² background activity, and autonomous gamma activity. Zero-lag synchrony and travelling waves occurred as complementary aspects of cortical transmission, and lead/lag relations between excitatory and inhibitory cell populations varied systematically around transition to autonomous gamma oscillation.

Properties of the simulations can account for generation and control of gamma activity. All factors acting on excitatory/inhibitory balance controlled the onset and offset of gamma oscillation. Autonomous gamma was initiated by focal excitation of excitatory cells, and suppressed by laterally spreading trans-cortical excitation, which acted on both excitatory and inhibitory cell populations. Consequently, although spatially extensive non-specific reticular activation tended to suppress autonomous gamma, spatial variation of reticular activation could preferentially select fields of synchrony.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

This paper specifies continuum state equations for simulation of local field potentials and population pulse rates in the cerebral cortex. The model thus obtained is used to help explain the origin and control of gamma synchrony in relation to global electrocortical activity.

In the attempt to capture essential aspects of cortical dynamics, models of neuronal interactions range from very detailed simulations of individual neurons then studied in interacting networks (e.g. Bower and Beeman (1998) and Traub, Whittington, Stanford, and Jeffereys (1996), to classical approaches utilizing highly simplified neurons (e.g. Amit (1989), Arbib (1995) and Buzsaki and Draguhn (2004)). Problems of numerical complexity and loss of analytical advantage, versus loss of physiological realism, arise at either end of this spectrum. A further, and complementary, approach has been motivated by studies of electroencephalogram (EEG) and local field potentials, and is variously described as mean-field, continuum, or population approximation (Freeman, 1975; Haken, 1996; Nunzi, 1981, 1995; Rotterdam, Lopes da Silva, van den Ende, Viergever, & Hermans, 1982; Wilson & Cowan, 1973).

Continuum models treat the cortical medium as a continuous field, and offer the possibility that the field equations might be applied at different spatial scale and resolution, by appropriate adjustment of parameters and connections. In the present work, for the first time, a single parameter set obtained a priori is applied to continuum simulations to test whether the results exhibit appropriate scale-dependent effects. Also for the first time, action potential retrograde propagation into the dendritic tree is included within continuum state equations – an inclusion motivated by the importance of retrograde propagation for the ascription of synaptic weight in computational models, in addition to the dynamic implications – and effects of receptor dynamics of three major neurotransmitter types are included. Parameter values and methods of application are drawn from various sources (Braitenberg & Schuz, 1991; Jirsa & Haken, 1996; Liley & Wright, 1994; Mountcastle, 1979; Rennie, Robinson, & Wright, 1999; Rennie, Wright, & Robinson, 2000; Rennie, Robinson, & Wright, 2002; Robinson, Rennie & Wright, 1997; Robinson et al., 2001, 2003; Robinson, Rennie, Rowe, & O’Connor, 2004; Scholl, 1956; Steyn-Ross et al., 2005; Suffczynski, Kalitzin, Pfurtscheller, & Lopes da Silva, 2001; Szentagothai, 1979; Wright, 1990; Wright & Liley, 1996; Wright, 1997, 1999; Wright, Bourke, & Chapman, 2000; Wright et al., 2003).

In a continuum description, stable, non-fluctuating states are equivalent to random fluctuations in individual cell-firing rates in a network description. Conversely, spontaneous oscillating states in a continuum indicate ordered firing in an equivalent neural population. Therefore, transitions between stable and
spontaneously oscillating states were studied in simulation, and manipulation of parameters was used to obtain insights into mechanism. Analogs of gamma oscillation and synchronous oscillation (Bressler, Coppola, & Nakamura, 1993; Eckhorn et al., 1988; Gray & Singer, 1989; Gray, Konig, Engel, & Singer, 1989; Gray, Engel, Konig, & Singer, 1992; Mittner, Braun, Arnold, Witte, & Taube, 1999; Neuenhewender & Singer, 1996; Singer, 1994; Singer & Gray, 1995) were targets of central importance, because field potentials and pulse activity are strongly correlated in the gamma frequency band (Stryker, 1989). There are theoretically important relations of synchrony to information storage and retrieval (Kay & Phillips, 1997; Malsburg, 1983; Phillips & Singer, 1997) including recent considerations of synaptic self-organization (Wright, Alexander, & Bourke, 2006; Wright & Bourke, 2008), as well as wave aspects of cortical signal transmission (Chapman, Bourke, & Wright, 2002; Freeman & Barrie, 2000; Freeman, Holmes, West, & Vanhatalo, 2006; O’Connor & Robinson, 2003; Robinson, Wright, & Rennie, 1998).

Freeman and co-workers have long considered gamma activity to involve transitions comparable to a thermodynamic phase transition, and emphasise the propagation of the phase transition from site to site in the cortex. Yet the mechanisms of origin and the control of gamma oscillation are yet to be clarified. During gamma oscillation an average phasic relation exists between local excitatory and inhibitory cells (Freeman, 1991; Hasenstaub et al., 2008). The properties of the simulations reported here suggest that recurrent inhibition from fast spiking inhibitory cells is largely responsible for maintaining the rhythmic drive, although the role played by excitatory processes in modulating or driving the oscillations remains uncertain. Without determination of these mechanisms it remains difficult to comprehensively define the link of gamma activity to information storage, retrieval and transmission.

Since cortical networks form a dilute and massively interconnected network, a satisfactory explanation for gamma activity and synchrony should not only consider local dynamics, but also explain the onset and offset of gamma activity in relation to events at more distant sites and at larger scale in the brain, including cortical regulation mediated via subcortical mechanisms. Many studies have indicated EEG correlations with cognition, and cortical/subcortical interactions are regarded as critical for the regulation of arousal and attention (e.g., Alexander, Crutcher, and DeLong (1990) and Scheibel and Scheibel (1970)) but the mechanisms remain elusive. The need to consider these relationships was highlighted by Stryker (1989), when, in contrast to gamma, he drew attention to the relatively weak relationship of other, non-gamma, EEG rhythms to variation of cortical pulse rates. It still remains unclear in what way large-scale EEG potentials are involved in cortical information processing, if at all.

The properties of the simulations reported here suggest resolutions of some of these problems.

2. Methods

Fig. 1 shows the qualitative physiological and anatomical features included in the simulations.

2.1. Conventions

State variables are average membrane potentials, $V_{p,q}(r,t)$, pulse densities, $Q_{p,q}(r,t)$, and afferent synaptic flux $\varphi_{p,q}(r,t)$. To enable a compact representation, the subscripts $p,q = e, i$ indicate either excitatory ($e$) or inhibitory ($i$) neuron populations, while $qp$ indicates synaptic connections from $p$ to $q$. Superscripts $[R] = [\text{NMDA}], [\text{AMPA}], [\text{GABA}_e]$ indicate neurotransmitter receptor types.

State equations are steady state functions of state variables, and lag response functions in $\tau = n\delta t$, where $\delta t$ is the time-step, and $n = 1, 2, \ldots$ Lag response functions are normalised so that $\int_0^\infty f(\tau) \cdot d\tau = 1$. This format reflects the way that the parameter set was derived from experimental data, as is recognisable in Fig. 2. Parameter values and notes on derivations are given in the Appendix, except for axonal ranges and conduction velocities, which are discussed in Section 2.9.

2.2. Afferent synaptic flux

The distribution of the neuron cell bodies giving rise to afferents at a cortical point $r$ is $f(r, r')$, where $(r')$ are all other points in the field. Connection densities are reciprocal for all $(r, r')$.

The afferent flux density, $\varphi_{p}(r, t)$, the population average input pulse rate per synapse is given by

$$\varphi_{p}(r, t) = \int f(r, r') Q_{p}(r', t) - |r - r'|/\nu_p) d^2 r'$$  \hspace{1cm} (1)$$

$Q_{p}(r', t)$ are mean pulse rates of neurons at $r'$ at $t$, also termed pulse density. $\nu_p$ is the velocity of axonal conduction.

$$f(r, r') \approx \text{Gaussian (e.g. Braitenberg and Schuz (1991)}$$

$$f(r, r') = \frac{1}{2\pi \gamma^2} \exp \left[ -\frac{|r - r'|^2}{2\gamma^2} \right]$$  \hspace{1cm} (2)$$

$\gamma$ is the standard deviation of axonal range.
2.3. Synaptic receptor dynamics

The postsynaptic impact of $\varphi_p(r, t)$ is modified by changes in the conformation of ion channels. The open channel steady state is

$$j^{\text{(o)}}(\varphi_p) = \exp[-\lambda^{\text{(o)}}(\varphi_p)]\varphi_p$$

(3)

(see Fig. 2(1b)) and $\Phi^{(o)}(\tau)$ (Fig. 2(1a)) describes rise and fall of receptor adaptation to a brief afferent stimulus

$$\Phi^{(o)}(\tau) = \left[ \sum_{n} B_n^{(o)}/\alpha_n^{(o)} - \sum_{m} A_m^{(o)}/\alpha_m^{(o)} \right]^{-1}$$
Fractional distributions, \( r_i[R] + r_f[R] = 1 \), of postsynaptic receptors of each type differ in near and far trees, so back-propagation also influences the efficacy of receptor types, and the voltage dependence on NMDA receptors requires that they be considered as essentially components of the distal tree, with \( r_i[NMDA] = 1 \).

2.7. Aggregate depolarisation

The voltage at the trigger points for action potential generation, \( \psi_q \), is obtained by convolution and summation over the receptor types, excitatory/inhibitory cell combinations, and fractions of quiescent and recently active cells, weighted by the average number of synaptic connections between cell types, \( N_q \). Where \( * \) indicates convolution in time

\[
\psi_q(t) = \sum_p \sum_j N_{qp} A_p^R (\{ (M_q^R \ast \Phi_q^R) \ast \psi_q^R \} \ast L^R).
\]

In the population average,

\[
V_q(t) \approx V_q^{0[R]} + \psi_q
\]

and Eq. (11) establishes \( V_q(t - \delta t) \) for the next time-step in Eq. (5).

2.8. Action potential generation

From Eq. (11) the mean firing rate is calculated from

\[
Q_q(t) = Q_q^{\text{max}} / (1 + \exp[-\pi (V_q - \theta_q) / \sqrt{3} \sigma_q])
\]

(Fig. 2(3b)), yielding the pulse densities of neurons required in Eq. (1).

\( \theta_q \) is the mean value of \( V_q \) at which 50% of neurons are above threshold for the emission of action potentials.

\( \sigma_q \) approximates one standard deviation of probability of emission of an action potential in a single cell, as a function of \( V_q \).

For comparison with standard EEG and local field potential (LFP) data, we also assume \( LFP \equiv V_q(t) \).

2.9. Application at different spatial scales

The above equations were applied numerically in spatially discrete form, in a 20 × 20 grid of “elements”, with periodic boundary conditions. Each element of the grid is situated at a position \( r \), surrounded by other elements at positions \( \{ r' \} \) and coupled as in Eq. (1) with delays, \( \delta_q = |r - r'| \). The grid can be used to represent the cortex at any chosen spatial scale, by choosing lengths for \( |r - r'| \), applying a physiologically appropriate value for \( \psi_q \), and setting excitatory/inhibitory axonal connection ranges appropriate to scale.

2.9.1. Centimetric scale

Because of the relatively short range of inhibitory axons, at large scales a coarse-grained approximation can be applied; \( \gamma_e \gg \gamma_i \), so for \( p = e \)

\[
\psi_e(r, t) = \int \frac{1}{2\pi\gamma_e^2} \exp \left[ -\frac{|r - r'|^2}{2\gamma_e^2} \right] Q_e(t - \delta_q) d^2 r'
\]

and for \( p = i \)

\[
\psi_i(r, t) = Q_i(r, t).
\]

An axonal conduction velocity was selected by considering the 20 × 20 grid approximate in total size to the human cortex. Assuming a cortical diameter of approximately 15 cm, and allowing a three-fold cortical folding factor, each grid element...
represents a square area of cortex 0.7 cm on the side. A 1.7 ms delay per element then implies an axonal conduction velocity of approximately 4 m/s. Nunez (1995) gives a conduction velocity of 6–9 m/s for myelinated cortico–cortical fibres, noting that a substantial proportion of unmyelinated fibres are present, implying a lower average conduction velocity.

The range parameter \( \gamma_e \) was set at 4 elements, in rough approximation of cortico–cortical connection range, and \( \gamma_i \) to zero.

### 2.9.2. Macrocolumnar scale

At the scale of a cortical macro-column (approximately 300 \( \mu \) diameter) (Braitenberg & Schuz, 1991) intra-cortical excitatory and inhibitory connections have roughly equal ranges, forming a “Mexican Hat” field (Kang, Shelley, & Sompolinsky, 2003). The action of long-range excitatory (cortico–cortical) connections provides an external source of afferent flux, here called the trans-cortical flux (TCF, see Fig. 1). With the additional subscripts ic or cc to indicate an intra-cortical or a cortico–cortical subclass, respectively, synaptic fluxes generated within the macro-column are

\[
\varphi_{ic}(\mathbf{r}, t) = \int \frac{1}{2\pi \gamma_{ic}^2} \exp \left[ -\frac{|\mathbf{r} - \mathbf{r}'|^2}{2\gamma_{ic}^2} \right] Q_{ic}(t - \delta_i) d^2 \mathbf{r}' \\
\varphi_{cc}(\mathbf{r}, t) = \int \frac{1}{2\pi \gamma_{cc}^2} \exp \left[ -\frac{|\mathbf{r} - \mathbf{r}'|^2}{2\gamma_{cc}^2} \right] Q_{cc}(t - \delta_i) d^2 \mathbf{r}'.
\]

Axonal conduction velocity in this case was chosen using 0.3 ms delay per element, which corresponds to a conduction velocity of 0.95 m/s, consistent with slow unmyelinated intracortical fibres. Range parameter \( \gamma_{ic} \) was 4.9, and \( \gamma_{cc} \) was 4.5 elements, so that both types of connection spanned the macro-column.

The trans-cortical flux, \( \varphi_{cc}(\mathbf{r}, t) \) generated outside the macro-column, and delivered to both excitatory and inhibitory neurons via terminal arborisation is roughly spatially uniform at the macro-column's scale and equivalent to the spatial-average excitatory pulse density at centimetric scale. That is;

\[
\varphi_{cc}(\mathbf{r}, t) = \varphi_{cc}(\mathbf{r}', t).
\]

To first approximation, the trans-cortical flux can be added to the intracortical fluxes, to emulate the effect of the macro-column's embedding within the extended cortex. Set at a constant value trans-cortical flux can then be used as a control parameter.

### 2.10. Non-specific cortical activation

A second control parameter, \( \psi_{exc} \) (less formally, \( \psi_m \)), the non-specific afferent flux (non-specific activation — NSF, see Fig. 1) from the reticular formation, is applicable at both scales. The net effect of non-specific activation is considered purely excitatory, and terminating on excitatory cortical neurons selectively, consistent with its predominant non-specific input to the upper layers of the cortex where the pyramidal cell dendritic trees predominate (Braitenberg & Schuz, 1991; Nunez, 1995; Scheibel & Scheibel, 1970). Values of \( \psi_m \) are arbitrary positive numbers since synaptic efficacies or transmitter classes are not specified here, for non-specific cortical afferents.

### 2.11. Numerical considerations

State equations were solved using integro-differential form. Programs were written in C +, and executed on an IBM mainframe computer. Simulation time-step was 0.1 ms, and it was shown that small variations of time-step were without significant effects on the results reported. Run times approximated real time. Individual simulation runs were considered to have reached a statistically stationary state at \( t = 200 \) s. Initialization was with state variables set to zero, and the final 0.8192 s of each simulation run was used to determine whether the final state was steady state (negligible power other than at DC) or one of oscillation. Single runs were used for all estimates in which external noise was not applied. With the application of noise-like driving signals, ensembles of 100 or more independently obtained 0.8192s final epochs were analysed for all spectral and correlation analyses. Noise inputs were applied as zero mean signals added to the applied constant values of non-specific flux, and unless otherwise stated, were applied to row zero of the simulation while excitatory cell potentials were recorded from the element at row 10, column 10.

### 3. Results

#### 3.1. Oscillation characteristics and comparison with gamma activity

An initial search was made to determine whether oscillation similar to gamma activity occurred at parameter values within, or close to, the standard parameter values.

Simulated gamma oscillation was defined as oscillation with a peak frequency in the 30–60 Hz band, associated with transition to autonomous oscillation occurring below a mean excitatory firing rate of 20 s\(^{-1}\), and not associated with excursions of membrane potential encountering reversal-potential bounds. Oscillations outside these limits were considered unlikely to have normal physiological analogues. Other normal firing patterns of cortical neurones (Steriade, Timofeev, & Grenier, 2001) were equated with stochastic background and identified with the simulation's non-oscillating states.

Progressive increments of non-specific afferent flux were applied in different runs, using values from zero to 1000 s\(^{-1}\), without other driving signals. In the case of the macrocolumnar scale simulations, trans-cortical synaptic flux was applied at levels from 0 to 20 s\(^{-1}\) in sets of individual runs. Runs were varied by application of a multiplication factor \((x_g^{\text{IC}})\) applied to the standard value of inhibitory gain, and by variation of axonal delay. These two parameters were selected for initial variation because they directly affected all aspects of excitatory/inhibitory balance and system time delays. Variation along those two dimensions was then used as a basis for variation of further parameters. (See Fig. 3 and Section 3.2.)

Increase of non-specific afferent flux always produced increasing cell firing rates. Either oscillation supervised at some critical level, or the system reached a state of high-firing-rate stability, in which oscillation did not occur. It was found that at both scales, and at all levels of trans-cortical synaptic flux, the threshold of oscillation increased systematically with axonal delay and decreased with multiplication of inhibition. Increase of the trans-cortical flux in the macrocolumnar scale simulation increased the threshold of oscillation, as is described in further results.

#### 3.2. Parameter sensitivities

Figs. 3a and 3b apply a classification to the classes of oscillation and high-firing-rate stability, observed with variation of parameters. The central graph of each figure is the plot for otherwise standard parameters, with variation of \((x_g^{\text{IC}})\) and axonal conduction velocity only. Surrounding this standard case, the superadded effects of variation of other parameters — those exerting effects on the distribution of receptors in the dendritic tree (\(x_{\text{AMPA}}, x_{\text{GABA}}\)), the impact of retrograde propagation \((d_2^r, d_1^r)\), and the contribution of NMDA and AMPA receptors, are shown. Less extreme variations of these parameters produced smoothly graded deviations from the central graph. In a few
instances, all involving drastic perturbation of parameters, small domains of high-firing-rate stability appear as speckling of the panels, against more consistent background trends. These speckles suggest that multiple regimes of oscillation may exist at these parameter ranges, as has been shown in systems similar to the present equations (Robinson, Rennie, Wright, & Bourke, 1998).
Variations of simulation grid size, numerical precision and time-step were also explored, and as expected, were found to produce small systematic variations from the central graph.

At centimetric scale large domains of gamma oscillation, non-gamma oscillations, and high-firing-rate stability are present. At macrocolumnar scale the same general form of the figures is present although the domain of gamma oscillation is reduced. At macrocolumnar scale little difference from the centimetric case was apparent when no trans-cortical flux was imposed. Suppression of the gamma domain was progressive, as seen in plots similar to Fig. 3b, as increasing levels of trans-cortical flux were applied.

### 3.3. Control of gamma activity

For standard parameter values, with \( g_i^{(R)} = 1 \) and axonal delays also set as described in Section 2.9 — at centimetric scale the cortex is stable to high levels of non-specific activation while at macrocolumnar scale gamma oscillation occurs at a threshold subject to both non-specific flux and trans-cortical flux. These values were used to generate all the subsequent results.

### 3.4. Power spectra

Figs. 5 and 6 show examples of characteristic power spectra. The time-series from which the spectra were derived were checked for the presence of spontaneous oscillation by repeating runs with and without the application of driving white noise.
3.4.1. Macrocolumnar scale

Fig. 5 left shows the impact of variation of trans-cortical flux while non-specific flux is constant, and Fig. 5 right shows the reverse situation. As can be seen, both trans-cortical flux and non-specific flux can regulate transition from a damped gamma resonance, to spontaneous oscillation in the gamma range. In all conditions the spectrum is strongly peaked near the gamma range, although variation from the high beta range to the high gamma occurs.

3.4.2. Centimetric scale

At centimetric scale non-specific inputs to cortex cannot be considered purely as white noise, since allowance must be made for non-stationary excursions associated with changing attention and alertness. The white noise input case is shown in Fig. 6, left, where spectral peaks in the alpha and gamma ranges are present, set against an approximately 1/f² background activity. At low levels of non-specific flux the low frequency peak predominates. With variation on the value of axonal delay upwards to 20 ms/element, the peak of the low frequency peak varies from the alpha range, as shown, to the theta range, suggesting that the value of the peak seen may depend on the dominant wavenumber sustained on the simulated cortical surface. With increasing non-specific flux (i.e., cortical activation) a progressive “shift to the right” occurs. As non-specific flux is increased to higher levels than shown in the figure, gamma peaking becomes predominant, but remains a damped resonance.

In Fig. 6, right, the input signals were band limited to very low frequency, imitating the changes in cortical tone induced during shifting focal attention. The spectrum is then dominated strongly by 1/f² activity.

3.5. Lag correlations and synchronous oscillation

3.5.1. Excitatory–excitatory timing relations

Comparison with physiological findings was made with respect to fields of synchrony and travelling waves. Two points on a simulated macrocolumn were driven with independent white noise inputs. Fig. 7 shows lag covariance between excitatory compartments in two reference elements, adjacent to each of the two driven elements. Amplitudes of the pair of inputs were adjusted in different runs so that the effect of reversing the ratio of the input amplitudes could be determined.

Below threshold of oscillation variation of the input amplitudes results in lead and lag between the reference elements, consistent with the passage of travelling waves outwards from the sites of input, while zero-lag synchrony is generated when the input magnitudes are at parity. In the presence of strong oscillation synchrony is widespread and there is no detectable occurrence of travelling waves despite disparity of input magnitudes.

Similar effects to those in the macrocolumn in the damped condition appear at centimetric scale except that the lag times of maximum covariance are correspondingly greater, due to the greater axonal conduction lags at the larger scale.

3.5.2. Excitatory–inhibitory timing relations

On average, phasic relations appear between excitatory and inhibitory cells (Hasenstaub et al., 2005) and a related comparison was made from the simulation.

Using time-series data generated in the construction of Fig. 4, lag correlations were calculated between the pulse densities of the excitatory and inhibitory compartments of a single element. As the threshold of gamma oscillation was approached and exceeded, lag correlations with a period between peaks corresponding to the gamma frequency were obtained. From the relative lead and lag of these peaks, the value \( \theta \) was calculated. This measure, rather than frequency-specific phase angle, was used since it is robust to changes in the peak frequency of oscillation occurring above and below transition to oscillation.

\[
\theta = 2\pi \frac{|l_{+}|}{|l_{+}| + |l_{-}|}
\]

where \( |l_{+}| \) and \( |l_{-}| \) were the respective lags from zero in the leading and lagging direction, to the first peaks in the lag correlation function.

Fig. 8 shows how \( \theta \) changes around the threshold of oscillation. The excitatory and inhibitory compartments have a zero-lag correlation close to 1 above threshold, and also well below threshold. As transition to spontaneous oscillation is approached, the excitatory and inhibitory components progressively reverse their lead/lag relationship, so that near threshold, excitatory and inhibitory cells are out of phase.

4. Conclusions

4.1. Comparability to experimental data

Simulation results show a good match to electro-cortical phenomena, despite uncertainty in the values of many parameters, and the use of simplified cortical architectures.

At centimetric scale the 1/f² background spectrum and the resonance peaks are in accord with EEG in a number of species (e.g. Buzsaki and Draguhn (2004) and Freeman et al. (2006)). The predominance of 1/f² background when very low frequency inputs are introduced mimics the slow fluctuations in cortical activation associated with variation of the field of attention. There are limitations in other respects. The spectrum lacks adequate definition of spectral peaks. Thalamo-cortical models (Rennie et al., 2002; Robinson et al., 2001, 2003; Rotterdam et al., 1982; Suffczynski et al., 2001) are able to generate better matches to observed spectra in this regard, consistent with the incompleteness of the present model which describes cortex in isolation. The occurrence of low-frequency activity conforms to experimentally observed wavenumber/frequency relations (O’Connor & Robinson, 2003) but the gamma resonance, observed at the larger scale, is greater than can be observed experimentally. This may be explained by volume-conduction filtering of real EEG, as well as by the fact that real cortex is seldom if ever held in a state of uniform and unvarying non-specific cortical activation.

At macrocolumnar scale there is good correspondence with gamma activity as widely observed experimentally (Eckhorn et al., 1988; Gray & Singer, 1989) including the strong covariance of pulses and waves (Stryker, 1989) as autonomous oscillation is approached, while mean cell firing rates are in accord with those found in awake cortex (Steriade et al., 2001).

The present results highlight the relationship between travelling waves and synchronous fields (Bressler et al., 1993; Eckhorn et al., 1988; Gray & Singer, 1985; O’Connor & Robinson, 2003) by showing that the apparent direction of travel of the waves depends on both the relative magnitude of signal inputs at any two cortical sites and whether or not autonomous, co-operative, oscillations have developed. Directed waves predominate at low levels of cortical excitation, but are no longer observable when swamped by a large zero-lag field, when autonomous oscillation supervenes.

The change in direction of the dominant travelling wave with changing input magnitudes may account for recently reported relations between travelling waves and bursts of gamma activity. In association with transition back and forth between cortical quiescence and bursts of gamma activity, foci of wave generation in the cortex occurring in roughly equal numbers of inwardly directed, and outwardly directed waves are observed, termed...
Fig. 6. Log power spectra at centimetric scale. In order of decrementing power at lower frequencies, non-specific afferent flux, $\psi_{ns}$, was set to 0, 8, 16 s$^{-1}$, respectively. Left: Synchronous white noise input to all elements in the driven row. Right: Synchronous noise band limited from 0.002 to 0.954 Hz input to all of the driven row of elements. The dashed line accompanying the $\psi_{ns} = 0$ s$^{-1}$ spectrum is the least-squares linear best-fit, of slope $-2.07 \log$power/$\log$Hz.

Fig. 7. Pulse density covariance (s$^{-2}$) versus conduction delay (ms) in the macrocolumnar simulation. Asynchronous white noise inputs are delivered to each of two elements, situated at row 10, column 16, and row 10, column 10. Lag covariances are between elements adjacent to the sites of input. RMS amplitudes of the input signals were in the ratios 1/4; 1/1; 4/1. Top row: Travelling waves apparent below threshold of oscillation. Trans-cortical flux, $\psi_{qe,cc} = 0$ s$^{-1}$. Non-specific afferent flux, $\psi_{ns} = 61$ s$^{-1}$. Bottom row: Synchronous fields apparent above threshold of oscillation. Trans-cortical flux, $\psi_{qe,cc} = 0$ s$^{-1}$. Non-specific afferent flux, $\psi_{ns} = 124$ s$^{-1}$.

“phase cones” by Freeman and Barrie (2000) and Freeman et al. (2006). Inwardly or outwardly directed waves, corresponding, respectively, to minima and maxima of input amplitudes, can be considered the two-dimensional equivalents of the unidirectional travelling waves shown in Fig. 7.

Excitatory/inhibitory timing relations at the intermediate range of cortical excitation shown in Fig. 8 are consistent with lag relations in average firing between excitatory and inhibitory cells in gamma oscillation (Hasenstaub et al., 2005; Morita et al., 2008).

4.2. Mechanisms of gamma synchrony and relations to large scale EEG

It appears that both local and global factors contribute to the control of gamma activity. Locally, all influences on excitatory/inhibitory balance, including back-propagation and receptor distributions in the near and far dendritic trees, voltage dependence of NMDA synapses and the average level of activity in the surrounding cortex, determine specific patterns of firing, while globally, the balance of excitatory tone to the excitatory and inhibitory components appears crucial to triggering and suppressing gamma activity. A principle aspect of global control is apparent when it is recalled that non-specific excitatory flux was delivered to excitatory cells only, while trans-cortical flux is delivered to both excitatory and inhibitory components.

A general account of gamma activity and information transfer in the cortex can be advanced. Patches of autonomous gamma activity inject information into the wider cortical field, thus generating traveling waves. Linkages between excitatory cells – including voltage-dependent transmitters – link together patches of gamma activity in synchronous fields, whereas long-range afferents to both excitatory and inhibitory cells (mediating the
average trans-cortical flux) act to suppress gamma oscillation, allowing the cortex to self-regulate the onset and offset of gamma activity in complex patterns. The spatial patterning of cortical activation, itself largely controlled by frontal and limbic connections (Alexander et al., 1990) can act to select particular synchronous fields, thus permitting a large set of possible states of attention. This fits Freeman’s concepts of gamma activity as akin to thermodynamic phase transitions and helps to explain some of the uncertainties in interpretation of experimental data (Hasenstaub et al., 2005; Morita et al., 2008; Robinson, 2007). Results in Fig. 8 suggest that gamma oscillation is not merely a simple pendulum-like to-and-fro exchange of excitatory and inhibitory pulses. While phase lagged activity occurs at intermediate states of excitatory tone, both at low levels of excitation and in states of autonomous oscillation, the excitatory and inhibitory components move into synchrony with each other. This suggests that normally cortex is poised near transition to oscillation.

As mentioned in Section 1, although the close relation of EEG to cognitive events has been widely accepted, the relationship is problematic, since signals of low frequency and long wavelength only have low informational content. This conundrum may be resolved by considering large-scale, but locally small, trends in neural activity – observable as EEG – as an adiabatic control of local cortical activity. If large-scale waves mediate the control of transition, rather than transfer substantial information, then it is the spatial patterning of thalamo-cortical inputs which permits a high channel capacity for the mediation of attention, and the low correlation of low frequency EEG with cell firing (Stryker, 1989) is to be expected. It has recently been reported that bursts of gamma occur at intervals concurrent with alpha or theta activity (Freeman & Barrie, 2000; Freeman et al., 2006), supporting this interpretation.

It may be asked which features of the state equations and specific parameter values are essential to the properties described. Within a wide envelope of factors affecting the excitatory/inhibitory balance, there is no precise dependence on any single individual parameter. Nor is the occurrence of synchronous fields or travelling waves dependent on specific parameter values, since earlier papers (Chapman et al., 2002; Robinson and Wright et al., 1998; Wright, 1997) show cortical synchrony may be explained as an example of a general property of wave transmission in summing junctions linked with conduction delays. Similarly, the stabilising effect of extended cortical activation depends on the longer conduction delays over long distances, rather than absolute values of conduction velocity. The sensitivity studies reported in Fig. 3 show that the introduction of channel kinetics for NMDA, AMPA and GABA, and the allowance made for back-propagation in the dendritic tree, are inessential for simulation of most EEG features so long as compensating changes are made in other parameter values. That is to say, if modelling of EEG fields only is required, then the state equations include redundancies. However, these added features are essential for more detailed considerations of dynamic linkage, and may be crucial if the model is to later include modifiable synapses and learning. Related considerations apply to the use of Gaussian homogeneous connection fields in the present results. It is expected that in simulations of this general class inhomogeneity of connections can be introduced, if aggregate gain is locally normalized (Wright, 1997, 1999).

The present treatment has been limited to low resolution of cortical events and separation of simulations according to scale. There appears no impediment to combining multiple scales within a single simulation, or of resolving events to a scale comparable to cellular models. This would enable the dynamics of gamma oscillation in complex connectivity to be related to synaptic modification (Kay & Phillips, 1997; Phillips & Singer, 1997; Tsuda & Kuroda, 2004; Tsukada, Aihara, & Saito, 1996; Tsukada, Yamazaki, & Kojima, 2007; Wright et al., 2006; Wright & Bourke, 2008).

Acknowledgements

The author thanks the Oakley Foundation (Auckland) for financial support; Professor Peter Hunter of the Bioengineering Institute, Auckland University, for access to large scale computing resources; Mr Nick Hawthorn for technical support; and Dr Chris Rennie, Dr Gordon Lees, Mr Paul Bourke, and Professor Rob Kydd for other assistance and support.

Appendix

A.1. Synaptic numbers and gain factors

Synaptic number densities, $N_i$, given in Table 1, are from secondary work (Liley & Wright, 1994; Rennie et al., 2000) based on primary data from Scholl (1956), Brandenberger and Schuz (1991) and Nunez (1995). The “ground state” estimates of synaptic gain, $g_i^{[0]}$, are derived from Thomson (1997) and Thomson, West, Hahn, and Deuchars (1996).

A.2. Threshold values

Maximum firing rates, and membrane reversal potentials shown in Table 2 are those given in Rennie et al. (2000) from primary sources. The estimate of $\theta_k$ is from Kandel, Schwartz, and Jessell (1991), and the value of $a_i$ is secondarily inferred, on the assumption that neurone firing probability has near Gaussian distribution on the range $0 \rightarrow V_m^{[0]}$.

A.3. Membrane temporal smoothing

Reciprocal average dendritic rise and fall constants in Table 3, are again from Rennie et al. (2000) and Thomson (1997), Thomson et al. (1996), while the parameters $a'$ are roughly in accord with the findings of Stuart and Sakmann (1994).

A.4. Receptor adaptation gains and time constants

Parameters in Table 4 were obtained by deriving steady-state and impulse response functions from three mass–action
models of receptor/transmitter interactions (Dominguez-Perrot, Feltz, & Poulter, 1996; Hauser & Roth, 1997; Lester & Jahr, 1992). Models provided by these authors were applied in simulations, assuming a background firing rate of 10 spikes s\(^{-1}\), and steady-state gain factors and normalised impulse functions were obtained by curve-fitting to time-responses of the models. Model fitting to the parameter order given is inessential, but is included for completeness.

### Table 2

| $N_{exc}$ | Excitatory to excitatory corticocortical synapses/cell | 3710 dimensionless |
| $N_{inh}$ | Excitatory to inhibitory corticocortical synapses/cell | 3710 dimensionless |
| $N_{exc}$ | Excitatory to excitatory intracortical synapses/cell | 410 dimensionless |
| $N_{inh}$ | Inhibitory to excitatory intracortical synapses/cell | 800 dimensionless |
| $N_{exc}$ | Excitatory to inhibitory intracortical synapses/cell | 410 dimensionless |
| $N_{inh}$ | Inhibitory to inhibitory intracortical synapses/cell | 800 dimensionless |
| $N_{syn}$ | Synapses per excitable cell from subcortical sources. | 100 dimensionless |
| $N_{syn}$ | Synapses per inhibitory cell from subcortical sources. | 0 dimensionless |

### Threshold values.

| $Q^{max}$ | Maximum firing rate of excitatory cells | 100 $s^{-1}$ |
| $Q^{min}$ | Maximum firing rate of inhibitory cells | 200 $s^{-1}$ |
| $V^{exc}$ | Excitatory reversal potential | 0 V |
| $V^{inh}$ | Inhibitory reversal potential | $-0.070 \text{ V}$ |
| $V^{rest}$ | Resting membrane potential | $-0.064 \text{ V}$ |
| $E_{syn}$ | Mean dendritic potential when 50% of neurones firing. | $-0.035 \text{ V}$ |
| $\sigma_e$ | Standard deviation of neuron firing probability. | 0.0145 V versus mean dendritic potential. |

### Receptor adaptation.

| $\lambda^{(a)}$ | Receptor adaptation constant | $[\text{AMPA}] = 0.012$, $[\text{NMDA}] = 0.037$ |
| $\beta^{(a)}$ | Pulse-efficacy decay constants | $[\text{AMPA}] = 0.005 \text{ s}$, $[\text{GABA}] = 1.0$ |
| $\alpha^{(a)}$ | Receptor onset coefficients | $[\text{AMPA}] = 1.0$, $[\text{GABA}] = 1.0$ |
| $\delta^{(a)}$ | Receptor offset coefficients | $[\text{AMPA}] = 0.50$, $[\text{NMDA}] = 0.25$, $[\text{GABA}] = 0.0005$ |
| $\epsilon^{(a)}$ | Receptor onset reciprocal time-constants | $[\text{AMPA}] = 280.0$ |
| $\gamma^{(a)}$ | Receptor offset reciprocal time-constants | $[\text{AMPA}] = 0.0004$, $[\text{GABA}] = 50.5$ |

### Table 5

| $r^{(a)}$ | Relative weighting of receptors on near dendritic field | $[\text{AMPA}] = 1 - r^{(a)}$, $[\text{NMDA}] = 1 - r^{(a)}$ |
| $r^{(b)}$ | Relative weighting of receptors on far dendritic field | $[\text{AMPA}] = 0.50$, $[\text{NMDA}] = 1.0$, $[\text{GABA}] = 0.375$ |

### A.5. Receptor distribution

Distribution of three receptor types were considered representative of a wider group of receptors. These were the principal fast excitatory glutamate receptor (AMPA), the principal fast inhibitory GABA receptor (GABA\(_a\)) and the principal slow and voltage-dependent glutamate receptor (NMDA). Total excitatory synaptic gain was divided equally between NMDA and AMPA (Rennie et al., 2000; Thomson, 1997; Thomson et al., 1996). NMDA is distributed predominantly in the distal dendritic tree (Phillips & Singer, 1997). To allow for NMDA’s voltage-dependent role, NMDA was ascribed entirely to the distal dendritic tree, in the present model. The other receptors may be more uniformly distributed, although most workers emphasise some bias toward inhibitory inputs to the proximal dendritic trees (Bower & Beeman, 1998), so this bias was introduced in the standard parameters (Table 5).

### References


### Table 3

| $\alpha^{(a)}$ | EPSP decay in excitatory cells | 68 $s^{-1}$ |
| $\alpha^{(b)}$ | EPSP decay in inhibitory cells | 176 $s^{-1}$ |
| $\alpha^{(f)}$ | EPSP decay in excitatory cells | 82 $s^{-1}$ |
| $\tau$ | Time-smoothing attributable to position of near and far synapses on dendritic tree. | $n^2 = 1000 s^{-2}$, $d^2 = 200 s^{-2}$ |

### Table 4

| $\lambda^{(a)}$ | Receptor adaptation | $[\text{AMPA}] = 0.0004$, $[\text{GABA}] = 0.0005$ |
| $\beta^{(a)}$ | Pulse-efficacy decay constants | $[\text{AMPA}] = 0.6339$, $[\text{GABA}] = 0.0060$ |
| $\alpha^{(a)}$ | Receptor onset coefficients | $[\text{AMPA}] = 0.3657$, $[\text{GABA}] = 0.0936$ |
| $\delta^{(a)}$ | Receptor offset coefficients | $[\text{AMPA}] = 0.00004$, $[\text{GABA}] = 50.5$ |
| $\epsilon^{(a)}$ | Receptor onset reciprocal time-constants | $[\text{AMPA}] = 0.50$ |
| $\gamma^{(a)}$ | Receptor offset reciprocal time-constants | $[\text{AMPA}] = 60.3$, $[\text{GABA}] = 280.0$ |

### Table 5

| $r^{(a)}$ | Relative weighting of receptors | $[\text{AMPA}] = 1 - r^{(a)}$, $[\text{NMDA}] = 1 - r^{(a)}$ |
| $r^{(b)}$ | Relative weighting of receptors | $[\text{AMPA}] = 0.50$, $[\text{NMDA}] = 1.0$, $[\text{GABA}] = 0.375$ |


Attractor Dynamics and Thermodynamic Analogies in the Cerebral Cortex: Synchronous Oscillation, the Background EEG, and the Regulation of Attention

J.J. Wright

Received: 8 March 2010 / Accepted: 17 June 2010 / Published online: 4 September 2010
© Society for Mathematical Biology 2010

Abstract  Ongoing changes in attention and cognition depend upon cortical/subcortical interactions, which select sequences of different spatial patterns of activation in the cortex.

It is proposed that each pattern of cortical activation permits evolution of electrocortical wave activity toward statistically stationary states, analogous to thermodynamic equilibrium. In each steady-state, neurons fire with an intrinsic Poisson spike probability and also with a bursting pattern related to network oscillations. Excitatory cell dendrites act as a regenerative reservoir in which pulse generation is balanced against dissipations.

Equilibria exhibit contrasting limits. One limit, at high cortical activation, generates widespread zero-lag synchrony among excitatory cells, with partial suppression of noise. Excitatory and inhibitory cells approach zero-lag local correlation, with 1/4 cycle lag-correlation at greater distances of separation. The high-activation limit defines a correlated system of attractor basins, capable of co-ordinating synaptic modifications and intracortical signal generation. Suppression of noise would enhance convergence about attractor basins in the manner of simulated annealing, while, conversely, the persistence of some noise prevents network paralysis by phase locking.

At the opposite limit—that of low activation—spikes and waves have low cross- and auto-correlation, but have wide-spectrum sensitivity to inputs. It is hypothesised that cortical regions, transiently at equilibrium near these extremes, engage in interaction with each other and with subcortical systems, to generate ongoing sequences of attention and cognition.

This account is compatible with classical and recently observed experimental phenomena. The principle features inferred from a simplified linear mathematical ac-
count are reproduced in a more physiologically realistic and non-linear numerical simulation.

**Keywords**  
Attention · Cognition · Cortical activation · Cerebral cortex · Gamma synchrony · Synchronous oscillation · EEG · LFP · Continuum neural field · Attractor neural networks · Thermodynamic analogies · Null-spikes · Phase-cones

1 Introduction

This paper attempts a preliminary unification of neural wave dynamics, attractor dynamics, and synchronous oscillation—including the role of cortical/subcortical interactions—with a view to linking these theoretic concepts to observable local field potential (LFP), electroencephalogram (EEG), and unit pulse relations.

An empirical relationship of the EEG—which is the large-scale LFP—to cognitive information processing is well known, since the waking state depends upon activation of the cortex from the brain-stem reticular formation, and this activation produces a dramatic change, from slow high amplitude EEG waves, to low amplitude waves in which gamma synchrony plays a prominent role (Stryker 1989). As well as the principle dimension of alertness, the direction of attention depends on continuous interplay between the cerebral cortex, the reticular formation, and intermediate structures in thalamus, limbic system, and basal ganglia, so that patterned domains of cortical excitation undergo sequential modification as attention and cognition proceeds (Scheibel and Scheibel 1970; Alexander et al. 1990; Kandel et al. 1991).

However, long-standing doubt hangs over the mechanism and significance of the relationship between EEG and cognitive processing. In 1989, Stryker highlighted the grounds of this controversy, at the same time heralding the publication of two papers describing the discovery of synchronous oscillation. The EEG’s absence of clear correlation with action potential spikes seemed to exclude its possible relevance to information processing—but in the gamma band, around 40–60 Hz—the relation of the LFP to spikes was apparent, and this activity was also associated with zero-lag synchronous cell firing at separated sites. What, Stryker asked, is the significance of this?

Since then synchronous oscillation has been extensively studied (Eckhorn et al. 1988; Gray and Singer 1989; Gray et al. 1989, 1992; Bressler et al. 1993; Singer 1994; Singer and Gray 1995; Miltner et al. 1999; Neuenschwander and Singer 1996, etc.). Much of this work has followed the original lead of von der Malsburg (1983) postulating synchrony as the mechanism mediating feature-linking and thus solving the “binding” problem. But the relationship of gamma activity to the background EEG—and particularly the wider problem of correlation of single unit activity and the LFP—has remained without full resolution.

Preceding and concurrent with this physiological research, theoretical work—from Pitts and McCulloch (1943) via the work of Hebb (1949) to later work on attractor neural networks (ANN) (e.g., Hopfield 1982; Amit 1989; Arbib 1995)—has generally assumed that neurons settle into organized patterns of firing—whether
“cell assemblies”, static attractor basins, or oscillation and resonance (e.g., Grossberg 1980). Absence of full understanding of the dynamics of real neural assemblies makes the realism of theoretical models uncertain, and for all their success, such models generally place little or no emphasis on the role of the subcortical mechanisms which have such a central bearing on the regulation of conscious states.

Particularly notably among those who have gone beyond the simplified neurons and dynamics of ANN theory, are Freeman and colleagues (Freeman 1975, 1991, 2003, 2004a, 2004b, 2005, 2006; Barrie et al. 1996; Freeman and Barrie 2000) and related workers (e.g., Steyn-Ross and Steyn-Ross 2009). They have emphasised non-linear and far-from-equilibrium neural dynamics—and have proposed that phase-changes in the cortical continuum, in the thermodynamic sense, are central to cognitive and perceptual processing. Present work focuses upon clarification of these relations (Freeman 2009). They have provided a rich experimental context for their theoretical modelling, and it is toward explanation of some central aspects of their results that the present paper aims to contribute, as well as seeking an answer to Stryker’s question.

The procedure followed here avoids analytic complexity. Equilibrium is assumed in a simple form of continuum cortical field equations, and properties are predicted in qualitative terms. Then numerical simulations, using more complex and physiologically realistic state equations applied in recent work (Wright 2009a, 2009b) are used to test for the effects predicted, also permitting the results to be compared to experimental data.

2 The Neural Field

2.1 General State Equations

The neural field is a cortical domain—either in a single patch or in several sub-domains in interaction—which can be considered comparatively autonomous from other cortical domains, during a sufficient period for analysis.

Neural field models (e.g., Wilson and Cowan 1973; Haken 1976; Nunez 1981, 1995; Jirsa and Haken 1996; Rennie et al. 2000, 2002; Wright 2009a, 2009b; Steyn-Ross et al. 2010; Potthast et al. 2010) treat the cortex as a continuum wave medium, and can be extended to the description of neural masses (e.g., Freeman 1975; van Rotterdam et al. 1982; Jirsa and Stefanescu 2010). All differ in the specific features incorporated in their state equations, but abstracting generally recognized features, neural field equations can be represented as

\[ \varphi_q(r, t) = \int f_q(r, r') Q_q \left( r', t - \frac{|r - r'|}{v} \right) dr' \]  
\[ V_q(r, t) = G \left( V_q(r), \sum_{q'=e, q' \neq q} \varphi_{q'}(r) \right) (t - \tau) * f_s(\tau) \]  
\[ Q_q(r, t) = f_S \left( V_q(r, t) \right) \]
\[ \psi; Q; V, \] are the state variables afferent synaptic flux, pulse density, and dendritic voltage (LFP), respectively. These are described separately for excitatory and for inhibitory cells and synapses, using the subscript \( q = e, i \), to indicate either excitatory or inhibitory cells, while \( \mathbf{r}, \mathbf{r}' \) are cortical positions. The functions \( f_{subscript} \) each indicate experimentally determined anatomical/physiological properties namely:

Equation (1) describes the transmission of either excitatory or inhibitory signals to synapses at a point on the cortical surface, by axons from cell bodies at other locations. The density of afferent axo-synaptic couplings, transmitting signals from elsewhere in the field, with axonal conduction velocity \( \nu \), are described by \( f_q(\mathbf{r}, \mathbf{r}') \).

Equation (2) describes the non-linear transformation of the summed excitatory and inhibitory synaptic flux into dendritic potentials via channel dynamics and dendritic cable properties incorporated in the convolution \( G(t - \tau) \ast f_s(\tau) \).

Equation (3) describes the local conversion of dendritic potentials into action potentials, via the function \( f_\Sigma(V_q(\mathbf{r}, t)) \)—commonly a squash function, which may include population variations in pulse emission threshold, pulse emission probability, and spontaneous pulse generation.

The field activity can be regulated by setting the average level of cell firings, \( \bar{Q}_q(\mathbf{r}) \)—the regulation being a function of subcortical input to the cortex. A crucial property for the following arguments is that solutions of the state equations include oscillating steady-states—oscillation becoming less damped with increasing, \( \bar{Q}_q(\mathbf{r}) \)—until at a critical level of excitation, it becomes autonomous (Wright 2009a, 2009b).

2.2 Definitions of Steady-State and Equilibrium

Here, these terms, usually considered synonyms, are used with separate meanings.

(a) “Steady-state” means that input from subcortical systems has set time-average excitatory firing rates, \( \{ \bar{Q}_e(\mathbf{r}, \mathbf{r}') \} \), over a particular time epoch, into spatial uniformity.

(b) “Equilibrium” means that the system has entered a stable state, with statistical stationarity of \( \{ \bar{Q}_e(\mathbf{r}, \mathbf{r}') \} \) about a particular steady state—where “stationarity” implies that variance, cross- and autocorrelation functions are constant over time. At equilibrium, signals exchanged from point to point in the cortex are symmetric and equal over time.

2.3 Virtual Pulse Representation of Steady-State and Equilibrium

Non-linearities in (1)–(3) render neural field models analytically difficult. A suitable simplification is to represent events in the dendrites not as voltages, but as transmission of “virtual pulses” with units equivalent to real pulses. Then at steady-state

\[
 w(\mathbf{r}) = I_e \int f_e(\mathbf{r}, \mathbf{r}') u(\mathbf{r}') d\mathbf{r}' + I_i \int f_i(\mathbf{r}, \mathbf{r}') d(\mathbf{r}') d\mathbf{r}'
\]

where \( u \) and \( d \) are rates of afferent excitatory and inhibitory pulses, transmitted from \( \mathbf{r}' \), reaching synapses on excitatory and inhibitory cells at \( \mathbf{r} \). For simplicity, pulse

\( \odot \) Springer
rates at steady-state for excitatory and inhibitory cells are treated as equal under an appropriate scaling. Axonal connection distributions \( f_e \) and \( f_i \) are normalized so \( \int_0^\infty f_q(\mathbf{r}, \mathbf{r'}) \, d\mathbf{r} = 1 \), and \( \Gamma_e > 0, \Gamma_i < 0 \) are aggregate gains for excitatory and inhibitory transmission, respectively.

Thus, \( w \) are rates of virtual pulses, generated by afferent pulses, and, in turn, transformed to efferent pulses. Therefore, temporal and spatial uniformity (steady-state without oscillation) requires that \( \forall \mathbf{r}, t \ w(\mathbf{r}) = u(\mathbf{r}) = d(\mathbf{r}) = \bar{x} \), where \( \bar{x} \) is the set point for \( u, d \) and \( w \) throughout the neural field. Comparison to full non-linear field equations requires \( \Gamma_e \) and \( \Gamma_i \) to be functions of \( \bar{x} \).

Equilibrium is defined as time-variation about the steady state when the time and space average of \([w - \bar{x}] (\mathbf{r}, t)\) is statistically stationary, so

\[
[w - \bar{x}] (\mathbf{r}, t) = \Gamma_e \int f_e(\mathbf{r}, \mathbf{r'}) [u - \bar{x}] \left( \mathbf{r'}, t - \frac{|\mathbf{r} - \mathbf{r'}|}{v} \right) \, d\mathbf{r'} + \Gamma_i \int f_i(\mathbf{r}, \mathbf{r'}) [d - \bar{x}] \left( \mathbf{r'}, t - \frac{|\mathbf{r} - \mathbf{r'}|}{v} \right) \, d\mathbf{r'}
\]

(5)

describes transformation of afferent pulse perturbations into virtual pulse perturbations, and

\[
[u - \bar{x}] (\mathbf{r}, t) = \left[ E_u - \bar{x} \right] (\mathbf{r}, t) + \left[ w - \bar{x} \right] (\mathbf{r}, t - \tau) * f_s(\tau)
\]

(6a)

\[
[d - \bar{x}] (\mathbf{r}, t) = \left[ E_d - \bar{x} \right] (\mathbf{r}, t) + \left[ w - \bar{x} \right] (\mathbf{r}, t - \tau) * f_s(\tau)
\]

(6b)

describes the transformation of virtual pulse perturbations into efferent pulse perturbations, including fluctuations in spontaneous pulse generation, \([E_u, d - \bar{x}] (\mathbf{r}, t)\). Delays in the conversion of virtual to real pulses are expressed in the convolutions \([w - \bar{x}] (\mathbf{r}, t - \tau) * f_s(\tau)\), where \( \int_0^\infty f_s(\tau) \, d\tau = 1 \) and \( f_s(\tau) \) differs between excitatory and inhibitory classes of neuron.

### 2.4 Synchronous Fields and Dynamical Frustration

A theoretical account of the origin of synchrony generated via long-range excitatory connections is given elsewhere (Wright and Liley 1996; Robinson et al. 1998; Wright et al. 2000; Chapman et al. 2002; Wright 2009a, 2009b). A simpler argument, based directly on assumption of equilibrium in the simple neural field follows.

Let \( E_{\rightarrow}(\mathbf{r}, \mathbf{r'}, t) \), and \( E_{\leftarrow}(\mathbf{r}, \mathbf{r'}, t) \), be the pulse-perturbations exchanged in opposite directions between any two points, via all paired combinations of excitatory-excitatory, inhibitory-inhibitory, and excitatory-inhibitory neuron connections. Writing \( u, d \) to apply to cells and synapses of either excitatory or inhibitory type, also with the appropriate subscripts of \( \Gamma_e, i \) and \( f_e, i \), the exchange of signals is

\[
E_{\rightarrow}(\mathbf{r}, \mathbf{r'}, t) = \Gamma_e,i \left( f_e,i(\mathbf{r}, \mathbf{r'}) [u(\mathbf{r}), d(\mathbf{r'}) - \bar{x}] \left( t - \frac{|\mathbf{r} - \mathbf{r'}|}{v} \right) \right)
\]

(7a)

\[
E_{\leftarrow}(\mathbf{r'}, \mathbf{r}, t) = \Gamma_e,i \left( f_e,i(\mathbf{r'}, \mathbf{r}) [u(\mathbf{r'}), d(\mathbf{r}) - \bar{x}] \left( t - \frac{|\mathbf{r} - \mathbf{r'}|}{v} \right) \right)
\]

(7b)
The exchange of signals must be symmetric and equal, to maintain equilibrium. Anatomically, bidirectional, roughly symmetric, coupling densities prevail at the macroscopic scale—i.e., $f_e(r, r') \approx f_e(r', r)$ and $f_i(r, r') \approx f_i(r', r)$ (Scholl 1956; Braitenberg and Schüz 1991; Liley and Wright 1994). So, equilibrium is approached for exchanges between excitatory cells when

$$[u - \bar{x}](r, t) \approx [u - \bar{x}](r', t), \quad \forall t, r, r'$$ (8a)

—that is, if all excitatory neurons approach zero lag synchrony. Similarly, equilibrium is approached for exchanges between inhibitory cells when

$$[d - \bar{x}](r, t) \approx [d - \bar{x}](r', t), \quad \forall t, r, r'$$ (8b)

—that is, if all inhibitory neurons approach zero-lag synchrony.

Interactions between excitatory and inhibitory cells permit no continuous and symmetric exchange of signals, but since coupling densities between closely adjacent cells are stronger than those of longer range, overall symmetry of exchange can be approached with the development of local oscillation, so that

$$\Gamma_i[d - \bar{x}](r, t) \approx -\Gamma_e[u - \bar{x}](r, t), \quad \forall r, t$$ (8c)

—that is, if excitatory and inhibitory pulses throughout the neural field approach zero-lag anti-synchrony.

Conditions (8a), (8b), (8c) can be concurrently met in the absence of noisy pulse generation.

On the other hand, even if spontaneous pulse generations in individual cells were completely stochastically independent, high positive local correlation of $[u - \bar{x}](r, t)$ and $[d - \bar{x}](r, t)$ is to be expected since inhibitory cells are driven predominantly by local excitatory neurons to which they are strongly coupled, thus imposing local excitatory signals, including noise, upon the inhibitory firing—in contrast to the smoothing of noise which will take place by integration from sources at longer range (5). Consequently,

$$[u - \bar{x}](r, t) \rightarrow [d - \bar{x}](r, t), \quad \forall r, t$$ (9)

—that is, noisy firing of cells promotes local synchrony of $u$- and $d$-fields.

Thus, in (5), conditions (8a), (8b) and (9), if met, imply $u(r, t) \rightarrow \min, \forall t, r$, while conditions (8a), (8b), (8c) if met, imply $u(r, t) \rightarrow \max, \forall t, r$. A form of dynamical frustration is present, with opposing tendencies toward, and away from, local excitatory/inhibitory synchrony. Approach to equilibrium will require compromise between synchrony and anti-synchrony among excitatory and inhibitory cells, but widespread zero-lag synchrony will be found among excitatory, and among inhibitory, cells, respectively, as no dynamical conflicts occur in the interactions of cells of the same type.

2.5 Oscillation and Optimum Resolution of Dynamical Frustration

In neural field models, oscillation may arise in several ways, as is further discussed in the Conclusions section. For present purposes, it is supposed that oscillations arise...
as action potentials circulate around all possible local excitatory-inhibitory loops, in a fashion similar to that proposed by Freeman (1975, 1991). Loops of maximum gain are two-way, typically bi-synaptic connections, emanating from, and returning to, any point, \( r \), along all radial paths with a particular total delay. Gain, \( \gamma \), over all loops of a particular length, \( |r - r'| \), is thus

\[
\gamma \propto 2\pi |r - r'| f_e(r, r') f_i(r', r)
\]  

(10)

Since connection strengths decline with distance, while the number of loops increases with the circumference of the system of recurrent connections, then loop gain must be maximal for circulation over some distance, \( |r - r'|_{res} \), and oscillation will develop with a centre frequency determined by delays over this characteristic distance. The oscillation is comprised of cycles of excitation-of-inhibitory-cells \( \rightarrow \) recurrent-inhibition-of excitatory-cells \( \rightarrow \) reduced-excitation-of-inhibitory-cells \( \rightarrow \) recurrent-disinhibition-of-excitation, etc. Local oscillation implies lag correlation between excitatory and surrounding inhibitory cells, maximum at separation \( |r - r'|_{res} \), at which distance a lead/lag relation between excitatory and inhibitory cells of 1/4 of the cycle of oscillation is expected. This combines overall symmetry of exchange between excitatory and inhibitory cells with positive local correlation of excitatory and inhibitory firing, resolving the dynamical frustration encountered in (8c) and (9).

2.6 Further Definitions of Terms

The perturbations of each of the three field variables, \( x \in \{ u, d, w \} \), at \( n \) small equal volumes of the cortical surface, in equilibrium at a steady-state, can be represented as

\[
F_x(t) = \{ x_i(r_i, t) - \bar{x} \} \text{ for } i = 1, \ldots, n
\]  

(11)

and, for subsequent use, the following terms can be defined in relation to \( F_x(t) \) over an epoch \( P \).

\( \mathbf{X} \): the covariance matrix associated with \( F_x(t) \).

\( \dot{X} \in \{ \dot{U}, \dot{D}, \dot{W} \} \): time average rates at which perturbations are generated, summed over the entire \( u, d \)- and \( w \)-fields, respectively. These rates can be treated as analogous to rates of supply/dissipation of energy about steady-state.

\( \{ \rho^{[X]}_i \} \): eigenvalues, each associated with a spatial eigenmode of \( \mathbf{X} \), where \( \sum n \rho^{[X]}_i \) is the total variance of \( \mathbf{X} \).

\( \dot{S}^{[X]} \): the rate of generation of entropy associated with \( \dot{X} \)—which can be shown for the differential entropy of Gaussian distributions (Wolfram Mathworld 2009) to be

\[
\dot{S}^{[X]} = \frac{1}{2} n (1 + \ln(2\pi)) + \frac{1}{2} \ln |\text{det} \mathbf{X}|
\]  

(12)

where \( \text{det} \mathbf{X} = \prod_n \rho^{[X]}_i \)  

(13)

\( \{ x_i \} \): a set of vectors in Hilbert space, identifying the basin(s) of attraction (if any) in \( F_x(t) \).

Figure 1 shows physiological relations among the patterns of pulse generation and dissipation, and suggests their associated entropies.
2.7 Pulse Variations—\( \dot{U} \)

\( \dot{U} \)—the rate of energy supplied by excitatory spike-rate perturbations—has two sources.

**Stochastic Background**

Consistent with the observations of Steriade et al. (2001), it is here assumed that the emission of pulses from excitatory cells have an inherently stochastic component, independent at all \( r \), with the Poisson probability density function

\[
    f = \frac{\lambda u e^{-\lambda}}{u!}
\]  

(14)

where \( u \) is the number of pulses emitted in unit time, at some one of the \( n \) small volumes of the cortical surface, with mean and variance, \( \bar{x} = \lambda \). The distribution is assumed to approach Gaussian.

**Oscillation**

Oscillation reflects co-incident, “bursting”, pulse firing (Steriade et al. 2001), approximated with probability density

\[
    f = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{(u - \bar{x})^2}{2\sigma^2}\right]
\]  

(15)
i.e., Gaussian with variance $\sigma^2$. Because the oscillation takes place in a linked field, pulse emissions of this type are not statistically independent at all $\mathbf{r}$.

By summing variances, then converting to RMS

$$\dot{U} \propto \sqrt{n(\lambda + \sigma^2)}$$

and equivalently, using the terms of (5),

$$\dot{U} = \left( \frac{1}{P} \int_{0}^{P} \int_{\mathbf{r}} \int_{\mathbf{r}'} \left[ \Gamma_e f_e(\mathbf{r}, \mathbf{r'}) \left[ u - \bar{x} \right] \left( \mathbf{r'}, t - \frac{\mathbf{r} - \mathbf{r'}}{v} \right) \right]^2 d\mathbf{r'} d\mathbf{r} dt \right)^{1/2}$$

(17)

2.8 Return-Inhibition and Dendritic Perturbations—$\dot{D}$ and $\dot{W}$

$\dot{D}$ is the rate of generation of perturbation due to feedback from inhibitory cells, themselves driven by excitatory cells, so, similarly to (17)

$$\dot{D} = \left( \frac{1}{P} \int_{0}^{P} \int_{\mathbf{r}} \int_{\mathbf{r}'} \left[ \Gamma_i f_i(\mathbf{r}, \mathbf{r'}) \left[ d - \bar{x} \right] \left( \mathbf{r'}, t - \frac{\mathbf{r} - \mathbf{r'}}{v} \right) \right]^2 d\mathbf{r'} d\mathbf{r} dt \right)^{1/2}$$

(18)

$\dot{W}$ is the rate of perturbation of excitatory dendritic potentials, expressed as virtual pulses, and is proportional to the LFP. With increasing time the history of perturbations is lost from the dendritic potentials, so $\dot{W}$ is also a rate of dissipation. From (5),

$$\dot{W} = \frac{\Gamma_e}{P} \int_{0}^{P} \int_{\mathbf{r}} \int_{\mathbf{r}'} f_e(\mathbf{r}, \mathbf{r'}) \left[ u - \bar{x} \right] \left( \mathbf{r'}, t - \frac{\mathbf{r} - \mathbf{r'}}{v} \right) d\mathbf{r'} d\mathbf{r} dt$$

$$+ \frac{\Gamma_i}{P} \int_{0}^{P} \int_{\mathbf{r}} \int_{\mathbf{r}'} f_i(\mathbf{r}, \mathbf{r'}) \left[ d - \bar{x} \right] \left( \mathbf{r'}, t - \frac{\mathbf{r} - \mathbf{r'}}{v} \right) d\mathbf{r'} d\mathbf{r} dt$$

(19)

$\dot{U}$ and $\dot{D}$ are RMS sums, but $\dot{W}$ is not equal to their sum, because of the cancellation of odd components at synapto-dendritic summation, as further considered in the following section. That is, at equilibrium:

$$\dot{W} = \dot{U} + \dot{D} - \text{[rate of signal loss at synapto -- dendritic summation]}$$

(20)

2.9 Loss of Entropy at Synapto-Dendritic Summation

Summation of perturbations at dendrites leads to the cancellation of “odd” components in the continuum field (Chapman et al. 2002), because trains of random synaptic impulses which sum to the mean are indistinguishable, and their signal content is lost in the summation. The loss by cancellation of input signal structure is analogous to the loss of heat from a reservoir to an external sink with a temperature of absolute zero. Thus,

$$T \dot{S} = \text{[rate of signal loss at synapto -- dendritic summation]}$$

(21)
where $T$ is dendritic “temperature” and $\dot{S} = \dot{S}^{[U]+[D]}$ is the rate of elimination of entropy occurring at synapto-dendritic summation. The superscript $[U] + [D]$ indicates that, since the $u$-field and $d$-field are not independent their joint entropy must be calculated by extension of (12), (13).

In analogy to the thermodynamic definition, the dendritic temperature can be defined as $T = W/S^{[W]}$, where $W(t)$ is the instantaneous state of LFP fluctuation (in virtual pulses) over the field, and $S^{[W]}(t)$ is the instantaneous entropy of the LFP. Both these running sums attributable to “leaky integration” can be obtained by applying the normalised convolution in (6a), to both $\dot{W}$, and to the corresponding entropy of $\dot{W}$ according to (12), (13)—i.e.,

$$W = \dot{W}(t - \tau) * f_s(\tau)$$

and

$$S^{[W]} = \dot{S}^{[W]}(t - \tau) * f_s(\tau)$$

Since $\dot{W}$ and $\dot{S}^{[W]}$ are time-constant at equilibrium, and, in (6a), $\int_0^\infty f_s(\tau) d\tau = 1$, then

$$T = \frac{\dot{W}}{S^{[W]}}$$

and, therefore,

$$T \dot{S} = \frac{\dot{W}}{S^{[W]}} \dot{S}^{[U]+[D]}$$

2.10 Equilibrium and the Relative Magnitudes of Oscillation and Stochastic Background

Combining (20) and (21) equates generation of perturbations against their dissipation at equilibrium

$$\dot{U} + \dot{D} = \dot{W} + T \dot{S}$$

The absolute magnitudes of all four terms in (26) vary with the position of the steady-state on a continuum arranged according to the average firing level, $\bar{x} = \lambda$.

The ratio $\dot{W}/T \dot{S}$ as a function of $\bar{x}$ requires consideration of non-linear relations. Equation (10) implies that, to first approximation, $\dot{W} \propto \Gamma_e \Gamma_i$, and in the general field equations the function $f_{\Sigma}(3)$ is a major determinant of the values of $\Gamma_e$ and $\Gamma_i$. Since $f_{\Sigma}$ is some squash function describing firing rates above and below a half-maximum-rate inflexion point, and the average firing rate of cells in cortex is less than half the maximum firing rate (Steriade et al. 2001), then $f_{\Sigma}$ is roughly exponential over the operating range—i.e., $\dot{W} \propto \Gamma_e \Gamma_i \propto A e^{B \bar{x}}$, where $A & B$ are positive-valued constants with physiologically realistic values.

On the other hand, $T \dot{S}$ is a measure of the elimination of incoherent, random synaptic pulses and, therefore, at equilibrium, is also a measure of the introduction of random pulses. Equation (14) determines the rate of introduction of random pulses,
so, approximately, \( T \dot{S} \propto \sqrt{\bar{x}} = \sqrt{\bar{\lambda}} \), and

\[
\frac{\dot{W}}{T\dot{S}} \propto \frac{\exp[B\bar{x}]}{\sqrt{\bar{x}}}
\]  

(27)

so \( \dot{W} / T \dot{S} \) increases monotonically with \( \bar{x} > 0 \). Thus, two limiting cases—dominant oscillation, versus predominance of the background state—can be contrasted, used to make predictions of properties expected in simulations, and compared to physiological observations.

### 2.10.1 Dominant Synchronous field (\( \dot{W} \gg T \dot{S} \))

Applying terms defined in Sect. 2.6:

Emergence of a synchronous field means that a principle component, weighted highly for one eigenvalue of the \( n \) spatial eigenmodes of oscillation, and with small weight on all other eigenvalues, has emerged.

\[ \therefore \text{For any given total LFP variance, } \sum \rho^{[W]}_i (13) \text{ requires that } \det W \text{ is small, compared to the case where all members of } \{\rho^{[W]}_i\} \text{ have equal weight.} \]

\[ \therefore \text{Equation (12) requires that } \dot{\bar{S}}^{[W]} \text{ has approached a minimum.} \]

\[ \therefore T = \dot{W} / \dot{\bar{S}}^{[W]} (24) \text{ has approached a maximum, since } \dot{W} \text{ is large.} \]

\[ \therefore \text{From (25) } \dot{\bar{S}}^{[U]+[D]} \text{ approaches a minimum, because } T \dot{S} \text{ is small.} \]

That is, a single mode zero-lag field of oscillation appears, with low entropy, and suppression of the disorder otherwise introduced by spontaneous noisy pulses.

Correspondingly, simulation of gamma oscillation should reveal widespread and persisting synchrony of LFP and of excitatory cells over the simulated cortical field. Resolution of dynamic frustration implies close-to zero-lag positive correlation between excitatory and inhibitory cell firing at any position on the cortical surface, and lag correlation of excitatory and inhibitory cell firing at more separated positions. EEG autocorrelation should have a long half-duration, reflecting the widespread and enduring synchrony.

### 2.10.2 Background Field (\( \dot{W} \ll T \dot{S} \))

Effects are the converse of those listed above. No principle component has emerged among the \( n \) possible spatial modes of oscillation.

\[ \therefore \dot{\bar{S}}^{[W]} \text{ approaches a maximum, since all members of } \{\rho^{[W]}_i\} \text{ have equal weight.} \]

\[ \therefore T \text{ approaches a minimum.} \]

\[ \therefore \dot{\bar{S}}^{[U]+[D]} \text{ approaches a maximum.} \]

That is, no mode of oscillation is dominant, and convergence to a single attractor does not take place.

Correspondingly, simulations should show low cross-correlations of LFP and excitatory pulses throughout the neural field, and among excitatory and inhibitory cells, except for high correlations between LFP, excitatory cells and inhibitory cells at a single point on the cortical surface, consequent to locally similar afferent bombardment. No such correlations should be apparent among near-neighbour sites, and EEG autocorrelation should show only short half-duration.

© Springer
3 Numerical Demonstrations

3.1 Method of Simulation

Simulations were performed using the state equations and parameters in Wright (2009a, 2009b) (Appendix).

A 20 × 20 grid of elements representing local groups of excitatory and inhibitory cells, linked by axonal connections, roughly mimic the connectivity of the human cerebrum discretized to centimetric scale. The simulation includes AMPA, NMDA, and GABA synaptic receptors, and the effects of back-propagation in the dendritic tree. The simulation’s mean-field properties permit reproduction of gamma oscillation and synchrony, travelling waves, and the $1/f^2$ electrocortical background.

For the present simulations, two levels of non-specific activation were chosen so that the system was clearly below, or clearly above, the critical threshold for the onset of autonomous oscillation in the gamma band. Each of the linked elements received an independent Gaussian white noise input, with variance equal to the mean excitatory pulse rate, to approximate pulse generation as in (14). The excitatory and inhibitory components within any single element received the same noise time-series, to imitate the local correlations described in Sect. 2.4.

In each simulation, after initialisation, 200 seconds of simulation-time were allowed to ensure a steady-state was approached (initial conditions being far from physiological conditions), and the following 200 seconds of output were analysed, to produce the spectra, time-series, cross and autocorrelation data, shown in Figs. 2 and 3. The same simulation was then used to demonstrate properties associated with dynamic frustration, and to seek evidence for the occurrence of first or second-order phase-transitions.

3.2 Results

Figure 2(a) and (b) show the LFP background $1/f^2$ spectrum and time-series, characteristic of the resting cortex. Figure 2(c) shows the correlation between the LFP and surrounding excitatory pulses. Although pulse emissions show high cross-correlation with the LFP at the same position, pulses from surrounding excitatory cell groups show no correlation.

In Fig. 2(d), cross-correlation of local excitatory cell pulses and inhibitory cells strongly and locally connected with them, is contrasted again with inhibitory cell pulses in surrounding areas. Again, no correlation is seen in the latter case, although locally, correlation close to zero-lag predominates. In Fig. 2(e), LFP cross-correlations from points very closely and distantly situated in the simulated cortex are shown, and all can be seen to have low or no correlation, whereas Fig. 2(f) shows that the summed LFP—the EEG—shows some lag autocorrelation, consistent with the presence of a small collective cross-correlation in the field, of short time correlation length.

Figure 3 shows matched results from the simulated cortex when the level of cortical excitation is high. The power spectrum and LFP time-series now show strong gamma oscillation. The LFP and excitatory pulse rates show zero lag synchrony both...
Fig. 2 Simulated LFP/EEG and pulse densities at a low level of cortical activation. Mean excitatory pulse rate 1.57 spikes/sec. (a) LFP—power spectrum. (b) LFP—time series. (c) Lag cross-correlation ($r_{V/E}$)—LFP versus pulse density of excitatory cells. Bold line—LFP and pulse density from the same cortical point. Dashed line—LFP versus pulse density from separate, but neighbouring, points. (d) Lag cross-correlation ($r_{E/I}$)—point pulse density of excitatory cells versus point pulse density of inhibitory cells. Bold line—excitatory and inhibitory pulse densities from the same point. Dashed line—excitatory and inhibitory pulse densities from separate, neighbouring, points. (e) Lag cross-correlation ($r_{V/V}$)—point LFP versus point LFP. Bold line—LFPs from neighbouring points. Dashed line—LFPs from points widely separated. (f) Lag autocorrelation—spatial sum of LFP (EEG) from all points of simulated cortical surface.
Fig. 3 Simulated LFP/EEG and pulse densities at a high level of cortical activation—1.05 × threshold for generation of autonomous gamma oscillation. Mean excitatory pulse rate 33.96 spikes/sec. (a) LFP—power spectrum. (b) LFP—time series. (c) Lag cross-correlation \((r_{VE})\)—LFP versus pulse density of excitatory cells. **Bold line**—LFP and pulse density from the same cortical point. **Dashed line**—LFP versus pulse density from separate, but neighbouring, points. (d) Lag cross-correlation \((r_{EI})\)—point pulse density of excitatory cells versus point pulse density of inhibitory cells. **Bold line**—excitatory and inhibitory pulse densities from the same point. **Dashed line**—excitatory and inhibitory pulse densities from separate, neighbouring, points. (e) Lag cross-correlation \((r_{VV})\)—point LFP versus point LFP. **Bold line**—LFPs from neighbouring points. **Dashed line**—LFPs from points widely separated. (f) Lag autocorrelation—spatial sum of LFP (EEG) from all points of simulated cortical surface.
Fig. 4 Lag correlations in simulations with and without noise applied, in conditions otherwise the same as in Fig. 3. **Solid lines**—no noise applied. **Dashed lines**—noise applied. **Top rows** shown auto- or cross-correlations of pulses at a single reference point on the cortical surface, and **middle and bottom rows** show pulse cross-correlations between that point and each of two other points, each located a quarter turn about the closed (toroidal) simulated cortical surface and respectively lying along orthogonal lines from the reference point. **Left-hand column**—auto- and cross-correlations of excitatory neurons. **Middle column**—cross-correlations of excitatory and inhibitory neurons. **Right-hand column**—auto- and cross-correlations of inhibitory neurons.

locally and with neighbouring cells. Excitatory and inhibitory pulses again show almost zero-lag correlation locally, but now a lag correlation consistent with active coupling of excitatory and inhibitory cells at greater distance is also apparent at approximately quarter-cycle-lag between excitatory and neighbouring inhibitory cells. The LFP is cross-correlated at zero lag in both near and distant cortical locations, and the summed LFP/EEG shows sustained synchrony is present throughout the field.

Thus, simulation results accord with the expectations in Sects. 2.10.1–2.10.2, and with physiological observations described in the Introduction.

The same simulation was next used to demonstrate the contribution of dynamic frustration to the attractor dynamics in states of synchronous oscillation. Figure 4 contrasts the stable equilibrium reached when the simulation oscillates spontaneously, but has no stochastic driving, with the equilibrium reached when noise is
applied. It can be seen that, without noise, zero-lag synchrony develops between all excitatory cell components and between all inhibitory cell components—but that the excitatory and inhibitory cell systems approach zero-lag anti-synchrony. These relations are uniform throughout the neural field, which has become phase-locked. The effect of the driving noise of random cell firing is to break the phase-locking. Zero-lag synchrony of excitatory cells, and of inhibitory cells, respectively, persists throughout the field, and locally, excitatory and inhibitory cells approach zero-lag synchrony—but, more widely, excitatory and inhibitory components exhibit reciprocal oscillation. Thus, the properties of the non-linear simulation reflect simple equilibrium considerations, and indicate the operation of dynamical frustration.

Finally, to obtain information on the order of phase transition in the thermodynamic sense, as defined by the onset of autonomous oscillation, the same simulation was tested for the occurrence of hysteresis when the level of non-specific cortical activation was raised by steps to the critical threshold, and then lowered, to estimate any difference in the critical level thus found. No evidence of hysteresis was found, and this was taken to indicate that the phase transition should be classified as second-order.

4 Conclusions

It seems that properties of equilibrium thermodynamics can be applied analogously to cortical neurodynamics, in conditions of steady input from subcortical sources. Essentially, via its reciprocal connections with subcortical systems, cortex may regulate the entropy of its own signals by control of oscillation amplitude.

Predicted properties of the simple neural field at equilibrium are found in the physiologically realistic simulation; consequently solutions can be offered for the problems outlined in the Introduction and suggestions made regarding the functional implications of these solutions.

4.1 Stryker’s Question

The predictions of Sects. 2.10.1–2.10.2 and the numerical results account for the properties of EEG commented on by Stryker—namely: the close relation of pulses and the electrocortical field when synchronous activity in the gamma band is prominent, and the absence of any such clear relationship in states of lower excitation. The basic reason for this effect seems to be that oscillation suppresses spontaneous noise in pulse emission, not only because of a change in signal/noise ratio, but by an increase in the rate of elimination of noise with increasing dendritic “temperature”.

The results also account for some experimental features and theoretical conflicts concerning the mechanism of synchrony and the timing of excitatory and inhibitory pulses. Since his early work, Freeman (1975) has asserted that the correlation of pulses and local field potentials in the olfactory system indicate a lagged relationship between excitatory and inhibitory cells. Experimental recording of pulses confirm this general relationship elsewhere (Hasenstaub et al. 2005). The model for synchrony cited in Sect. 2.5 depends upon excitatory-excitatory interactions, and yields results...
consistent with Freeman’s view. In contrast, other simulations Schillen and Konig 1994; Traub et al. 1996; Whittington et al. 2000, supported by recordings made from hippocampus, account for synchronous oscillation by resonance via inter-inhibitory linkages, and also by effects of asymmetrical excitatory and inhibitory couplings. Notably, this alternative account predicts not a lagged relationship between excitatory and inhibitory cells, but their zero-lag synchrony (Traub et al. 1996). An important distinction is the range over which excitatory/inhibitory interactions take place, as observed in the numerical results above. Both mechanisms could operate, one embedded within the other. On the other hand, the local correlation of excitatory and inhibitory firing may be explained as an aspect of dynamical frustration, without requiring a tuned inter-inhibitory resonance.

4.2 Freeman’s Findings

The present findings account for further important aspects of Freeman’s experimental data. Emphasis can be placed on stationarity to explain the data Freeman describes as phase transition. This hinges upon the measure Freeman describes as a “null spike”.

A null spike is a zero, or a relative approach to zero, of the sum of the squares of the real and imaginary parts of the Hilbert transform, computed for some epoch on the EEG time-series. The imaginary part of the Hilbert transform is given by

\[
H(z)(t) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_{-\infty}^{\infty} z(t + \tau) - z(t - \tau) \frac{d\tau}{\tau}
\]

where \(\tau\) is temporal lead/lag and \(z(t)\) is stationary, continuous, and of infinite duration—while the real part is \(z(t)\) itself.

The real part of the Hilbert transform reaches zero where the signal undergoes a zero-crossing. The imaginary part is precisely zero when the signal is symmetrical about time \(t\), and approximates zero when the signal is randomly distributed about \(t\). Freeman’s data show a null spike sometimes associated with a zero crossing, and sometimes not, while also showing null spikes recur around the theta frequency—the frequency with dominant amplitude in most electrocortical signals. This frequency band offers opportunity for an approach to symmetry about a signal peak, and thus a null spike may occur although not associated with a zero-crossing. Thus, both the limiting cases considered in Sect. 2.10 imply the necessary occurrence of null spikes. In this interpretation, the null spike marks not the onset of phase transition, but the midpoint of a period of activity at one or other phase of a second-order phase transition.

Freeman has also emphasised the null spike as a key indicator of the outburst of spreading, rotating or contracting fields of electrocortical waves (“phase cones”). Such spreading, contracting, or rotating fields are to be expected as consequences of interaction of multiple sites of generation of synchrony, as has been argued elsewhere (Wright 2009b).

4.3 Functional Implications

These results provide mechanisms by which areas of cortex can utilize noise to remain open to new input, yet suppress noise, travel in phase-space trajectories about
basins of attraction, and emit information stored in synaptic configurations as patterns of pulse trains generated by oscillation, as required at different times, and in different parts of the cortex.

The attractor dynamics implied are not, in detail, like those of standard feed-forward or attractor neural networks. Instead, equilibrium in a system with symmetry of connections at macroscopic scale implies that in any given steady state, a single, global synchronous field can exist, and thus only a single attractor is defined. Yet, at microscopic level, asymmetry of connections predominates, so movements in phase space about the global attractor could be very complicated, as well as time-varying. The global attractor may be better defined by a non-orthogonal vector set, \( \{x_i\} \) (Sect. 2.6), rather than a single attractor. Furthermore, the dynamic properties of synapses, and the onset of autonomous firing of neurons means that the basins of attraction would themselves be time-varying. The suppression of noise as oscillation approaches the synchronous limit would increase the reliability of system performance in the manner of the simulated annealing algorithm (Kirkpatrick et al. 1983), while the persistence of some noise component prevents the system entering phase-locked paralysis. Conversely, at the background limit, there would be no clearly defined attractor states, but the reception of signals from excited cortex, or external sources, would drive the cortical surface toward a particular attractor.

A general mode of overall cortical self-control is also suggested. Inputs from subcortical systems adjust excitation in the cortical field so that mixed fields, partly excited to the threshold for oscillation and emitting specific signals, and partly in the “open”, receptive, background state, approach equilibrium collectively, and also in interaction with fresh specific sensory inputs. As overall equilibrium is attained, signals from the cortex to subcortical systems then trigger a shift of subcortical inputs to cortex, so that a new “next state”, for which a different set of cortical equilibria are specified, is generated. Such a cyclic process would provide a mechanism for the direction of attention, and the ongoing sequence of cognition—incidentally offering an embodiment of stimulus/context multiplicative interaction, shown by Mizraji and Lin (2010) to be central to information search strategies. Subcortical neural mechanisms mediating the role of the hippocampus in orientation and episodic memory are the subject of ongoing research (Tsuda and Kuroda 2004; Fukushima et al. 2007), and might act as an intermediary between cortex and reticulo-thalamic neurons. Also, a cyclic process, generating sequences in which entropy currently decreases and increases in EEG measures, has been observed by Saddy et al. (2004).

Periods of stationarity of electrocortical activity also imply continuity of activity at individual synapses, in turn allowing synaptic learning rules (e.g., Tsukada and Fukushima 2010) to operate smoothly—and when stationarity is associated with synchrony, learning could then proceed in accordance with the “coherent infomax” principle (Kay and Phillips 2010). Localized excitatory-inhibitory correlation—the consequence of dynamical frustration—might also play an important role in maintaining precise timing relations in local firing, and regulating the time and activity-dependent operations of Hebbian and STLR learning and reinforcement (Tsukada and Fukushima 2010).

Some 30 years ago, Amari and colleagues (Amari 1974; Amari et al. 1977; Amari and Maginu 1988) described general computational and memory storage properties
available in large-scale random networks, using notions drawn from thermodynamics. Recently, beim Graben et al. (2009) extended these findings more generally, within their concept of “contextual emergence”. The present results suggest application of these ideas within a framework of physiologically realistic continuum and neural network models may help unite theory and experiment.

Acknowledgements Special thanks are due to Adrienne Wright. I thank Walter Freeman, Peter beim Graben, and Alistair and Moira Steyn-Ross for discussion and criticism—the Oakley Foundation, Auckland, for financial support—the Bioengineering Institute of the University of Auckland for large-scale computing resources—and Paul Bourke and Nick Hawthorn for graphical and technical support.

Appendix: State Equations for Physiological Simulations (After Wright 2009a)

$$\phi_p(r, t) = \int f_{qp}(r, r') Q_p \left( r', t - \frac{|r - r'|}{v} \right) d^2r'$$  \hspace{1cm} (A.1)

$$M^{[R]}(V_q, \phi_p) = g^{[R]}_p \left( \frac{V^{rev}_p - V_q}{V^{rev}_p - V_{[0]}^q} \right) e^{-\kappa^{[R]}_p \phi_p}$$  \hspace{1cm} (A.2)

$$V_q(r, t) = V_{[0]}^q + \sum_p \sum_j \sum_{[R]} N_{qp} A_j^{[R]} \left( \left( (M^{[R]} \ast \phi^{[R]}_p) \ast \psi^{[R]}_p \right) \ast L^j \right)$$  \hspace{1cm} (A.3)

$$Q_q(r, t) = \left( Q_{q_{max}}^q / \left( 1 + \exp \left[ -\pi \left( V_q(r, t) - \theta_q / \sqrt{3}\sigma_q \right) \right] \right) \right) + E_q$$  \hspace{1cm} (A.4)

$\phi_p$, $V_q$, $Q_q$ are the state variables afferent synaptic flux, dendritic voltage and efferent pulse density. Here, $p = e, i$ indicate presynaptic excitatory or inhibitory cells at locations $\{r'\}$, and $q = e, i$ indicate post-synaptic cells at locations $\{r\}$. Subscript $qp$ indicates connections from $p$ to $q$, and the summations in (A.3) are over the categories of excitation and inhibition ($p$), near and far dendritic trees ($j$), and types of neurotransmitter ($[R]$).

$E_q(r, t)$ is stochastic variation in pulse emissions. $V_{[0]}^q$, $N_{qp}$, $A_j^{[R]}$, $g^{[R]}_p$, $V^{rev}_p$, $\kappa^{[R]}_p$, $v$, $Q_{q_{max}}^q$, $\theta_q$, $\sigma_q$, are constants—resting membrane potential, synaptic numbers, fractions of synapses on proximal and distal synaptic trees, “ground state” synaptic gain, reversal potentials, steady-state constants for transmitter-receptor kinetics, axonal conduction velocity, maximum firing rates, mean threshold potential and standard deviation of threshold potential, respectively.

$$\int_{\tau=0}^{\tau=\infty} \phi^{[R]}_p(\tau) d\tau = 1,$$  \hspace{1cm} (A.5a)

$$\int_{\tau=0}^{\tau=\infty} \psi^{[R]}_p(\tau) d\tau = 1$$  \hspace{1cm} (A.5b)

are normalized impulse response functions for rise/fall of channel opening in synaptic receptors, and rise/fall of dendritic voltage in response to synaptic ionic flux

$$\int_{\tau=0}^{\tau=\infty} L^j(\tau) d\tau = 1$$  \hspace{1cm} (A.6)
are normalized distribution of conduction delays in the near and distal parts of dendritic trees, and
\[
\int f(\mathbf{r}, \mathbf{r}') \, d\mathbf{r}' = 1 \quad (A.7)
\]
is normalized distribution of axo-synaptic connections.

\(A^f\) are fractions of the cell population which are currently quiescent, or actively firing at any time:

\[
A^n = 1 - \frac{Q_q}{Q_{\text{max}}} \quad (A.8a)
\]
\[
A^f = \frac{Q_q}{Q_{\text{max}}} \quad (A.8b)
\]

Estimates for physiological values of all parameters are listed in Wright (2009a, 2009b).

References


Jirsa, V. K., & Stefanescu, R. A. (2010). Neural population codes capture biologically realistic large scale network dynamics. This Special Issue, BMB.
Kay, J. W., & Phillips, W. A. (2010). Coherent informax as a computational goal for neural systems. This Special Issue, BMB.
Mizraji, E., & Lin, J. (2010). Logic in a dynamic brain. This Special Issue, BMB.
Potthast, R. P., beim Graben, P., & Wright, J. J. (2010). Emergence of cortical maps through synaptic competition and cooperation dynamics. This Special Issue, BMB.


Part III. The embryogenesis of the cortex

Home to Auckland, with more freedom for research than most, I felt that my colleagues and I had made some inroads into the neurodynamics of the cortex, but major concerns remained.

We were treating the cortex as isotropic – that is, as if its connections in all directions from a particular point were more or less similar – when anatomical evidence indicated this was certainly not so. A very large body of anatomical work, begun by Hubel and Wiesel, and built upon the minicolumnar hypothesis of Mountcastle, analysed the functional connectivity of cortical neurons in terms of their responses to specific “features” such as preferential orientation (OP) for response to a line in the visual field. The functional implication of this tuning to “features” was uncertain. Explanations for the response organisation depended on Hebbian learning in response to visual inputs, yet in mammals with long gestation periods the organisation was apparent at birth, before visual experience. Neither did neurons show singular responses to feature categories as if these were the elementary properties – instead, responses to categories were interdependent.

Another problem persisted. Since the wavelengths of electrocortical waves are very long, their information capacity must be limited, but the cortex can respond to a large number of input states, and generate a wide variety of output states, so how could the waves mediate the transformation of input to output states? It seemed possible that the pulse and wave dynamics of the electrocortical field might interact with factors affecting synaptic efficacy and growth. Long, attenuated neurons, while providing large synaptic contact area of each neuron, place demand on transport systems within the cell to supply essential metabolites throughout the cell. Competition between synapses for resources such as calcium ions takes place, and probably for many other essential substrates. If there is only enough of some critical metabolite for half the synapses to operate at maximum capacity, and if pre-synapses must compete with each other on a small local scale, then competition would maximise the Shannon entropy of the possible synaptic activity levels, and thus, if the local supply of metabolites fluctuated with firing states of the network – particularly with multiple time-scales of supply – then conditional Markov processes of great complexity, generating a wide variety of pulse states, would be possible even though the electrocortical waves were not capable of great complexity. Furthermore, during approach to synchronous equilibrium of the electrocortical field the pulse and wave dynamics would necessarily guide the distribution of active synapses towards one of many possible synaptic equilibria so that many different pulse states would each appear, in the spatially smoothed recordings, as indistinguishable.

In studying the anatomical order, one of my colleagues in Melbourne, David Alexander, had developed an idea he called “the local/global model”. He argued that cortical patchy connections mapped wide areas of surrounding cortex onto each small, local area of the cortex, taking the orientation-preference (OP) columns of the visual cortex as exemplar. The mapping could be expressed as

\[ \mathbf{P} \rightarrow \mathbf{p}^2 \]

where \( \mathbf{P} \) are complex vector positions on the cerebral cortex, and \( \mathbf{p} \) are the corresponding positions within a macrocolumn. The complex mapping described the way OP from 0 – 180 degrees appeared to be laid out over 0 – 360 degrees, about OP singularities. To explain this effect of angle doubling, we wrote

This paper suggested that the global/local maps arose because the patchy connections mapped the wider cortex (and thus part of the visual field) onto the local map in the manner of the projection of a Euclidean plane onto a Möbius strip, and that synchronous oscillation and synaptic competition mediated the evolution of the connections. I later continued the work in company with Paul Bourke, who was by then in Perth.

The earlier outline for Möbius/local/global anatomical organisation was further developed into a more complete theory of the regulation of embryonic cortical growth and the emergence of mature functional connections. This was published as


The theory draws upon the work of others on genetic diversification during neurogenesis of the cortex, and on apoptosis (cellular death) during the growth process. The main assumptions made are, firstly, that the cells that survive to maturity are those that form into an “ultra-small world” configuration, thus minimising the metabolic consumption of scarce resources needed to support their axonal connections, and, secondly, that the cells that survive are those that maximise their collective uptake of a critical synaptic metabolite, as a consequence of their synchronous firing. The cells that survive and the synaptic connections of those cells must therefore be spatially arranged to maximise their joint synchrony, while minimising the total length of their mutual connections. The theory implies the neural system evolves with the most economic deployment of resources per connections possible, while simultaneously offering the maximum possible synaptic information representation capacity.

The first of two requirements to be met is given in equation 8 as the case for two different types of cortical neurons that collectively must meet the ultra-small-world definition of Cohen and Havlin

\[
p(q,r) = N_\alpha \lambda_\alpha e^{-\lambda_\alpha |q-r|} + N_\beta \lambda_\beta e^{-\lambda_\beta |q-r|}
\]

This describes the average density of synaptic connections \( p \), generated with increasing distance of separation of cells at \( q \) and \( r \). The power function (of exponent \( A \)) for the ideal synaptic density is approximated by simplifying to two types the many populations of neurons intrinsic to the cortex – one type with long axons, of axonal inverse length constant \( \lambda_\alpha \), corresponding to patch neurons, and comprising a fraction \( N_\alpha \) of all neurons, and the other type with short axons representing the common types of intracortical neurons, characterised by \( \lambda_\beta \) and \( N_\beta \). Thus equation 8 provides terms for \( f_\omega \) in equation 1.

The second requirement is that \( J \), the RMS amplitude of total pre-synaptic pulse activity between excitatory neurons, tends to a maximum, as measured by

\[
J = \left[ \frac{1}{T} \int \int \int (\Phi_\omega - \Phi_\omega')^2 dt dr dq \right]^{1/2}
\]
where \( q_r \) is the pulse flux between \( q \) and \( r \), \( \overline{q_r} \) is the mean pulse flux, and \( T \) is the duration of a short epoch.

If synaptic resources are sufficient to enable only 50% of synapses to be active, so that pre-synaptic competition applies at all scales within neurons, and synapses must also compete for limited post-synaptic space on the afferent neurons, then the geometry of cell-body positions and the networks created by the strongest synaptic connections that lead to maximum \( J \) must be as follows:

(a) Cells with short axons must be gathered into clusters, to minimise their distance of separation and the cells with long axons are thus forced to form patchy connections at longer range. In some circumstances where patch connections are distributed over considerable distances, notably in visual cortex, this implies organisation into clearly defined columns.

(b) The patchy synaptic connections created by the cells with long-range axons must become arranged to map activity in the surrounding cortex so that the long range and short range cell systems have maximum resonance with each other. This requires that the large scale is mapped to each cluster as a continuous 1:1 map onto each of the clusters. Small-scale pre-synaptic competition prevents the 1:1 map being a simple Euclidian representation. Instead, chains of the most resource-rich synapses are forced to create a winding, spiral-like system of connections within the cluster, closed after travelling twice around the cluster centre, the 1:1 map therefore becoming analogous to representation on a Möbius strip.

(c) Adjacent clusters must array their maps in approximately mirror-image arrays, with synaptic connections bridging adjacent clusters, to maximise co-resonance between clusters.

In other cortical locations, where values of \( \lambda_{\alpha,\beta} \) differ from those of visual cortex, structures may remain diffuse, but the same functional connection system would occur. Selection for maximisation of synchrony thus explains the emergence of anatomically realistic cortical columns and column variation, and explains the appearance of response maps prior to visual experience. Consideration of the effect of axonal delays in such a network also explains the experimental relationship between stimulus properties of speed, angle of attack and object orientation\(^{12}\) that is not explained by orthodox models.

The basic assumptions of the theory were later supported by the discovery, by UK colleagues\(^{13}\), of synchronous firing and emerging small-world connectivity in neural in-vitro cultures, and by an earlier experimental finding\(^{14}\) that cultured neurons prevented from firing in synchrony were destroyed by apoptosis.

A recent test of the theory’s applicability in primary somatosensory cortex also gave confirmatory results.

---


Further experimental confirmation/refutation is planned.

An important aspect of this embryological anatomical model is that it suggests a schema for signal transmission in the mature cortex, in which the exchange of signals between cortical areas falls systematically into spatial and temporal order. Within a single cortical area, surrounding cortical areas are mapped by patchy connections into the Möbius-configuration clusters so that inputs are radially arrayed about central singularities according to distance and delay from source, and associations can be made within the cluster. Between cortical areas, signals conveyed by cortico-cortical fibres from many such clusters/maps converge and are superimposed in similar column-like organisations in the next area, enabling specific spatio-temporal features of a complex stimulus to be associated. Such relays can be continued between areas until they act upon motor outputs, and signals travelling over similar relays in the reverse sense could filter signals originating in sensory areas, thus permitting complex sensory motor sequences and attentional filtering. After birth, overwriting in Hebbian fashion can take place, with the embryonic order providing a reference framework upon which the later learning is deployed.

It now seemed possible that the solutions we had proposed for a number of problems previously considered separate might be combined to provide a sketch account of the operation of the brain – as I will outline in the next section.

In 2009, while we were visiting Reading University, my dear Adrienne was found to have developed ovarian cancer. There was no chance of extended survival. She responded to this challenge with unflinching realism, never showing depression, false hope or disillusion, but thinking of me, of our children, and our young grandchildren. This, I came to realise, was the way she had always been, so she was ready when the demand was placed on her. In her last three months she wrote a memoir – “Sorry I’ve taken so long to write to you …” – that is a thing of enduring beauty. In a way, I intend this thesis to accompany her book, as the other side of a coin.
Contribution of lateral interactions in V1 to organization of response properties

J.J. Wright a,b,* , D.M. Alexander c , P.D. Bourke d

a Liggins Institute, University of Auckland, Auckland, New Zealand
b Brain Dynamics Centre, University of Sydney, Sydney, Australia
c Faculty of Information Technology, University of Technology, Sydney, Australia
d Centre for Astrophysics and Supercomputing, Swinburne University, Melbourne, Australia

Received 15 August 2005; received in revised form 4 December 2005

Abstract
We propose a model of self-organization of synaptic connections in V1, emphasizing lateral interactions. Subject to Hebbian learning with decay, evolution of synaptic strengths proceeds to a stable state in which all synapses are either saturated, or have minimum pre/post-synaptic coincidence. The most stable configuration gives rise to anatomically realistic “local maps”, each of macro-columnar size, and each organized as Mobius projections of retinotopic space. A tiling of V1, constructed of approximately mirror-image reflections of each local map by its neighbors is formed, accounting for orientation-preference singularities, linear zones, and saddle points—with each map linked by connections between sites of common orientation preference. Ocular dominance columns are partly explained as a special case of the same process. The occurrence of direction preference fractures always in odd numbers around singularities is a specific feature explained by the Mobius configuration of the local map. Effects of stimulus velocity, orientation relative to direction of motion, and extension, upon orientation preference, which are not accounted for by spatial filtering, are explained by interactions between the classic receptive field and global V1.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Primary visual cortex; Lateral interactions; Neural networks; Orientation; Neuroanatomy

1. Introduction
This paper addresses three areas of controversy in visual neuroscience. First, the role of lateral cortical connections in visual processing within visual cortex (V1); secondly, the relationship of Hebbian learning to dimension-reduction explanations of V1 architecture; thirdly, the role played by synchronous oscillation. Considering these in order:
Stimulation of a small locale in the visual field (the receptive field (RF)) brings forth a response in the 40–60 ms lag range, in a small cortical locale corresponding to the field of projection of the RF to V1 (Angelucci & Bullier, 2003; Li, Their, & Wehrhahn, 2000), and V1 has been shown to process information from correspondingly small locales of the visual field (De Valois, De Valois, & Yund, 1979; Movshon, Thompson, & Tolhurst, 1978; Schiller, Finlay, & Volman, 1976). The classical view of feed-forward processes in the direct visual pathway is that surround inhibitory fields produce spatial filtering of input signals, enhancing movement, lines and edges, and giving rise to orientation preference (OP). Several groups have modified this view. Ringach, Hawken, and Shapley (1997) showed that intracortical feedback appears to change OP on a dynamic basis. Chavane et al. (2000) drew attention to the interaction of receptive field responses and activity in the visual cortical surround, and Series, Georges, Lorenceau, and Fregnac (2002) have produced a neural network model of lateral interactions in V1 to account for interactions of stimulus orientation and apparent speed.
The majority of synapses on V1 neurons are formed by lateral connections arising from neurons with somata distributed over a wide range of cortical separations (Braitenberg & Schüz, 1998) and modeling supports the propagation of signals in a wave-like manner via polysynaptic pathways (e.g., Jirsa & Haken, 1996; Robinson, Rennie, & Wright, 1998). When V1 is stimulated by direct pathway projections, some of the evoked responses occur at greater delays than 40–60 ms, and at greater distances than the classical receptive field (Slovín, Arieli, Hildesheim, & Grinvald, 2002), consistent with lateral spread of information via polysynaptic routes.

A number of recent papers have related the electrophysiological responses of neurons to the spatial ranges of various connection systems (Angelucci & Bullier, 2003; Angelucci et al., 2002; Levitt & Lund, 2002). A recent critical review (Alexander & Wright, 2006) shows the maximum scale of excitatory modulation in V1 may be the size of V1 itself. There is clear evidence of the importance of lateral, as well as descending, influences as major controls of activity throughout the visual pathway (e.g., Angelucci et al., 2002; Lamme, Super, & Spekreijse, 1998; Lee, 2002; Stein & Sarnthein, 2000). In some situations a neuron can be driven solely by extra-RF stimulation: foveal V1 cells can be driven by a bar whose entire extent within the RF is obscured by a patch that is in the foreground relative to the bar (Sugita, 1999); long moving bars can still fire V1 neurons when artificial scotomas cover the RF ( Fiorani, Rosa, Gattass, & Rocha-Miranda, 1992); 26% of V1 neurons will respond to the illusory contours of Kanizsa squares (Lee, 2002); over half of V1 neurons can be driven by moving gratings presented within an annulus (Cavanaugh, Bair, & Movshon, 2002); a textured field with a large hole over the RF fires 93% of V1 neurons tested (Rossi, Desimone, & Ungerleider, 2001).

Mathematical models of connections in V1 began with the “ice-cube” model (Hubel & Wiesel, 1968, 1977). This classic model has subsequently proved to have considerable limitations (Bressloff, 2002; Bressloff & Cowan, 2002). Explanatory models fall generally into two categories.

The first category includes models depending upon neural network principles, which originate from the work of von der Marlsburg (e.g., Bressloff, 2002; Bressloff & Cowan, 2002; Goodhill, 1993; Grossberg & Williamson, 2001; Linsker, 1986; Miller, Keller, & Stryker, 1989; Obermayer, Ritter, & Schulten, 1990; von der Marlsburg, 1973; Tanaka, 1989), showing that a variety of realistic spatial ordering of response properties emerges in two-dimensional arrays of neurons under Hebbian learning, with “Mexican Hat” local inhibitory surround architectures.

Models of the second category give an account of the same response organization in terms of dimension reduction (Durbin & Mitchison, 1990; Durbin & Willshaw, 1987; Kohonen, 1982; Mitchison & Durbin, 1986). Here, the higher dimensional complexity of information in the afferents is reduced to the two-dimensional cortical surface. Dimension reduction is achieved by a mapping endeavoring to satisfy goals of continuity and completeness—that is, the need to maintain local smoothness of response properties within the cortical sheet, and the need to ensure the sheet includes a compact, representative selection of points in stimulus space. Dimension-reduction mapping solutions yield results similar to the neural network models, but the interrelationship of the Hebbian and dimension-reduction accounts is unclear.

Both these classes of models encounter difficulties with some recent experimental findings—notably the sensitivity of OP to velocity, orientation relative to direction of motion, and extension, of moving textures (Basole, White, & Fitzpatrick, 2003).

At both long and short ranges, synchronous oscillation can be induced by concurrent stimulation of cortical sites (Gray, Engel, Konig, & Singer, 1992; Gray, Konig, Engel, & Singer, 1989; Singer & Gray, 1995). Consistent with experimental observations, theoretical modeling of synchrony (Chapman, Bourke, & Wright, 2002; Robinson et al., 1998; Wright, 1997; Wright, Bourke, & Chapman, 2000) shows that synchrony is a broadband phenomenon not restricted to the gamma range, is induced over bi-directional couplings, and develops in response to changes in stimuli after a lag comparable to the axonal delay between the sites. Synchrony offers a mechanism for the co-ordination of learning under Hebbian synaptic rules. Its relevance to the maximization of information stored in connections evolving under Hebbian learning has been demonstrated (Kay & Phillips, 1997; Phillips & Singer, 1997) but it is not known how this “coherent infomax” principle might become manifest in real neural organization.

We have earlier suggested (Alexander, Bourke, Sheridan, Konstandatos, & Wright, 1998, 2004) that the organization of OP arises from a tiling of the surface of V1 with local synaptic maps, each representing activity in the global map—that is, the extended surface of V1—and having a singularity within each local map corresponding to the position of the local map within the global map. In this paper, we justify the earlier hypothesis by a theoretical treatment of the development of synaptic connections consequent to lateral interactions. We will show that continuity and completeness arises in local maps as a consequence of Hebbian learning with decay, mediated by fields of synchronous oscillation, and that the most stable synaptic configuration requires that the local map has a form analogous to a Mobius strip representation of the global map.

2. Theory

2.1. Mathematical conventions

\[ X_j, Y_k; \] elements of the complex planes \( \{ X \}, \{ Y \} \)

\[ Y_j \rightarrow X_k; \] one to one map

\[ Y_j \Rightarrow X_k; \] N to one map

Scalar distances, e.g., \( |X_j - X_k| \) may be rendered as \( |j - k| \) in context.
\( A \propto B \) indicates a monotonic relation, rather than linear proportionality.

2.2. Simplified two-disk model

2.2.1. Configuration

An abstract system (see Fig. 1) which we will later apply to V1 in several modified forms, is comprised of two disks, respectively situated in two planes, \( \{ X \} \) and \( \{ Y \} \). The disks are two-dimensional continuum representations of the cortical surface, with positions within each disk given by complex numbers with ordered subscripts \( 1,2,\ldots,j,\ldots,k,\ldots,l,\ldots,n \). No explicit size constraint is required in the abstract representation. The two disks may be taken to represent areas each the size of macrocolumns, or one disk may be of macrocolumnar size, while the other may be as large as the entire extent of V1.

Symmetric bi-directional axo-synaptic couplings form uniform all-to-all connections between planes, and bi-directional all-to-all connections with synaptic density decreasing as a function of distance within each plane. That is, where \( \sigma \) indicates synaptic density; \( \sigma(X,Y) = \text{constant} \), and \( \sigma(X,X) \propto |X_j - X_k|^{-1} \).

The general decline of the density of synapses generated by cortical pyramidal cells as a function of distance from the soma (e.g., Braitenberg & Schuz, 1998) is a crucial property upon which this model depends, because of the relation of coupling density to the magnitude of synchronous oscillation, as described in the next section. The uniform coupling density between planes is an initial simplification, which will be amended in application to real anatomical systems, in accord with Section 2.4.

Between disks, all connections are excitatory. The population of neurons within each disk is of mixed inhibitory and excitatory cells, and within each disk couplings via excitatory and inhibitory synapses create “Mexican Hat” surrounds.

2.2.2. Neural wave dynamics and synchrony

Theoretical work (Chapman et al., 2002; Freeman, 1975; Haken, 1996; Jirsa & Haken, 1996; Rennie, Wright, & Robinson, 2000; Robinson et al., 1998, 2003; Wright, 1997; Wright et al., 2000, 2003) (See Appendix A) enables neurons within each disk to be considered as a polysynaptic medium for the propagation of electrocortical waves, and to act as a medium for synchronous oscillation. Explicit use of these dynamic models is not required for the simulations and results that follow, but their account of synchronous oscillation (Chapman et al., 2002; Robinson et al., 1998; Wright, 1997) provides a mechanism for the co-ordination of synaptic modification in the present model.

Each disk is driven by externally imposed or internally generated synapto-dendritic activity approximating spatio-temporal white, or brown, noise. Mean pulse rates, \( \overline{Q}_{i,j} \), are spatially uniform. Synchronous oscillation develops between reciprocally connected neurons. The mean and standard deviation of pulse trains and dendritic potentials between any two points is proportional to the sum (with appropriate sign and weight) of the densities of excitatory and inhibitory connections between the points, modified by gain factors, \( H_u \) attributable to Hebbian modification of excitatory synapses.

Where \( R \) is a measure of synchrony (which may be measured by either spike co-incidence rate, or local field potential covariance), then between disks

\[ R(X,Y) \propto \sigma(X,Y) \times \overline{P}_S(X,Y) \]  

and within planes

\[ R(X,X) \propto [\sigma_e(X,X) \times \overline{P}_S(X,X) - \sigma_i(X,X)] \]  

2.2.3. Learning rule

\( R(X,Y) \) and \( R(X,X) \) in Eqs. (1,2) are measures that are directly proportional to the aggregate of the prepost-synaptic coincidence, \( r_{Q_{i,j}} \), at individual synapses connecting \( X,Y \), \( X,X \), over short epochs, \( n \Delta t \). Synchronous oscillation acts to co-ordinate synaptic gain modification, over bi-directional couplings, according to the following learning rule. Where pre/post-synaptic coincidence is defined by

\[ r_{Q_e} = \frac{Q_{max}}{n} \sum Q_e(t) \times \varphi_e(t) \quad t=t_0,\ldots,n \Delta t, \]  

where \( Q_{max} \) is the maximum action potential firing rate of neurons. In approximately steady-state conditions the mean post-synaptic pulse rate, \( \overline{Q}_{e} \), is constant, equal to the pre-synaptic pulse rate, or afferent pulse density, \( \varphi_e \), and \( Q_{max} \gg \overline{Q}_{e} \).

Synaptic gain increases to saturation with high pre- and post-synaptic coincidence of depolarization, but decays if pre- and post-synaptic depolarization no longer coincide,
thus approximating rules of the “floating hook” type (Artola, Brocher, & Singer, 1990; Bienenstock, Cooper, & Monro, 1982; Hancock, Smith, & Phillips, 1991; Kay & Phillips, 1997). The steady-state Hebbian gain factor is
\[ H_t = H_{\text{max}} \exp[-\lambda/r_{\text{Q}_t}], \quad \lambda + \varepsilon_t, \]  
where \( \lambda \) is a suitable constant.

The rise and fall of synaptic gain in response to a brief period of high pre/post-synaptic coincidence is described by the normalized impulse response function
\[ H(t) = H_s(t - \tau) \exp \left( \frac{\beta - \alpha}{\alpha} \right) \exp[-\alpha t] - \exp[-\beta t], \]
where \( \alpha \) and \( \beta \) are rise and fall time constants. This simple time-response can incorporate different mechanisms and rates of synaptic consolidation (and thus long and short-term memory) by defining a set of \( \lambda, \alpha, \) and \( \beta \) applicable at different time-scales.

2.3. Conditions of synaptic stability in the two-disk model

2.3.1. Stability at individual synapses

Since pre- and post-synaptic firing rates approximate Poisson processes, the variance, \( \zeta^2(t_{\text{Q}_t}, \tau) \), of pre/post-synaptic coincidence measured during short epochs, \( \tau \), is proportional to the mean value, \( r_{\text{Q}_t} \), over a set of epochs, i.e.,
\[ \zeta^2(t_{\text{Q}_t}, \tau) \propto r_{\text{Q}_t}. \]  

Differentiating the steady-state synaptic gain (Eq. (4)) gives
\[ \frac{dH_s}{dr_{\text{Q}_t}r_{\text{Q}_t}} = H_{\text{max}} \frac{\lambda}{r_{\text{Q}_t}} \exp[-\lambda/r_{\text{Q}_t}]. \]  

Therefore \( S(H_s, \tau) \), the RMS deviation of synaptic gain as a function of \( r_{\text{Q}_t} \), is given by
\[ S = \frac{dH_s}{dr_{\text{Q}_t}} \times \zeta(r_{\text{Q}_t}, \tau) \]
\[ = H_{\text{max}} \frac{\lambda}{r_{\text{Q}_t}} \exp[-\lambda/r_{\text{Q}_t}] \times \sqrt{k r_{\text{Q}_t}}, \]  
where \( k \) is a constant of proportionality, so
\[ \frac{dS}{dr_{\text{Q}_t}} = H_{\text{max}} \lambda \sqrt{k} \exp[-\lambda/r_{\text{Q}_t}] (\lambda r_{\text{Q}_t}^{-7/2} - 2r_{\text{Q}_t}^{-5/2}). \]  

From Eq. (8); since \( \lambda, k, H_{\text{max}} \) are all \(+ve\), then
\[ r_{\text{Q}_t} = 0 \quad \text{then} \quad S = 0; \quad r_{\text{Q}_t} = \infty \quad \text{then} \quad S = 0; \]
\[ 0 < r_{\text{Q}_t} < \infty \quad \text{then} \quad S + ve. \]  

From Eq. (9);

setting \( \frac{dS}{dr_{\text{Q}_t}} = 0 \), solutions are \( r_{\text{Q}_t} = 0, \infty, 2\lambda/3 \).  

That is, \( S(r_{\text{Q}_t}) \) has a single maximum at \( r_{\text{Q}_t} = 2\lambda/3 \), and zero minima at \( r_{\text{Q}_t} = 0, \infty \), and therefore stable states of the synapse can occur only at either maximum saturation or zero saturation. We will subsequently refer to maximally saturated synapses as \textit{saturated} and zero-saturated synapses as \textit{sensitive}. (The latter term arises from the maximum slope of Eq. (4) which occurs at \( r_{\text{Q}_0} = 0 \), and the term is intended to emphasize that the zero-saturated synapse retains the capacity for relatively rapid and flexible change of state.)

2.3.2. Population synaptic stability

The time-decay of individual synaptic gains (Eq. (5)) favors evolution in the population toward global stability, analogous to the operation of the Metropolis algorithm in related thermodynamic systems (Kirkpatrick, Gelatt, & Vecchi, 1983). We will not prove that global stability must be attained, but will instead consider the necessary consequences if all synapses closely approach stability.

Absolute population stability (\textit{maximum} or \textit{global stability}) is reached when all synapses are either saturated or sensitive. Where \( \sigma(X_tY_k), \sigma(X_tX_k) \) are numbers of synaptic connections between \( X_tY_k, X_tX_k \), these can be partitioned into fractional terms for saturated, \( \sigma_{\text{SAT}} \), and sensitive, \( \sigma_{\text{SENS}} \), synaptic numbers, depending on the stable end point to which the individual synapses tend in their evolution. Consequently, if \( \sigma(X_tY_k), \sigma(X_tX_k) \) are now used to represent the mean RMS variations of synaptic gain in all synapses connecting \( Y_k \) to \( X_t \) and \( X_k \) to \( X_t \), then

\[ \Psi(XY) = \sum_{j=1}^{j=n} \sum_{k=1}^{k=n} \sigma_{\text{SAT}}(X_tY_k) \sigma_{\text{SAT}}(X_tY_k) \]
\[ + \sum_{j=1}^{j=n} \sum_{k=1}^{k=n} \sigma_{\text{SENS}}(X_tY_k) \sigma_{\text{SENS}}(X_tY_k) \approx 0, \]
\[ \Psi(XY) = \sum_{j=1}^{j=n} \sum_{k=1}^{k=n} \sigma_{\text{SAT}}(X_tX_k) \sigma_{\text{SAT}}(X_tX_k) \]
\[ + \sum_{j=1}^{j=n} \sum_{k=1}^{k=n} \sigma_{\text{SENS}}(X_tX_k) \sigma_{\text{SENS}}(X_tX_k) \approx 0 \]

defines global stability as the sum of RMS variation of all synaptic gains, which must reach zero at maximum stability.

2.3.3. Maintenance of uniform metabolic load in axons, and the densities of saturated and sensitive synapses

Grossberg and Williamson (2001), in work modeling the development of lateral and feed-forward connections within layers of the cortex, drew attention to the added constraint of metabolic requirements upon the kinds of connections which can be formed under Hebbian rules. We apply a related constraint here.

Competition occurs for metabolic resources within axons, so metabolic rate at all parts of the axonal system are approximately equal. If metabolic demands for saturated synapses are much greater than for sensitive synapses, then all cells must receive and give rise to similar proportions of saturated and sensitive synapses, the densities of
saturated and sensitive synapses respectively, must decline with distance from their cell bodies of origin, and must maintain a constant ratio. Given also that different members of \{R(X,Y_k), R(Y_kX)\} range in value at stability from zero to some sufficiently positive value, then each cell must give rise to both saturated and sensitive synapses.

These metabolic limitations place further constraints on (11) and (12), so that

\[ \sigma_{\text{SAT}}(X,Y_k) \sigma_{\text{SENS}}(X,Y_k), \sigma_{\text{SAT}}(X,X_k), \sigma_{\text{SENS}}(X,X_k) > 0, \]
\[ \sigma_{\text{SAT}}(X,Y_k) \propto |j-k|^{-1}, \]
\[ \sigma_{\text{SENS}}(X,X_k) \propto |j-k|^{-1}. \]

2.3.4. Overall requirements for maximum stability

As a simple property of sums of products, for any total sum of RMS variation of synaptic gain, the terms in (12), will sum more closely to zero if dense couplings of high synaptic number are matched to low values of gain variation. That is, stability is more closely approached where

\[ \sigma_{\text{SAT}}(X,Y_k) \propto [\sigma_{\text{SAT}}(X,X_k)]^{-1} \]
\[ \sigma_{\text{SENS}}(X,X_k) \propto [\sigma_{\text{SENS}}(X,X_k)]^{-1}. \]

However, this leads to contrasting requirements at saturated versus sensitive synapses—i.e.,

\[ \sigma_{\text{SAT}}(X,Y_k) \propto R(X,X_k) \]
\[ \sigma_{\text{SENS}}(X,X_k) \propto R(X,X_k)^{-1}. \]

This means that as global stability is approached, saturated synapses linking positions more densely connected (and therefore closer together) must exhibit higher pre/post-synaptic coincidence. Conversely, sensitive synapses linking positions more densely connected (and therefore also closer together) must have lower pre/post-synaptic coincidence. These contrasting requirements impose an order upon both saturated and sensitive synapses, mapping activity in each disk to the other, in a form analogous to representation of a flat surface on a Mobius strip, as will now be shown.

2.3.5. Consequences of maximized stability at saturated synapses

We will term the configuration of saturated synapses connecting \{Y\} to \{X\} when stability is approached, the input map, and the corresponding configuration of saturated synapses within \{X\} the local map. The input and local maps must converge to mutually compatible stable forms.

2.3.5.1. Within each disk. Saturated connections form as required by Eqs. (14) and (16), to produce the 1:1 connections.

\[ X_j \rightarrow X_k, \quad \text{where} \quad \sigma_{\text{SAT}}(X_kX_j) \propto |j-k|^{-1}, \]

which act to enhance the initial condition Eq. (2) of inverse relationship between distance of and amplitude of synchronous oscillation.

2.3.5.2. Between disks. As saturation develops in the connections mapping \{Y\} to \{X\}, the formation of saturated connections within \{X\} as required by (20) would be disrupted, unless the connections between the disks transmitted signals with cross-correlation declining with distance within the input map. Therefore, for overall stability to be approached, the input map must belong to the group of maps given by

\[ \frac{Y^N}{|Y|^N} \rightarrow X_k, \quad \text{where} \quad \sigma_{\text{SAT}}(X_kY_j) \propto |j-k|^{-1}. \]

Eq. (21) maps normalized distance in \{Y\} to \{X\}, with relative angles determined by the value of \( N \). The mapping is continuous, and also complete, since all \( \sigma_{\text{SAT}}(X_kY_j) \) are non-zero, as given in (13).

In an abstract mathematical sense, the input map is analogous to mapping a Euclidean plane to a Mobius plane of order \( N \) (See Fig. 2 and Appendix B). As Figs. 2 and 3 show, this abstract \( N \):1 map corresponds to the observable form of OP response organization—but to give anatomical meaning to the connections thus described, the subscripts \( j = 1, 2, \ldots, n \) must be renumbered as \( jm = 1, 2, \ldots, n \) where \( nm = 2N \), since all \( \sigma_{\text{SAT}}(X_kY_j) \) are non-zero, as given in (13).

In an abstract mathematical sense, the input map is analogous to mapping a Euclidean plane to a Mobius plane of order \( N \) (See Fig. 2 and Appendix B). As Figs. 2 and 3 show, this abstract \( N \):1 map corresponds to the observable form of OP response organization—but to give anatomical meaning to the connections thus described, the subscripts \( j = 1, 2, \ldots, n \) must be renumbered as \( jm = 1, 2, \ldots, n \) where \( nm = 2N \), since all \( \sigma_{\text{SAT}}(X_kY_j) \) are non-zero, as given in (13).

The value of \( N \) is determined by the requirements for maximum stability at the sensitive synapses. This requires that the local map be compressed so as to bring the most separated points on \{Y\} into closest conjunction in both the input and local maps, so that the \( N \) distinct groups of neurons at \( X_{jm} \) can then each be linked by the densest connection of sensitive synapses. Therefore the local map and the input map must have that value of \( N \) which maximizes \( \delta \), the distance over which positions \( jm \) and \( jm + 1 \) in the Euclidean plane are translated to become coincident in projection to a Mobius plane of order \( N \) (See Appendix B). This is given by

\[ \delta = |Y_{jm} - Y_{jm+1}| = 2|Y_j| \sin \left( \frac{\pi}{N} \right), \quad m = 1, \ldots, N \]

indicating that there are \( N \) distinct groups of neurons at all positions \( X_{jm} \), each group receiving afferents from a distinct domain defined by angular positions in \{Y\}.

2.3.6. Consequences of maximized stability at sensitive synapses

The value of \( N \) is determined by the requirements for maximum stability at the sensitive synapses. This requires that the local map be compressed so as to bring the most separated points on \{Y\} into closest conjunction in both the input and local maps, so that the \( N \) distinct groups of neurons at \( X_{jm} \) can then each be linked by the densest connection of sensitive synapses. Therefore the local map and the input map must have that value of \( N \) which maximizes \( \delta \), the distance over which positions \( jm \) and \( jm + 1 \) in the Euclidean plane are translated to become coincident in projection to a Mobius plane of order \( N \) (See Appendix B). This is given by

\[ \delta = |Y_{jm} - Y_{jm+1}| = 2|Y_j| \sin \left( \frac{\pi}{N} \right), \quad m = 1, \ldots, N \]

indicating that there are \( N \) distinct groups of neurons at all positions \( X_{jm} \), each group receiving afferents from a distinct domain defined by angular positions in \{Y\}.

The value of \( N \) is determined by the requirements for maximum stability at the sensitive synapses. This requires that the local map be compressed so as to bring the most separated points on \{Y\} into closest conjunction in both the input and local maps, so that the \( N \) distinct groups of neurons at \( X_{jm} \) can then each be linked by the densest connection of sensitive synapses. Therefore the local map and the input map must have that value of \( N \) which maximizes \( \delta \), the distance over which positions \( jm \) and \( jm + 1 \) in the Euclidean plane are translated to become coincident in projection to a Mobius plane of order \( N \) (See Appendix B). This is given by

\[ \delta = |Y_{jm} - Y_{jm+1}| = 2|Y_j| \sin \left( \frac{\pi}{N} \right), \quad m = 1, \ldots, N \]
Since $\delta$ is at a maximum for $N = 2$, this form of folding of the input and local maps maximizes stability at the sensitive synapses.

2.3.7. The emergent form of the input and local maps

Substituting the maximum of (23) into (22), and revising (20) to allow for the revised numbering convention, the input map and local maps take the form closest to global stability when

$$\frac{Y^2_{jm}}{Y_{jm}} \rightarrow X_{km}, \quad m = 1, 2 \tag{24}$$

and

$$X_{jm} \rightarrow X_{km}, \quad m = 1, 2 \tag{25}$$

where the density of saturated synaptic connections decreases as $|1 - k1|$ and $|2 - k2|$, while the density of sensitive couplings decreases as $|2 - k1|$ and $|1 - k2|$.

Fig. 2 visualizes the above relations. The local map is formed as an intertwined mesh of saturated couplings, closed after passing twice around the local map’s centre, with sensitive synapses locally link the two turns of the mesh together. The input and local maps can, in principle, emerge with any orientation, and with either left or right handed chirality, since these properties are not constrained by Eqs. (24) and (25).

2.3.8. Mutual organization of disks

Symmetrical and all-to-all couplings between disks permit the mutual and symmetrical organization of input and local maps of each disk with respect to the other, since the requisite initial condition—the inverse relation of covariance with distance within disks, required to drive the organization of the input maps—is maintained within the evolved local maps.

2.4. Effects of variation of constraints

The two-disk model is applicable to connection systems in V1, as will next be shown—but its application requires
that the basic assumptions be relaxed in ways appropriate to each context. Changes of some of the assumptions have the following consequences:

(a) If spatial covariance is negligible, or follows a rule other than decline with distance of separation in one or both disks, the mappings described by (24) and (25) cannot emerge.

(b) If connections between disks are unilateral, then reciprocal re-modeling of each map will not take place. The efferent disk will retain its prior spatial covariance, whereas the input map and connections within the afferent disk will form as described.

(c) Where connections between \{Y\} and \{X\} are not uniformly dense, but are biased in density toward some sector of \{X\}, \{Y\}, this effect will skew the orientation and/or the mirroring of the map of \{Y\} relative to \{X\}.

(d) Where signals input from \{Y\} to \{X\} are not spatially isotropic, but are biased toward some orientation, this effect will increase representation of lines of that orientation in \{X\}, and will skew the orientation and/or the mirroring of the map of \{Y\} relative to \{X\}.

Consequences (c) and (d) can be derived by considering the effects needed to minimize all terms in (11) and (12) by similar reasoning to that applied in the standard case.

2.5. Application to V1

2.5.1. Range of application

The two-disk model is applicable to the evolution of synaptic strengths during the interaction of V1 at macroscopic scale with each macrolaminar area, and also to interactions between respective macrolaminar areas, as will now be described.

Application of the model in piece-meal fashion leads to solutions maximizing synaptic stability in each case, and as no contradiction arises in the various applications made, then the collective system must also achieve stability. Inter-areal cortico-cortical connections linking V1 to higher visual areas, connections between cortical layers and reciprocal interactions with the visual pathway are excluded from specific consideration in the present account. The consequences of these exclusions are deferred to the Conclusion.

2.5.2. Initial connections

We consider only projection of the visual field onto V1 via the visual pathway, and lateral connections within V1.

Two scales are naturally imposed upon this system. The ramifications of axonal branches from the lateral geniculate nucleus terminating in layer 4C\(\alpha\) are 300 μm diameter, matching the average interpatch distance of the long-range intrinsic connections within V1 (Blasdel & Lund, 1983; Mountcastle, 1979; Nunez, 1995). This enables us to identify a local scale—approximately 300 microns in extent, and roughly equivalent to a macrolaminar, or the breadth of an oculomotoric column—approximating a disk in the two-disk model. The global scale—V1—is subdivided into elementary units of local scale. Within the local scale surround excitation and inhibition give rise to “Mexican Hat” effects at several different scales, from a fraction of a macrolaminar to fibers spanning several macrolaminar columns (Bosking, Zhang, Schofield, & Fitzpatrick, 1997; Das & Gilbert, 1999; Fitzpatrick, Lund, & Blasdel, 1985; Kang, Shelley, & Sompolinsky, 2003; Liley & Wright, 1994). At all scales these systems are capable of supporting fields of synchronous oscillation. At the global scale, polysynaptic connections link individual macrolaminar columns to the activity over much of V1 (Braitenberg, 1978; Braitenberg & Schuz, 1998).

The initial condition for the system (corresponding to that during embryogenesis) is one of diffuse connections. Subsystems required for description are:

\{Q\}; \(j = 1 \ldots n\); receptive fields in the visual plane.

\{P_{jm}\}; \(jm = 1,2,\ldots,j1,\ldots,k1,\ldots,n/2,\ldots,j2,\ldots,k2,\ldots,n\); the correspondence organization of projections of the receptive fields to V1.

\{p_{jm}\}; elements of \(P_0 \in \{P_{jm}\}\), which will become organized into local maps.

\{p'_{jm}\}; elements of local maps at \(P'\), the neighbors of \(P_0\).

2.5.3. Evolved connections

Final maps of saturated connections are listed in Table 1, and the processes leading to the emergence of these maps are discussed in the following sections.

2.5.4. The classical receptive field (cRF) and local maps

Direct axonal projections of the visual pathway cannot be assumed to transfer signals in which spatial covariance declines with distance. The two-disk model is not applicable, in accord with (a) in Section 2.4. Feed-forward spatial filter models appear more appropriate, such as the work of Ernst, Pawelzik, Sahar-Pikielny, and Tsoyds (2001) which shows that feed-forward Hebbian processes can impose OP on cortex.

Feed-forward models and the present model have an important correspondence in one respect. Spatial filter models treat the rotation of a straight line through 0–2π radians as aliased by rotation through π–2π, and in dimensional models a similar effect gives rise to the angle-doubling of OP about a singularity. The same angle-doubling effect is seen in the two-disk model. Consequently co-incident synaptic representation of OP via feed-forward connections and OP responses mediated via lateral connections (as described in the following sections) is topologically possible. While cRF fields need not correspond precisely to emergent local maps, overlap must occur if connection scales are similar. Co-incident representations also imply that cRF fields are modulated by lateral interactions.
2.5.5. Interaction of V1 with each local map via polysynaptic routes

2.5.5.1. Relation of local and global maps. Beginning from the initial conditions, the two-disk model can be applied to extended V1 (the \{ Y \} disk) acting upon any given cortical area of approximately 300 microns diameter (the \{ X \} disk). We first consider the interaction of the whole of V1 with a single local map, confining attention to the unilocular case. Activity throughout the extent of V1 conforms to the distance/covariance relations assumed in the two-disk model, for an additional reason to the variation of synaptic density with range—viz: spatial covariance in visual stimuli generally declines with increasing distance. Synchronous fields generated in V1 can spread over polysynaptic pathways to the local map, so the two-disk model is applicable under the restriction (b) of Section 2.4. The spread of signals over polysynaptic pathways is considered, at first approximation, to meet the assumption of a uniform and all-to-all input, from V1 to the macrocular sized area, which will become the local map. Therefore we expect emergence of a retinotopic map folded into the Mobius configuration within each macrocolumn. Input from the global to the local map is approximated by

\[
p_m^j \rightarrow p_{km} \quad (26)
\]

and return connections throughout V1 must follow the inverse map

\[
p_{km}^{1/2} \rightarrow p_{km} \quad (27)
\]

The return connections are necessarily weak, and cannot alter the retinotopic ordering of V1.

Extending consideration to the binocular case, where visual input from both eyes reaches areas within V1, the assumption of stochastic dependence of spatial covariance versus distance in the stimulus field breaks down, leading to failure of the model in accord with (a) of Section 2.4. Further consideration of this special case is deferred to Section 3.2.

2.5.5.2. Object representations, effect of lags, OP, and DP. The local map permits moving objects in the visual field to produce specific moving patterns of neuronal activity within each local map (See Fig. 3). This effect follows from the retinotopic organization, in Mobius configuration, which arises within the local map purely from the inverse distance/covariance relation in the inputs, and requires the inputs to have only white or brown spatial noise structure. Once this retinotopic projection has developed, further evolution of local connections under the influence of specifically structured visual stimuli will permit conditional probabilities to be represented in the local map, as follows.

The retinotopic ordering requires that some point on the local map must correspond to the position of the local map upon the global map. As indicated by the circular graphical origins in Fig. 2, the position \( p_0 \) corresponds to \( P_0 \) on the global map. Lines radiating from \( p_0 \) must correspond to lines oriented at all angles in the cRF. Therefore spatial filtering on the input pathway will lead to the production of an OP singularity at \( p_0 \), by transfer of covariances in the global map to the local map. In the same way, continuity of OP will extend from the singularity outwards into the local map.

Against this primary spatial ordering, specific object information can then be relayed. The map \( O(q,t) \rightarrow O(P,t) \) describes relay of information concerning a visual object, \( O(q,t) \) in the visual pathway, where \( r \) is a spatially uniform delay. Introducing time lags required for synchronous relations to be established between points on the global and local maps, and substituting \( jm = km \) in Eq. (26) so as to indicate the pattern of highest density input connections to each local map, we can write:

\[
O \left( \frac{\hat{P}}{P} \right)^2, t - \frac{\hat{P}}{v} \Rightarrow O(p,t),
\]

where \( \frac{\hat{P}}{v} \) is the axonal conduction delay (essentially equal to the time to establish synchrony) for signals relayed from position \( P \) in V1, to the local map situated at \( P_0 = 0 \), and \( v \) is axonal conduction velocity. Note that this object mapping is 2:1 as a consequence of the folding of the image of V1 into Mobius form, as shown in Fig. 3, and thus corresponds to the form observable by voltage-sensitive dyes and related experimental methods.

An object representation, \( O(p,t) \), can be generated by either of visual objects traveling in roughly opposite direc-
tions, which relay signals to the same position on the local map. Designating these dual objects as \( O^+(P,t) \), and \( O^-(P,t) \), the relay of their images to the local map are described by

\[
O^+(\frac{P^2}{|P|}, t - \frac{|P|}{v}) \Rightarrow O^+(p,t), \quad O^-(\frac{P^2}{|P|}, t - \frac{|P|}{v}) \Rightarrow O^-(p,t).
\]

Local waves generated within the local map by each object’s passage differ, and consequently evolution of distinct patterns of local connectivity will then develop, thus representing contingent probabilities associated with selective responses to particular objects and motions. Depending on linkages of neurons in the local map associated with \( O^+(p,t) \) or \( O^-(p,t) \) neurons may respond selectively to objects with separate directions of movement. Direction preference (DP) can thus be represented as

\[
DP(p) = O^+(p) - O^-(p).
\]

2.5.6. Interaction between neighboring local maps via monosynaptic routes

In isolation local maps would arise with random chirality and orientation. However, the consolidation of their input synapses is also subject to interactions between the local maps, mediated by local intra-cortical axons of greater length than the diameter of the local map. For the long intra-cortical connections an inverse synaptic density/range relation holds, similar to distance and densities within each local map. Inter-map connections conform to the two-disk model, apart from the added feature that adjacent maps form interconnections with each other with synaptic densities which decline with distance, rather than being uniform density all-to-all connections, so connections between local maps have the form

\[
P_{jm} \rightarrow P'_{km}, \quad \text{where} \ \sigma_{sat}(p_{km}p_{jm}) \propto \frac{|p_{km} - p_{jm}|^{\gamma}}{2}
\]

in accordance with condition (c) of Section 2.4. Consequently, synaptic stability is maximized when adjacent local maps are arranged so that each local map forms a mirror-image of its neighbor and the local maps pack in an hexagonal array. Since each local map has six neighbors, their orientations and chirality must minimize the departure from mirror orientation in each direction. Fig. 4 shows the principles of alignment of adjacent local maps.

The couplings between local maps, once they have emerged, are well suited to mediate relay of large-scale synchronous fields throughout V1 by polysynaptic pathways.

Long-range inhibitory connections have been neglected in this account. We suppose that inhibitory connections are not subject to Hebbian modification, but vary in strength conversely to the strength of excitatory connections. Thus, long-range inhibitory connections would link neurons in each local map having low coincidence with each other.

\[
J.J. \ Wright \ et \ al. / \ Vision \ Research \ 46 \ (2006) \ 2703–2720
\]

\[
\text{WORK \ TOWARD \ A \ THEORY \ OF \ BRAIN \ FUNCTION } \ I \ \ 481
\]
Figs. 5 A–C, show stages in the simulated evolution of OP in V1.

As the initial condition (Fig. 5 A) pairs of adjacent mirror-image color-wheels were placed in random positions in the frame, as indicated by the solid black lines. These initial pairs represent local maps that have developed early, under strictly local influence, in conformity with the processes diagrammed in Figs. 2 and 4. These pairs then seed subsequent organization amongst later evolving local maps, via a two-phase process, as initially discussed in Section 2.5.7. (See Appendix C for further details..) The first phase is the induction of preference for a particular orientation and chirality in emerging maps in the surrounding field, induced by the average of the activity in the wider field. This field effect is given by vector summation of OP in the completed local maps, conveyed throughout the field of V1. The second phase is the final allocation of orientation and chirality of later emerging maps, placed sequentially about each of the seeding pairs, with their chirality and orientation chosen to minimize disparity of OP at the interfaces of each local map. Since the two phases correspond roughly to the earlier and the later stages of local map evolution, temporal smoothing of the evolving field provided continuity. Propagation of newly completed local maps outwards from initial seed-pairs maps ultimately results in intersection of sets of local maps for which no very close mirroring of adjacent maps was possible, as is seen occurring in Figs. 5 B and C, and in final form in Fig. 6 A.

This procedure reproduces typical features of orientation preference maps, as can be seen by comparison of simulated (Fig. 6 A) and real data (Fig. 6 B). Linear zones, saddle points, and occasional additional aberrant singularities (which do not correspond to centers of local maps) appear as junctions where differently seeded zones intersect, while preservation of some of the idealized orderly mirroring diagrammed in Fig. 4 is seen in both real and simulated data.

As discussed earlier, mutual organization among local maps as diagrammed in Fig. 4 results in saturated synaptic connections between homologous points. In accord with this expectation, linkage between cells of similar orientation preference should occur between local maps (Malach, Amir, Harel, & Grinvald, 1993; Yoshioka, Blasdel, Levitt, & Lund, 1996), and is known to do so in fact, as is shown in Fig. 6 B.

The emerging form of OP in Fig. 5 is plausibly comparable to experimental data. In newborn ferrets, Chapman, Stryker, and Bonhoeffer (1996) showed that the detail of OP resolution increased rapidly in early post-natal life. The earliest maps seen were low-contrast, with regions of orientation-specific activity that were difficult to distinguish from noise. In our simulations the large areas of continuous OP seen in Fig. 5 A are associated with low vector magnitudes, and would be obscured by the presence of noise. The early maps matured over a period of several days into the high-contrast, patchy maps typical of adult animals, and the structure of the orientation maps was remarkably constant over time. The indistinct features in the earliest maps were always patches of the same sizes and shapes and at the same locations as in the maps obtained in subsequent recording sessions. Details of the more mature maps, including the relative intensities of individual iso-orientation domains, were also constant from one recording session to another over periods of several weeks. This relative constancy of local maps once their emergence has begun is also seen in our simulations.

Our simulations are limited, in that they do not describe the full process of synaptic consolidation, but instead rely on temporal smoothing of early and late dynamical states, and on predicted topological relations of the emergent maps. Also they do not include features of evolution of feed-forward connections in the visual pathway. However, the comparative realism of the results may indicate that
evolution of feed-forward connections is synergic with that of lateral connections.

### 3.2. Ocular dominance columns

Organization of OP into ocular dominance (OD) columns has been described by Swindale, Shoham, Grinvald, Bonhoeffer, and Hubener (2000), among others. Adjacent macrocolumns responsive to either the left or right eye alternate, organized into parallel columns for each eye. Singularities tend to lie at the centers of each OD column, and OP linear zones cross the boundary between OD columns orthogonally to the boundary.

The visual information delivered to V1 from the two eyes arises from overlapping projections of the visual field because of the lateral separation of the eyes. Their co-ordinated visual scanning also introduces a spatio-temporal lag covariance in the signals transmitted from each eye. Consequently, the assumption made for the uniocular case, that spatial cross-correlation is a declining function of distance of separation, is not valid. In this altered circumstance maximum stability may be reached by suppressing input from opposite eyes in adjacent local maps, each sharing closely similar inputs, and individually conforming to the inverse distance/covariance relation. Synergic linkage between adjacent maps will then follow a modification of Eq. (31) —i.e.,:

\[ p_{jm} \rightarrow p'_{km}, \quad \text{where } \sigma_{\text{SAT}}(p'_{km}, p_{jm}) \propto |jm - (k + d)m|^{-1}, \]

(32)

where \( d \) is a distance determined by lagged covariance of the twin signals. Maximization of stability is then achieved if all the local maps are now arranged as mirror images of their neighbors, directed in radial bands outward from the foveal point, since the disparities \( d \) are oriented in the major axes of visual scanning. Pairs of adjacent maps with opposite OD preference will form with radial symmetry —and synergy between homologous points will now occur optimally within a square, rather than an hexagonal symmetry.

The same algorithms as applied to form Fig. 6A have been used to form Fig. 7A—but in this case dual local map pairs with mirror image symmetry were placed in horizontal positions with square symmetry, creating columns. Subsequent development of intervening local maps was then simulated in the same way. The result is a more orderly and systematic mirroring of local maps in alternate columns throughout—with alignment of singularities down each OD column, and orthogonal crossing of OP between OD columns, as described experimentally by Obermayer and Blasdel (1993).

Support for a contribution to the formation of OD columns by this mechanism comes from the finding of enhanced definition of OD columns in animals with severe strabismus imposed during visual development, since this would increase ocular disparity. However, the incompleteness of this explanation is apparent from the persistence of some OD structure in animals raised with bilateral eye closure (Kandel & Schwartz, 1985).

### 3.3. Object velocity, object extension, and orientation preference

Using moving texture stimuli and optical imaging, Basole et al. (2003) recorded OP in populations of cells and single cells, as a function of stimulus velocity, direction of motion relative to orientation, and length, of small texture elements. All these variables caused apparent OP to differ, in contrast to findings with conventional, laterally extended, grating stimuli. We propose that these findings can be explained by a modulation of the cRF OP response by conditioning signals transferred to the local map from surrounding V1.

In Fig. 8, \( P_o \) and \( P_s \) are two points in V1 which are activated by a texture element, \( P'_oP'_s \), moving with a velocity \( v \), with an orientation \( \omega \) relative to the direction of motion, such that, by continued element movement, \( P'_oP'_s \) will straddle the local map situated at \( P_0 \), thus facilitating a maximal
response to the stimulus received from the cRF, which arrives at $P_0$ soon after the relay of signals to $P_0$ from $P_a$ and $P_b$. Because of conduction delays $\Delta t$, activity in the local map at $p_{aP_0}(t)$ is generated by signals at $P_a$ at $t - \frac{|P_a-P_0|}{v}$, and at $P_b$ at $t - \frac{|P_b-P_0|}{v}$. Apparent OP experimentally observed at $p_{aP_0}$ is measured as functions of the velocity, orientation and extension of $P_aP_b$ that is required to deliver simultaneous signals to $p_{aP_0}$.

Appendix D gives a geometrical argument based on Fig. 8, to show that $\Delta \psi$, the change in apparent OP, is given by

$$\Delta \psi = \sin^{-1}[\sin(\pi - \omega) \frac{|P_a-P_0| - |P_b-P_0|}{|P_a-P_b|}]$$

(33)

The results of Basole et al. can then be summarized and explained by application of (33):

- When texture elements are oriented at right angles to their direction of motion, no effects on OP is observed, and results resemble those for conventional grating stimuli. This is the case where $\omega = \pi/2$, so $P_aP_b$ and $P_0$ are parallel, therefore $|P_a-P_0| = |P_b-P_0|$, and $\Delta \psi = 0$.

- When texture elements are oriented at opposite senses to the axis of motion, the changes in OP are equal but opposite. Equivalently, with change of the angle of attack from $+\omega$ to $-\omega$, since $\sin(\pi - \omega) = -\sin(\pi + \omega)$, $\Delta \psi$ will have equal magnitude, but opposite sense.

- As the length of texture elements is increased, the magnitude of changes in OP is diminished, and results again resemble those for conventional gratings. This is the case where $|P_a-P_b|$ tends large. Consequently, $\frac{|P_a-P_0| - |P_b-P_0|}{|P_a-P_b|}$ decreases in magnitude, therefore $\Delta \psi$ tends to zero.

- As the velocity of movement of the texture elements is increased, change in OP increases in magnitude roughly proportionally to the increase in velocity. Since $\Delta \psi$ is roughly proportional to $|v|$ in (33) this result is also as expected.

The applicability of Eq. (33) has a limit where $|v| = \frac{|P_a-P_0| - |P_b-P_0|}{|P_a-P_b|} \sin(\pi - \omega)$, since $\sin \Delta \psi$ cannot exceed unity. This limit of applicability is reached when the stimulus bar passes at such a high angle that it does not pass over the receptive field of $P_0$. Further, where $|v|$ approaches $|\omega|$, the concept becomes invalid, since the stimulus representation in V1 would overtake the propagating wave front. The findings of Basole et al. suggest an asymptotic limit is approached at high stimulus velocities, but this effect has not yet been systematically studied.

3.4. Direction preference and fractures

The occurrence of “fractures”—sudden reversals in DP—has been observed by Weliky, Bosking, and Fitzpatrick (1996), and by Swindale, Grinvald, and Shmuel (2003). DP fractures have been demonstrated in the primary visual cortex of the cat and the ferret, although in the monkey strong preferences for direction of movement are not apparent in optical imaging maps (Swindale et al., 2003; Weliky et al., 1996). The species specificity may reflect the functional importance of movement.
A related topological argument for the occurrence of only odd numbers of fractures is discussed in Swindale et al. (2003). Their argument rests upon continuity properties and requires a priori assumption of the form of OP around singularities, and is therefore consistent with feed-forward spatial-filter models. Our argument shows that lateral interactions can order DP in a form compatible with the feed-forward account.

4. Conclusion

We have shown that V1 response properties may arise from Hebbian learning with decay when a stable configuration is achieved via lateral interactions. Effects of object velocity, orientation and extension which are not explained by feed-forward models are accounted for, yet there is no contradiction with the principles of feed-forward modeling and the two approaches appear to be complementary.

Our explanation generates continuity and completeness in a similar way to dimensional models, since these topological properties are transferred from the global to the local maps, and is consistent with theoretical considerations regarding synchrony, contextual learning and the coherent infomax criterion (Kay & Phillips, 1997; Phillips & Singer, 1997). That is, that learning rules of this type, coupled to a mechanism of synchrony, can lead to maximization of information storage in neural networks. It now appears that this maximization of information storage implies the development of realistic anatomical features, further validating the realism of such types of learning rule. Of central importance to this model is the way in which synapses are assumed to achieve stability with a substantial proportion of synapses remaining in the sensitive state. Through this assumption is essential for the model to be applicable to short-term learning, it is not essential for long-term consolidation of synaptic couplings. If short-term learning has established the framework of synaptic

Fig. 9. Proposed origin of direction preference fractures. Left: a field of V1 projecting to a local map located at the centre of the field. Preferential response to moving objects is reversed along an axis of the V1 field, indicated by oppositely directed arrows. A visual stimulus, $O^+$, transiting the axis, exhibits reversed changes in preferential response on opposite sides of the V1 field. Middle: projection of direction preference to the local map. White and black sides of the V1 field occur at the vertical black line, forming a DP fracture. Right: multiple fractures are possible only if the total number fractures is odd.
strengths, slow consolidation could proceed in saturated synapses, leading to permanent structural change, while sensitive synapses might wither to nothing. Biochemical and anatomical differences in long-term versus short-term memory, and synaptic pruning, might therefore be introduced into the model.

A novel aspect of our proposal is that a meshwork of saturated synapses analogous to a Mobius strip emerges within each local map. This pattern of connectivity has not been anatomically or physiologically directly detected, but since the connection meshwork would necessarily be very fine-grained there is no reason to expect it to be readily demonstrable.

According to our account, local maps achieve their distance metric from statistics of the stimulus field and of the fields of synchronous oscillation, and secondarily develop more detailed structure from details of specific visual objects. This may explain why the visual cortex appears poorly organized at birth and when deprived of visual input and exhibits greater resolution with increased exposure to visual stimuli (Blasdel & Lund, 1983; Chapman et al., 1996; Kandel & Schwartz, 1985; LeVay, Wiesel, & Hubel, 1981; Ruthazer & Stryker, 1996). The structure of input signals would need to be only white or brown spatial noise, to begin the organization of maps similar to those later observed in the mature, trained, cortex. This explanation is also compatible with the results of more selective visual deprivation (Blakemore, 1976; Blakemore, Movshon, & Slueters, 1978; Peck & Blakemore, 1975). If no visual stimuli constructed of horizontal lines were ever seen, then no local map intra-cortical connections corresponding to horizontal lines would result. Conversely, it is apparent from other experiments that innate factors contribute, outside the scope of this and other learning based models. For instance, binocular columns appear even in the absence of visual inputs (LeVay et al., 1981) and so the mechanism for OD column formation we have suggested can play only a supportive role in the emergence of the adult structure.

Our account is, in principle, subject to further experimental tests. Although analysis of connections using electrophysiological estimates of synaptic gain might reveal the Mobius meshwork connection pattern, this appears a demanding task. It may be more practical to attempt demonstration of the symmetrical folding of the V1 field within the local map, as diagrammed by solid and dashed lines in Fig. 3. Cells closely located to one another, but responsive to stimuli positioned on opposite sides of the cRF, ought to be demonstrable. Eq. (33) suggests the possibility of quantitative testing of the effects of object velocity on OP in conditions in which all terms in the equation are precisely specified. Our account of the findings of Basole et al. (2003) also has features in common with the model advanced by Series et al. (2002) to account for orientation-dependent modulation of apparent speed and unification may be possible. Third, our proposal raises the possibility that other features in the global V1 map may find representation in local maps. We have earlier suggested (Alexander et al., 1998) that CO blobs might fill the role of representation of the macular region. However, present evidence is not sufficient to determine the extent of V1 represented in local maps.

The self-organizing principles applied here to V1 may have generality beyond the primary visual cortex to that of interaction between cells at other stages of cortical vision, and to cortical processing in general. The relationship between spatial separation and covariance pertaining to visual images may be transformed to other metrics than physical distance in other sensory modalities, and at other hierarchical levels, but since axo-synaptic connection densities decline with distance on almost all spatial scales of the brain (Braitenberg, 1978; Braitenberg & Schuz, 1998) then the covariance/distance relations from which the synaptic maps emerge may have wide currency. Similar principles may apply to the relationship of V1 with higher visual areas—a problem deferred from consideration in this paper. Since all visual areas are reciprocally and richly interconnected, their joint synaptic stability might emerge by processes analogous to, but on a larger scale than, those leading to the mutual organization of local maps.

The discrete macrocolumnar structure of V1 is not apparent more widely in cortex, and this may be because of the need for the local maps in V1 and the discrete cRF connections to conform, while elsewhere, the equivalents of local maps might intermingle. The occurrence of topologically identical representations at local and global levels may also be relevant to the occurrence of interactions between scales, and between small local neural networks operating co-operatively at wide separations in the brain.

Acknowledgments

We thank the Oakley Foundation for financial support, and our colleague Dr C. J. Rennie, of the Department of Physics and the Brain Dynamics Centre, University of Sydney, for helpful discussion.

Appendix A. Wave properties underlying synchronous oscillation

Electrotonic and pulse activity in the cortex can be treated as activity in a wave medium (Chapman et al., 2002; Freeman, 1975; Haken, 1996; Jirsa & Haken, 1996; Rennie et al., 2000; Robinson et al., 1998; Wright, 1997; Wright et al., 2000). A suitable wave equation (Robinson et al., 1998) is

$$\left( \frac{\partial^2}{\partial t^2} + 2\gamma_i \frac{\partial}{\partial t} + \gamma_i^2 - v^2 \nabla^2 \right) \phi_{c,i}(p, t) = \gamma_{c,i} Q_{c,i}(p, t), \quad (A1)$$

$\gamma_{c,i}$ is the characteristic range of excitatory and inhibitory intracortical axons, respectively, $v$ is the local velocity of axonal conduction, $p$ is vector position on the cortical surface, $t$ is time, $\phi_{c,i}$ is the normalized synaptic flux at
excitatory and inhibitory synapses, and \( Q = Q_{e,i} \) represents
the firing rate of excitatory and inhibitory cells. Auxiliary
equations are

\[
Q_{e,i} = \frac{Q_{e,i}^{\text{max}}}{(1 + e^{-\pi(V_{e,i} - \theta_{e,i})/\sqrt{3}v_{e,i})}},
\]
(A2)

\[
V_{e,i}(p, t) = \sum_{(q-p)/y_i} g_q \left[ \frac{ab}{b-a} (e^{-\alpha t} - e^{-\beta t}) \right]
\]
\[
\otimes \left[ \Phi_q + \Phi_\theta \right](p, t - \tau), \ t \geq 0,
\]

where \( Q_{e,i}^{\text{max}} \) are maximum action potential rates, \( V_{e,i} \) are
dendritic field potentials, \( \theta_{e,i} \) and \( \sigma_{e,i} \) are distribution
parameters of firing-probability squash functions, \( g_{e,i} \) are
aggregate synaptic gains, \( \alpha, \beta \) are time-constants, and \( \Phi_{e,i} \) are
synaptic flux inputs from external sources.

Experimental observations (Steriade, Timofeev, & Gre-
nier, 2001) show action potential rates in alert cortex aver-
arage about 20 sps, and vary in rate much less than the full
range of firing rates possible for driven neurons. Simulation
of activated cortex associates the comparatively stable
firing rate with a transitional state, balanced between dissi-
pative and generative modes of neural activity (Wright
et al., 2003).

In common with all additive delay networks, the cortical
wave medium selectively rejects inverse-phase ("odd")
components in spontaneous or induced activity in all bi-di-
rectionally coupled excitatory elements, and selectively
retains in-phase ("even") components (Robinson et al.,
1998; Chapman et al., 2002), producing synchronous osci-
lation, and at the local map scale the relative density of
excitatory and inhibitory synapses produces greater syn-
chrony magnitudes at shorter ranges.

Appendix B. Properties of the Mobius plane

B.1. Definition of Mobius plane

We define a Mobius plane, \( \{ \hat{X}_j \} \), of order \( N(3 + ve) \) as
an abstract plane which is a generalization of a Mobius
strip, and is a simple example of a Riemann surface. Posi-
tions on the plane are given by the complex variable \( \hat{X} \),
which has common reference zero to a Euclidean plane
\( \{ X_j \} \), where

\[
j = 1, 2, \ldots, n,
\]
\[
|X_j| = |\hat{X}_j|,
\]
\[
\pm N \arg X_j = \arg \hat{X}_j,
\]
(A3)

where \( \arg X_j \) has the range \(-2\pi \), \( \arg \hat{X}_j \) has the range
\(-2N\pi \) and \( \pm \) indicates that the Mobius plane may have left
or right chirality.

B.2. Projection to the Mobius plane

The \( N:1 \) map

\[
\frac{X^N_i}{|X^N_i|} \Rightarrow Y_i
\]

is a projection of the Euclidean \( \{ X_j \} \) upon a second Eucli-
dean plane \( \{ Y_i \} \). Renumbering \( j = 1, \ldots, n \) as
\( jm = j \times \pi, j = 1, 2, \ldots, j, \ldots, n/N \) and \( m = 1, 2, \ldots, N \) so that
the subscript \( m \) defines an angular sector about the origin
of \( \{ X_j \} \), then

\[
\frac{X^N_{jm}}{|X^N_{jm}|} \Rightarrow Y_{jm}
\]

is the projection of a Euclidean plane to a Mobius plane.

The map (A5) compresses the representation of points
\( X_{jm} \) and \( X_{jm+1} \) within \( \{ Y \} \) to co-incidence. The distance
of translation to co-incidence, \( \delta \), is found by plane trigo-
nometry to be

\[
\delta = |X_{jm} - X_{jm+1}| = 2|X_{jm}| \sin \left( \frac{\pi}{N} \right).
\]

Appendix C. Simulation of maturation of OP

Simulation of the evolution of OP was performed on a
virtual plane of pixels, with hexagonal elements partition-
ing the plane. Each element ultimately becomes the locus
of a local map, with a color pinwheel, centered at the sin-
gularity within each element.

For simplicity, we assumed an initial condition arising
from the random placement of several pinwheels, in mir-
ror-image pairs, as indicated by the black bars in
Fig. 5A. These "seeding" map pairs were located randomly
and varied between simulations.

From the initial condition, two distinct, but complemen-
tary processes, which approximate early and late phases of
interaction among evolving local maps, were applied.

C.1. Early phase: Field effects on the surround exerted by
completed maps

Seeding maps were ascribed a vector of unity magni-
tude, and an orientation angle (appropriately colored)
and a mirror state as follows: The \( \theta \)th "seeding" map \( M_i \)
is characterized by:

\[
(x, y) \quad \text{the position of the centre of the pinwheel on the}
\]
\[
\text{continuous plane}
\]
\[
\psi_j \quad \text{the angle of the pinwheel, } 0, \ldots, 2\pi
\]
\[
m_i \quad \text{the mirror state, } \pm 1.
\]

From the initial condition a vector was calculated for
every other pixel in the field, which was the vector average
of the seeding maps, weighted by scalar value \( 1/d^2 \), where \( d \)
is the distance between each pixel and the ascribed pixels,
and \( n \) is a parameter which was varied from 1 to 3 in suc-
cessive simulations, to ensure that no essential property of
the simulation was sensitive to the degree of weighting.
The unit vector at any point (pixel) on the plane \((x, y)\) due to the \(i\)th pinwheel/map is
\[
P_i(x, y) = P(m, \cos(\theta_i + z_i), \sin(\theta_i + z_i)),
\]
where \(z_i\) is the angle of the vector \((x - x_i, y - y_i)\) in the range \(0, \ldots, 2\pi\). The vector \(P\) at position \((x, y)\) is the sum of the weighted contribution over all \(p\) pinwheels/maps
\[
P(x, y) = \sum_i P_i(x, y)/(x - x_i)^2 + (y - y_i)^2)^{n/2},
\]
and thus \(\theta(P(x, y)) = \arg[P(x, y)]\) gives induced average OP preference at \((x, y)\).

This process reflects the influence of all completed local maps upon the surrounding field, and the calculations were iterated at each time step, with new local maps added in accord with the procedure next described.

### C.2. Late phase: Placement of completed maps

At the subsequent time-step, values were then ascribed to a single further color wheel within an hexagonal element adjacent to one of the “seeding” maps. These newly ascribed values were those required to optimize the positioning of the new map in relation to its neighbors, minimizing the maximum angular disparity between the new map and the mirror-image of all its neighbors. Magnitude of the vector for the new map was set to unity. This process reproduces the settling into final stable states of adjacent local maps.

### C.3. Temporal smoothing

The two phasic processes were each iterated at each time-step until the evolution of the field was complete. Temporal smoothing of successive time-steps then merged the early effects produced by existing local maps, with the final consolidated positions of new maps, into a continuous growth process.

Temporal smoothing was given by
\[
P(x, y, t) = \sum_\tau P(x, y, t, \tau)/(1 + |\tau|)^n, \tag{A9}
\]
where \(\tau\) is lag from the present time-step, \(t\), and \(n\) is again a parameter varied from 1 to 3 in successive simulations to check that sensitivity of convergence to the final state was low.

### Appendix D. Effect of velocity on orientation preference

Figs. 3 and 8 are schematic representations showing how the image of a moving object projected by the direct visual pathway to global V1, becomes relayed to a local map. Geometric relations arising from Fig. 8 are:

If \(O^+\) is traveling at a velocity \(v\) at angle \(\omega\) with respect to \(P_aP_b\) and angle \(\psi\) with respect to the reference frame of \(\{P\}\), then
\[
P_a = P_b + v\|a\| \quad \text{and} \quad P_a = P_b + v\|a\|,
\]
where
\[
\Delta\psi = \sin^{-1}\left[\frac{\sin(\pi - \omega)}{\|b\|} \left(|P_b - P_o| - |P_b - P_0|\right) / \|P_a - P_b\|\right].
\]

### References


We describe a model for cortical development that resolves long-standing difficulties of earlier models. It is proposed that, during embryonic development, synchronous firing of neurons and their competition for limited metabolic resources leads to selection of an array of neurons with ultra-small-world characteristics. Consequently, in the visual cortex, macrocolumns linked by superficial patchy connections emerge in anatomically realistic patterns, with an ante-natal arrangement which projects signals from the surrounding cortex onto each macrocolumn in a form analogous to the projection of a Euclidean plane onto a Möbius strip. This configuration reproduces typical cortical response maps, and simulations of signal flow explain cortical responses to moving lines as functions of stimulus velocity, length, and orientation. With the introduction of direct visual inputs, under the operation of Hebbian learning, development of mature selective response “tuning” to stimuli of given orientation, spatial frequency, and temporal frequency would then take place, overwriting the earlier ante-natal configuration. The model is provisionally extended to hierarchical interactions of the visual cortex with higher centers, and a general principle for cortical processing of spatio-temporal images is sketched.

Keywords: synchronous oscillation, cortical development, synaptic organization, cortical response properties, cortical information flow

INTRODUCTION

During its embryological development the mammalian brain differentiates from a group of stem cells into an organized form ready to begin a life-long adaptive interaction with signals from the sensory environment. At the beginning of extra-uterine life, despite exposure to a limited milieu, it is somehow already organized to begin this engagement, as tho a matrix of connections has formed in which signal flows are pre-adapted to learn specific recurring patterns of the experiential world. A large body of work, following the pioneering work of Hubel and Wiesel (1959), has addressed just this issue, taking as the main target for research the primary visual cortex (V1). The majority of this work has sought to understand the emerging connections in terms of stimulus “features”—that is, elementary properties of sensory stimuli—rather than as a process independent of sensation until the post-natal stage. Our approach depends on alternative assumptions. Here we summarize and extend our earlier work (Wright et al., 2006; Wright and Bourke, 2008, 2013; Wright, 2009, 2010) relating the basic dynamics of neuron firing and competition among developing neurons for the resources needed for their growth, to the emergent connections at birth.

Our model draws on two recent experimental observations. Firstly, neurons in neonatal cerebral cortical slices show increased apoptosis when their capacity to enter into synchronous firing is disrupted by pharmacological means (Heck et al., 2008). Secondly, embryonic neurons developing in vitro develop synchronous firing, and as their growth proceeds, show self-organization into “small world” networks (Downes et al., 2012).

We propose that the synchronous firing and protection from apoptosis are directly causally related, because during cortical embryogenesis there is competition among developing neurons and synapses, which, although mediated by trophic factors (Harris et al., 1997; van Ooyen and Willshaw, 1999; van Ooyen, 2001) is ultimately a competition for available metabolic energy and/or some other scarce resource needed to promote metabolism (Montague, 1996; Thomaidou et al., 1997). We suppose that pre-synaptic pulse synchrony increases uptake of critical metabolic resources by some action not presently specified, and we argue that the assembly of cells that maximizes synchronous firing, and thus energy uptake, is also that which has the minimum metabolic cost per neuron in the length of axonal connections—the combination optimum for their survival.

Synchronous oscillation of pulses and local field potentials is a ubiquitous aspect of cortical activity (Eckhorn et al., 1988, 1990; Gray et al., 1989; Bressler et al., 1993; Singer, 1999) and has been proposed as a mechanism solving the “binding problem” of perceptual grouping and cognitive processing (Eckhorn et al., 1990; Singer, 1999; Crick and Koch, 2003). “Synchrony” refers to the broadband cross-correlation of neuron firing and field potentials at zero time-lag. The mechanism of origin of synchrony itself is controversial. In this paper we rely on an explanation that appears best applicable to the synchrony seen in neuron cultures, brain slices, or the early embryonic brain, and depends on a universal property of networks with summing junctions, including dendrites (Robinson et al., 1998; Wright et al., 2000; Chapman et al., 2003).
GEOMETRY OF RESPONSE ORGANIZATION IN THE DEVELOPED BRAIN
Since the discovery that individual cells in the primary visual cortex (V1) respond with an orientation preference (OP) to visual lines of differing orientation (Hubel and Wiesel, 1959), analysis of the response organization and its relationship to cortical function has remained both conceptually influential and controversial (von der Malsburg, 1973; Willshaw and von der Malsburg, 1976; Swindale, 1996). The surface organization of OP in V1 has recently been compared with appropriate random surrogates, and shown to exhibit significant hexagonal periodicity, in which each roughly delineated macrocolumn unit exhibits all values of OP arrayed around a pinwheel. Varying chirality and orientation of the pinwheels achieves continuity of OP at the columnar margins, thus producing zones of irregular but continuously varying OP, known as linear zones and saddles.

Some species exhibit little or no sign of this hexagonal and continuous ordering, and because of the marked interspecies variation, serious doubt has been expressed that the pattern of OP singularities (the singularity taken as demarcating the center of a macrocolumn) show this distance to be relatively constant over a 40-fold variation of body size, and related size of V1 (Kaschube et al., 2010; Keil et al., 2012). Models using symmetry arguments indicate that macrocolumns must undergo divisions during cortical development to maintain uniform surface density of singularities (Wolf and Geisel, 1998; Oster and Bressloff, 2006). Kaschube and colleagues conclude that self-organization has canalized the evolution of the underlying OP maps into a single common design—subject to the proviso that, from further symmetry arguments, this can only be the case where long-range interactions between developing macrocolumns, suppressing some possible connections, can take place. Thus, in animals with very small V1, this organization breaks down, creating a “pepper and salt” OP map pattern (Meng et al., 2012).

THE SUPERFICIAL PATCH SYSTEM
A related puzzle of V1 organization is posed by the superficial patch system. This system, composed of relatively long-range, largely excitatory (Hirsch and Gilbert, 1991; McGuire et al., 1991) patchy connections (Gilbert and Wiesel, 1979; Rockland and Lund, 1983) is ubiquitous in cortex (Muir and Douglas, 2011) and has a functional relationship to OP. Patchy connections develop before sensory afferents reach the cortex (Price, 1986; Callaway and Katz, 1990; Durack and Katz, 1996; Ruthazer and Stryker, 1996) but do not arise or terminate in the vicinity of OP singularities. Instead, near singularities, connections are apparently diffuse and local (Sharma et al., 1995; Yousef et al., 2001; Marín et al., 2003; Buzás et al., 2006; Muir and Douglas, 2011). Patchy connections link areas of common OP ("like-to-like") over distances several times the diameter of a macrocolumn (Gilbert and Wiesel, 1989; Buzás et al., 2006; Muir et al., 2011), and are periodic on roughly the same interval as OP, and are largely patch-reciprocal (Rockland and Lund, 1983; Angelucci et al., 2002). It has been shown that development of patchy connections must depend on the supply of organizing information from the neural field, and is not explicable from considerations of local neural growth per se (Muir and Douglas, 2011). Just as for maps of response properties, there is variation of patchy connection orderliness between species. Muir et al. (2011) have pointed out that those species with less orderliness have smaller visual cortices and/or less defined organization of “like-to-like” connections—an argument congruent with the findings on brain size, orderliness of response maps, and surface density of OP singularities cited above (viz. Kaschube et al., 2010; Keil et al., 2012, etc.).

PROBLEMS OF STANDARD MODELS OF FEATURE RESPONSES
Explanations of organization of OP have been undertaken in a group of now-classical theories, which we will refer to as "standard models," following the comparative description of Swindale (Swindale, 1996). Descriptive dimension reduction methods (Kohonen, 1982; Durbin and Willshaw, 1987; Durbin and Mitchison, 1990) show that the response maps of OP, eye preference (OC), direction preference (DP), and spatial frequency preference (SF) are consequences of requiring continuity and completeness of representation of each response property, in a two-dimensional representation in which every type of response property occurs within any small area on the surface of V1 (Swindale, 1996; Carriera-Perpiñán et al., 2005). The same ordering is also explained as a consequence of competitive Hebbian learning among small neighborhood assemblies of excitatory neurons, driven by spatially filtered cortical noise. Separate spatial filters each distinguish a type of response, and total synaptic gain is conserved during the training (Grossberg and Olson, 1994).

Classical standard models depend on seeding with oriented lines, in one way or another (von der Malsburg, 1973; Swindale, 1982, 1992; Durbin and Mitchison, 1990; Obermayer et al., 1990, 1992; Tanaka, 1990; Miyashita and Tanaka, 1992; Grossberg and Olson, 1994) and recently, initial belief that primary response to static oriented lines in the visual field forms the basis of OP maps has been undermined in two ways:

Firstly, in large species particularly, maps of OP appear in the cortex prior to visual experience (Wiesel and Hubel, 1974; Blakemore and Van Sluyters, 1975; Sherk and Stryker, 1976). This problem has been addressed by arguments for the normal occurrence of line-like structure in ante-natal retinal input (Albert et al., 2008; Ringach, 2007; Paik and Ringach, 2011). In contrast to all the above models, Kang et al. (2003) have proposed a model which breaks with the traditional dependence on the primacy of lines, and depends instead on time-invariant correlations in cortical “Mexican Hat” inhibitory surrounds. This model accounts

Frontiers in Computational Neuroscience www.frontiersin.org February 2013 | Volume 7 | Article 4 | 2
successfully for the apparent isotropy of local intracortical connections and the observed uniformity of sharpness of definition of OP independent of proximity to singularities, and provides a mechanism which might plausibly operate before eye-opening. It requires instead, that LGM inputs to cortex become tuned according to orientation. A further model avoiding the problem of ascription of OP as a primary, stimulus dependent property, explains the conjoint development of OP and ocular dominance columns as a consequence of Hebbian connection formation driven by correlation of visual inputs as a declining function of retinotopic distance of separation at short distances, and reversed correlation of activity in ON and OFF V1 simple cells at greater distances (Erwin and Miller, 1998). All these models however, result in the emergence of OP as a property of line orientation alone, rather than as one attribute of some more complex mechanism of feature response.

Secondly, and more recently, Basole and colleagues, who tested OP using stimulus lines moving at different speeds, and oriented at differing angles to the line of movement of the stimulus, found OP to be a function of these variables to such a degree that for lines oriented non-orthogonally to the direction of movement, OP could vary progressively with increments of speed to an asymptotic limit of 90° (Basole et al., 2003, 2006). Longer lines showed less variation of OP with increasing speed. This finding challenged all models which depended on OP being a fixed "feature" of cortical response, whether or not direct visual stimuli was required to prime the process of self-organization. Basole and colleagues at first concluded that the primal stimulus characteristics are not isolated features such as orientation, direction and speed, but a single characteristic—the "spatio-temporal energy"—that is, the combined spatial and temporal Fourier components of the moving visual stimulus’ projection to V1. Subsequent workers explained these results by retaining OP as a primary characteristic, and adding separate consideration of the temporal and spatial frequencies associated with the moving stimuli (Baker and Issa, 2005; Mante and Carandini, 2005; Basole et al., 2006). This analysis was consistent with earlier single unit results, in which tuning of V1 neurons to spatial and temporal frequencies was demonstrated (DeAngelis et al., 1993). Issa and colleagues (Baker and Issa, 2005; Issa et al., 2008) reported that a total of six parameters are required to explain response maps—OP, SF preference, and temporal frequency preference, and the tuning bandwidths of all three. This account is referred to as the spatio-temporal filter model. How these response characteristics arise during cortical development and how neurons become tuned to just those features is the subject of continuing research (Rosenberg et al., 2010), and of this paper.

In common with the model of Erwin and Miller (1998) and that of Kang et al. (2003) the model reviewed here depends upon time-average correlations—that is, the common occurrence of synchronous oscillation in the cortex—although it does not share their other assumptions or conclusions. It seeks to avoid the ascription of “features” as primary characteristics, and to explain both the findings of Basole et al. (2003) and the empirical reduction to alternative feature attributes used in the spatio-temporal model, as well as explaining the emergence of the anatomical features described above.

**DESCRIPTION OF MODEL**

**NEURAL FIELD EQUATIONS**

As alternatives to neural network models, lumped neural models and neural field equations have been expressed in many forms (e.g., Wilson and Cowan, 1973; Freeman, 1975; Haken, 1996; Amari, 1977; Nunez, 1981; van Rotterdam et al., 1982; Jirsa and Haken, 1996; Robinson et al., 2001; Wright et al., 2003; beim Graben, 2008; Bressloff, 2012). These offer means of approximating the properties of ensembles of cells on a larger scale than neural networks per se. Here we have used a generic form of neural field equations to represent an idealized, isotropic, neural field, representing the developing cortex as if it were not subject to apoptosis—a potentiality from which connections are selected during development. The scale of the field is that of a cortical area such as V1, representing intracortical connections rather than cortico-cortical. Thus, the density of connection between neurons declines with increasing separation of their cell bodies (Braitenberg and Schüz, 1991). The high non-linearity of synapto-dendritic summations are linearized at the field level, and axonal conduction speed is considered single-valued. Subject to these strictures, the following equations include features relevant to the present context:

\[
\psi_p^q(t) = j_p^q \times Q_p \left( r, t - \frac{q - r}{v} \right) \tag{1}
\]

\[
\psi_p^q(t) = M_p^q \ast \psi_p^q(t) \tag{2}
\]

\[
\Psi_p^q(q, t) = \int_{r'} \psi_p^q(t') dr' \tag{3}
\]

\[
V_p(q, t) = \sum_{p = e, i, p = i} G_p \ast \Psi_p^q(q, t) \tag{4}
\]

\[
Q_p^q(q, t) = f_2(V_p(q, t)) + E_p(q, t) \tag{5}
\]

Subscript \( p = e, i \) refers to excitatory or inhibitory neurons; superscript \( qr \) refers to synaptic connection from \( r \) to \( q \) where \( q, r \) are cortical positions occupied by single neurons.

\( \psi_p^q(t) \) is the flux of pulses reaching pre-synapses at the neuron at \( q \), from the neuron at \( r \).

\( \psi_p^{qr}(t) \) is the synaptic current generated by \( \psi_p^q(t) \).

\( \Psi_p^q(q, t) \) is the aggregate synaptic current of type \( p \) generated at \( q \).

\( V_p(q, t) \) is the soma membrane potential (relative to the resting potential) generated at \( q \).

\( Q_p^q(q, t) \) is the pulse emission rate at \( q \).

\( j_p^{qr} \) is the probability density of occurrence of pre-synapses generated by axons of the neuron at \( r \) terminating at \( q \).

\( v \) is axonal conduction speed.

\( M_p^{qr} \) is the steady-state term in a convolution transforming pre-synaptic flux to synaptic current.

\( G_p \) is the steady-state term in a convolution transforming synaptic flux into dendritic potentials.

\( f_2(V_p(q, t)) \) is a sigmoid function describing the local conversion of dendritic potentials into the rate of generation of action potentials.
$E_p(q, t)$ is a driving signal noise, arising from intrinsic random cell action potentials.

Restriction of the field to the scale of a cortical area carries several implications important for the model, all because the probability of connections between any two neurons declines with distance of separation. Firstly, descriptively we can consider “reciprocal couplings” as an idealization/representation of field coupling symmetry, and in many instances reciprocal couplings will in fact exist. Secondly because of more generally field coupling symmetry, and in many instances reciprocal couplings will in fact exist. Firstly, descriptively we can consider “reciprocal couplings” as an idealization/representation of with distance of separation. Secondly because of more generally dense connections among near neighbors, smoothing at dendritic summation requires that $Q_p(q, t)$ is spatially and temporally “brown”—i.e., has high correlation at short distances and times of separation. Thirdly, the average “degree” of separation—i.e., the average number of neighboring cells traversed by synaptic connections linking one cell to another—will also increase in proportion to physical distance of separation.

Experimental observations (Freeman, 1975, 1991; Hassenstaub et al., 2005) show intrinsic cortical oscillation arises from alternating excitatory cell and inhibitory cell firing at lags $\frac{1}{4}$ of the period of oscillation. Simulations of the oscillations (Wright, 2009, 2010) show that traveling waves are thus generated, the intersection of which produces broadband synchrony. In conditions of uniform cortical excitation without strong perturbation from external inputs the exchange of pulses between all cells reaches an equilibrium—that is, a steady-state of changes in synaptically connected pairs—of synaptic connections linking one cell to another—will increase in proportion to physical distance of separation.

$\bar{\phi}_q$ is the time-average presynaptic flux, uniform throughout the cortical field. The equilibrium reached implies differences in timing between the firing of excitatory and inhibitory cells. The interaction of excitatory and inhibitory cells ($p_1 \lor p_2 = e$, and $p_1 \land p_2 = i$) leads to closely correlated firing of both cells if they are very closely situated, as a consequence the similar local values of $E(q, t)$ equation (5), while $\frac{1}{4}$-cycle-out-of-phase oscillation develops between more separated excitatory and inhibitory cells. Inhibitory/excitatory or excitatory/inhibitory interactions ($p_1 \land p_2 = e$, or $p_1 \lor p_2 = i$) between reciprocally connected neurons lead to zero-lag synchrony, and since conduction delays are short compared to the period of oscillation, the equality of equation (6) is generally approached even when $T$ is smaller than the period of oscillation (Chapman et al., 2002). As there are equal time-lags in both directions of conduction excitatory pulse trains throughout the cortex have maximum correlation at zero lag—i.e., where $\bar{\Omega}_r$ is the time-average firing rate—also uniform throughout the cortical field -

$$ (Q_e - \bar{\Omega}_r)(r', t) \approx (Q_e - \bar{\Omega}_r)(q, t) $$

Figures 1 and 2 show these properties generated in a simulation of cortical dynamics with physiologically realistic parameters (Wright, 2009, 2010). In conditions of strong cortical excitation local oscillation is autonomous and corresponds to cortical gamma rhythm, while in conditions of lower cortical excitation, damped gamma oscillation, and a predominance of background $1/f^2$ is seen.

**MAGNITUDE OF PRE-SYNAPTIC PULSE SYNCHRONY**

Zero-lag synchronous oscillation thus entails presynaptic pulse synchrony, with a magnitude of presynaptic flux variation which can be defined respectively for individual synapses, individual cells, and in aggregate, as

$$ J_{qr} = \left[ \frac{1}{\tau} \int (\psi_{qr} - \bar{\psi}_e)^2 dt \right]^{1/2} $$

$$ J_q = \left[ \frac{1}{\tau} \int (\psi_e - \bar{\psi}_e)^2 dt \right]^{1/2} $$

$$ J = \left[ \frac{1}{\tau} \int \int (\psi_{qr} - \bar{\psi}_e)^2 dt dr' dq \right]^{1/2} $$

$J_{qr}$ is RMS presynaptic flux variation between $q$ and $r$, $J_q$ is the sum of $J_{qr}$ at a single excitatory neuron, and $J$ is the aggregate of $J_q$ over the cortex.

**SELECTION OF SCALE-FREE SMALL-WORLD CONFIGURATIONS OF NEURONS**

For any given level of cortical excitation, $J$ is greatest for that ensemble of $C$ connected neurons, in which excitatory pulses arrive at dendrites, from all sources at differing distances of separation, as closely in-phase as possible, so as to maximize their summation. Axonal delays, small compared to the period of gamma oscillation, contribute a phase difference between cell firing at $r'$ and the arrival of presynaptic pulses at $q$, of

$$ \Delta \Phi_{qr} = 2\pi \left| q - r' \right| / P_v $$

where $P$ is the period of oscillation. Therefore that ensemble selected by its capacity to maximize presynaptic synchrony must approach minimal total axonal length, $L = \int_r \int_q \int |q - r'| dq dr', dq'$, and minimization of this length minimizes the metabolic requirements of the axons.

It has been shown generally (Cohen and Havlin, 2003) for all systems of connected elements, the path length in a topological sense is at a minimum where degree distribution follows a power law. As was pointed out in conjunction with equations (1–5), in our idealized neural field, average degree of separation, in the topological sense, increases linearly as metric distance of separation of the cell bodies, so that if $L$, their total length of axonal connections, is minimal, then the path length in the topological sense is also minimal, and the degree distribution is that of a scale-free, or ultra-small world. Therefore, the connection density between cells vs. their metric distance of separation should also be approximated by a power-law distribution. Further, according to Cohen and Havlin

$$ L \sim \log \log C $$
so the metabolic efficiency of the connection system is further enhanced if the surviving cells are linked into a continuum, as opposed to separate pools of neurons.

In accord with equation (1), the number of neighboring excitatory cells connected to a given excitatory neuron, as a function of distance of separation, is proportional to $2\pi \times f_{qr}^{\alpha} (|q - r|)$—so the ensemble of neurons selected by greatest synchrony must have a connection density function of the form:

$$f_{qr}^{\alpha} \sim (2\pi |q - r|)^{-A} A > 0$$

(13)

Intracortical axonal trees have approximately exponential density/range relations (Scholl, 1956; Braithenberg and Schüz, 1991), and a power function is fitted exactly by an infinite sum of exponential functions—i.e.:

$$(2\pi |q - r'|)^{-A} = \frac{1}{\Gamma(A)} \int_{0}^{\infty} u^{A-1} \exp[-u2\pi |q - r'|] du$$

(14)

so an ultra-small-world connectivity can be achieved by sets of populations of cells with differing axonal characteristic lengths. During embryogenesis primal cells divide sequentially by layer (Rakic, 1988; Shi et al., 2012) with differences in growth pattern and characteristic axonal length programmed in sequential cell divisions. For simplicity, we consider only two populations of excitatory cells, with cell bodies partially separated by layer, but with intermingled axonal and dendritic trees, and axonal tree connection probabilities described by:

$$f_{qr}^{\alpha} = \frac{N_{\alpha}}{N} 2\pi \lambda_{\alpha} \exp[-\lambda_{\alpha}2\pi |q - R|]$$

(15)

$$f_{qr}^{\beta} = \frac{N_{\beta}}{N} 2\pi \lambda_{\beta} \exp[-\lambda_{\beta}2\pi |q - r|]$$

(16)

$$f_{e}^{\alpha} = f_{qR}^{\alpha} + f_{qr}^{\alpha}$$

$$f_{e}^{\beta} = f_{qR}^{\beta} + f_{qr}^{\beta}$$

$f_{qR}^{\alpha}$ refers to the axonal trees with longest axonal extensions, and $f_{qr}^{\beta}$ refers to the axonal trees with short axonal extension, thus $\lambda_{\alpha} < \lambda_{\beta}$. $N_{\alpha} + N_{\beta}$ is the number of synapses received/generated by each cell. Distances from $r'$ to $q$ are substituted as $r, R$ to indicate equal distances, $|q - r|$ and $|q - R|$, measured along the axonal trees of the respective populations.

The further defining characteristic of small-world connectivity—the occurrence of connection nodes—emerges as a consequence of the formation of the superficial patch system, as follows.
FIGURE 2 | Simulated background electrocortical activity, in conditions of low cortical excitation. Graphical format is the same as in Figure 1.

THE ORIGIN OF THE SUPERFICIAL PATCH SYSTEM

The two populations of cells described by equations (15) and (16), and the synapses they give rise to can be referred to as α-cells and synapses, and β-cells and synapses. We first make a provisional assumption (later justified on a species-specific basis) that $N\beta >> N\alpha$, so that α-cells with long-range axons are embedded among much more numerous β-cells, all with sparse connectivity. Equation (10) can be written by separately summing contributions from α-cells at positions $\{q\alpha\}$ and β-cells at positions $\{q\beta\}$, to give:

$$J = \int_{q\alpha} \int_{R} J^{qR} dq\alpha dR + \int_{q\beta} \int_{r} J^{q\beta} dq\beta dr$$

so $J$ is at a maximum if $\int_{q\alpha} \int_{R} J^{qR} dq\alpha dR$ and $\int_{q\beta} \int_{r} J^{q\beta} dq\beta dr$ are individually at maxima. Applying equations (15) and (16) via equation (1) to find values of $J^{q\beta}$ in equation (8) as functions of $|q - r, R|$, shows that:

$$J^{q\beta} = J^{qR} \quad \text{if} \quad |q - r, R| = x$$

$$J^{q\beta} > J^{qR} \quad \text{if} \quad |q - r, R| < x$$

$$J^{q\beta} < J^{qR} \quad \text{if} \quad |q - r, R| > x$$

Consequently $\int_{q\beta} \int_{r} J^{q\beta} dq\beta dr$ is at a maximum if β-cells are clustered so they make reciprocal connections at minimum distance and maximum density (β-clusters). β-cells at the center of β-clusters, for which $J^{q\beta}$ attains the maximum possible value, must give and receive all their connections as β-connections to a radial distance of $x$.

Since β-clusters are clustered, α-cells necessarily are also clustered (α-clusters), and since maximization of reciprocal β-connections excludes formation of short-range reciprocal α-connections, α-cells must form reciprocal synaptic connections at distances greater than $x$, to maximize $\int_{q\alpha} \int_{R} J^{qR} dq\alpha dR$. Similarly, reciprocal connections between α- and β-cells must occur at cluster margins, over distances approximate to $x$. Since we made the provisional assumption that $N\beta >> N\alpha$, then fitting the sum of equations (15) and (16) to a power function requires $\lambda_\alpha << \lambda_\beta$. Consequently α-cells may form multiple patches of synaptic connections, skipping from α-cluster to α-cluster.

Since β-clusters have radius $x$ and α-clusters are separated by distance $x$, α-clusters are necessarily placed at the vertices of hexagons tiling the cortical surface, with each hexagon embracing a β-cluster. Analogy to the superficial patch system in some species is apparent.

As noted earlier, hexagonal symmetry of OP and the superficial patch system is an idealization that is roughly approached in some species, while in others it is effectively absent (Horton...
and Adams, 2005). Since approximation of a power law distribution by two populations of neurons requires \( N_\beta < < N_\alpha \) if \( \lambda_\alpha < < \lambda_\beta \), this case is more closely approached for larger cortical sizes, and the patchy connection system will have higher orderliness and hexagonal rotational symmetry. If \( \lambda_\alpha < \lambda_\beta \) by only a small amount, as in animals with small cortical size, then \( N_\beta \) is not necessarily greater than \( N_\alpha \), and an ordered hexagonal structure need not be apparent. Such reduction of the apparent orderliness does not imply the absence of “small world” connectivity, nor imply impairment of function. The comparative invariance of distance between OP singularities across species reported by Kaschube et al. (2010) and Keil et al. (2012) implies invariance of distance between OP singularities across species and pre-synaptic synergy, including the limiting rate of metabolic energy supplied to excitatory synapses—viz:

\[
\psi_e^{qr}(t) = \Gamma^{qr} M_e^{qr} \ast \psi_e^{qr}(t) \tag{19}
\]

\[
M_e^{qr} = D \times S \tag{20}
\]

\( \Gamma^{qr} \) is the available fraction of the metabolic supply rate needed to attain maximum current flow. Since we have assumed increasing synaptico current in synchronously activated synapses increases the available metabolic supply, the value of \( \Gamma^{qr} \) must follow that of \( \psi_e^{qr} \).

\[
D = \frac{1}{B} \exp[-Bt] \quad B > 0 \tag{21}
\]

represents impulse decay following delivery of an afferent action potential, with time-integral of 1 (after Rennie et al., 2000).

\[
S = 1/(1 + \exp[-g(\mathcal{J}(t))]) \tag{22}
\]

is a sigmoid function with range 0–1, representing synaptic adaptation to the afferent pulse rate, and including the effect of pre-synaptic co-operation (Tsukada and Fukushima, 2010) upon individual synaptico current flow as \( g(\mathcal{J}) \)—a suitable ascending function in \( \mathcal{J} \), such that if \( \mathcal{J} = 0 \), there is no current flow at the synapse.

As well as inter-cellular competition between assemblies of neurons, we assume competition takes place between adjacent individual synapses arising from the same neuron. Therefore those neurons that survive apoptosis must have found an efficient deployment of resource to the synapses best positioned to maximize the magnitude of synchrony. Any two adjacent synapses arising from the same pre-synaptic neuron may terminate on the same, or different, post-synaptic neurons. If they terminate on the same neuron their conditions are essentially identical in terms of equations (19–22). If they terminate on different neurons, then the relevant values of \( \mathcal{J} \) need not be identical—and their competition for resources would lead, via the feedback between \( \psi_e^{qr} \) and \( \Gamma^{qr} \), to low synaptic current at one synapse, and high current at the other. Just what the physiological corollary of these opposite high and low-activity states is, and the critical metabolic component for which the synapses compete, we do not specify. A likely, but by no means unique contributing factor is the supply of extracellular calcium (Montague, 1996). Whatever the critical component(s), the important consequence is that, at synchronous equilibrium, closely situated neurons have either high, or low, pulse correlations with each other.

**ORGANIZATION OF PRE-VISION RESPONSE PROPERTIES**

We can now term those synapses that are transmitting impulses more strongly near equilibrium “saturated” synapses, and those which are more quiescent, but potentially able to be activated, “sensitive” synapses, and can consider what spatial patterns of saturated connections would best meet the requirement to maximize synchrony. Here a further property of the neural field commented on in relation to equations (1–5)—higher spatial cross-correlation of pulses and field potentials at shorter range — has a decisive impact on the equilibrium pattern of synaptic saturations. These emergent patterns, diagrammed in Figure 3, arise for the following reasons:

(a) Maximum synchrony generation with highest cross-correlation among near-neighbors in each \( \beta \)-cluster requires saturated couplings link near-neighbor cells—but sensitive connections must also form between closely adjacent \( \beta \)-cells. Both requirements are met when saturated connections within each \( \beta \)-cluster form a re-entrant network analogous to a Möbius strip. A similar argument regarding connections formed within macrocolumns has been advanced earlier (Wright et al., 2006; Wright and Bourke, 2008).

(b) The \( \alpha \)-cluster system and each of the \( \beta \)-clusters must enter into maximum joint resonance. This requires the formation of a homeomorphic projection between scales. The projection must be homeomorphic, since spatial cross-correlation is constrained to decline with distance at both scales, and so if resonance is at a maximum, the projection map must be one preserving topological identity between scales. This is possible because a disk can be mapped to a Möbius strip. Thus saturated \( \alpha \)-cell to \( \beta \)-cell synapses must systematically map limited angular ranges of the surrounding \( \alpha \)-system onto limited angular ranges on the margins of each \( \beta \)-cluster, and receive reciprocal saturated \( \beta \)-cell to \( \alpha \)-cell synapses. Such a mapping requires specification of an orientation and chirality for each \( \beta \)-cluster, and requires a reciprocal distribution of saturated and sensitive synapses from opposite sides of the \( \alpha \)-system to neurons in a limited angular range within each \( \beta \)-cluster.

(c) Maximum synchrony generation with high cross-correlation among near-neighbors in the \( \alpha \)-system requires \( \alpha \)-cells be...
Synchrony and synaptic organization

Further analogy between the hypothetical α- and β-systems and real anatomical structures can now be drawn. As well as the α-system’s congruence with the superficial patch system, the β-systems, each with a dense system of local connections that are centrally spared from patchy connections, are analogous to macrocolumns each centred about an OP singularity. The distribution of OP for lines of orientation $0 - \pi$ to angles $0 - 2\pi$ in pinwheels about a singularity finds analogy in the wrapping of a Euclidean plane onto a Möbius strip. It has also been earlier shown that arrangements of adjacent pinwheels in broken mirror symmetry match classical OP maps (Wright et al., 2006).

The structure of real patchy connections and classical OP response maps, contrasted with the results of simulating the arrangement of adjacent macrocolumnar structures in accord with the description above, are shown in Figure 3, while Figure 4 shows diagrammatically the proposed arrangement of saturated and sensitive synapses, and foreshadows the effect of structured visual stimuli, once the post-natal phase of development begins—to be described in the next section.

Figure 5 shows a further impact upon response map organization—the emergence of OD columns.

Just as OP organization in some species is apparent before eye opening, so too is the organization into OD columns (Blakemore and Van Sluyters, 1975; Erwin and Miller, 1998). Explanation of this can be included in the present model by an argument similar to that of Erwin and Miller, who suppose the correlation of cell firing at short distances of separation of V1 cells to be greater than the correlation of visual inputs over a similar distance. This forces a columnar OD organization because of instability—in the present model’s terms, the resulting disruption of the synchronous field at equilibrium produced by binocular inputs to the same cells—resolved by formation of columns in Turing patterns. A corollary of this effect is impact on the hexagonal arrangement, with broken mirror symmetry of OP organization, predicted in (d) above. The required alternation of OD columns would impose a frustration on the approach to hexagonal tiling of the cortical surface—forcing approach closer to a square tiling. The occurrence of mirror symmetry within a square tiling accounts for the way that lines of OP cross orthogonally between OP columns. (Obermayer and Blasdel, 1993). Following eyeopening inputs from the two eyes transmit images which are necessarily cross-correlated at a small spatial lag, because of angular disparity in their line of focus. Spatial lag correlation in their inputs at V1 level could then help maintain the columnar organization (Wright and Bourke, 2008).

**WAVE TRANSMISSION OF VISUAL INFORMATION, FOLLOWING EYE-OPENING**

We compactly express the emergent map by which the patchy connections over a part of V1 link to positions within each cluster.
FIGURE 4 | Equilibrium distribution of synaptic activity, and the impact of visual inputs disrupting equilibrium from Wright and Bourke (2013). Top: equilibrium disposition of saturated and sensitive synapses. Black circles represent cell bodies and dendrites. Synapses are indicated as saturated (solid) or sensitive (dashed) terminations of axons. Reciprocal connections between \( \alpha \)-patches (patchy connections) form the hexagonal array. (Other connections, although shown as unidirectional, are also reciprocal.) A representative pair of connections from \( \alpha \)-cells to the \( \beta \)-patch is displayed in the upper and lower aspects of the figure. At the center of the figure, saturated and sensitive synapses show the network's analogy to a Möbius-strip within a \( \beta \)-patch (macrocolumn). To the right, representative links from the central macrocolumn to cells at homologous positions in neighboring macrocolumns. Bottom: exposed to strong transient signals conveyed over the superficial patch system, summing with direct visual inputs conveyed to the cRF, the equilibrium configuration breaks down. The green bar represents the field of excitation of cells by the contextual signals, within which cells also directly excited in the cRF, fire at high rates.

FIGURE 5 | Top: Simulation of OD columns in accord with Wright and Bourke (2008). Bottom: Real OD columns, visualized by Obermayer and Blasdel (1993). Color coding of OP and scale as for Figure 3. Black lines demarcate alternation of OD between columns. Fine black lines in the lower figure trace the way OP is aligned so it matches orthogonally across OD column boundary.

macrocolumn, as an homeomorphic projection from a disk on a Euclidean plane, \( P \), to a Möbius strip, \( p^{[2]} \)- the square brackets [2] indicating the map's resemblance, if viewed from a third dimension, to a 2:1 map formed by squaring a complex vector. Defined in polar co-ordinates,

\[
P \left( |R - C_j|, \theta \right) \rightarrow p^{[2]} \left( |r - C_j|, \pm \theta + \phi \right)
\]

where \( C_j \) is the origin of both \( P \) and \( p^{[2]} \) for the \( j \)-th local map, and corresponds to the position of the OP singularity in that macrocolumn. \( \theta \) is the polar angle of \( R \), chirality of the local...
map is indicated by $\pm \theta$, and $\varphi$ is the orientation of the local map relative to the global map. $\theta + \varphi$ is defined on the range $0 - 2\pi$ in both local and global maps, but is represented with apparent angle doubling in the local map. This describes a topology for “contextual” connections (Li et al., 2000; Angelucci and Bullier, 2003) to each macrocolumn.

Visual input after eye opening will cause departures from the equilibrium condition. Let $O(P, t)$ be a visual image projected to V1 by the direct visual pathway. Laterally traveling waves of pulses and local field potentials relayed by the patchy connections can transmit that image to each local map with a point to point delay, $|\mathbf{R} - \mathbf{r}|$, where $v$ now represents wave speed, so that

$$O(P, t) \rightarrow O\left( p^{[2]}, t + \frac{|\mathbf{R} - \mathbf{r}|}{v}\right) \quad (24)$$

Suppose $O(P, t)$ is a segment of the image of a visual line, traveling with uniform velocity, $V_x$, along an $x$–axis directed toward a macrocolumn with its singularity at $O_C$. $O$ has a component of its extension on the $x$–axis, $O_x$, and an orthogonal component of extension, on the $y$–axis, $O_y$, $O_x$, $K_x$ is the dominant spatial frequency of $O_x$, and $K_y$ is the dominant spatial frequency of $O_y$. Then the local map projection of $O$ has a transformed spatial frequency in the $x$–axis but not in the $y$–axis—i.e.:

$$K_x \propto \frac{v}{\pm V_x} K_x \quad (25)$$

$$K_y \propto K_y \quad (26)$$

where $K_x$, $K_y$ are the spatial frequencies in the local map projection of $O$, and the sign $\pm$ in equation (25) depends on whether $O$ is approaching or departing from $O_C$. That is, $O$’s orientation in the global map is transformed to its projection to corresponding areas in the local map, by Doppler shift, with a difference in orientation, $\delta \theta$;

$$\delta \theta = \left[ \tan^{-1}[K_y/K_x] - \tan^{-1}[K_y/K_x] \right] \quad (27)$$

**INTERACTION OF CONTEXTUAL SIGNALS AND THE CLASSIC RECEPTIVE FIELD**

Laterally transmitted contextual signals generally do not trigger cell firing, until the classic receptive field (cRF) is directly stimulated (Li et al., 2000; Angelucci and Bullier, 2003) via the visual pathway. Those cells that then fire within a macrocolumn are those that reflect the supra-threshold summations of sub-threshold signals conveyed over the contextual, patchy, connections, and the direct pathway. We next assume that the summation of contextual and direct cRF inputs acts as an impulse causing a transient breakdown of equilibrium, during which synapses that were in the sensitive state in equilibrium briefly generate substantial synaptic currents [See Figure 4 (Bottom) and Figure 6]. Action potentials are triggered in surrounding cells, and subsequently there is a restoration toward the equilibrium state on withdrawal of the stimulus. During the breakdown the mapping of activity from the global to the local map becomes:

$$O(P, t) \rightarrow O\left( p^{2}, t + \frac{|\mathbf{R} - \mathbf{r}|}{v}\right) \quad (28)$$

The change from equation (24) made by removal of the square brackets from $p^{[2]}$ represents the breakdown’s form, as itself a map from global to local scale, resembling a 2:1 complex-multiplication map, as initially described by Alexander et al. (2004). The 2:1 map implies that single cells would show similar responses to a stimulus moving in either direction, but because firing is initiated over contextual connections in a 1:1 mapping, multi-cellular recordings would show that the spatial and temporal order of firing of neurons was unique for a given stimulus form and velocity.

**POST-NATAL EFFECTS OF LEARNING, THE SPATIO-TEMPORAL FILTER MODEL, DIMENSION REDUCTION, AND “LIKE TO LIKE” CONNECTIONS**

Equations (2,3,4, 19–22) contain state-variables required by mathematical expressions of physiological versions of the Hebb rule, and the spatio-temporal learning rule (Elliott and Shadbolt, 2002; O’Connor et al., 2005a,b; Enoki et al., 2009; Tsukada and Fukushima, 2010; Elliott, 2011). Following eye opening, stimuli with regularly repeated spatial and temporal structure reach V1, so we assume that exposure to a repeated stimulus leads to permanent synaptic consolidation of connections, overlaying those formed in the ante-natal, equilibrium condition. As remarked in the Introduction, Baker and Issa (2005) have shown that all V1 response features can be described in terms of six variables—optimal values of OP, spatial frequency preference, and temporal

![FIGURE 6](image)

**FIGURE 6** | The effect of increasing stimulus speed on apparent OP, for a bar of length 6 units, oriented at 45° to its direction of motion, and traveling left to right. Examples shown are freeze-frames, from separate simulation movies, at similar positions in the visual stimulus’ transit across the macrocolumn. From left to right, in each example, the bar speed/wave speed is 0.1, 0.5, 1.0, 1.5, respectively.
frequency preference, each associated with a Gaussian bandwidth of tuning of the cortical response to these features. These define three hypothetical filter processes. However, stimulus variables in the present model have equivalents to those used in the spatio-temporal filter model. These are:

<table>
<thead>
<tr>
<th>Spatio-temporal model</th>
<th>Present model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object orientation</td>
<td>Orientation relative to the y-axis defined for equations (25, 26)</td>
</tr>
<tr>
<td>Object velocity</td>
<td>( V_x )</td>
</tr>
<tr>
<td>Object drift angle</td>
<td>( \tan^{-1}\left(\frac{K_y}{K_x}\right) )</td>
</tr>
<tr>
<td>Object spatial frequency</td>
<td>( \frac{K_y}{\cos(\tan^{-1}\left(\frac{K_y}{K_x}\right))} )</td>
</tr>
<tr>
<td>Object temporal frequency</td>
<td>( V_x K_x )</td>
</tr>
</tbody>
</table>

Repeated stimulation with a particular stimulus will therefore lead, under Hebbian learning, to maximization of the response to that stimulus, thus creating an apparent “tuning” of particular neurons to that particular combination of stimulus features. Thus, the spatio-temporal model can be regarded as a consequence of the present model. Optimization by learning of the parameters for each of the three filters must be competitive between adjacent cells, providing the necessary condition for fitting response maps with continuity and completeness, by dimension-reduction methods (Kohonen, 1982; Durbin and Willshaw, 1987; Durbin and Mitchison, 1990).

Finally, the consolidation of saturated long-range patchy connections by Hebbian learning would result in mature “like to like” connections.

SIMULATIONS—A CRITICAL TEST

A critical test of our model, then, is whether we can reproduce in simulation the results of Basole et al. (2003), without a priori feature-specific responses to orientation, spatial frequency, or temporal frequency. Our simulations assume the steady-state presence of the Mobius synaptic configuration and its perturbation by visual signals, intended to reflect the state of the visual cortex shortly after birth, when first exposed to visual stimuli.

Equation (28) was applied in simulations of an hexagonal array of seven adjacent macrocolumns. Results reported here are for the central macrocolumn of the array of 7. Examples are shown in Figure 6, which shows the orthogonal transformation of apparent OP from the lowest to the highest bar speed for a moving line stimulus oriented at 45° to its line of passage.

Diameter for each macrocolumn is 300 microns, and wave speed for transcortical polysynaptic propagation 0.1 m/s (Bringuier et al., 1999). Units of length subsequently referred to, are multiples of the radius of a macrocolumn—150 microns. Simulation time-step was 0.1 ms.

A moving line in the visual field, relayed by the direct visual pathway to the cRF of each macrocolumn is represented as a red bar. In a single simulation the red bar traveled across the entire hexagonal array from left to right, with constant speed, direction and orientation. The orientation of the red bar to the line of passage is measured as bar angle from 0°, where the bar is oriented orthogonally to the direction of travel, to ±90°, where the bar is oriented in the direction of travel.

The lag-transmitted image of the red bar, relayed as subthreshold activation to each macrocolumn via the superficial patch system, is shown in green, with illumination about the zone of subthreshold activation, to indicate that input to the cRF from the direct visual pathway and contextual signals caused triggering of action potentials. The average angle from the macrocolumn singularity to the centers of action potential generation (i.e., all points on the green line with illumination) was calculated at each time-step, and shown as a black arrow, thus indicating the part of the macrocolumn with a response preference (apparent OP) for the particular bar movement. (A change in the sector of the macrocolumn that is maximally stimulated is equivalent to an equal change in the angle of approach of the bar needed to maintain stimulation of the same sector). The black arrow angle was averaged over a window beginning after the red bar had passed the center of the macrocolumn by a distance equal to 10% of macrocolumn radius, and extending from the 10th percentile to the 20th percentile of that radius, thus obtaining an estimate of the apparent OP during the cRF activation time. The standard error (SE) of the black arrow angles was calculated from 11 equally spaced time steps through the averaging window.

Combinations of bar-length, orientation of the bar to the direction of movement, and bar speed, were then systematically varied in separate simulations, results of which are supplied as supplementary animated movies. Their effects on OP, measured at the central local map of the hexagonal group, were obtained as OP difference, \( \Delta \phi \)—a measure of the change in OP as a function of these variables—calculated as

\[
\Delta \phi = \begin{cases} 
\phi_1 - \phi_0 - \pi & \text{when } \pi/2 < \phi_1 - \phi_0 \\
\phi_1 - \phi_0 & \text{when } -\pi/2 \leq \phi_1 - \phi_0 \leq \pi/2 \\
\phi_1 - \phi_0 + \pi & \text{when } \phi_1 - \phi_0 < -\pi/2 
\end{cases}
\]

(29)

The reference OP, \( \phi_0 \in [0, \pi] \), was the OP found at the lowest bar speed applied (bar speed/wave speed = 0.1) and the apparent OP, \( \phi_1 \in [0, \pi] \), was the OP found at higher speeds.

Systematic results are shown in Figure 7, which graphs OP difference vs. bar speed/wave speed, for bar angles 0 to ±90°, calculated for a bar length of 6 units. Variation of bar length showed progressive lessening of the effect of velocity on OP for greater bar lengths.

For the case of bar-angle zero degrees (a line oriented orthogonally to its direction of passage, as in classical measurements of OP) no OP difference is seen until, as bar speed approaches wave speed, a 90° change in apparent OP takes place at a single increment in speed. This corresponds to transition to a “motion streak,” as object movement blurs resolution in the direction of motion. Increasing OP difference with bar speed at other bar angles is a more gradual development of the same effect—that is, mixing of responses to object speed and to object orientation. The illuminated field of supra-threshold excitation generated is not that expected to accompany a Gaussian-shaped tuning
mature animals studied by Basole et al. (2003) and Issa et al. concurrently strengthen responses to signals which are close to enhancing the peak response to the optimal visual signal, would in the prior section, subsequent post-natal Hebbian learning processes of OP, but with conduction delays of “like to like” fibers included. Then, systematic variation of OP with bar velocity did not occur.

These results match the findings of Basole et al. (2003) with respect to variation of OP peak responses as a function of line velocity and length. They do not reproduce the form of the experimentally observed Gaussian tuning curves, but as argued in the prior section, subsequent post-natal Hebbian learning progressively over-writing the Möbius configuration, and strengthening the peak response to the optimal visual signal, would concurrently strengthen responses to signals which are close to the optimum, resulting in Gaussian tuning curves in the more mature animals studied by Basole et al. (2003) and Issa et al. (2008).

**INTER-AREAL INTERACTIONS OF V1 AND HIGHER VISUAL AREAS**

The principle underlying the development of connections between macrocolumns and the superficial patch system may be generalized to the emergence of inter-areal connections. To recap, taking V1 as an example we have argued above that, because co-variance of activity declines with metric distance at both the scale of the patchy connections and within a macrocolumn, a homeotypic mapping between scales can emerge. This requires that relative distances on the maps at each scale must be in the ratio of correlation lengths of synchronous oscillation at the two scales, and adjacent maps must themselves have a correlated structure over a distance approximate to the correlation length of the patch system. It then follows that superposition of adjacent local maps, with appropriate rotation and correction to a common chirality, would result in a further map with co-variance of activity declining with metric distance, over the correlation length of the patch system.

Inter-areal connections, made by cortico-cortical axonal projections, could permit maps of this type to arise during antenatal development, with the composite map at the higher cortical-area level itself folded into the Möbius configuration. The selection of saturated connections, projecting between areas with normalization of rotations and chiralities, would be possible by selection from the larger set of possible connections made by branches of cortico-cortical axons, diverging from their cells of origin to their cells of termination, overlapping as they terminate, and generally reciprocal between areas (Braitenberg and Schüz, 1991; Boucsein et al., 2011). Thus, antenatally, sets of macrocolumns at both the lower, V1, level and higher levels, could resonate with, and form preferential connections with, superimposed and overlapping groups at the other level, in accord with the developmental selection requirement to maximize joint synchrony. With the occurrence of eye-opening, Hebbian learning would then begin to overwrite the equilibrium resonance configuration between areas, in analogy to the process at intra-areal level—with the added property of associating concurrent patterns of activity in the V1 macrocolumns.

Illustrating this effect, Figure 8 shows, at the bottom, a system of seven macrocolumns at V1 level, driven via the direct visual pathway by a pair of intersecting lines in the visual field.

The top part of the figure shows a projection of activity in conjointly activated macrocolumns in V1, to a higher visual area, in which responses in the seven macrocolumns in V1 have been superimposed, with disparities in their orientation and chirality eliminated. Summations of points stimulated by both lines, shown by highlighted white points, occur frequently in the forward projection—much more so than at the level of V1 itself. These indicate response to angles of intersection of the lines at the lower level, and, commonly at the higher level, summing responses to time-lagged correlations between disparate positions of the moving visual stimulus. This effect is consistent with the preferential responses to angular and complex stimuli, characteristic of higher cortical levels (Merigan and Maunsell, 1993).

Conversely, since connections between higher and lower levels are generally reciprocal, a possible mechanism permitting control of attention (Rao and Ballard, 1997; Kveraga et al., 2007;
CONCLUSION

The model of cortical development we have outlined above is efficient from both energetic and information-processing perspectives, and has considerable anatomical and physiological explanatory power. It leads to an explanation of the spatial organization of signal flow in the cortex that differs from any other model. The proposed antenatal self-organization of cortical synapses leads to the creation of a *tabula rasa* on which homeomorphic maps, in a form disguised by the Möbius-strip-like folding of connections, occur in lateral connections at the millimetric scale, embedding the statistics of spatial organization of the sensory world to first approximation, before any detailed sensory inputs are received. The assumptions and findings of the model overlap with, and although not necessarily contradictory to, are not identical to, those of other models (Erwin and Miller, 1998; Wolf and Geisel, 1998; Kang et al., 2003; Oster and Bressloff, 2006). Distinguishing features include the explanation of the relationship of superficial patch connections to macrocolumn centers, and their hexagonal rotational symmetry, and crucially, the findings of Basole et al. (2003), which cannot be explained by any model dependent on “like-to-like” connections between feature-specific neurons. Nor can any model with otherwise similar assumptions about the self-organizing effect of synchrony be formulated without introducing a Möbius configuration to the connections, since an equivalent model utilizing only Euclidian conformations would represent a given OP twice, rather than around a singularity.

In review, the assumptions and conclusions reached, were as follows. By assuming that cells surviving apoptosis are selected by competition for metabolic substrates, and that synchronous oscillation mediates the uptake of metabolic substrates, we showed the outcome was a neural system with ultra-small world axonal configuration. Further assuming the small world connections were necessarily constructed from neuron populations characterized by respective axonal length, we showed that long range patchy connections and regular macro-column-like areas with central sparing of patchy connections emerge, with some degree of hexagonal rotational symmetry, with species variation in orderliness according to cortical size, and were able to show that this result was consistent with anatomical observations of limited interspecies variation of singularity density. A crucial further assumption made, was that metabolic competition between synapses from the same neuron leads to particular configuration of synaptic current flows at equilibrium, in which active connection networks within each macrocolumn are arranged in a Möbius-strip-like configuration. Then, with the introduction of visual inputs, signals conveyed by contextual fibers transfer a visual image from the global map to each local map, determining the pattern of neuron firing induced by activation of the cRF, and synaptic consolidation on Hebbian principles begins—thus storing information based on visual experience—explaining how response maps for OP, SF, and TF become organized in accord with the spatio-temporal filter model (Baker and Issa, 2005; Issa et al., 2008), and how “like to like” anatomical connections emerge, as well as providing conditions for dimension-reduction description of response features. The model is also compatible with explanation of ocular columns and direction preference fractures, as proposed in our earlier work (Wright et al., 2006).

The resulting synaptic storage of learned information in local topological maps of Möbius configuration offers a further compression of format, adding to the efficiency of the “small world” arrangement, by minimizing the distance which need be spanned by connections between positions on the local map representing positions widely separated on the global map. The development of cross-links also offers large potential information storage, since the regular spatial organization of links in the Möbius configuration implies the synaptic connections have low joint entropy in their ante-natal state. With visual experience, and the storage of image information in cross-links, joint entropy could, in
principle, increase to a limit where all synaptic states are independent, and equally distributed about some mean connection strength, as implied by Montague’s (1996) resource consumption principle. In effect, before eye-opening, the cortex has “learnt” the underlying statistical structure of visual space—that of cross-correlation declining with metric distance—and subsequently stores information about departures from this “first component” of structure in the visual world.

The antenatal development of response maps (Wiesel and Hubel, 1974; Blakemore and Van Sluyters, 1975; Sherk and Stryker, 1976) presents no paradox in this model, since emergence of organized response properties within the Möbius configuration does not depend upon structured visual stimuli. EEG activity progressively matures toward alternating alert and sleeping states in the later antenatal period (Marks et al., 1995; Mirmiran, 1995) providing the widespread co-ordination of pre-synaptic activity required for initial synaptic self-organization. Conversely, over-writing by learning in the immediate and later post-natal periods explains why representation in adult response maps of stimuli to which the subject as has not been exposed would not be present—as also seen experimentally (e.g., Blakemore and Van Sluyters, 1975).

No direct evidence yet exists of Möbius-like patterns of connections in cortex, yet this is scarcely surprising if transient dynamic couplings, present only in the equilibrium state, are overwritten by post-natal learning-related changes. However, relaxation toward the equilibrium condition is still to be expected in the mature state, so it is important whether or not some anatomical substrate exists in which the dynamic state of synapses may be capable of transient assembly into Möbius patterns. Markram and colleagues (Perin et al., 2011) found that pyramidal neuron networks cluster into multiple groups of a few dozen neurons each, with the neurons composing each group typically more than 100 µm apart, allowing for multiple groups to be interlaced in the same space. Connections within groups were largely reciprocal, and those between groups relatively sparse. Transient interlinkages between such interwoven linked groups could form Möbius-like networks. The temporal plasticity of synaptic connections near singularities (Dragoi et al., 2001) is also consistent with this interpretation. As well as plasticity of responses near OP singularities, Dragoi and colleagues found lack of plasticity in linear zones—the areas of strong patch connection termination. This is to be expected if the patch system is composed of well consolidated connections suitable for consistent transmission with delay from fixed points in V1, while of the other hand, more complex, continually modified, information processing goes on in the areas around singularities. Consequently, an anatomical test of the model may be possible, in regard to the terminations of patchy connections in the periphery of the patch-free areas about singularities. As indicated in Figure 3, two populations of synaptic connections should be demonstrable in principle, by double injection/staining methods, near the singularity/patch edge. If some Hebbian consolidation occurs both antenatally and postnatally, then, in principle it should be possible to observe Möbius-like connections within macrocolumns antenatally, and the overwriting of these connections during post-natal learning.

If later testing supports this model, current conceptions of cortical information processing will require modification. Synchronous oscillation has been regarded as a mechanism for feature-binding—requiring that groups of cells in synchrony stand out in some way against a non-synchronized background. Instead, this model emphasizes synchrony as the organizer of a matrix of connections within which each macrocolumn gathers information from its surround, and organizes these connections systematically according to spatial position and time-lag, as functions of distance from each singularity. The topology of signal organization is markedly different to that of the association of “feature” neurons embedded in neural connections that are deployed on a Euclidean plane, as it implies that sensory images are not broken up into “features” which are subsequently associated in an abstract feature space, but retain, in modified form, an organization representing sensory space and time. Upon this more complicated matrix of connections, moment-by-moment states of autonomous local firing could interact with each other via traveling waves, generating internal images adding to those arising from sensation—all selectively strengthening preferred pathways by Hebbian learning, under the supervision of motivational systems. This gives a modified basis to Sherrington’s “enchanted loom” (Sherrington, 1906, 1940), and a stage for the kind of neuro-dynamic events progressively observed and envisaged by Freeman for many years (Freeman, 1975; Freeman and Quiroga, 2013).

Hierarchical interaction of V1 with higher visual areas, by superposition of spatio-temporal images transmitted over convergent and divergent pathways might proceed to higher levels of abstraction, at higher cortical levels, and feedback interactions of ascending and descending signals in such a system might permit very complex image manipulation. Analogous processes may apply to other modalities throughout the cortex in general, since all sensory input systems are analogous to the visual system, in as much as they encode the sensory world by imposing a topological order to inputs as they arrive at the sensory cortices. Again, the ubiquitous distribution of patchy connections throughout the cortex, and the basic modular similarity of the paleo- and neo-cortex throughout, supports the notion that a single schema of information flow may be characteristic of all. The principle of organization might even extend to the motor cortex, with the efferent pyramidal motor neurons simply reversing the role of neurons in the direct visual pathway.

ACKNOWLEDGMENTS
A special debt of gratitude is owed to Adrienne Edith Wright.

This work was supported by the Frank P. Hixon Fund of the California Institute of Technology and the Oakley Foundation of New Zealand, and by iVEC, through the use of advanced computing resources located at the University of Western Australia.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at http://www.frontiersin.org/Computational_Neuroscience/10.3389/fncom.2013.00004/abstract


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Möbius-strip-like columnar functional connections are revealed in somato-sensory receptive field centroids

James Joseph Wright1*, Paul David Bourke2 and Oleg Vyacheslavovich Favorov3

1 Department of Psychological Medicine, Faculty of Medicine, The University of Auckland, Auckland, New Zealand
2 IVEC@UWA, University of Western Australia, Perth, WA, Australia
3 Department of Biomedical Engineering, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

INTRODUCTION

It is widely acknowledged that Mountcastle's (Mountcastle et al., 1955; Mountcastle, 1997) concept of the cortical column was critical to subsequent development in the theory of the cerebral cortex (Hubel, 1981; DeFelipe et al., 2012). However, the concept has proved problematic (Purves et al., 1992; Horton and Adams, 2005). The terminology describing supposed variants and orderings into columns has become confused (Rakic, 2008; Da Costa and Martin, 2010). DeFelipe et al. (2012) have concluded that a comprehensive framework of the function and structure of columns remains elusive, the classical criteria (shared response properties, shared input and common output) may need to be modified, and area-specific and species-specific variations more adequately defined, to achieve fundamental understanding of what columns are, and how they are used in cortical processes.

Opinions diverge on whether or not structural modularity is of functional importance at all. The utility of modularity, both for the actual operation of cognitive processes, and for our task of analysis of brain function, has been emphasized by Szentagothai (1983). A contrary opinion was presented explicitly by Purves et al. (1992) when they argued that “[columnar] patterns arise not because the functional organization of the brain demands them, but as an incidental consequence of the rules of synapse formation.” The latter interpretation, forced to its limit, suggests that no functional modularity exists in the cortex, and therefore the task of its analysis is greatly multiplied. However, there might exist modularity in the patterns of cortical synaptic connections, that become manifest as cellular anatomical order only under some circumstances. That is the position that will be argued here.

In this paper we apply a recent model for the organization of columnar structure in primary visual cortex (V1) to primary somatosensory cortex (S1), in areas anatomically not obviously columnar, to attempt to solve some of the dilemmas posed above. We will place emphasis on horizontal connectivity within cortex (Boucsein et al., 2011) and on organized, modular, patterns of synaptic connections, rather than on spatial clustering of cell bodies into defined columns. To avoid the many terminological difficulties, we will simply define “macrocolumns” asvariably resolved clusters of neurons sharing relatively dense short-range lateral connectivity, and distinguish separately the long-range horizontal patch connections, which also have clustered cell bodies, but have patchy areas of synaptic termination, both patch-reciprocal and also terminating on the peripheries of macrocolumns. Reasonably clear macrocolumns in this sense are apparent only in V1 and in some other locales such as the barrel-based cortical columns representing vibrissae in S1 (Woolsey and Van der Loos, 1970)—sensory systems in which the organization of inputs is clear-cut—although even this degree of orderliness is only apparent in some species, and not in all situations with clear input demarcation (Purves et al., 1992).

Further adding to the difficulties of analysis, attempts to understand the relation between columnar organization and the response properties of single neurons have also struck problems (Blakemore and Cooper, 1970; Blakemore and Van Sluyters, 1975; Swindale, 1996). In V1, visual line orientation has long been...
considered a primary property (Hubel and Wiesel, 1959), but Fitzpatrick et al. (Basole et al., 2003) have shown that orientation preference (OP) is a function not only of visual line orientation, but also of the angle of the line to the direction of motion, the speed of motion, and the length of the line. So long as stimulus properties are clearly defined, individual neurons show clearly tuned specific responses, whether or not response maps can be clearly resolved, as can be shown when OP tuning is contrasted in macaque and rat (Girman et al., 1999). In V1, where response maps are relatively well resolved and continuous, Issa et al. (2008) have shown that the combined cell responses to orientation, spatial frequency, temporal frequency, and the bandwidths of each, define the response maps. In S1, although macrocolumns are not clearly defined anatomically outside the cortical barrels, organization of responses into combinations of stimulus features is also indicated by the way that receptive field (RF) boundaries form “segregates,” as next discussed.

Via dorsal-column-medial-lemniscus-cortical pathways, a homotypic representation of the skin surface is projected to the horizontal plane of the S1 cortex (e.g., Marshall et al., 1941) and in Mountcastle’s conception of the columnar organization the homotypic map has depth from the pial surface to the white matter, with different afferent submodalities in a mosaic ordering. Exploring this concept in S1 forelimb representations, Favorov et al. (1987); Favorov and Whitel (1988a,b); Favorov and Diamond (1990) advanced extracellular recording micro-electrodes completely through the cortex at near-radial angles deviating on average about 22.5 degrees from the vertical. As expected, the more closely the path of descent approached the vertical (radial) track, the more similar became the RFs of the sequentially recorded single units. With the slight oblique drift of the electrode path they commonly encountered sequences of RFs that approximated a homotypic map, but there were also sudden discontinuities, where the RF of the next neuron found in the electrode’s progress lead to a relatively distant position on the skin surface—sometimes returning back closer to the prior localities as the electrode was still further advanced.

Using the skin locations of the maximum response (the “minRF”) to tactile stimulation of local groups of neurons in a common skin locus in their maxRFs and thus belonged to a single nest, whereas only about half of neurons sampled in a near-radial penetration across a segregate boundary shared a common skin point. Thus, it appears that RF nests are centered on segregates but are larger than a single segregate.

We intend to show that the breaks in continuity of minRF positions found in S1 by Favorov et al. are explicable by a model of embryogenesis and synaptic self-organization independently developed to account for V1 columnar organization, and the nested RFs are an overlay of this primary order.

**THE MÖBIUS MODEL OF CORTICAL DEVELOPMENT AND ITS APPLICATION IN V1**

During embryogenesis cells becoming cortical neurons divide and migrate to their mature positions while undergoing apoptosis (Rakic, 1988, 2009). Wright and Bourke (2013) devised a theory for embryonic selection of neurons and their strongest synaptic connections that leads to a modular description of cortical organization—one in which the modular organization is that of the pattern of synaptic connections, but in which cellular order need be only variably apparent. The basic assumptions of the proposal rested on two experimental findings—firstly, in vitro, embryonic neurons fire synchronously and self-organize into “small worlds” (Downes et al., 2012) and secondly, synchronous firing of neurons prevents their apoptosis (Heck et al., 2008). We assumed synchrony and cell survival are directly causally linked—that by some mechanism, synchrony of pre-synaptic activity or of action potentials promotes the uptake of a critical resource. Calcium is a likely candidate as that resource (Montague, 1996) but many different metabolites might be critical at different stages of neurogenesis. But overall, as a consequence of the balance between uptake of, and demand for, the critical resources, the emergent cell network would be that selection of cell types, and their arrangement, maximizing the amplitude of synchrony while minimizing the metabolic demand per cell. Consequent steps in our argument were as follows:

1. **Synchrony,** measured as either amplitude of local field potentials, or as variance of pulse rates of excitatory cells, we considered generated as described in our simulations of electrocortical activity (Wright, 2009, 2010). These simulations show gamma oscillation to arise from an oscillating equilibrium of excitatory/inhibitory interactions, and synchrony of the excitatory cells from summation of in-phase pre-synaptic signals vs. cancelation of out-of-phase pre-synaptic signals.

2. **Uptake** of the resource from the extracellular environment we take to be proportional to the magnitude of presynaptic synchrony, while consumption of the resource is proportional to neuron size/length. Competition thus leads to selection of an ensemble of cells with the maximum number of pre-synapses and the shortest possible set of connections permitting interaction of all cells in the ensemble—that is, the ensemble must have the topology of a “small world” (Watts and Strogatz, 1998).

3. In the dilute network of neuronal connections, metric distance of some separation is analogous to “degree of separation” in the topological sense, whereas range of fiber...
connections, and thus the number of cells reached by connections from a given cell, is analogous to degree distribution. Therefore, on average, the probability density of synaptic connectivity between excitatory cells should decline with distance of their somas as a power function, as required in “ultra-small-worlds” (Cohen and Havlin, 2003).

(4) A power function is the sum of exponential functions, and pre-synaptic densities of cortical neurons decline roughly exponentially (Braitenberg and Schüz, 1991). Therefore, ultra-small-world connectivity can be approximated by selection of subpopulations of neurons with suitable axonal ranges. In reality a considerable variety of excitatory cell types is encountered. For simplicity we treat these as only two ideal populations—a local type, with dense short-range horizontal connectivity, roughly corresponding to spiny stellate and pyramidal neurons with short axonal trees, and a type with long intracortical axons—the superficial patch cells. Cortico-cortical connections are excluded from present consideration, but can be accommodated within the same theory. With only two subpopulations, approximation of a power function takes the form

$$\rho(\|q - r\|^{-A}) \approx N_\alpha \lambda_\alpha e^{-\lambda_\alpha \|q - r\|} + N_\beta \lambda_\beta e^{-\lambda_\beta \|q - r\|}$$  \hspace{1cm} (1)

where A is an exponent between 2 and 3 (Cohen and Havlin, 2003),

$$\rho$$ is the optimal pre-synaptic density/distance relation, $$\|q - r\|$$ are positions of excitatory cortical neurons, $$N_\alpha$$ and $$N_\beta$$ are numbers of patch cells and local neurons respectively, and $$\lambda_\alpha$$ and $$\lambda_\beta$$ are the corresponding inverse length constants of their axonal trees. For given $$\lambda_\alpha$$ and $$\lambda_\beta$$, selection of cells in a specific ratio $$N_\alpha / N_\beta$$ must occur, to best approximate a power function.

(5) Any two populations of neurons conforming to Equation (1) sustain a magnitude of synchronous oscillation, $J$

$$J \propto \iiint (N_\alpha \lambda_\alpha e^{-\lambda_\alpha \|q - r\|} + N_\beta \lambda_\beta e^{-\lambda_\beta \|q - r\|}) dqdr$$  \hspace{1cm} (2)

Maximization of $J$ requires cells occur in clusters of each type. Two limiting cases arise: At the limit where one cell type has markedly longer axons than the other, $$\lambda_\alpha << \lambda_\beta$$ Equation (1) requires $$N_\beta >> N_\alpha$$, and therefore maximization of $J$ in Equation (2) requires cells with relatively short axons be closely situated, causing clustering into a macrocolumn. Consequently the cells with long-range axons must form connections at longer range, enforcing a “patchy” connection system. In the limiting case arrangement into a fully demarcated hexagonal periodic system must emerge, where

$$\frac{\text{local cells}}{\text{local cells} + \text{patch cells}} = \frac{N_\beta}{N_\alpha + N_\beta} \geq \frac{\pi}{2\sqrt{3}}$$  \hspace{1cm} (3)

The relation in Equation (3) is that of the most efficient packing, and follows from the ratio of area of a circle to a hexagon.

At an opposite limit, where $$\lambda_\alpha \approx \lambda_\beta$$ Equation (1) requires $$N_\beta \approx N_\alpha$$, and there is effectively only one cell type. Cell bodies can be intermingled, although Equations (1) and (2) are still met, and the absence of a clearly columnar arrangement does not imply a loss of the small world organization. The left hand images in Figure 1 show these limiting cases, suggesting how variations of macrocolumnar orderliness might arise as intermediates between the limiting cases, in different species and cortical areas.

(6) Maximization of $J$ also requires that the patch connections project a 1:1 map of activity in the surrounding cortex onto each macrocolumn. Competition between closely situated presynaptic terminals arising from the same cell but terminating on different cells, would force strong synaptic links to some neighbors and weak links to others. The requirement to form a 1:1 map can then be met only if the strong local connections within the macrocolumn form a closed network analogous to a Möbius strip, in the sense that a series of strong connections tracing a continuous loop circling the center of the macrocolumn would need to circle the center twice before the loop was closed, as does a line traced on the familiar Möbius strip of a paper strip connected at its ends with a half-twist. The long-range patchy connections would then project the surrounding cortex onto this system as a Euclidean plane projected onto a Möbius strip. The right hand images in Figure 1 indicate how, in V1, this would force classical OP from 0 to 180° to be arranged from 0 to 360° about a singularity, and Figure 2 (top) shows how the strong and weak synaptic connections must be organized.

FIGURE 1 Variation of the structure of macrocolumns at limiting extremes of the axonal lengths and cell numbers. Top left: With large long/short axon length ratio, clearly resolved hexagonal organization emerges, with macrocolumns (white) surrounded by groups of patch cells (black dots). Top right: long (red line) patch connections link “like to like” OP (shown as background color wheel) in contrast to highly clustered short intracortical axons (blue lines). Bottom left and right: near-complete loss of resolution when long/short axon ratio approaches 1.
Wright et al. Columnar functional connections

Figure 2 | Maximization of synchrony with local synaptic competition leads to Möbius ordering, within macrocolumns. Top: Disposition of strong (solid) and weak (dashed) synapses in the developing neocortex. Bottom: “Like to like” patchy connections map the same part of the surrounding cortical field onto homologous cell positions on the Möbius configuration. Adjacent macrocolumns tend to mirror organization of OP.

(7) Interactions between local maps must also maximize $J$. Therefore, patch connections must link “like to like” OP while forming multiple 1:1 maps, and adjacent macrocolumnar “maps” must be arranged so “like” map positions on adjacent macrocolumns are as closely situated as possible within the hexagonal frame, thus approximating a mirroring with continuity among adjacent maps. Figure 2 (bottom) shows these effects.

Applied to V1 before eye-opening, this model is able to explain:

- The emergence of macrocolumns (in some species) in V1 and OP response maps, their singularities, and continuity at column margins (e.g., Bosking et al., 1997).
- V1 approximation to hexagonal rotational periodicity (Muir et al., 2011; Paik and Ringach, 2011).
- Interspecies and inter-areal variation in superficial patch organization (Horton and Adams, 2005; Muir et al., 2011) and relative species invariance of intersingularity distances (Kaschube et al., 2010; Keil et al., 2012).
- Emergence (in some circumstances) of OD columns (Obermayer and Blasdel, 1993).
- Emergence of superficial patch connections, “like to like” (Gilbert and Wiesel, 1989; Muir et al., 2011) with patch-sparing of macrocolumn centers (Sharma et al., 1995; Muir and Douglas, 2011).
- Variation of apparent OP with stimulus orientation, angle of movement relative to orientation, object speed, and object length (Basole et al., 2003).

And after eye-opening, consequent to visual stimuli and Hebbian learning:

- Later absence of cortical responses to stimuli of which the subject is deprived (Blakemore and Cooper, 1970).
- The consolidation of learned cortical responses, producing selective tuning of cells stimulus orientation, spatial frequency, and temporal frequency (Issa et al., 2008).
- Complex stimulus responses in higher visual areas.

The theory implies both the retention of pre-natal Möbius ordering into post-natal life (evidenced by reproduction of the results of Basole et al.) and the overwriting of the earlier Möbius order by later learning, (accounting for the bandwidth tuning to classes of features shown in the work of Issa et al.).

It is not proposed that variation of genetically prescribed cell types is the only source of variation of columnar order. Spatial and temporal resolution of signals in the input pathways, degree of approximation to “brown” cross-correlation in inputs, amount of convergence and divergence of axo-synaptic connections, and Turing pattern formation can all be expected to contribute to the apparent order, and might in principle be included in more developed forms of this model.

We turn now to the application of the model to S1, in areas where no clear anatomical macrocolumnar order has been shown.

APPLICATION TO S1

MATHEMATICAL CONSIDERATIONS

Viewed from above the V1 cortical surface, the Möbius-like arrangement of OP resembles a 2:1 map formed by squaring a complex vector, creating a local representation of global positions,

$$\pm P \rightarrow \frac{P^2}{|P|}$$

where $P$ is a position in the visual field (and on the surface of V1), $p$ is its mapped representation in the macrocolumn, the global and local maps have a common origin ($P_0$ and $p_0$) at the OP singularity, and are equivalently scaled.

However, this apparent 2:1 map conceals a 1:1 map in three dimensions, and application to sequential extracellular recordings on a trajectory passing from surface to depth in S1 requires consideration of three dimensions.

Applied in 3D, Equation (1) and maximization of $J$ in Equation (2) requires local neurons with strong short-range...
synaptic connections to be packed closely together, forcing separation of the somata into two groups, one superficial and one deep, which we term "skeins"—the analog of each side of the paper in a paper Möbius strip. Onto these skeins of local cells projections conveyed by patch connections wind clockwise at one cortical level, and anti-clockwise at the other, with the projections circling a central zone, analogous to an OP singularity in V1.

Since complete skein separation could only be approached as ideal, at an intermediate level mixing and intertwining of the skeins is anticipated. The skeins are necessarily continuous at some angle directed from the central singularity, so mixing at the skeins would be lesser at the continuity. (See Figure 3, Top panels).

The interface between the skeins can be represented by a 3D representation of a Möbius strip in Euclidean space, as is used in the top images of Figures 3–5, and is given by:

\[
\begin{align*}
x &= |p| \cos(\pm \vartheta + \varphi) \\
y &= |p| \sin(\pm \vartheta + \varphi) \\
z_M &= K |p| \sin((\pm \vartheta + \varphi)/2)
\end{align*}
\]

where \(x, y\) are Euclidean horizontal co-ordinates on the cortical surface, and \(z_M\) is the depth of the interface between the skeins. \(\vartheta\) is polar angle, \(\varphi\) the angle of orientation of the macrocolumnar map vs. the global coordinates; \((\vartheta + \varphi) \in [0, 4\pi]\), and \(\pm\) indicates chirality of the macrocolumnar map. \(K\) determines the stretching of the skin interface in cortical depth.

Interpenetration and entanglement of the upper and lower skeins is presumed at the intermediate zone of cortical depth, between \(+z_M\) to \(-z_M\), with greater separation of the skeins above and below.

The track followed by a recording electrode is:

\[
z_D = ax + by + c
\]

where \(z_D\) is the depth in the cortex reached by the probe, and \(a, b, c\) parameters of the track.

Thus, the positions at which maximum RF responses (subsequently treated as centroids of maxRFs) are discovered on the skin as the recording electrode advances follow a reverse mapping to that in Equation (4)—viz:

\[
|\sqrt{p}| \sqrt{p} \rightarrow \pm P
\]

The sign (\(\pm\)) in Equation (7) is assigned positive if \(z_D(x, y) > +z_M(x, y)\), negative if \(z_D(x, y) < -z_M(x, y)\), and randomly assigned if \(+z_M \geq z_D \geq -z_M\). That is to say, according to where the cell lies in relation to the skins.

Specifying Equation (7) for a pair of adjacent macrocolumns

\[
|\sqrt{p} - p_1| \sqrt{p} - p_1 \rightarrow \pm(P - p_1) \quad \text{for } \mu_1(p) \ldots \mu_{trans}(p)
\]

and

\[
|\sqrt{p} - p_2| \sqrt{p} - p_2 \rightarrow \mp(P - p_2) \quad \text{for } \mu_{trans}(p) \ldots \mu_n(p)
\]

\(\mu_1 \ldots \mu_{trans}\) and \(\mu_{trans} \ldots \mu_n\) are positions on the line of advance of the probe Equation (6), and in cases where the track is transitioning between the adjacent macrocolumns, the position of transit is designated \(\mu_{trans}\). Positions \(p_1\) and \(p_2\) are the centers of the two maps, separated by a distance equal to the diameter of a macrocolumn. Adjacent macrocolumns each have representation, via the patchy connections, of extensive areas of the skin, and overlap of the skin area representation is considerable, and so to first approximation the skin area representations can be treated as identically centered, so that \(P_1 \equiv P_2\).
FIGURE 4 | High horizontal passage of a recording electrode, crossing the junction of two adjacent macrocolumns, from \((x, y, z) = (3, 2, 2)\) to \((-3, 1, 2)\). RF centroids form columns of progressing sequentially in one direction, until the junction of the two macrocolumns is crossed, after which point the direction of sequential progression is reversed.

The signs ± and \(\mp\) in Equations (8a,b) carry the same meanings as in Equation (7) with the reversal ± to \(\mp\) indicating the macrocolumns are so arranged that the entire Möbius mesh of one has opposite chirality to the other, forming mirror images, each of the other. For any pair of adjacent macrocolumns randomly selected, the axis of mirroring will be random relative to \(\vartheta = 0\), but the patterns of RF centroids arising with variation of verticality of the recording electrode would be little affected. So, for simplicity, only mirroring along the \(\vartheta = 0\) axis is considered in the next section.

FIGURE 5 | Oblique passage of the electrode, from \((x, y, z) = (2, 2, 1)\) to \((-2, -0.5, -1)\). The path of penetration both crosses the junction of adjacent macrocolumns and passes through zones in which clockwise and anticlockwise skeins of synaptic connections are poorly demarcated. Centroids of RFs associated with clockwise and anticlockwise skeins are entangled, but sequential progression of RFs associated the clockwise skein is still in the reverse direction to those associated with the anticlockwise skein.

POSITIONS OF SEQUENTIAL RF CENTROIDS, PREDICTED FOR ELECTRODE TRAJECTORIES THROUGH S1
Figures 3–5, show expected results for the RFs sequentially discovered as an extracellular recording electrode passes through S1, from superficial to deep, through clockwise and anticlockwise skeins. Results vary depending on whether the electrode descends vertically, near-horizontally, or obliquely, and on whether the electrode remains within a single macrocolumn, or passes from one macrocolumn into its neighbor.

In each of the three figures, the top frame shows the skein interfaces in a pair of macrocolumns, viewed obliquely from above the cortical surface. Colors on the color wheel indicate equivalent polar angles on the clockwise and anticlockwise winding of the skeins. The passage of the recording electrode is shown as a line of dots passing through the cortex, and thus penetrating the connection systems from top to bottom. Red dots represent positions in which the patch projections are winding clockwise, and blue dots positions in which the projections are winding anticlockwise.

The middle frames show the pair of macrocolumns and the electrode in horizontal plane view. Although the individual macrocolumns are shown as circular, we have assumed columnar continuity at their margins.

The bottom frame shows the positions of the centroids of RFs of the neurons found sequentially by the penetrating electrode.
(Supplementary movies provided with this paper show the sequential development, as the electrode is advanced.) RF centroids associated with positive-valued solutions on the right-hand side in Equations (8a,b) are marked red, while negative-valued solutions are marked blue. In the intermediate region where the skeins may mix, RF centroids are randomly assigned red or blue.

In these images no allowance has been made for cortical anisotropy, nor recording noise, and the bottom images have been normalized in scale, and centered.

Figure 3 shows results typical of a near-vertical electrode penetration remaining within a single macrocolumn (a type I penetration). Sequential RF centroids as the electrode descends are tightly clustered until a “jump” occurs in the sequence, as the electrode passes from one skein to the other, and a second cluster of RF centroids begins and persists until the cortex is completely penetrated. If the electrode track is slightly off vertical, the RF centroids may be less tightly clustered, and progress in one direction until the jump occurs, and in the opposite direction after the jump. If the electrode track passes through the less demarcated zone between skeins sequential centroids may jump back and forth between the clusters.

Figure 4 shows results typical of the passage of the recording electrode at an angle approaching the horizontal, remaining high enough in cortex to pass partly across the border of a pair of macrocolumns (a type II penetration), with only the upper skeins of both columns penetrated. In this case the sequential progression of RF centroids is in well-differentiated steps in one direction, then the reverse direction, as the junction between macrocolumns is crossed, encountering the mirror-image reversal of skeins in adjacent macrocolumns. Were the electrode to pass horizontally through deep cortex, a similar result (but with reversed positions of sequential RF centroids) would be obtained.

Figure 5 shows an intermediate result, also a type II penetration, where the electrode has passed obliquely through the system, encountering a substantial fraction of neurons lying in the poorly demarcated zones between skeins. A mixture of effects arises, all producing “jumps” backwards and forwards.

This continuum of outcomes permits model fitting to experimental data.

METHODS
All data were obtained by digitizing the original MaxRF drawings of single neurons studied by Favorov, Whitsel and Diamond in 1980s at the University of North Carolina at Chapel Hill and published in Favorov et al. (1987), Favorov and Whitsel (1988a) and Favorov and Diamond (1990).

PHYSIOLOGICAL RF RECORDINGS
Experimental methods are described for monkeys in Favorov and Whitsel (1988a) and for cats in Favorov and Diamond (1990). All experiments were conducted in accord with ethical standards then current at the host institutions.

A recording chamber was installed over a small opening in the skull over the S1 forelimb region of Mucaca fascicularis monkeys or domestic cats. Dura was removed under general anesthesia, surgical sites infiltrated with local anesthetic, then halothane anesthesia was discontinued, and the subjects were immobilized with gallamine triethiodide, placed on positive pressure ventilation, and EEG monitored.

Using electrolytically sharpened, glass-insulated tungsten electrodes in arrays, closely spaced radial penetrations each yielded extracellular records from 11–43 neurons. The effective sampling radius of the electrodes was estimated at <50 μm. Single units were isolated on the basis of the amplitude and the shape of the recorded action potentials; multi-unit recordings were excluded. For each isolated single unit demonstrated to receive input from the skin, the modality (hair movement, skin contact, or movement over the skin) of the effective peripheral stimuli, and the rate of adaptation of the neuronal response to maintained adequate stimulation were determined. The stimulus that evoked the most vigorous response when applied to the most sensitive part of the RF was used to map RF boundaries. Cells responding primarily to deep tissue stimuli were excluded.

Maps of skin RF boundaries were transferred to a standard forelimb drawing and subsequently digitized for analysis. Electrolytic lesions were placed in selected penetrations to enable reconstruction of electrode tracks in the serial histological sections prepared for each brain.

Centroids of RFs were computed from their boundaries, and the centroids plotted on standard forelimb drawings. In a few instances divided RFs were found, that could not be defined within a single enclosing boundary, so in these cases the centroid was calculated for the overall RF.

PRELIMINARY IDENTIFICATION OF RF CENTROID PATTERNS BEFORE MODEL FITTING
In each dataset, progressions of the centroid positions in the order they were recorded were noted. Three typical classes of results were found, corresponding to those described by Figures 3–5, with an overlay of spatial noise.

1. Cases in which there is one cluster of points, and a few outliers or a smaller cluster. For model fitting, centroids in the first cluster (in numerical order of the neurons recorded) were designated red, and those of the outliers/second cluster blue, in accord with colors ascribed in Figure 3.
2. Cases in which there is a clear sequence of points going continuously in one direction, then reversing back in the other direction. The initial direction of progress of the centroid progress in order was designated red, the reverse progress blue, in accord with the colors ascribed in Figure 4.
3. Cases in which the points in sequence “jump” back and forth irregularly, without distinct clusters. In accord with Figure 5, points following a jump in one direction were designated red, jumps in the opposite direction blue.

Since these data were to be fitted to a theoretical model utilizing a maximum of two adjacent macrocolumns, the model can be fitted only to datasets in which there is a maximum of three distinct clusters or sequential runs of red and blue points in a single direction. More than three clusters or runs were apparent in some datasets obtained in a single electrode penetration. Then the dataset was split into sections each of which could be fitted to the
model individually, and these subsets are subsequently referred to as “splits.”

A few datasets were obtained in which the centroids progressed in jumps both proximal and distal as well as ulnar and radial, preventing orderly numbering. These data were set aside, considered as artefactual or arising from situations outside the theoretical model to be used in fitting—e.g., they might arise from electrode penetration of cortical areas on junctions between three macrocolumns.

Calculations were then made to allow for effects of cortical anisotropy. Figure 6 shows a typical RF against a cat’s forepaw outline. The size and elliptical shape of the RF reflects the anisotropic representation of the skin surface in S1 (e.g., Kaas et al., 1979) in which cortical representation is of roughly inverse size to skin extension. Length of the longest axis of each RF and its angle to a reference frame on the animal’s limb, and length on the broadest axis orthogonal to the long axis, were measured and averaged over the RFs in a dataset. Subsequently the average anisotropy for the dataset was applied to the spatial distribution of theoretical points at each trial of fit.

MODEL FITTING
Best fit of theoretical model to experimental data was found for each experimental dataset of RF centroid positions, by random matching of experimental data to results possible according to the theoretical model. This was done by:

(a) Translating the experimental centroid points so that their combined center-of-mass was located at the origin, and normalizing the scale of the translated plot to the unit circle.

(b) Computing sets of theoretical RF centroid point sets

\[ P_1 \ldots P_n \]

—each set associated with one of all possible probe tracks, and of all equidistant neuron positions from \( \mu_1 \ldots \mu_n \), along each possible subset of each track—thus choosing random neuron point sets over the full range of possibilities for parameters \( a, b, \) and \( c \), in Equation (6), and generating theoretical centroid point sets for each of these, in accord with Equations (8a,b).

(c) Transforming each theoretical centroid point set to eliminate the effects of cortical anisotropy, as described in the prior section, normalizing their scaling and translation as for the experimental data and randomly rotating each set.

(d) For each of these anisotropy-corrected, normalized, scaled and translated theoretical point sets, calculating the sum of squared deviations of theoretical \( P_1 \ldots P_n \) from their experimental equivalents in each trial.

(e) Selecting the best fit. For each random match, goodness of fit was calculated as RMS noise/signal ratio—that is, the sum of squared distances of separation of each of the matched theoretical and experimental centroids was divided by the sum of squared distances of the experimental centroids from the origin, and the square root of the result then taken.

For each dataset 120 billion random matches were tried, and their best-fit results tabulated.

CONSTRUCTION OF NESTED RFs
The RFs were superimposed and the number of RFs overlapping at each position on the standard forelimb counted, to obtain local maximum counts. Local maxima thus indicated nest centers, as those with the largest number of contributing RFs. We then superimposed upon each of the centers all those RFs within which that center lay, thus obtaining all nested RFs for the dataset. Each nest was plotted on the standard forelimb outlines, and the color spectrum from red to blue used to indicate their density of overlap.

RESULTS
DATASETS AND EXCLUSIONS
From the existing monkey data, we excluded experiments in which RF outlines were confined to fingers, as this distribution made calculation of RF centroids problematic. From cats we excluded datasets in which a large fraction (about 50%) of RFs were non-sequential, because interpolated neurons had not had RF boundaries recorded but were used for a different purpose. We retained one dataset (CAT8615p1) that had only two neurons’ recordings missing in sequence, and in that instance interpolated blank, unweighted positions during model fitting.

The only other exclusions were two monkey datasets in which the plot of centroids progressed in steps proximal and distal as well as radial and ulnar, preventing orderly numbering, and were therefore considered artefactual or obtained from an irregular cortical area.

This left 22 datasets, composed of either complete sequences passing through the cortex (10 datasets) and 12 “splits,” divided as stated in Methods (6 complete sequences, each divided into 2 split datasets).
The 22 datasets were obtained from 7 animals (3 cats and 4 monkeys, 7 datasets from cats and 15 from monkeys) providing RF outline data from 310 neurons.

**ELECTRODE TRACK HISTOLOGY**

Penetrations were never perfectly perpendicular to the cortical surface. A 200 µm advance of the electrode resulted in an average displacement of 50 µm in the horizontal plane. In comparing experimental data with the theoretical model, we have made no attempt to directly relate histological measurement of obliquity of the electrode track to the outcomes, because track variation was markedly overshadowed by the effect of part of the limb in which the measured RFs were clustered. If located distally in the limb, results approximated expectation for a vertical electrode passage. If in the proximal limb, the results approached expectation for oblique or almost horizontal electrode passages. That is, cortical anisotropy, by affecting relative size of cortical representation of equal skin areas, predominated in determining the apparent obliquity of the electrode track, because the larger the cortical representation the less the apparent lateral drift across the skin representation.

**REPRESENTATIVE RESULTS OF MODEL FITTING**

Figures 7–12 show representative results. Complete analyses of all datasets are supplementary to this paper, and are commented on further at the end of this section. An example from each of cat and monkey is shown, for each of the three types on the continuum of possibilities forecast in Figures 3–5.

In the top left corner, each figure has the label given the data in the original experiments of Favorov et al. Top left is the plot of RF centroids, and top right the plot is referred to the cat or monkey forelimb outline. Centroids are colored red or blue according to the conventions described in Methods. Bottom left is the best-fit theoretical result, with, shown in green, the corresponding connected line through the experimental centroids. Bottom right, the two circles show the pair of theoretical macrocolumns and needle track in plane view.

The values shown in each figure as “RF anisotropy” give the average ratio of the long to the short axes of the RFs in a dataset, and the angle of the long axis to the reference axes. “Rotation” is the angle through which the theoretical best fit was rotated during matching of model to experiment. It is, therefore, a measure of the orientation of the hypothetical local map.

RMS noise/signal ratios (goodness of fit) ranged from 0.23 to 1.06, the lower values tending to occur with more proximal RF datasets. Mean RMS noise/signal over all datasets was 0.62. Tabulated results are provided as Supplementary to this paper.

Considerations relevant to interpreting these results are as follows:

(i) No other theoretical model exists, to the best of our knowledge, which can explain the positions of the RF centroids found in these experiments. The earlier segregate analysis of Favorov et al., cited in the Introduction, is based on different experimental methods, and is related to the RF nests described in a subsequent section.

(ii) The experimental datasets vary markedly in their apparent orderliness, but are not random patterns. Each dataset resembles one or other of the patterns predicted in Figures 3–5.

(iii) The theoretical model finds best fits in accord with the patterns predicted in Figures 3–5, and the model could not be fitted adequately to most geometric configurations—with the proviso that interference by spatial noise, and ambiguity in some centroid configurations makes alternative fits of the model, with similar goodness of fit, possible in some cases. Trials were made in which datasets where fitted with red and blue designations other than as described in section Preliminary Identification of RF Centroid Patterns Before Model Fitting, to illustrate these properties and limitations (See Supplementary Material).

(iv) There is mutual consistency of the findings of the model-fits across separate datasets. Over all datasets, including those provided as supplementary material, the following trends were apparent:

- In 9/22 datasets, a pattern of RFs similar to the “vertical” outcome exemplified in Figure 3 was found, all with a type I
best-fit configuration. In 8 of these, the RFs were located in the
distal limb. This is the type of outcome expected in distal skin
areas, which have larger cortical representation.
• In 6/22 datasets, a pattern of RFs similar to the “horizontal”
outcome exemplified in Figure 4 was found, all with a type II
best-fit configuration. In 5 of these the RFs were located in the
proximal limb. Conversely, this is the type of outcome expected
in cortical areas with smaller cortical representation.
• In 7/22 datasets, a pattern of RFs similar to the “oblique” out-
come exemplified in Figure 5 was found, 5 of which were of
type II, and one of type I. All were in middle dispositions
on the limb, except for the type I instance, which was prox-
imally located. They are thus deployed in the skin area that has
intermediate levels of cortical representation.
• Where sequentially discovered RFs in a single electrode
penetration were “split” so as to be able to be fitted to the
two-macrocolumn theoretical model, then in 4/6 cases, a type
II fit was found in the first “split” followed by a type II fit to
the second split, and in 2/6 a type II fit was followed by a type
I fit—showing the model gives consistent results in all cases,
indicative of continued horizontal straight-line passage of the
electrode.
• Best-fit noise/signal ratios tended to increase proximally to
distally. This is also consistent with expectation for the effects
of cortical anisotropy, since larger cortical representation
compresses the distribution of the RFs, without reducing
recording uncertainty (spatial noise) regarding the exact
position of the receiving neuron.

• The mean RMS noise/signal of 0.62, given the effective elec-
trode sampling radius of <50 µm, implies mean horizontal
drift of the recording electrode in passage through some or all
of the cortical depth, was approximately 80 µm. This appears
close to actual electrode drift.

COMPARISON OF MÖBIUS ORDERING WITH NESTED RFs
Nested RFs were obtained for all 22 datasets, and also for
all 16 complete full-cortical thickness sequential recordings, so
that nests in the “splits” were considered both combined and
separately.
In one dataset we found 4 nests, in one 3 nests, in four 2 nests,
and only 1 nest in all others. The datasets that had been divided
into “splits” for model fitting, in all cases but one, had only a sin-
gle nest, and the exceptional case was composed of 3 nests, 2 in
one split, and 1 in the other.
There was no one-to-one correspondence between nested RFs
and Möbius model constructions. However, where multiple nests
were found in a dataset, the RF centroids of the set were relatively
distally located on the limb—the first and largest nest more prox-
imaly, and the smaller nests more distally on the cat’s paw, or
on individual fingers of the monkey. In the proximal forearm of
both species, large single nests predominated. The proximal/distal
differentiation into fewer and more nests corresponded to the
proximal/distal trend of type II vs. type I model fits.
Figures 13A,B show two cases at these extremes.
Figure 13A shows the entire overlapping RF set, and the four
nests obtained from this dataset. Two of the four nests are each

FIGURE 8 | A “vertical” penetration of cortex in a monkey.

FIGURE 9 | A “horizontal” penetration in a cat.
FIGURE 10 | A “horizontal” penetration in a monkey.

FIGURE 11 | An “oblique” penetration in a cat.

FIGURE 12 | An “oblique” penetration in a monkey.

FIGURE 13B shows a result from the forearm that is all one nest. The Möbius model had to be fitted in two “splits” implying the dataset was obtained from penetration of three macrocolumns.

DISCUSSION
As originally recognized by Favorov et al., the RFs do not conform to an homotypic map, nor are they wholly random, nor a mixture of homotypic map with uniformly distributed spatial noise. Other than the present model, we know of no other candidate geometry that could explain these results. The expected variant patterns with path of descent of the recording electrode were all encountered in roughly equal numbers, average noise/signal ratios conformed to expectation for spatial noise of recordings, other expectations arising from the effects of position on the limb, and cortical anisotropy were consistently obtained, including mutually consistent modeling of recording sequences that were split for analysis.
organized for all modalities. We have also argued previously (Wright and Bourke, 2013) that similar principles may enable mutual self-organization of synapses and interactions between cortical areas.

Continued successful application of this model would enable reformulation of the column concept, not as an anatomical entity dependent on the positions of cell bodies and the overall deployment of axons, but as a functional entity, constructed from the patterning of synaptic connections, whether or not the cells composing the system are separated into obvious columns, or intertwined more completely. This places emphasis much more strongly on horizontal connections (Boucein et al., 2011) and moves away from Mountcastle’s minicolumn as a structural entity. The minicolumn, in this view, is the necessary result of order in the horizontal connections, and in this revised concept, the minicolumn does not retain response identity throughout surface to depth of the cortex.

Further experimental testing of the model appears practicable. Using other sensory modalities, testing, similar to the present work, using multiple or single unit analyses, and electrode trajectories exploring small cortical areas, could seek to determine positions of sudden “jumps”—the breaches of continuity, or reversal of RF centroids, that are the crucial model signature. These are likely to be found most clearly with electrode trajectories very close to vertical, and nearly horizontal. Primary sensory cortices of all modalities might be thus explored, as may other cortical areas for which the sources of inputs can be defined. More accurate quantitative modeling would need to take into account the local distributions of local cell axonal lengths, and the spatial resolution in inputs, relating these to the relative clarity of definition of anatomical columnar definition. Direct anatomical demonstration of Möbius-strip-like organization of synaptic connections would place high demands on the capacity of existing techniques to specify detailed synaptic connections by the thousands. The task is made even more difficult where primary embryonic order is overwritten during subsequent learning, as must be the case with continuing pre-synaptic pruning and apoptosis in the early post-natal period. The discovery, made by Markram and colleagues, of “lego” assemblies of strongly connected pyramidal cells that form interdigitated networks (Perin et al., 2011) may reflect the establishment of the cross-correlations revealed in RF nests, and the “lego sets” might also be regarded as subassemblies capable of connection into Möbius meshes. The very large-scale analysis of detailed synaptic connectivity intended in the Human Brain Project and related proposals (Kandel et al., 2013) might definitively confirm or refute the model.

Theoretical development of this connectivity concept might exploit the long-sought advantages of decomposition of the cortex into modular information processing units. There are other implicit and attractive properties. The model explains how neural connections, self-organized into ultra-small-world configuration, may approach a maximum for speed and energy-efficiency, and indicates how spatio-temporal order may underpin the organization of signal traffic and learning, since local maps within each macrocolumn have inputs organized in accord with distance and delay from central points—of which the OP singularity is archetypal. The principle of synaptic resource competition,
upon which the model is based, also suggests a form of synaptic metabolic entanglement, permitting signal processing complexity, and synaptic information storage, to approach theoretical maxima.

ACKNOWLEDGMENTS

A further debt of gratitude is owed to Adrienne Edith Wright. This work was supported by the Frank P.ixon Fund of the California Institute of Technology and the Oakley Foundation of New Zealand, and by IVEC, through the use of advanced computing resources located at the University of Western Australia. The authors are also indebted to Betsy Houston for technical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fnana.2014.00119/abstract

REFERENCES


Frontiers in Neuroanatomy www.frontiersin.org October 2014 | Volume 8 | Article 119 | 13


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 June 2014; accepted: 06 October 2014; published online: 31 October 2014.


This article was submitted to the journal Frontiers in Neuroanatomy. Copyright © 2014 Wright, Bourke and Favorov. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Conclusion
Towards a theory of brain function

It remains for me to summarise the extent to which this work might contribute to our understanding of the brain’s basic operation – of how the brain stores, accesses, and avoids mutual interference of information flows, while learning selectively.

As I remarked in the first section, when I began in Caltech I had studied the work of Sherrington, who by 1906 recognised and addressed essential elements of these problems. He identified reflex systems, innately wired at lower brain levels, and operating with reciprocal inhibition, so as to prevent confusion of separate responses. He thought similar principles might be extended into the cortex, creating “a magic loom” (Sherrington, 1940). I had also read Donald Hebb’s influential Organization of Behaviour (1949), which envisaged transient functional “cell assemblies” of cortical neurons, blending into the next assembly in a “phase sequence”. His learning rule for the development of synapses suggested how cell assemblies could be strengthened with practice, and phase sequences learned in order. Similarities between the operation of the brain and of the computer had been emphasised by Alan Turing, and to the extent that Hebb’s account may be equated with properties of a Turing Machine, the problem is how to match the Machine’s components to the components of brains. Also, however the components are organised, newly evolved parts of the brain must be built upon, and interact with, the lower reflex processes, as was elucidated by MacLean who argued aspects of reptilian brains carry over into mammalian brains, as if the older brain was enveloped within the newer, with old and new acting in concert.

Those were my principal influences in the 1970s, as well as that of Sperry, and later, particularly of Freeman, and of Nunez, as I have stated. That of Nunez is of mainly historical importance now, since the explanation of cerebral rhythms as resonant modes dependent on brain size has proved incorrect. The work of Freeman is ongoing, has emphasised far-from equilibrium thermodynamic properties and physical analogies applicable to whole neuron complexes, rather than synaptic dynamics, and thus my own work has diverged from this.

Since 1971, in the wider world, no predominant general model of the brain has emerged despite gigantic technical advances and accumulation of data. While it is impossible to do justice to these bodies of work, their weaknesses are easier to point out, and I do so to indicate the ways in which the present work may help. Beginning from a classification in Wikipedia, present brain theories can be divided between modular theories, and distributed process theories. These are only loosely termed “theories” and are better described as methods.

Modular theories seek to determine which brain components are most active in, or crucial to, particular cognitive processes. This approach dates from the 19th century or earlier, and functional scanning by fMRI, PET, EEG and MEG are the major tools for current investigations. The limited temporal resolution of fMRI and PET, and the limited spatial resolution of EEG and MEG, have restricted the significance of this work to that of improving the precision of classical neuropsychological findings, without adding greatly to understanding of the details of cellular-level information processing.

---

Distributed process, or connectionist, theories can be separated into “wet” and “dry” branches.

The “wet” branch is that of cellular neurophysiology. Still dominated by influences of Hubel and Wiesel, this covers not only attempts to determine function by studying single cell responses to stimuli using very sophisticated techniques, but a wealth of synaptic and cell biology, channel biophysics and genomics. This huge body of results has not yet found a framework about which it can be adequately integrated.

The “dry” branch is that of computer science. There are simulations of simplified neurons in computer programs – feed-forward and attractor neural network models – in which the simulated neurons do not closely approach the complexity of real neurons, although the abstract principles revealed are of vital importance. That approach perhaps reached a zenith in the SyNAPSE project – an IBM Project embodying millions of neuron-like units in computer chips. Beyond effects of scale no new property not already implicit in the architecture was revealed. The mimicry of cognitive processes on computers as “Artificial Intelligence” increasingly dominates much of daily life, but tells us little about how the brain works. More formidably, the field of quantum computing under current development funded by the Google Corporation (“D-Wave”) offers the prospect of very fast parallel processing of problems that are decidedly more like the capacities of the human brain. This prospect is further discussed, below.

At the time of writing, the failure of fusion of the enormous body of neuroscientific work into a model of the brain has called forth two current projects that aim to address the difficulty – the one billion euro Human Brain Project (HBP), and the one billion US dollar BRAIN (Brain Research through Advancing Innovative Neurotechnologies). Controversy between and within these projects illustrates the crisis. The HBP has as a major goal the simulation on a conventional megacomputer of a very large neuronal ensemble, approximating the scale of the brain, with highly specified cellular detail to the extent to which it is presently understood. The BRAIN project eschews such an approach as premature, and is aimed at developing still more precise investigative tools so as to be able to specify the brain’s “connectome” to the required level of detail (whatever that is). Both projects, but particularly the HBP, claim analogy to the Human Genome Project, and are consequently criticised as lacking any comparable clarity of purpose to that provided by Watson and Crick’s solution of the structure of DNA.

This being the current climate, it seems that any reasonable attempt at a bridging synthesis might be worthwhile, and I will now state properties I believe are those needed for a preliminary model of the brain.

Embodiment. The system must be continuously engaged in interaction with the environment, through the medium of the body. Its structure must have emerged by processes of organic evolution, and individual development. The information processed must be rooted in spatio-temporal sensory and motor sequences.

State sequentiality. The system must be able to change modes to deal with quite separate parts of the environment and different demands for responses, as it moves through sequential aspects of an environment that itself imposes much of the orderliness on brain activity.

There must be a “cybernaut”. That is, some continuously operating principle of steerage – a homeostatic mechanism – that is not dependent on the specifics of the information being processed, but has a general organising property like that of attractor basins in simulated neural networks.

Computation. The brain must be capable of very general and highly efficient parallel processing on a very large scale.
Memory and recall. There must be a large memory store, in which separate packages of memory are segregated so as to become available as required by context, without mutual interference. Memory must be ordered into some contextual filing system that interacts with, and corresponds to, the form of the ongoing sensory-motor sequences.

Selective self-reinforcement. The brain must preferentially store, as new memory, those sequences that favour survival.

I will take these requirements one by one, and state findings by my colleagues and me along with other bodies of research. I here treat these assertions of our findings and our closely related hypotheses as truths, for purposes of argument:

**Embodiment**
Cortical/subcortical interactions (and thus interaction with body and sensory environment) necessarily provide the raw information of ongoing cognition, and of memory. The functional specialisation of cortical areas is taken to be a reflection of intermediate relations in the flow of information from sensory inputs to motor outputs, and memory storage and access must be related to position within that ordered flow (see further on memory and recall below).

**Sequentiality**
Complex pathways, with a long evolutionary history, descending from cortex to the brain stem, passing through the hippocampus in part, and including connections to the reticular and non-specific thalamus may set cortical activation into one of many possible spatial patterns – these regulating attention and cognition, and ensuring the adoption of a stimulus-appropriate set. The discovery of hippocampal place neurons\(^\text{16}\) may indicate the existence of specialised networks governing the stepping from one to another attentional set.

Such a system, itself subject to Hebbian learning, could learn to switch the pattern of cortical activation into a new, and relevant, state, based upon learned experience of locales in physical space, and later generalise this capacity to places in “mental” space. Something of this kind was shown in our early work on visual attention in split-brain monkeys.

**The brain’s cybernaut**
This role is played by synchronous oscillation. The electrocortical field in a steady-state is physically a near-equilibrium, multilinear system. Constructive and destructive interference of travelling waves in this medium leads to equilibrium of signal exchange in which excitatory neurons approach maximum zero lag synchrony and excitatory and inhibitory cells approach overall 1/4 cycle lag-correlation. Synchronous oscillation tends to a maximum for given cortical inputs, and the free energy manifested in the travelling wave component tends to a minimum. Thus, the oscillating synchronous field provides global attractors for the electrocortical field, in analogy to the attractors of point-state neural networks.

---

Computation

The assumption that synapses compete for resources in the brain’s mature state, as they do during embryogenesis, means that as activated cortical areas approach equilibrium with maximum synchrony, synaptic metabolic resources will become distributed to maximise the synchrony. There are very many ways the synaptic resources can be distributed, each of which corresponds to a unique pattern of pulse transmission, yet all of which are associated with long-wavelengths of electrocortical activity and zero-lag synchrony of pulses. This resolves the apparent paradox of the EEG’s limited information capacity but close association with cognition. Potentiality to perform large-scale parallel computation, analogous in mathematical form to quantum computation, is implicit – as follows:

In equation 1 the supply of critical metabolites to synapses can be introduced by writing

\[ f_p^{qr} = e_p^{qr} \times g_p^{qr} \]  

where \( e_p^{qr}(t) \) is the pre-synaptic rate of metabolic supply, determining each synaptic efficacy – a rate which follows the recent firing rate of a synapse connecting a neuron at \( r \) to a neuron at \( q \). The total rate of supply is assumed to be sufficient to sustain only 50% of a given neuron’s synapses at maximum efficacy, thus ensuring maximum information representation capacity

\( g_p^{qr} \) is the synaptic density function for intra-cortical cells, composed of separate species of cells of differing characteristic axonal length, with summed range/synaptic density approximating ultra-small world connectivity.

In equation 5 subsets of cell pulse rates/densities arising from sensory and somatic inputs, groups of burst-firing cortical cells, subcortical input cells, and cortical cells firing slowly in their background state, can be separately specified as

\[ Q_p(q,t) = Q(\text{cortical bursting cells}) + Q(\text{cortical low firing rate cells}) + Q(\text{sensory inputs}) + Q(\text{somatic inputs}) + Q(\text{nonspecific subcortical inputs}) \]

Since \( e_p^{qr}(t) \) follows \( q_p^{qr}(t) \) with a short lag, and all synapses compete for resources sufficient for a maximum firing rate of half of them, then by normalising all \( e_p^{qr}(t) \) and \( q_p^{qr}(t) \) on an operating range of 0–1, and representing their product as \( q_i^j \), we can write an function

\[ E(\Psi) = \sum_{i,j} (e_i^j + B_i^j q_i^j) \]

where \( B_i^j \) also on the range 0–1, measures the degree to which synapses share joint success in the acquisition of the critical metabolites, and \( N \) is the total number of synapses.

At synchronous equilibrium pre-synaptic pulse rates and synaptic supply will be aligned, providing maximum efficacy at the most active synapses, and vice versa – i.e. \( e_i^j = 1 \) with a probability of 0.5 that \( e_i^j = 1 \). As all synaptic efficacies are interdependent in the competition, we can also set \( B_i^j = 1 \) for all \( i,j \). Therefore the maximum values \( E(\Psi) \) can attain are binary approximations corresponding to maxima of \( J \) in equation 9.
Maximum for this binary measure is

$$E(q) = \frac{(N^2 + N)}{4}$$

and there are $$2^{N2/4}$$ ways that the pre-synaptic pulses could be arranged to reach maximum, a number restricted to a lesser (but still enormous) number by the possible firing patterns of the pre-synaptic neurons.

Each of these synaptic arrangements corresponds to either a specific pattern of organisation arrived at during embryonic self-organisation, or one of these patterns subsequently partially, or wholly overwritten by post-natal learned modifications. Therefore, each defines a basin of attraction, and an individual cerebral cortex can develop a very large number of basins of attraction, described in an N-dimensional state space.

Under the influence of a sudden change in the external inputs, and subsequent triggering of cortical burst firing at various sites – that is, the inputs listed in equation 11 – solutions of the state equations (1–5) will determine subsequent approach to one of the basins of attraction. Depending on the Mahalanobis distance of the closest approach of different trajectories through the state space, trajectories will avoid mutual interference, or will merge on close approach.

The descriptive usefulness of expressing equation 9 in the normalised and binary form of equation 13 lies in the analogy of equation 12 to the quantum energy function

$$E(S) = \sum_{i,j=1}^{N} (h_i S_i + J_{ij} S_i S_j)$$

where $$E(S)$$ is the energy function to be minimised by adjustment of h and J in the execution of quantum computation, S are quantum probabilities, and h and J are weights applied to probabilities and joint probabilities. At the conclusion of such computations uncertainty in the quantum states is removed – i.e. $$S_{i,j} \in \{0,1\}$$. Quantum computation is considered to offer efficient solution of the large group of problems known as NP-complete, including pattern recognition and other problems performed relatively slowly and poorly by standard computers, but relatively well by humans.

The analogy between quantum probabilities with entanglement, and pulse probabilities with synaptic metabolic entanglement, indicates a cortical capacity to perform very large scale parallel computations, operating on the raw material of sensory-motor sequences. There are parallels and differences in the analogy. In quantum computation, the quantum probabilities are wholly random initially. In the cortex, after a sudden “reset” of subcortical inputs, pre-synaptic pulse rates and efficacies would become pseudo-random, including scattered combinations of pre-synaptic pulses and metabolic supply, permitting effects similar to quantum tunneling in quantum computation – a very rapid means of converging towards optimum, or near-optimum, solutions.

The values h and J in the quantum energy function are imposed by the programmer, whereas in our hypothetical brain their analogues are discovered in the process of convergence to an equilibrium state in an attractor basin; a state which is, in turn, the analogue of the solution to the problem solved by the

---

17 Equation (14) is taken from the elementary introduction to quantum computing given at dwavesys.com
quantum computation. Obviously, without appropriate selection of the possible trajectories through the brain’s state-space – without motivation and reinforcement – the brain’s computations would remain disorganised and ineffective for organism survival (see further below).

**Memory and recall**

On longer time-scales – that of the consolidation of long-term memory – permanent, Hebbian synaptic learning would emerge consequent to repetitions in the cortical computations, effectively charting preferred courses through the cortical state space. Hebb’s principle seems naturally applicable when synchronous firing naturally chooses those sets of cells that “fire together” so they can “wire together” in accord with the standard cliché. Thus, there is analogy to the setting of program values, $h$ and $J$ in the quantum energy function, and the execution of an overlearned sensory-motor or cognitive sequence.

A problem with Hebbian learning is the limited capacity of any small neural network to store more than a limited number of attractors without mutual interference. In the present proposal only a limited role need be played by any small set of neurons and synapses, yet overall integration is possible, since Hebbian learning applies to two relatively distinct groups of synapses.

Firstly, it applies to those synapses determining the sequential steps of the pattern of cortical activation triggered by cortical-subcortical interactions, resetting and biasing the global basin of attraction and determining present context. Some of these synapses may be at subcortical levels, but within cortex they are synapses on the proximal dendritic trees, on which the Hebbian principle provides bias of the matrix of values $B_{ij}$ in equation 12. This permits global, long-term patterns of synaptic connection to guide state trajectories towards particular equilibrium attractor basins, from a given range of initial conditions.

Secondly, it applies to synapses in the distal dendritic trees. As travelling waves, propagating at relatively low pulse rates, and transmitted primarily by synapses located in proximal dendritic trees, intersect, and the system moves towards its global attractor, transient additive superpositions of inputs trigger small cortical locales in foci of bursting activity. As described in the preceding papers, the non-linear bursts are self-stabilising and localised. Each bursting focus transmits travelling waves triggering and/or suppressing other loci of oscillation, allowing complex sequences of activity to take place, under the overall guidance towards convergence to synchronous equilibrium. During strong burst firing, retrograde propagation of action potentials into the dendritic trees of the neurons, so that synapses distal to the somas can undergo Hebbian learning, would permit forming small local attractor systems, thus classifying input patterns and ensuring the release of a similar bursting pattern for a range of similar input conditions. Thus, information could be internally stored and released, concurrent with, and modifying, the system responses to external stimuli.

The mechanisms for maintenance of overall stability, restricting spread of focal activity, segregates separate “packets” of memory, while the embryonic cortical connectivity provides the crucial contextual reference framework for the packets of stored memory, which are thus always stored and released in relationship to sensory-motor sequences. Thus cognition is embodied, and operates within anatomical order.

Hebb’s “cell assemblies” would roughly correspond to set of synapses and cells most engaged in the approach to a global attractor, and his “phase sequences” to the ongoing shifts of cortical activation mediated by ongoing cortical-subcortical interactions.
Selective self-reinforcement
The brain’s inherited core structures mediating primitive survival behaviour, emphasised by MacLean, provide a means for self-supervision of learning. The powerful motivational effects exerted by pathways travelling to and from cortex via the lateral hypothalamus trigger storage of memory traces by a diffuse action, distinct from the localised storage of specific cognitive sequences in cortex itself – as is illustrated by my experiments in “split-brain” animals.

It appears, therefore, that cognitive computations performed by entangled synaptic states and pulses might be selectively stored on the basis of their cooperative consistency with hard-wired activity of lower-level survival circuits. If so, what the cortex learns is the application of ancient patterns of survival behaviour in ever increasingly complicated environments – similarly to adjustments of cortical non-specific activation in sequences. Neuromodulation effects upon Hebbian consolidation would preferentially select synaptic computations performed by the interactions of the electrocortical field and changing synaptic efficacies – storing those linked with the separate subcortical systems mediating approach behaviours or avoidance behaviours respectively. Computation sequences would not be consolidated if they enlisted neither system, and conflict would arise until resolution by experience resolved complicated situations that partially activated both systems.

In conclusion
So there you have it. This account goes some way to indicating how the brain might store, access, and avoid mutual interference of information flows, while learning selectively. The proposal is highly provisional, and requires extensive testing in competition with alternative models, wherever, and whenever, these are available. No claim for completeness is made, but the principles presented seem sufficient for future instantiation in hardware (possibly in the context of quantum computing) and a conceptual framework for the assembly of increasingly exact cellular detail.

I hope I have indicated the many ways in which the work is indebted to others, and is consistent with many ideas widely understood elsewhere in neuroscience. A claim to originality lies in the analysis of synchronous oscillation, and in the theory of cortical embryogenesis.

I am most grateful to all the colleagues with whom I have worked over the years, and those seniors and colleagues who helped me when in need – but most especially I am grateful for my family, my children’s loyalty, and above all, to Adrienne.
Appendix: Curriculum Vitae

Personal Details
James Joseph Wright, MB, ChB, MD (Otago), FRACP, MRCPsych, FRANZCP.
Born Christchurch, NZ, 1942.
Married to Adrienne Edith Wright (nee Greer), BA (Otago), JP.
Born Ranfurly, NZ, 1943.

Children
Jason John Wright, BSc (Auckland).
Born Christchurch, NZ, 1968.
Tanya May Wright, BSc Hons (Otago), BHB, MB, ChB (Auckland), FRANZCP.
Born Christchurch, NZ, 1970.
Conrad Greer Wright, BA Hons (Otago), LLB (Auckland).

Academic Record

Degrees
1966 MB ChB (Otago)
1972 MD (Otago)

Professional qualifications
1970 (March) MRACP Royal Australasian College of Physicians (Fellow 1973)
1974 MRC Psych Royal College of Psychiatrists (London)
1975 MRANZCP Royal Australian & New Zealand College of Psychiatrists (Fellow, 1986)

Prizes and Competitive Fellowships
1966. Ardagh Memorial Prize (First in class of final year medical students, Christchurch Branch Faculty, University of Otago School of Medicine).
1971–72. Research Fellow under the Frank Hixon Memorial Fund of the California Institute of Technology, Pasadena, California. (Professor R.W. Sperry.)
1973–75. Mental Health Trust and Research Fund Fellow, Dept. of Physiology, Institute of Psychiatry, Kings College, London (Professor G. Brindley.)

Positions Appointed
1969. Medical Registrar, North Canterbury Hospital Board.
1970. Research Registrar, The Medical Unit (Dr D.W. Beaven), Princess Margaret Hospital, Christchurch.
1973 (Jan–June). Registrar (Psychiatry), St Georges Hospital, London (Professor A. Crisp).
1973 (July)–1976 (March). Research Fellow, Department of Physiology, Institute of Psychiatry, Kings College, London.
1976–1978. Senior Lecturer, Department of Psychiatry, University of Auckland School of Medicine.
1978–October 1993. Associate Professor, Department of Psychiatry, University of Auckland School of Medicine.
October 1993–1999. Senior Scientist and Professorial Fellow, Mental Health Research Institute, Melbourne.
February 1999–present. Consultant Psychiatrist, Waitemata District Health Board, and Honorary Professor, University of Auckland.

**Doctoral Level Supervisions**

**MD Degree (Otago)**
- Dr S.W. Miles

**PhD Degrees (Auckland)**
- Dr R.R. Kydd
- Dr A.A. Sergejew
- Dr D.T.J. Liley

**PhD Degree (Melbourne)**
- Dr Peter Line

**PhD Degrees (Sydney)**
- Dr Clare Chapman
- Dr C.J. Rennie
Refereed Publications


J.J. Wright (1989). Linearity and non-linearity in electrocortical waves, and their elementary statistical

J.J. Wright, R.R. Kydd & A.A. Sergejew (1990). Autoregressive models of EEG. Results compared with
expectation for a high-order multilinear near-equilibrium biophysical process. Biological Cybernetics.


J.J. Wright, A.A. Sergejew & H.G. Stampfer (1990). Inverse filter computation of the neural impulse giving
rise to the auditory evoked potential. Brain Topography. 2: 293–302.


activity. Psycholoquy. 4(60).


J.J. Wright, D.T.J. Liley (1994). A millimetric scale simulation of electrocortical wave dynamics based on


Programme Models of the Mind. (Open Education).

J.J. Wright, D.T.J. Liley (1996). Dynamics of the brain at global and microscopic scales: neural networks and


P. beim Graben, J.J. Wright (2011). From McCulloch-Pitts neurons toward Biology. (Guest Editorial, Special edition of Bulletin of Mathematical Biology, honouring the publication in BMB of the founding papers of Neural Network Theory, by McCulloch and Pitts.)


