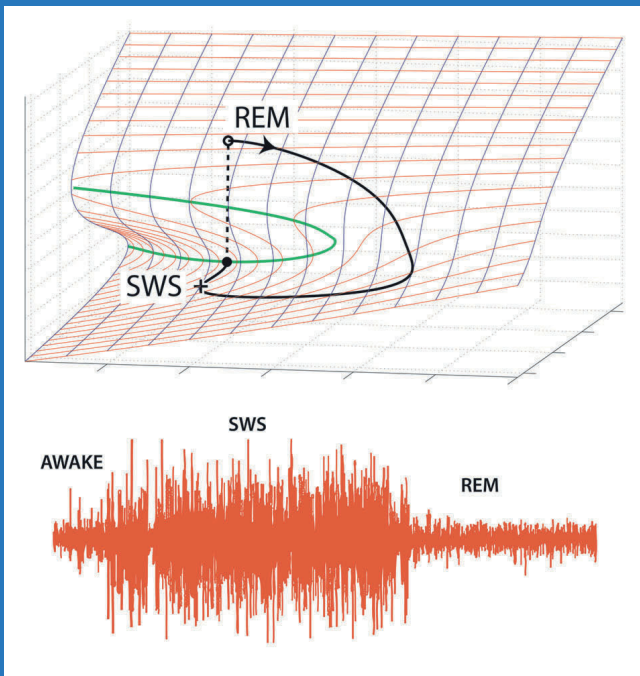


SPRINGER SERIES IN COMPUTATIONAL NEUROSCIENCE

Modeling Phase Transitions in the Brain



D. Alistair Steyn-Ross
Moira Steyn-Ross
Editors

 Springer

Springer Series in Computational Neuroscience

Volume 4

Series Editors

Alain Destexhe

Unité de Neurosciences Intégratives et Computationnelles (UNIC)

CNRS

Gif-sur-Yvette

France

Romain Brette

Equipe Audition (ENS/CNRS)

Département d'Études Cognitives

École Normale Supérieure

Paris

France

For other titles published in this series, go to
<http://www.springer.com/series/8164>

D. Alistair Steyn-Ross · Moira Steyn-Ross
Editors

Modeling Phase Transitions in the Brain

 Springer

Editors

D. Alistair Steyn-Ross
Department of Engineering
Hillcrest Road
University of Waikato
Hamilton 3240
Gate 8
New Zealand
asr@waikato.ac.nz

Moira Steyn-Ross
Department of Engineering
Hillcrest Road
University of Waikato
Hamilton 3240
Gate 8
New Zealand
msr@waikato.ac.nz

Cover design shows brain electrical activity recorded from the cortex of a cat transiting from slow-wave sleep (SWS) into rapid-eye-movement (REM) sleep. The folded grid illustrates a mathematical model for the transition, with SWS and REM phases corresponding to quiescent (lower branch) and activated (upper branch) brain states. [Cat time-series adapted from Destexhe, A., Contreras, D., Steriade, M.: *J. Neurosci.* **19**, 4595–4608 (1999), reproduced with permission, Society for Neuroscience. Sleep manifold adapted from Steyn-Ross, D.A., Steyn-Ross, M.L., Sleigh, J.W., Wilson, M.T., Gillies, I.P., Wright, J.J.: *J. Biol. Phys.* **31**, 547–569 (2005).]

ISBN 978-1-4419-0795-0

e-ISBN 978-1-4419-0796-7

DOI 10.1007/978-1-4419-0796-7

Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2009941302

© Springer Science+Business Media, LLC 2010

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of going to press, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

Early in the 19th century debates on Darwinian theory of evolution, William James asked the question whether consciousness had biological survival value, such that it might be subject to natural selection. The alternatives he considered were widely held notions that consciousness was either an epiphenomenon or a celestial gift of the capacity to conceive and know a Creator. He answered in fluid Victorian prose:

A priori analysis of both brain and conscious action shows us that if the latter were efficacious it would, by its selective emphasis, make amends for the indeterminacy of the former; whilst the study *à posteriori* of the distribution of consciousness shows it to be exactly such as we might expect in an organ added for the sake of steering a nervous system grown too complex to regulate itself.¹

In raising and answering the question this way, James penetrated to the essential role of the brain in behavior. The brain simplifies. We and other animals cannot fully know the world, Kant's *Ding an sich*, as it is in its infinite complexity. Instead, we make finite educated guesses about the world that Kant called "categories" and that we now call "representations" or "world models". We test these hypotheses by taking action into the world and refining our guesses into formal theories. We learn to know our world by accommodating and adapting to the sensory consequences of our own and others' actions through trial-and-error reinforcement learning [Freeman (2001)].² Thereby we achieve the simplicity that makes it possible for each of us, immersed in a sea of uncertainty, to take effective action lit by flashes of insight.

Neurodynamicists model this self-organized, self-educating process by constructing mathematical descriptions of the motor systems that thrust the body into and through the world. They postulate that the sensory systems maintain attractor landscapes that are constructed by Hebbian and other forms of synaptic modification in cortical networks, which are the structural repository of experience. Each act of observation is a test of the world, and the multiple attractors are predictions of possible outcomes of the test, giving evidence for sustenance, companionship, danger,

¹ See p. 18 of James, W: Are we automata? *Mind* 4, 1–21 (1879)

² Freeman, W.J.: *How Brains Make Up Their Minds*. Columbia University Press, New York (2001)

nothing new, or something novel. The basins of attraction are generalization gradients from prior receptions of stimuli. A stimulus places cortical dynamics in one of the basins of attraction. Convergence to an attractor is an inductive generalization by which the stimulus is categorized. The attractor manifests a spatiotemporal pattern of neural activity to which the cortical trajectory converges [see Chaps 1, 2, 7 of this book; also Freeman (2001)], and which the sensory cortex transmits to its targets by well-known networks and pathways [Chap. 5 of this book].

Here is the crux of perception. The sensory input is a representation of the stimulus; the cortical output is not. Based on the memories of the stimulus, the output is the mobilized knowledge about the meaning of the stimulus [Freeman (2001)]. The experience is familiar to everyone; a whiff of perfume, a few notes of a tune, or a glimpse of a face can trigger a cascade of recollection and emotion. Whereas the pattern of the sensory-driven cortical activity is defined by the parameters of the physical world, and by the neural operations of the sensory systems, the self-organized pattern of cortical activity is defined by the modified synapses that store the accumulated experience of the perceiver [Chap. 11]. Hence the critical event in each act of perception is the reorganization of a stimulus-driven activity pattern in cortex, which embodies the unique and unknowable impact from the world into an endogenous pattern of self-organized activity. The neurons are the same; their anatomical connections are the same; even their level of energy may be the same; what differs is the spatiotemporal organization of their interactions.

The process of sudden reorganization of neural masses in the brain is the subject matter of this book. It is the *phase transition* [Freeman (1999)]³ that is modeled by use of differential equations [Chap. 8; Freeman and Vitiello (2006)]⁴ or random graph theory [Chap. 5; *neuropercolation*, Kozma *et al.* (2005)].⁵ In its simplest form it is the succinct, localized transition in the state of a sensory cortex from a receiving state to a transmitting state. Cortex transforms a receipt into a percept by constructing knowledge from information. That is the first step in the transition by the brain from an expectant state to a knowing state, the elusive “Aha!” experience. It is also the transition from body into mind, from a pattern determined by the physics of matter in the world to a self-organized pattern that exists only in the perceiver as a mental state. Abrupt global reorganizations by phase transitions in larger brain systems implement a wide variety of intellectual and intentional brain functions, ranging from simple go/no-go choices, switching from rest to action and back [Chap. 4], from prodrome to epilepsy [Chaps 2, 5], from sleep to wake or REM [Chap. 9], and, far beyond our current reach, from Heidegger’s *thrownness* in childhood through

³ Freeman, W.J.: Noise-induced first-order phase transitions in chaotic brain activity. *Internat. J. Bifur. Chaos* **9**(11), 2215–2218 (1999)

⁴ Freeman, W.J., Vitiello, G.: Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. *Physics of Life Reviews* **3**, 93–118 (2006), doi:10.1016/j.plrev.2006.02.001, <http://repositories.cdlib.org/postprints/1515>

⁵ Kozma, R., Puljic, M., Balister, P., Bollabás, B., Freeman, W.J.: Phase transitions in the neuropercolation model of neural populations with mixed local and non-local interactions. *Biol. Cybern.* **92**, 367–379 (2005), <http://repositories.cdlib.org/postprints/999>

adolescence to mid-life crises, military, religious or political conversions, and all other forms of social bonding.

Physicists and engineers are familiar with state changes, charting them as discontinuities in trajectories of state variables through state space. Neurologists and psychiatrists well understand states of mind and altered states of consciousness. What is to be gained by calling brain states “phases”, which gives the title of this book?

On its face the usage appears to be no more than a treacherous analogy. On the one hand the classical thermodynamic definition holds for closed systems at equilibrium, whereas brains are open, dissipative systems operating far from equilibrium. The classical phases and their boundaries are unequivocally defined in terms of temperature and pressure, whereas brains homeostatically regulate temperature, pressure, volume, and mass. Conventional phase transitions involve latent heat, so that the Ehrenfest classification by discontinuities of derivatives has been largely discarded by physicists, but as yet no comparable transition energies have been seen or postulated in cortical phase transitions, so discontinuities must suffice for neurodynamicists.

On the other hand, the several fields of condensed-matter physics have evolved in diverse directions such as nonequilibrium thermodynamics, ferromagnetics, optics, and computational fluid dynamics, but with commonality in important aspects [Schroeder (1991)].⁶ Phase now is defined as a state of aggregation of particles [Schwabl (2006)],⁷ whether they are atoms, molecules or neurons. In each complex system there are multiple types of state. In the brain, families of attractor landscapes in sensory cortical dynamics define the phase space [Freeman and Vitiello (2006)]. In each aggregate there are certain conditions that specify a critical point in the phase space at which the system is particularly susceptible to transit from one phase to another phase [Chap. 1]. The transition involves a change in the degree of order, as when the neurons in sensory cortex transit from a disorganized state of expectancy to an organized state of categorization, from noise to signal, from the symmetry of uniformity of the background activity at rest to the asymmetry of spatiotemporal structure in action. This is symmetry breaking, which is described using bifurcation theory [Chaps 1, 10, 11].

Most importantly, the order emerges by *spontaneous* symmetry breaking within and among populations of cortical neurons. Order in the form of gamma synchrony [Chaps 7, 8, 11, 12] is not imposed by sensory receptors or pacemaker neurons. It is constructed by broadly distributed synaptic interactions by which neurons constrain or “enslave” themselves and each other in circular causality [Haken (1983)].⁸ Modeling symmetry breaking requires the introduction of an extra variable, an *order parameter* [Chap. 3], which serves to evaluate the strength of interaction by which

⁶ Schroeder, M.R.: *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*. W.H. Freeman, New York (1991)

⁷ Schwabl, F.: *Statistical Mechanics*, 2nd ed. Ch. 7. Phase transitions, scale invariance, renormalization group theory, and percolation. 331–333, Springer (2006)

⁸ Haken, H.: *Synergetics: An Introduction*. Springer, Berlin (1983)

the order is achieved [Sethna (2009)].⁹ The variable must be evaluated by measuring the summed activity of the aggregate, minimally from practical experience on the order of 10,000 neurons [Freeman, (2001)]. The mesoscopic order [Chap. 7] is undefined in microscopic activity, in much the way that molecules do not have pressure and temperature. Furthermore, it is undetectable in microelectrode recordings of action potentials except by prolonged time-averaging in histograms, which precludes measuring rapid changes in the degrees of freedom or patterns of order. Herein lies the value of the dendritic potentials recorded from cortex extracellularly and referentially in various forms of electroencephalogram (EEG) and local field potentials [Chaps 1, 6, 8, 11], which are averages of potential fields from local neighborhoods of neural populations. The EEG order parameter (derived from field potential measurements) is not the order, nor is it the agency of the order; it is an index of the distributed, self-organized and self-organizing interaction strength among the neurons.

The perceptual phase transition is many-to-one by convergence to an attractor, so it is irreversible, non-Abelian and non-commutative with no inverse. Unlike the holographic transformation, which is information-preserving and non-categorizing, the phase transition destroys information in categorizing as the prelude to decision-making. To these properties are added the characteristic amplification and slowing of fluctuations as criticality is approached [Chaps 1, 8]; the emergence of power-law distributions of spectral energy and functional connectivity [Chaps 1, 3, 4, 8]; long correlation lengths reflecting emergence of truly immense domains [Freeman (2003)]¹⁰ of coherent gamma oscillations; and reorganization/resynchronization of phase and amplitude modulations of the transmission frequencies at rates in the theta and alpha ranges [Freeman (2009)].¹¹

Perhaps the most compelling reason to model the dynamics of perception as a phase transition is the reduction in degrees of freedom owing to augmented interaction [Freeman and Vitiello (2006)],¹² which resembles the increase in density as gas condenses to liquid. The condensation of neural activity is manifested in the long-range spatiotemporal coherence of gamma oscillations (Chaps 1, 4, 12), and the conic phase gradients resembling vortices that accompany the EEG amplitude patterns that are correlated with behavior [Freeman (2001)]. The phase transition begins at a singularity (Chaps 1, 8), which in cortex is demarcated spatially by the

⁹ Sethna, J.P.: *Statistical Mechanics. Entropy, Order Parameters, and Complexity*. Clarendon Press, Oxford (2009), <http://pages.physics.cornell.edu/sethna/StatMech/EntropyOrderParametersComplexity.pdf>

¹⁰ Freeman, W.J., Burke, B.C., Holmes, M.D.: Aperiodic phase re-setting in scalp EEG of beta-gamma oscillations by state transitions at alpha–theta rates. *Hum. Brain Mapp.* **19**(4), 248–272 (2003), <http://repositories.cdlib.org/postprints/3347>

¹¹ Freeman, W.J.: Deep analysis of perception through dynamic structures that emerge in cortical activity from self-regulated noise. *Cognit. Neurodynamics* **3**(1), 105–116 (2009), <http://www.springerlink.com/content/v375t514t065m48q/>

¹² Freeman, W.J., Vitiello, G. Dissipative neurodynamics in perception forms cortical patterns that are stabilized by vortices. *J. Physics Conf. Series* **174**, 012011 (2009), <http://www.iop.org/EJ/toc/1742-6596/174/1>, <http://repositories.cdlib.org/postprints/3379>

apex of the cone. It is marked temporally by a downward spike in power in the pass band of the transmission frequency [Freeman (2009)].

Given these properties in brain dynamics, the analogy is exceedingly attractive and likely to persist and grow, because it provides matrices of educated guesses by which further progress can be made in making sense of diverse data. The phase transition establishes a link between energy and order. Brains are profligate in the dissipation of metabolic energy, yet by their own feedback controls they keep constant a vast reservoir of electrochemical energy in the ionic concentration gradients that empower the neural activity of the brain. The major thermodynamic variables are in steady state, owing to provision by arterial blood flow of free energy and the disposal by the venous blood flow of waste heat, except one: there is a continual decrease in entropy [Chap. 4], which is paid for by the throughput of energy. Initially the patterns are solely functional, the creation of chaotic dynamics. Owing to the plasticity of cortical connectivity [Chaps 2, 4, 7, 9, 11] the functional patterns guide the structural connectivity into more or less permanent brain patterns, which constitute the neural foundation for long-term memory.

Despite these properties and the powerful tools used to derive and describe them, the hypothesis that phase transitions underlie perception and other brain functions remains unproven. Asserting it is like signing a promissory note. There are immediate intellectual gains from access to the capital of others' ideas, but they bring unsolved problems, salient among them defining the relation between metabolic brain energy and neural activity, in which both excitation and inhibition dissipate energy. The debt will not be paid until a detailed theory of nonlinear neurodynamics is constructed that can stand on its own, in company with other major branches of physics devoted to the study of condensed matter. Considering the saliency of its subject matter, a successful theory of neurodynamics is likely to outshine all others.

University of California at Berkeley
August 2009

Walter J. Freeman

List of Contributors

Ingo Bojak

Department of Cognitive Neuroscience (126), Donders Institute for Neuroscience, Radboud University Nijmegen Medical Centre, Postbus 9101, 6500 HB Nijmegen, The Netherlands, e-mail: i.bojak@science.ru.nl

Mathew Dafilis

Brain Sciences Institute (BSI), Swinburne University of Technology, P.O. Box 218, Victoria 3122, Australia, e-mail: 3154874@swin.edu.au

Brett Foster

Brain Sciences Institute (BSI), Swinburne University of Technology, P.O. Box 218, Victoria 3122, Australia, e-mail: bfoster@bsi.swin.edu.au

Federico Frascoli

Brain Sciences Institute (BSI), Swinburne University of Technology, P.O. Box 218, Victoria 3122, Australia, e-mail: ffrascoli@swin.edu.au

Andreas Galka

Department of Neurology, University of Kiel, Schittenhelmstrasse 10, 24105 Kiel, Germany, e-mail: a.galka@neurologie.uni-kiel.de

Anandamohan Ghosh

Theoretical Neuroscience Group, Institut des Sciences du Mouvement, Etienne-Jules Marey UMR 6233, Université de la Méditerranée, 163 Avenue de Luminy, CP 910 13288 Marseille cedex 9 France, e-mail: anandamohan@gmail.com

Iain-Patrick Gillies

Dept of Engineering, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand

David Hailstone

Dept of Engineering, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand

Axel Hutt

LORIA, Campus Scientifique-BP 239, 54506 Vandoeuvre-Les-Nancy Cedex, France, e-mail: axel.hutt@loria.fr

Viktor Jirsa

Theoretical Neuroscience Group, Institut des Sciences du Mouvement, Etienne-Jules Marey UMR 6233, Université de la Méditerranée, 163 Avenue de Luminy, CP 910 13288 Marseille cedex 9 France, e-mail: viktor.jirsa@univmed.fr

Marcus Kaiser

School of Computing Science, Newcastle University, Newcastle-upon-Tyne NE1 7RU, U.K.; Institute of Neuroscience, Newcastle University, Newcastle-upon-Tyne NE2 4HH, U.K., e-mail: m.kaiser@newcastle.ac.uk

Jong Kim

School of Physics, University of Sydney, NSW 2006, Australia; Brain Dynamics Centre, Westmead Millennium Institute and Westmead Hospital, Westmead, NSW 2145, Australia, e-mail: jwkim@physics.usyd.edu.au

Xiaoli Li

School of Computer Science, The University of Birmingham, Edgbaston, Birmingham, B15 2TT, U.K., e-mail: X.Li@cs.bham.ac.uk

David Liley

Brain Sciences Institute (BSI), Swinburne University of Technology, P.O. Box 218, Victoria 3122, Australia, e-mail: DLiley@groupwise.swin.edu.au

Hans Liljenström

Research Group of Biometry and Systems Analysis, Department of Energy and Technology, SLU, SE-75007 Uppsala, Sweden; Agora for Biosystems, SE-19322 Sigtuna, Sweden, e-mail: hans.liljenstrom@et.slu.se

Tohru Ozaki

Tohoku University, 28 Kawauchi, Aoba-ku, Sendai 980-8576, Japan

Andrew Phillips

School of Physics, University of Sydney, NSW 2006, Australia; Brain Dynamics Centre, Westmead Millennium Institute and Westmead Hospital, Westmead, NSW 2145, Australia, e-mail: ajp@physics.usyd.edu.au

Christopher Rennie

School of Physics, University of Sydney, NSW 2006, Australia; Department of Medical Physics, Westmead Hospital, Westmead, NSW 2145, Australia; Brain Dynamics Centre, Westmead Millennium Institute and Westmead Hospital, Westmead, NSW 2145, Australia, e-mail: chris.rennie@wmi.usyd.edu.au

James Roberts

School of Physics, University of Sydney, NSW 2006, Australia; Brain Dynamics Centre, Westmead Millennium Institute and Westmead Hospital, Westmead, NSW 2145, Australia, e-mail: jamesr@physics.usyd.edu.au

Peter Robinson

School of Physics, University of Sydney, NSW 2006, Australia; Brain Dynamics Centre, Westmead Millennium Institute and Westmead Hospital, Westmead, NSW

2145, Australia; Faculty of Medicine, University of Sydney, NSW 2006, Australia,
e-mail: robinson@physics.usyd.edu.au

Jennifer Simonotto

School of Computing Science, Newcastle University, Newcastle-upon-Tyne NE1
7RU, U.K.; Institute of Neuroscience, Newcastle University, Newcastle-upon-Tyne
NE2 4HH, U.K., e-mail: jennifer.simonotto@newcastle.ac.uk

Jamie Sleigh

Department of Anaesthesia, Waikato Clinical School, University of
Auckland, Waikato Hospital, Hamilton 3204, New Zealand, e-mail:
sleighj@waikatodhb.govt.nz

Alistair Steyn-Ross

Department of Engineering, University of Waikato, P.B. 3105, Hamilton 3240,
New Zealand, e-mail: asr@waikato.ac.nz

Moira Steyn-Ross

Department of Engineering, University of Waikato, P.B. 3105, Hamilton 3240,
New Zealand, e-mail: msr@waikato.ac.nz

Lennaert van Veen

Department of Mathematics and Statistics, Faculty of Arts and Sciences, Concordia
University, 1455 de Maisonneuve Blvd. W., H3G 1M8 Montreal, Quebec, Canada,
e-mail: lvanveen@mathstat.concordia.ca

Logan Voss

Department of Anaesthesia, Waikato Clinical School, University of
Auckland, Waikato Hospital, Hamilton 3204, New Zealand, e-mail:
VossL@waikatodhb.govt.nz

Marcus Wilson

Department of Engineering, University of Waikato, P.B. 3105, Hamilton 3240,
New Zealand, e-mail: m.wilson@waikato.ac.nz

Kevin Kin Foon Wong

Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114,
USA

James Wright

Liggins Institute, and Department of Psychological Medicine, University of
Auckland, Auckland, New Zealand; Brain Dynamics Centre, University of Sydney,
Sydney, Australia, e-mail: jj.w@xtra.co.nz

Acronyms

ACh	acetylcholine
ARMA	autoregressive moving average
BOLD	blood-oxygen-level dependent
ECoG	electrocorticogram
ECT	electroconvulsive therapy
EEG	electroencephalogram
EPSP	excitatory postsynaptic potential
fMRI	functional magnetic resonance imaging
GABA	gamma-aminobutyric acid
HH	Hodgkin–Huxley
IDE	integro-differential equation
IPSP	inhibitory postsynaptic potential
IS	intermediate sleep (pre-REM) in rats
LFP	local field potential
LOC	loss of consciousness
MEG	magnetoencephalogram
NSF	nonspecific flux
PCA	principal components analysis
PDE	partial differential equation
PET	positron-emission tomography
PSP	postsynaptic potential
REM	rapid-eye-movement sleep
ROC	recovery of consciousness
SN	saddle–node
SR	stochastic resonance
SWS	slow-wave sleep
TCF	transcortical flux

(See also the list of anatomical brain abbreviations on p. 97)

Contents

Foreword	v
List of Contributors	xi
Acronyms	xv
Introduction	xxiii
1 Phase transitions in single neurons and neural populations: Critical slowing, anesthesia, and sleep cycles	1
D.A. Steyn-Ross, M.L. Steyn-Ross, M.T. Wilson, and J.W. Sleigh	
1.1 Introduction	1
1.2 Phase transitions in single neurons	2
1.2.1 H.R. Wilson spiking neuron model	3
1.2.2 Type-I and type-II subthreshold fluctuations	5
1.2.3 Theoretical fluctuation statistics for approach to criticality	7
1.3 The anesthesia state	11
1.3.1 Effect of anesthetics on bioluminescence	11
1.3.2 Effect of propofol anesthetic on EEG	13
1.4 SWS–REM sleep transition	15
1.4.1 Modeling the SWS–REM sleep transition	17
1.5 The hypnic jerk and the wake–sleep transition	20
1.6 Discussion	23
References	24
2 Generalized state-space models for modeling nonstationary EEG time-series	27
A. Galka, K.K.F. Wong, and T. Ozaki	
2.1 Introduction	27
2.2 Innovation approach to time-series modeling	28
2.3 Maximum-likelihood estimation of parameters	28

2.4	State-space modeling	30
2.4.1	State-space representation of ARMA models	30
2.4.2	Modal representation of state-space models	32
2.4.3	The dynamics of AR(1) and ARMA(2,1) processes	33
2.4.4	State-space models with component structure	35
2.5	State-space GARCH modeling	36
2.5.1	State prediction error estimate	36
2.5.2	State-space GARCH dynamical equation	37
2.5.3	Interface to Kalman filtering	38
2.5.4	Some remarks on practical model fitting	38
2.6	Application examples	40
2.6.1	Transition to anesthesia	41
2.6.2	Sleep stage transition	43
2.6.3	Temporal-lobe epilepsy	45
2.7	Discussion and summary	48
	References	51
3	Spatiotemporal instabilities in neural fields and the effects of additive noise	53
	Axel Hutt	
3.1	Introduction	53
3.1.1	The basic model	54
3.1.2	Model properties and the extended model	57
3.2	Linear stability in the deterministic system	58
3.2.1	Specific model	60
3.2.2	Stationary (Turing) instability	61
3.2.3	Oscillatory instability	63
3.3	External noise	66
3.3.1	Stochastic stability	68
3.3.2	Noise-induced critical fluctuations	70
3.4	Nonlinear analysis of the Turing instability	71
3.4.1	Deterministic analysis	71
3.4.2	Stochastic analysis at order $O(\varepsilon^{3/2})$	74
3.4.3	Stochastic analysis at order $O(\varepsilon^{5/2})$	76
3.5	Conclusion	77
	References	78
4	Spontaneous brain dynamics emerges at the edge of instability	81
	V.K. Jirsa and A. Ghosh	
4.1	Introduction	81
4.2	Concept of instability, noise, and dynamic repertoire	82
4.3	Exploration of the brain's instabilities during rest	86
4.4	Dynamical invariants of the human resting-state EEG	89
4.4.1	Time-series analysis	90
4.4.2	Spatiotemporal analysis	93
4.5	Final remarks	94
	References	97

5 Limited spreading: How hierarchical networks prevent the transition to the epileptic state 99
M. Kaiser and J. Simonotto

5.1 Introduction 99

5.1.1 Self-organized criticality and avalanches 100

5.1.2 Epilepsy as large-scale critical synchronized event 101

5.1.3 Hierarchical cluster organization of neural systems 101

5.2 Phase transition to the epileptic state 103

5.2.1 Information flow model for brain/hippocampus 103

5.2.2 Change during epileptogenesis 104

5.3 Spreading in hierarchical cluster networks 105

5.3.1 Model of hierarchical cluster networks 105

5.3.2 Model of activity spreading 107

5.3.3 Spreading simulation outcomes 107

5.4 Discussion 111

5.5 Outlook 112

References 114

6 Bifurcations and state changes in the human alpha rhythm: Theory and experiment 117
D.T.J. Liley, I. Bojak, M.P. Dafilis, L. van Veen, F. Frascoli, and B.L. Foster

6.1 Introduction 117

6.2 An overview of alpha activity 118

6.2.1 Basic phenomenology of alpha activity 119

6.2.2 Genesis of alpha activity 120

6.2.3 Modeling alpha activity 121

6.3 Mean-field models of brain activity 122

6.3.1 Outline of the extended Liley model 124

6.3.2 Linearization and numerical solutions 128

6.3.3 Obtaining physiologically plausible dynamics 129

6.3.4 Characteristics of the model dynamics 130

6.4 Determination of state transitions in experimental EEG 136

6.4.1 Surrogate data generation and nonlinear statistics 137

6.4.2 Nonlinear time-series analysis of real EEG 137

6.5 Discussion 138

6.5.1 Metastability and brain dynamics 140

References 141

7 Inducing transitions in mesoscopic brain dynamics 147
Hans Liljenström

7.1 Introduction 147

7.1.1 Mesoscopic brain dynamics 148

7.1.2 Computational methods 149

7.2 Internally-induced phase transitions 150

7.2.1 Noise-induced transitions 150

7.2.2	Neuromodulatory-induced phase transitions	155
7.2.3	Attention-induced transitions	156
7.3	Externally-induced phase transitions	162
7.3.1	Electrical stimulation	162
7.3.2	Anesthetic-induced phase transitions	167
7.4	Discussion	170
	References	173
8	Phase transitions in physiologically-based multiscale mean-field brain models	179
	P.A. Robinson, C.J. Rennie, A.J.K. Phillips, J.W. Kim, and J.A. Roberts	
8.1	Introduction	179
8.2	Mean-field theory	181
8.2.1	Mean-field modeling	181
8.2.2	Measurements	184
8.3	Corticothalamic mean-field modeling and phase transitions	184
8.3.1	Corticothalamic connectivities	184
8.3.2	Corticothalamic parameters	185
8.3.3	Specific equations	187
8.3.4	Steady states	187
8.3.5	Transfer functions and linear waves	189
8.3.6	Spectra	189
8.3.7	Stability zone, instabilities, seizures, and phase transitions	191
8.4	Mean-field modeling of the brainstem and hypothalamus, and sleep transitions	194
8.4.1	Ascending Arousal System model	194
8.5	Summary and discussion	198
	References	198
9	A continuum model for the dynamics of the phase transition from slow-wave sleep to REM sleep	203
	J.W. Sleigh, M.T. Wilson, L.J. Voss, D.A. Steyn-Ross, M.L. Steyn-Ross, and X. Li	
9.1	Introduction	203
9.2	Methods	204
9.2.1	Continuum model of cortical activity	204
9.2.2	Modeling the transition to REM sleep	207
9.2.3	Modeling the slow oscillation of SWS	208
9.2.4	Experimental Methods	209
9.3	Results	210
9.4	Discussion	212
9.5	Appendix	215
9.5.1	Mean-field cortical equations	215
9.5.2	Comparison of model mean-soma potential and experimentally-measured local-field potential	217

9.5.3 Spectrogram and coscalogram analysis 217

References 219

10 What can a mean-field model tell us about the dynamics of the cortex? 223

M.T. Wilson, M.L. Steyn-Ross, D.A. Steyn-Ross, J.W. Sleigh, I.P. Gillies, and D.J. Hailstone

10.1 Introduction 223

10.2 A mean-field model of the cortex 224

10.3 Stationary states 226

10.4 Hopf bifurcations 227

10.4.1 Stability analysis 227

10.4.2 Stability of the stationary states 228

10.5 Dynamic simulations 229

10.5.1 Breathing modes 230

10.5.2 Response to localized perturbations 233

10.5.3 K-complex revisited 237

10.5.4 Spiral waves 240

10.6 Conclusions 241

References 241

11 Phase transitions, cortical gamma, and the selection and read-out of information stored in synapses 243

J.J. Wright

11.1 Introduction 243

11.2 Basis of simulations 244

11.3 Results 245

11.3.1 Nonspecific flux, transcortical flux, and control of gamma activity 245

11.3.2 Transition to autonomous gamma 246

11.3.3 Power spectra 248

11.3.4 Selective resonance near the threshold for gamma oscillation 248

11.3.5 Synchronous oscillation and traveling waves 251

11.4 Comparisons to experimental results, and an overview of cortical dynamics 252

11.4.1 Comparability to classic experimental data 253

11.4.2 Intracortical regulation of gamma synchrony 253

11.4.3 Synchrony, traveling waves, and phase cones 254

11.4.4 Phase transitions and null spikes 255

11.5 Implications for cortical information processing 257

11.6 Appendix 260

11.6.1 Model equations 260

11.6.2 Hilbert transform and null spikes 264

References 265

12 Cortical patterns and gamma genesis are modulated by reversal potentials and gap-junction diffusion	271
M.L. Steyn-Ross, D.A. Steyn-Ross, M.T. Wilson, and J.W. Sleigh	
12.1 Introduction	271
12.1.1 Continuum modeling of the cortex	272
12.1.2 Reversal potentials	272
12.1.3 Gap-junction diffusion	273
12.2 Theory	274
12.2.1 Input from chemical synapses	274
12.2.2 Input from electrical synapses	280
12.3 Results	282
12.3.1 Stability predictions	282
12.3.2 Slow-soma stability	284
12.3.3 Fast-soma stability	284
12.3.4 Grid simulations	287
12.3.5 Slow-soma simulations	288
12.3.6 Fast-soma simulations	290
12.3.7 Response to inhibitory diffusion and subcortical excitation	290
12.4 Discussion	294
Appendix	297
References	298
Index	301

Introduction

Historical context

The study of phase transitions—changes in the states of matter—was a leading area of research in the 19th and 20th centuries. Physicists categorized these state changes as being either *first-* or *second-order*. First-order or *discontinuous* transitions are characterized by an abrupt change of phase when a control variable, such as temperature, is smoothly and monotonically varied across the transition point, causing, for example, ice to melt, and water to vaporize. Reversing the direction of the temperature change reverses the transition, causing steam to abruptly condense into liquid water, and water to freeze into a crystalline solid. But for first-order transitions the particular state of matter can depend on the history of the control variable. Thus, in the absence of nucleation centers, pure water can remain liquid when cooled below its normal ice–water transition point, or when heated above its normal water–steam point. This state dependence on history is called *hysteresis*.

In contrast, second-order or *continuous* phase transitions show a smooth change of state with no evidence of hysteresis. For example, when a ferromagnet in a zero magnetic field is heated so that its temperature crosses a critical temperature (the Curie point), its magnetic state smoothly changes from aligned (ferromagnetic) to random (paramagnetic).

Early studies focused on so-called *equilibrium* transitions in which the behavior of state variables, such as gas pressure, temperature, and volume, is governed by a thermodynamic equation of steady state whose mathematical form is determined by the locations of the local minima attractors within a free-energy potential landscape.

Experiments indicated that phase transitions exhibit a set of universal properties—notably power-law divergence of bulk parameters such as heat capacity, susceptibility, compressibility—as the critical point is approached. The quest to understand the origin of these unifying principles led to advances such as the Kadanoff and Fisher *scaling laws*, and culminated in the development of renormalisation group theory. These theoretical advances also introduced the notion of an *order parameter*, assigned a non-zero value in the more ordered (e.g., ferromagnetic) phase.

In the 1960s, the pioneering work of the Brussels group led by Nicolis and Prigogine revealed another type of transition—the pattern-forming phase transitions exhibited by particular types of chemical reaction in fluids. Another fundamental advance at this time was Haken’s treatment of the laser as a self-organizing tran-

sition from uncorrelated light fields into directed photon emissions that are coherent in time and space. The fluid and laser transitions both belong to the family of *nonequilibrium* phase transitions, so called because the underlying reactions require a continuous flux of reactants and energy, so are far from thermodynamic equilibrium and cannot be described in terms of a free-energy potential function. Such phase transitions are of a *spatiotemporal* kind, characterized by spontaneous spatial pattern formation (the original homogeneous steady-state becomes destabilized by a *Turing instability*) and temporal oscillations (destabilization via a *Hopf instability*). The identification of an order parameter for such transitions remains unclear, with the amplitude of the critical, slow mode that emerges at the transition point being a promising candidate.

Book overview

In this book, we put forward the perhaps controversial idea that phase transitions can and do occur in the brain. Like all “living” biological components, the brain never operates in closed thermodynamic equilibrium, and yet we find that on approach to a neural change of state (e.g., moving from wake to anesthetic sleep, or from slow-sleep into REM sleep), its bulk electrical signals can display divergent correlated fluctuations that are tantalizingly similar to first- and second-order thermodynamic phase transitions. Further, we argue that the emergence of spatiotemporal patterns in the brain (e.g., epileptic seizure, alpha and gamma oscillations, the ultraslow oscillations of BOLD fMRI patterns) provides strong evidence of nonequilibrium transitions in brain state.

This idea for this book arose from discussions at the CNS*2007 Computational Neuroscience Meeting held in Toronto, Ontario in July 2007. Joseph Burns, then Senior Editor for Life Sciences, Springer, suggested to ASR and MSR that they construct a book proposal for a contributed volume of chapters written by senior researchers in computational brain modeling. *Phase Transitions in the Brain* was chosen as the unifying foundation for the book, and this proved to be an attractive theme that was enthusiastically adopted by the coterie of invited authors.

Brain activity can be modeled either by a discrete *network of active nodes*, or by a *mean-field continuum*—both approaches are illustrated in this book. In the *network* approach, each node could be a conductance-based spiking neuron, or an idealized neuron, or could represent a cluster of neurons, with the biological fidelity (and mathematical complexity) of each node being determined by the purpose and scale of the model. The topology of the connections *between* nodes is another modeling choice, and can be, for example, all-to-all, random, small-world, or distance-dependent. Chapters 4, 5, and 7 illustrate the network-based approach to brain modeling.

In the *continuum* approach, the brain is described in terms of *populations* of excitatory and inhibitory neurons that interact via chemical synapses over both short and long ranges. Differential (or integro-differential) forms are derived to give the equations of motion for spatially-averaged (i.e., *mean-field*) activity subject to external

(subcortical drive) and internal (neurotransmitter) influences. Although the authors of chapters 3, 6, and 8–12 all use a mean-field philosophy, their model details can and often do differ in subtle but important ways. Some of these distinctions will be outlined below.

We now give a brief summary of each of the 12 chapters. With the exception of Chaps 1 and 12 (by the editors), the chapters are organized alphabetically by author.

In **Chap. 1**, Alistair Steyn-Ross and colleagues examine theoretical and experimental evidence for phase transitions in single neurons (onset of spiking), and in neural populations (induction of anesthesia, SWS–REM sleep cycling, transition from wake to sleep). The type-I and II spiking-neuron models due to H.R. Wilson are biophysical simplifications of the conductance-based Hodgkin–Huxley “gold standard”. A spiking instability can be induced by increasing the dc stimulus current entering the neuron. By adding a subtle white-noise “tickle” to the dc bias, the nearness to spiking transition can be quantified by allowing the neuron to explore its near-equilibrium state space, exercising what Jirsa and Ghosh (Chap. 4) call its *dynamic repertoire*. The subthreshold fluctuations grow in amplitude while becoming critically slowed (type I) or critically resonant (type II) as the bifurcation point is approached. Similar critical changes in fluctuation statistics are seen in the EEG activity recorded from patients undergoing, then recovering from, anesthesia; and in the brain activity of mammals (cat, fetal sheep, human) transiting from slow-wave to REM sleep. Near the point of falling asleep, the phase transition conjecture predicts a nonlinear increase in neural “irritability” (susceptibility to small stimulus); this critical effect may explain the puzzling, yet commonly experienced whole-body hypnic jerk at sleep onset.

We argue that there appears to be ample evidence of phase transition-like behavior in the brain. But can we quantify these qualitative state changes by extracting the critical exponents underlying the power-law growth of the fluctuations in cortical activity? Given the inherently nonstationary nature of the signal at the point of transition, this would be a highly challenging ambition. In **Chap. 2**, Andreas Galka and colleagues present a possible way forward. These authors describes a state-of-the art *generalized autoregressive conditional heteroscedastic* (GARCH) method, first used in financial modeling, which allows the dynamical noise covariance to change with time. They apply this GARCH technique to three distinct dynamical state transitions captured by EEG recordings: induction of general anesthesia in a human patient, emergence of epileptic seizure in a human, and transition from slow-wave to REM sleep in a fetal sheep. They demonstrate that the GARCH variance can accurately locate the point of phase transition without any prior information on the timing of the nonstationary event.

In **Chap. 3**, Axel Hutt investigates a 1-D continuum model of the cortex that is expressed in terms of a single neural population whose effective membrane voltage V is written as the signed summation of the excitatory and inhibitory postsynaptic potentials that combine at the soma: $V = V^e - V^i$. Using Mexican-hat axonal distribution functions, he is able to establish analytic conditions for the existence of stationary (Turing) and oscillatory (Hopf and traveling wave) instabilities. The sta-

bility of the Turing bifurcation in the presence of noise is investigated. Close to the instability point, a linear analysis predicts *critical fluctuations*, that is, the emergence of long-lived zero-wavenumber fluctuations of large variance. When the system becomes unstable, a nonlinear stability analysis shows that the presence of global noise can restore stability to the homogeneous state by suppressing the stochastic Turing instability. The fact that noise can both stimulate *and* suppress formation of spatiotemporal activity patterns in the cortex may have significant implications for information processing in the brain.

In **Chap. 4**, Viktor Jirsa and Anandamohan Ghosh emphasize the fundamental importance of random noise in allowing an excitable system to explore and exercise the repertoire of dynamic behaviors that can be accessed from its resting, equilibrium state. This idea is illustrated first in simple bifurcation (saddle–node and Hopf) models, and then applied to a network simulation of a brain at background rest. The simulation utilizes a biologically realistic primate (macaque) connectivity matrix with 38 nodes, and includes time delays via signal propagation between brain areas, and intrinsic noise. The authors argue that the working point of the brain during wakeful rest is often close to the critical boundary between stable and unstable regions. From the network simulation results they are able to identify the correlated and anticorrelated subnetworks that are active during the ultra-slow BOLD signal oscillations, and demonstrate excellent agreement with experimental observations. The authors offer some insights into the rich default-mode dynamics of the idling brain, suggesting that undirected, spontaneous thoughts can only arise because of the presence of noise (else the rest-state would be truly at rest), but that the response to this random stimulus is tuned by the brain’s deterministic “skeleton” (anatomical connectivity, time-delay structure, internal dynamics) delicately poised close to an instability.

In **Chap. 5**, Marcus Kaiser and Jennifer Simonotto picture the cortex as a complex neural network whose overall level of activity is delicately balanced between the unhealthy extremes of quenched silence, and runaway excitation of the entire network—as seen in the transition to the epileptic state. How is the desired state of persistent, yet contained, network activation sustained, and what prevents uncontrolled spreading of activity to ignite the entire network? The authors argue that, in addition to neuronal inhibition via inhibitory interneurons, network *architecture* plays a crucial role. In particular, their simulations demonstrate that a network composed of hierarchical clusters—with denser connectivity *within* clusters than *between* clusters—provides a form of *topological inhibition* that tends to suppress runaway spreading, even in the absence of inhibitory units. This topological protection against transition into epilepsy arises because of the sparser connectivity between clusters, while the higher density of connections within clusters allows sustained levels of activity. Non-hierarchical topologies, such as small-world and random networks, are shown to be not only less protective against global spreading, but also more susceptible to quenching. Since the anatomy of the brain displays a modular, hierarchical architecture (from microcircuits at the lowest level, to cortical areas then brain areas at the global level), these network insights are likely to have biological significance.

In **Chap. 6**, David Liley and colleagues focus attention on the alpha rhythm (8–13 Hz activity) ubiquitous in human brain EEG. They present the constitutive partial differential equations for a 2-D continuum model of a cortex driven by noisy extracortical sources. With appropriate changes to the cortical parameters, the model can exhibit linear noise-driven dynamics, and nonlinear deterministic (chaotic and limit-cycle) oscillations. The authors demonstrate that the cortical model can undergo Hopf and wave bifurcations in both the alpha and gamma (~ 40 Hz) bands, and that the transition from subthreshold alpha to limit-cycle gamma can be achieved by reducing the strength of the inhibitory postsynaptic potential. They argue that a marginally stable alpha rhythm could provide a “readiness” substrate for neural activity, enabling rapid transitions to the higher frequency cortical oscillations required for information processing.

In **Chap. 7**, Hans Liljenström presents a range of network models, of varying complexity, designed to investigate phase transitions in *mesoscale* brain dynamics, arguing that this intermediate scale, lying somewhere between *micro* (ion channels and single neurons) and *macro* (whole brain), provides a bridge between neural and mental processes. His network models include a three-layer paleocortex (olfactory cortex and hippocampus) of nonspiking nodes; a five-layer neocortex of either conductance-based or simplified spiking neurons; and a monolayer grid of spiking neurons. By modulating an appropriate control parameter, each network model exhibits a characteristic phase transition behavior; thus an increase in intrinsic noise can cause the paleocortex to form spatiotemporal patterns of activity; an electrical stimulus applied to the neocortex can generate seizure-like oscillations; and a reduction in ion-channel conductance can cause a dramatic slowing in EEG dynamics similar to that seen during the induction of general anesthesia.

In **Chap. 8**, Peter Robinson and colleagues outline a philosophy for the construction of continuum models for the brain, and correct some misconceptions about mean-field theory and its applications. The authors present two biophysically-motivated mean-field models. The first model describes the EEG signals generated by the neuronal interactions between the cortex and thalamus, and the second is focused on the slow interactions between brainstem and hypothalamic structures that control the wake–sleep diurnal cycle. Loop resonances in the corticothalamic model produces EEG spectral peaks at the alpha and beta frequencies. By varying loop gains and synaptic strengths, four distinct saddle–node and Hopf bifurcations to limit-cycle oscillations (slow spike–wave, 3-Hz theta, 10-Hz alpha, 10–15-Hz spindle) are identified. These neurodynamic phase transitions may correspond to the genesis of seizure activity. Their sleep–wake model is based on mutual inhibition between wake-active (MA) and sleep-active (VLPO) brainstem nuclei, resulting in a flip-flop dynamics that is controlled by the circadian (*C*) and homeostatic (*H*) drives. The model predicts first-order phase transitions between sleep and wake, with state stability enhanced by a hysteresis separation between transition points.

In **Chap. 9**, Jamie Sleight and colleagues investigate the phase transition between slow-wave and REM sleep, comparing continuum-modeling predictions against ECoG activity recorded from sleep-cycling rats. They argue that the primary effector of the cortical change into the REM state is a progressive linear increase in

cholinergic excitation of the cortex which they model in terms of an increase of neuron resting voltage V_e^{rest} , coupled with a simultaneous decrease in λ , the strength of the excitatory postsynaptic potential (EPSP). The distribution of equilibrium soma voltages, plotted as a function of resting voltage and EPSP gain, defines a sleep manifold featuring multiple steady states and a region of instability that extends beyond the fold, into the upper (active) and lower (quiescent) cortical states. The upper-branch displays a Hopf bifurcation to ~ 8 -Hz limit-cycle oscillations; this dynamic instability may explain the theta-band oscillations observed in ECoG recordings of rats transiting from SWS into REM sleep. The authors are able to demonstrate good agreement between modeled and measured spectrograms across the transition event.

In **Chap. 10**, Marcus Wilson and colleagues present a thorough investigation of the dynamical properties of the mean-field sleep model alluded to in Chap. 1.4, and tuned for rat sleep-cycling in Chap. 9. A linear stability analysis predicts that the homogeneous steady-state cortex will be destabilized by a sufficient reduction in γ_i , the rate-constant for the inhibitory postsynaptic potential (IPSP). This prediction is verified by numerical simulations on a 2-D grid for a range of $(\lambda, V_e^{\text{rest}})$ sleep-manifold coordinates; these simulations show a range of supercritical and sub-cortical Hopf bifurcations to slow (~ 1 -Hz), spatially-symmetric, whole-of-cortex limit-cycle oscillations which the authors identify with seizure. If the homogeneous cortex is stimulated with a brief, spatially-localized impulse, traveling-wave instabilities—reminiscent of K-complexes and delta-waves of slow-wave sleep—can also be evoked.

In **Chap. 11**, Jim Wright describes a closely detailed continuum model of the cortex that attempts to capture, via a hierarchy of nested integral convolutions, the contributions from multiple scales of neural activity, integrating up from the microscale (ion channels and receptors), to the mesoscale (cortical macrocolumns), and the centimetric macroscale sensed by EEG electrodes. His model includes the effects of ion-channel conformation, receptor adaptation, reversal potentials, and back-propagation of action potentials. Numerical simulations of a 2-D cortex demonstrate that, at a critical level of input flux, the homogeneous cortex undergoes a phase transition into autonomous gamma oscillations. The nature and strength of this bifurcation is controlled by both local (excitatory–inhibitory balance) and global (excitatory tone) flux sources. The author discusses possible implications of gamma synchrony and traveling waves on the recall and processing of information by the brain.

Although the **Chap. 12** continuum model of Moira Steyn-Ross and colleagues shares features with the models of Wright (Chap. 11), Robinson *et al.* (Chap. 8), and Liley *et al.* (Chap. 4), it differentiates itself with the inclusion of electrical synapses or *gap junctions*. Inhibitory gap-junction diffusion is found to modulate the onset of Turing and Hopf instabilities, leading to the appearance of spatiotemporal patterns and waves. The authors show that the nature of the feedback between soma and dendrite strongly influences the dynamics of the cortex. If the soma responds slowly to dendritic input, Turing and low-frequency Hopf bifurcations are predicted, but if the soma response is fast, a gamma-band wave instability emerges. These mutually

exclusive extremes may correspond respectively to the awake brain states of *idling rest* (discussed in Chap. 4) and *active cognition* (Chap. 11).

Final comments

The brain can sustain an extremely rich diversity of brain states, some of which are unhealthy and pathological. Of the universe of states available to the brain, these chapters have focused attention on just a small subset—anesthetized, slow-wave sleeping, REM sleeping, awake, idle daydreaming, active thinking, epileptic seizing. Transitions between these gross states can be driven by a range of “control” parameters such as altered neurotransmitter and neuromodulator concentrations, gap-junction diffusion, synaptic efficiency, electrical stimulus. Phase transitions in matter arise because of changes in the nature of the bonding between atomic constituents. Perhaps for the brain, one analog of “atomic bonding” might be the strength of coupling between different neural populations, with stronger coupling leading to enhanced synchrony and qualitatively distinct neural states.

The phase transition approach to brain modeling seems to provide natural explanations for some of the counterintuitive fluctuation divergences seen on approach to state change, such as the EEG power surge seen during induction of general anesthesia, and during the natural progression from slow-wave sleep into REM sleep. But as yet, there is no general theory of neural phase transitions. The noise-evoked fluctuation statistics from single-neuron models might provide some guidance, with critical slowing at a saddle–node annihilation, and critical ringing at a Hopf bifurcation, corresponding to increased correlation times and correlation lengths.

Left open and unanswered are many challenging and intriguing questions: How do single-neuron transitions scale up to influence the behavior and response of neural populations and brain areas? Are critical fluctuations “biologically useful”? Might they have a role in the facilitation—or suppression—of phase transitions? What are the general principles that underlie neural changes of state? Can we find universal scaling laws for brain phase transitions?

Acknowledgments

We are very grateful to librarian Cheryl Ward for her professional attention to detail as she repaired incomplete and sometimes inconsistent bibliographic details in the lists of references, and to Jennifer Campion who, with unflagging good humor and energy, polished the references and improved their utility by locating their online DOI descriptors, and constructed the list of indexing terms while simultaneously mastering the \LaTeX idiom with remarkable efficiency.

University of Waikato
Hamilton, New Zealand
August 2009

Alistair Steyn-Ross
Moiria Steyn-Ross

Chapter 1

Phase transitions in single neurons and neural populations: Critical slowing, anesthesia, and sleep cycles

D.A. Steyn-Ross, M.L. Steyn-Ross, M.T. Wilson, and J.W. Sleigh

1.1 Introduction

It is a matter of common experience that the brain can move between many different major states of vigilance: wakefulness; sleep; trauma- and anesthetic-induced quiet unconsciousness; disease- and drug-induced delirium; epileptic and electrically-induced seizure. By monitoring cortical brain activity with EEG (electroencephalogram) electrodes, it becomes possible to detect more subtle alterations within these major states; for example, we find that natural sleep consists of periodic cyclings between inactive, quiet slow-wave sleep (SWS) and a paradoxically active phase—characterized by rapid eye movements and reduced muscle tone—named REM (rapid-eye-movement) or paradoxical sleep.

The existence of these contrasting brain states motivates us to ask: How does the brain move between states? Is the changing of states a smooth, graduated motion along a trajectory of similar states? Or is the transition more like an abrupt switching choice between two (or more) mutually-exclusive cortical destinations? If the state-change can be thought of as a switching choice, then we might envision a hills-and-valleys cortical landscape in which the crest of a hill represents a decision point, and the two valleys falling away to either side are the alternative destination states. In this picture, we could expect a cortex, delicately poised at a decision point, to exhibit signature behaviors in the statistical properties of its fluctuations as it “ponders” its choices. This notion—that decision points can be identified from critical changes in fluctuation statistics—is a unifying theme that we will return to several times in this chapter.

D. Alistair Steyn-Ross · Moira L. Steyn-Ross · Marcus T. Wilson
Department of Engineering, University of Waikato, P.B. 3105, Hamilton 3240, New Zealand.
e-mail: asr@waikato.ac.nz msr@waikato.ac.nz m.wilson@waikato.ac.nz
<http://phys.waikato.ac.nz/cortex>

Jamie W. Sleigh
Waikato Clinical School, University of Auckland, Waikato Hospital, Hamilton 3204, New Zealand.
e-mail: sleighj@waikatodhb.govt.nz

The chapter is structured as follows. We examine first a simplified single-neuron model due to Hugh R. Wilson [29, 30] that is able to produce either the arbitrarily slow firing rates (so-called type-I behavior) observed in cortical neurons in the mammalian cortex, or, with a minor change in parameter values, exhibits the sudden-onset firing rates (type-II behavior) that characterize both the squid giant-axon excitable membrane, and the neurons in the mammalian auditory cortex. Our interest here is *not* the above-threshold behaviors such as the shape and time-course of the action potential, nor the functional form of spike-rate on stimulus current; instead, we focus on the sensitivity of the non-firing, but *near-threshold*, resting membrane to small noisy perturbations about its equilibrium resting state as the neuron makes a stochastic exploration of its nearby state-space, exercising what Jirsa and Ghosh¹ describe as its *dynamic repertoire*.

Gross changes in states of brain vigilance, such as from awake to asleep, and from anesthetized to aware, reflect alterations in the coordinated, emergent activity of entire populations of neurons, rather than a simple “scaling up” of single-neuron properties. In Sect. 1.3 we examine historical support for the notion that induction of anesthesia can be viewed as a first-order “anesthetodynamic” neural phase transition, comparing biological response to an “obsolete” drug (ether) with a very commonly used modern drug (propofol). We describe EEG response predictions using a noise-driven mean-field cortical model, and identify an explanation for the paradoxical observation that inhibitory agents (such as anesthetics) can have an excitatory effect at low concentrations.

Section 1.4 investigates the SWS–REM sleep cycle, finding similarities in the EEG sleep patterns of the human, the cat, and the fetal sheep. We suggest that the species-independent surge in correlated low-frequency brain activity prior to transition into REM sleep can be explained in terms of a first-order jump from a hyperpolarized quiescent state (SWS) to a depolarized active state (REM).

In Sect. 1.5 we examine the recently published Fulcher–Phillips–Robinson model [18, 19] for the wake–sleep cycle, demonstrating a divergent increase in brain sensitivity at the transition point: the occurrence of a peak in neural susceptibility may provide a natural explanation for the so-called “hypnagogic jerk”—the falling or jolting sensation frequently experienced at the point of falling asleep.

We summarize the common threads running through these neuron and neural population models in Sect. 1.6.

1.2 Phase transitions in single neurons

In the absence of noise, a single neuron is bistable: it is either at rest or generating an action potential. As noted by Freeman [6], the approach to firing threshold is heralded by an increasing sensitivity to stimulus:

¹ See Chap. 4 of this volume.

When a depolarizing current is applied in very small steps far from threshold, the neural dynamics is linear; responses to current steps are additive and proportional to the input [...] As threshold is approached, a nonlinear domain is encountered in which local responses occur that are greater than expected by proportionality.

The fact that a biological neuron is constantly buffeted by a background wash of low-level noisy currents allows the neuron to explore its local state space. These stochastic explorations can be tracked by monitoring the voltage fluctuations at the soma. We will show that the statistics of these fluctuations change—in characteristic ways—as the critical point for transition to firing is approached.

1.2.1 H.R. Wilson spiking neuron model

The H.R. Wilson equations [29, 30] describe neuron spiking dynamics in terms of a pair of first-order coupled differential equations,

$$C \frac{dV}{dt} = I_{\text{Na}}(t) + I_{\text{K}}(t) + I_{\text{dc}} + I_{\text{noise}}(t), \quad (1.1)$$

$$\tau \frac{dR}{dt} = -R(t) + R_{\infty}(V) + R_{\text{noise}}(t). \quad (1.2)$$

The neuron is pictured as a “leaky” capacitance C whose interior voltage V is determined by sum of ionic (I_{Na} , I_{K}) and injected (I_{dc}) currents entering the lipid membrane. Here we have supplemented the original Wilson form by adding white-noise perturbations (I_{noise} , R_{noise}) to the current (1.1) and recovery-variable (1.2) equations.

The sodium (Na) and potassium (K) ionic currents are determined by their respective conductances (g_{Na} , g_{K}) and reversal potentials (E_{Na} , E_{K}),

$$I_{\text{Na}}(t) = -g_{\text{Na}}(V)(V - E_{\text{Na}}), \quad I_{\text{K}}(t) = -g_{\text{K}}R \cdot (V - E_{\text{K}}), \quad (1.3)$$

where R is the recovery variable that approximates the combined effects of potassium activation and sodium inactivation that dominate the slower neuron dynamics for the return to rest following the fast up-stroke of an action potential. Definitions and constants for the H.R. Wilson model are listed in Table 1.1.

Comparing Eqs (1.1–1.3) against Hodgkin and Huxley’s (HH) classic four-variable model for the excitable membrane of the squid giant-axon [11], we see the significant simplifications Wilson has made to the complicated HH forms for the time- and voltage-dependence of sodium and potassium conductances: the sodium conductance g_{Na} is now a quadratic function of membrane voltage; the potassium conductance g_{K} becomes a constant; and the steady-state for the combined potassium activation/sodium inactivation is either a quadratic (for type-I spiking behavior), or linear (type-II spiking), function of voltage. These simplifications reduce the dimensionality of the neuron from four dynamic variables to two—while preserving

Table 1.1 Definitions and constants for stochastic implementation of the H. R. Wilson [30] model for type-I (mammalian) and type-II (squid) excitable membrane. I_{dc}^{crit} is the threshold input current for spike generation.

Description	Symbol	Type-I (mammal)	Type-II (squid)	Unit
Capacitance	C	1.0	0.8	$\mu\text{F cm}^{-2}$
Time-constant	τ	5.6	1.9	ms
Reversal potentials	$E_{\text{Na}}, E_{\text{K}}$	+48, -95	+55, -92	mV
K^+ conductance	g_{K}	26.0	26.0	mS cm^{-2}
Noise-scale (current)	σ_I	1.0	0.1	$\mu\text{A cm}^{-2} (\text{ms})^{1/2}$
Noise-scale (recovery)	σ_R	1.0	0.1	$(\text{ms})^{1/2}$
Threshold current	I_{dc}^{crit}	~ 21.4752886	~ 7.77327142	$\mu\text{A cm}^{-2}$
Na^+ conductance, $g_{\text{Na}}(V) = a_2 V^2 + a_1 V + a_0$				
	a_2	33.80×10^{-4}	32.63×10^{-4}	$\text{mS cm}^{-2} \text{mV}^{-2}$
	a_1	47.58×10^{-2}	47.71×10^{-2}	$\text{mS cm}^{-2} \text{mV}^{-1}$
	a_0	17.81	17.81	mS cm^{-2}
Recovery steady-state, $R_{\infty}(V) = b_2 V^2 + b_1 V + b_0$				
	b_2	3.30	0	mV^{-2}
	b_1	3.798	1.35	mV^{-1}
	b_0	1.26652	1.03	–

some essential biophysics²—making the model much more amenable to mathematical analysis and insight.

The additive noises appearing on the right of Eqs (1.1) and (1.2) are two independent time-series of white-noise perturbations that are supposed to represent the continuous random buffeting of the soma and recovery processes within a living, biological neuron. The noises are defined as,

$$I_{\text{noise}}(t) = \sigma_I \xi_I(t), \quad R_{\text{noise}}(t) = \sigma_R \xi_R(t), \quad (1.4)$$

where σ_I , σ_R are the rms noise scale-factors for current and recovery respectively, and ξ_I , ξ_R are zero-mean, Gaussian-distributed delta-correlated white-noise sources with statistics,

$$\langle \xi(t) \rangle = 0, \quad \langle \xi_i(t) \xi_j(t') \rangle = \delta_{ij} \delta(t - t'). \quad (1.5)$$

Here, δ_{ij} is the dimensionless Kronecker delta and $\delta(t)$ is the Dirac delta function carrying dimensions of inverse time. The $\xi(t)$ are approximated in simulation by the construction

$$\xi(t) = \mathcal{N}(0, 1) / \sqrt{\Delta t}, \quad (1.6)$$

² Notably: ionic reversal potentials, and the implicit Ohm's-law dependence of ionic current on the signed displacement of the membrane voltage from the reversal values.

where Δt is the size of the time-step, and $\mathcal{N}(0, 1)$ denotes a normally-distributed random-number sequence of mean zero, variance unity. In the numerical experiments described below, the noise amplitudes are set at a sufficiently small value to ensure that the neuron is allowed to explore its near steady-state subthreshold (i.e., non-firing) state space; in this regime we will find that, as firing threshold is approached from below, the subthreshold fluctuations become *critically slowed*, exactly as predicted by small-noise linear stochastic theory.

1.2.2 Type-I and type-II subthreshold fluctuations

Excitable membranes are classified according to the nature of their spiking onset. For the squid axon and for auditory nerve cells, action potential oscillations emerge at a non-zero frequency when an injected dc stimulus current exceeds threshold; such membranes are classified as being type-II or *resonator* [12]. In contrast, for type-I or *integrator* membranes (e.g., human cortical neurons), spike oscillations emerge at zero frequency as the current stimulus crosses threshold—that is, the firing frequency in a type-I neuron can be arbitrarily slow. By altering the voltage dependence of R_∞ (the steady-state value for the recovery variable in Eq. (1.2)) from linear to quadratic, the H.R. Wilson model neuron can be transformed from a resonator into an integrator (see Table 1.1 for polynomial coefficients).

Figure 1.1 compares the near-threshold behavior of the Wilson resonator neuron (Fig. 1.1(a)) with that of the integrator neuron (Fig. 1.1(b)) for white-noise perturbations superimposed on five different levels of constant stimulus current I_{dc} . For the squid-axon type-II resonator, the voltage fluctuations show an increasing tendency to “ring” at a characteristic frequency, with the ringing events becoming more

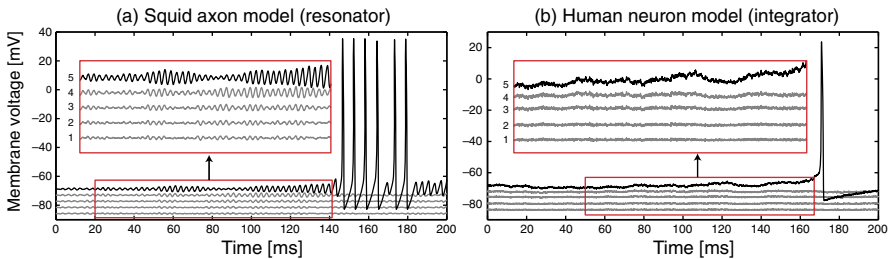


Fig. 1.1 Stochastic simulations for the H.R. Wilson models for (a) squid axon (type-II) and (b) human cortical (type-I) neuron. Framed insets show detail of the subthreshold voltage fluctuations prior to spike onset. (a) Numbered from bottom to top, the five squid stimulation currents are $I_{dc} = 0, 2, 4, 6, 7.7 \mu\text{A}/\text{cm}^2$. (To improve visibility, the squid traces have been displaced vertically by $(4m - 20)$ mV where $m = 1 \dots 5$ is the curve number.) (b) Cortical neuron stimulation currents are (bottom to top) $I_{dc} = -100, -40, 0, +16, +21.4752 \mu\text{A}/\text{cm}^2$. Integration algorithm is semi-implicit Euler-trapezium with timestep $\Delta t = 0.005$ ms. All runs within a given figure used the same sequence of 40 000 Gaussian-distributed random-number pairs to generate the white-noise perturbations. (Reproduced from [22].)

prolonged and pronounced as the critical level of drive current $I_{dc}^{\text{crit}} \approx 7.7732 \mu\text{A}/\text{cm}^2$ is approached from below.

In contrast, the mammalian type-I integrator shows voltage fluctuations that become simultaneously larger and slower as the drive current approaches the critical value $I_{dc}^{\text{crit}} \approx 21.4752 \mu\text{A}/\text{cm}^2$. One is reminded of Carmichael's eloquent description of a state change in quantum optics in which the process [1],

...amplifies the initial fluctuations up to the macroscopic scale, making it impossible to disentangle a mean motion from the fluctuations.

Prior to spike onset, is the slowly varying trend a fluctuation about the mean, or the mean motion itself? At the critical point leading to the birth of an action potential in an integrator neuron, the mean motion *is* the fluctuation.

In order to better appreciate the underlying statistical trends in fluctuation variability as the critical stimulus current is approached, we repeat the 200-ms numerical simulations of the stochastic Wilson equations (1.1)–(1.2) a total of 2000 times, each run using a different constant value of I_{dc} . These I_{dc} stimulus values, in $\mu\text{A}/\text{cm}^2$, are evenly distributed over the range -10 to $+7.77$ for the resonator experiments (see Fig. 1.2(a)), and -10 to $+21.475$ for the integrator experiments (Fig. 1.3(a)). Despite the fact that the variances (σ_I^2, σ_R^2) of the white-noise perturbations remained unchanged throughout these series of experiments, it is very clear that—for both classes of excitable membrane—the variance of the resulting fluctuations *increases strongly and nonlinearly* as the critical value of dc control current is approached, confirming Freeman's earlier observation of growing non-proportionality of response for a neuron near spiking threshold.

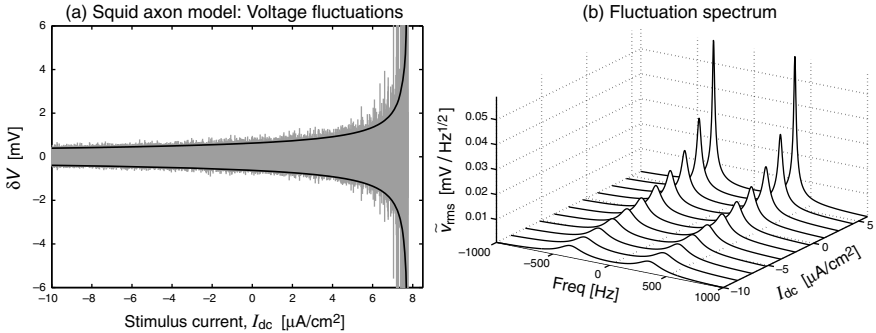


Fig. 1.2 H.R. Wilson type-II (resonator) response to white-noise perturbation as a function of subthreshold stimulus current I_{dc} . (a) Each vertical gray stripe shows maximum voltage excursions recorded in a 200-ms stochastic simulation of Eqs (1.1), (1.2) at each of 2000 settings for stimulus current ranging from -10.0 to $+7.77 \mu\text{A}/\text{cm}^2$. Solid black curves show theoretical $\pm 3\sigma$ limits for voltage excursions away from equilibrium. (b) Theoretical spectral response to white-noise driving for the squid-axon model. The double-sided spectrum develops a pronounced and increasingly narrow resonance at $\sim \pm 360$ Hz as the critical current is approached from below. (Reproduced from [22].)

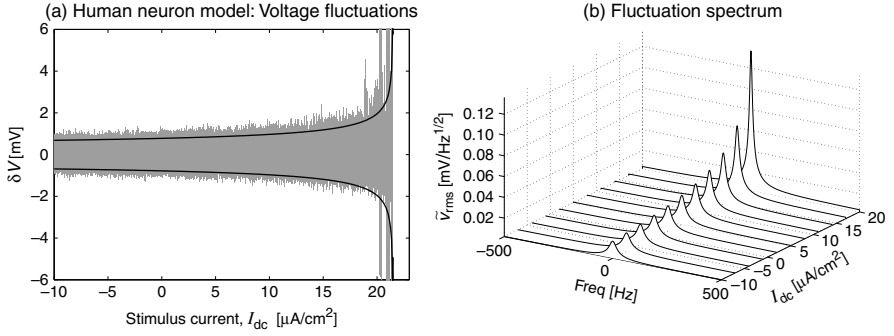


Fig. 1.3 H.R. Wilson type-I (integrator) response to white-noise perturbation as a function of subthreshold stimulus current I_{dc} . (a) Caption as for Fig. 1.2(a), but here stimulus current ranges from -10.0 to $+21.475 \mu\text{A}/\text{cm}^2$. Black curves are $\pm 3\sigma$ predictions; gray background verticals indicate fluctuation extrema recorded from 2000 independent numerical experiments. (b) Theoretical spectrum for subthreshold cortical neuron shows a strong resonance developing at *zero* frequency as threshold current for spiking is approached from below. (Reproduced from [22].)

1.2.3 Theoretical fluctuation statistics for approach to criticality

Provided the white-noise perturbations are kept sufficiently small, it is possible to compute *exact* expressions for the variance, power spectrum, and correlation function of the voltage and recovery-variable fluctuations. By “sufficiently small”, we mean that the neuron remains subthreshold (i.e., does not generate an action potential spike), so can be accurately described using linear Ornstein–Uhlenbeck (Brownian motion) stochastic theory.

The analysis was detailed in Ref. [22], but in outline, proceeds as follows. For a given (subthreshold) value of stimulus current I_{dc} , compute the steady-state coordinate (V^0, R^0) . For the H.R. Wilson resonator, V^0 is a monotonic increasing function of I_{dc} (see Fig. 1.4(a)), whereas for the Wilson integrator the graph of V^0 vs I_{dc} maps out an S-shaped curve (Fig. 1.4(b)), so there can be up to three steady states for a given value of stimulus current [30]—in which case, select the steady state with the lowest voltage.

We rewrite the Wilson equations (1.1), (1.2) in their deterministic (noise-free) form,

$$F_1(V, R) \equiv (I_{Na} + I_K + I_{dc})/C, \quad (1.7)$$

$$F_2(V, R) \equiv (-R(t) + R_\infty(V))/\tau, \quad (1.8)$$

and linearize these by expressing the fluctuations (v, r) as small deviations away from steady state (V^0, R^0) ,

$$v(t) = V(t) - V^0, \quad r(t) = R(t) - R^0, \quad (1.9)$$

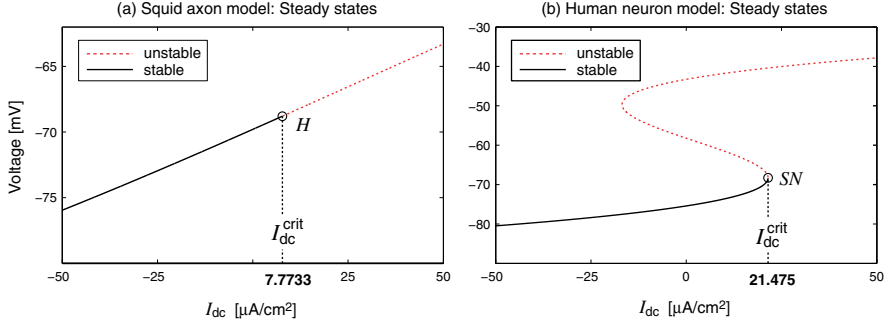


Fig. 1.4 Distribution of steady-state membrane voltages as a function of dc stimulus current I_{dc} for (a) Wilson type-II resonator model; and (b) Wilson type-I integrator neuron. The critical current I_{dc}^{crit} is determined by the point at which the (real part of the) dominant eigenvalue becomes positive, heralding emergence of instability (generation of action potentials). Transition occurs (a) via subcritical Hopf bifurcation at H for the resonator [30]; and (b) via saddle–node annihilation at SN for the integrator. (Modified from [22].)

Reinstating the additive noise terms, the linearized Wilson equations become coupled Brownian motions of the form,

$$\frac{d}{dt} \begin{bmatrix} v \\ r \end{bmatrix} = \mathbf{J} \begin{bmatrix} v \\ r \end{bmatrix} + \sqrt{\mathbf{D}} \begin{bmatrix} \xi_I \\ \xi_R \end{bmatrix}, \quad (1.10)$$

with Jacobian and diffusion matrices defined respectively by,

$$\mathbf{J} = \left[\begin{array}{cc} \frac{\partial F_1}{\partial V} & \frac{\partial F_1}{\partial R} \\ \frac{\partial F_2}{\partial V} & \frac{\partial F_2}{\partial R} \end{array} \right] \Bigg|_{(V^o, R^o)}, \quad \mathbf{D} = \begin{bmatrix} \left(\frac{\sigma_I}{C}\right)^2 & 0 \\ 0 & \left(\frac{\sigma_R}{\tau}\right)^2 \end{bmatrix}, \quad (1.11)$$

where \mathbf{J} is evaluated at the selected equilibrium point.

In the vicinity of an equilibrium point, the deterministic behavior of the two-variable Ornstein–Uhlenbeck system is completely defined by the two eigenvalues, λ_1 and λ_2 , belonging to the Jacobian matrix.

For the subthreshold Wilson *resonator*, the eigenvalues are complex, $\lambda_{1,2} = -\alpha \pm i\omega_0$, with the damping $\alpha = -\text{Re}(\lambda)$ being *positive* for a decaying impulse response and a stable equilibrium. If the damping becomes negative (i.e., $\text{Re}(\lambda) > 0$), a minor disturbance will grow exponentially, signaling onset of nonlinear *super-threshold* behavior (generation of a spike). But if the drive current matches the critical value I_{dc}^{crit} *exactly*, the damping will be precisely zero, thus a small disturbance will provoke a resonant response at frequency ω_0 whose oscillations will neither decay nor grow over time, but will persist “forever”.

For the Wilson *integrator* neuron, both eigenvalues are purely real, with $\lambda_2 < \lambda_1 < 0$ for a stable equilibrium. Exponential growth leading to spike onset is predicted if the dominant eigenvalue λ_1 becomes positive. At the critical current for the integrator (lower-right turning point in Fig. 1.4(b) labeled SN), the unstable

mid-branch saddle equilibrium meets the stable lower-branch node at a saddle–node bifurcation. At this bifurcation point the dominant eigenvalue is precisely zero, leading to a delicate point of balance in which small perturbations are sustained indefinitely, neither decaying back to steady state nor growing inexorably into nonlinearity and thence to a spike. At this point, the neuron response will become *critically slowed*.

1.2.3.1 Fluctuation variance

Following Gardiner’s analysis of the multivariate Ornstein–Uhlenbeck process [8], we can write theoretical expressions of the white-noise evoked fluctuation variance and spectrum, and deduce scaling laws for the divergences that manifest at the critical point.

The steady-state variances of the fluctuations developed in the Wilson excitable membrane depend explicitly on the elements of the diffusion matrix and the Jacobian matrix, and on the Jacobian eigenvalues. For the H.R. Wilson type-I integrator, the variance of the voltage fluctuations reads [22],

$$\text{var}\{v\} = \frac{(\lambda_1\lambda_2 + J_{22}^2)D_{11} + J_{12}^2D_{22}}{-2(\lambda_1 + \lambda_2)\lambda_1\lambda_2} \xrightarrow{\lambda_1 \uparrow 0^-} \sim \frac{1}{-\lambda_1} \sim \frac{1}{\sqrt{\varepsilon}}. \quad (1.12)$$

Here, λ_1 is the dominant (i.e., least negative) eigenvalue, and both eigenvalues are real. As the dc stimulus current approaches its critical value, λ_1 approaches zero from below. Thus, at the threshold for spiking, the integrator neuron becomes infinitely responsive to white-noise perturbation with the fluctuation power diverging to infinity. The scaling for this divergence follows an $\varepsilon^{-1/2}$ power-law, where $\varepsilon = (I_{\text{dc}}^{\text{crit}} - I_{\text{dc}})/I_{\text{dc}}^{\text{crit}}$ is a dimensionless measure of distance from criticality. This is the case because in the vicinity of the saddle–node bifurcation point, the dominant eigenvalue scales as $\sqrt{\varepsilon}$ in a locally parabolic relationship. Since the inverse of the dominant eigenvalue defines the dominant time-scale T for system response, it follows that the characteristic times (correlation time, passage time) will obey the same scaling law: $T \sim \varepsilon^{-1/2}$. We note that this inverse square-root scaling law is a very general feature of systems that are close to a saddle–node bifurcation [26].

For the case of the Wilson type-II resonator, the eigenvalues form a complex conjugate pair, $\lambda_{1,2} = -\alpha \pm i\omega_0$, so the expression for voltage variance becomes [22],

$$\text{var}\{v\} = \frac{(\alpha^2 + \omega_0^2 + J_{22}^2)D_{11} + J_{12}^2D_{22}}{4\alpha(\alpha^2 + \omega_0^2)} \xrightarrow{\alpha \downarrow 0^+} \sim \frac{1}{\alpha} \sim \frac{1}{\varepsilon}. \quad (1.13)$$

As the critical point is approached, the damping $\alpha = -\text{Re}(\lambda)$ goes to zero from above, leading to a prediction of a divergent power increase that scales as ε^{-1} (for the Wilson resonator close to threshold, α scales *linearly* with ε).

Equations (1.12) and (1.13) were used to compute the theoretical $\pm 3\sigma$ voltage fluctuation limits plotted in Figs 1.3(a) and 1.2(a) respectively; we note excellent agreement between simulation (gray stripes) and small-noise linear theory (black curves).

1.2.3.2 Fluctuation spectrum

The stationary spectrum for the membrane-voltage fluctuations in the stochastic H.R. Wilson neuron is given by the S_{11} entry of the 2×2 spectrum matrix of the two-variable Ornstein–Uhlenbeck process [8, 22]. For the Wilson integrator neuron,

$$S_{11}(\omega) = \frac{1}{2\pi} \frac{J_{22}^2 D_{11} + J_{12}^2 D_{22} + D_{11} \omega^2}{(\lambda_1 \lambda_2 - \omega^2)^2 + (\lambda_1 + \lambda_2)^2 \omega^2} \xrightarrow[\omega \rightarrow 0]{\lambda_1 = 0} \sim \frac{1}{\omega^2}. \quad (1.14)$$

The spectral character of the fluctuations changes as the I_{dc} stimulus current increases towards the critical value I_{dc}^{crit} , and the corresponding lower-branch steady state moves closer to the saddle–node critical point (marked SN in Fig. 1.4(b)): the dominant eigenvalue λ_1 tends to zero from below, causing the power spectral density to diverge at zero frequency, obeying an asymptotic power-law $\sim 1/\omega^2$. Thus, at the critical saddle–node annihilation point, the Wilson integrator is predicted to become “resonant at dc”. This spectral tuning of fluctuation energy towards zero frequency is illustrated in the plots of Eq. (1.14) graphed in Fig. 1.3(b).

The noise-driven time-series for the squid-axon model illustrated in Fig. 1.1(a) shows a strongly increasing tendency for the voltage trace to “ring” at a characteristic frequency as the drive current is increased towards the threshold for spiking. This ringing behavior is precisely consistent with the spectrum predicted from Ornstein–Uhlenbeck theory for the Wilson resonator neuron [22],

$$S_{11}(\omega) = \frac{1}{2\pi} \frac{J_{22}^2 D_{11} + J_{12}^2 D_{22} + D_{11} \omega^2}{(\alpha^2 + \omega_0^2 - \omega^2)^2 + 4\alpha^2 \omega^2} \xrightarrow[\omega \rightarrow \omega_0]{\alpha = 0} \sim \frac{1}{(\omega - \omega_0)^2}, \quad (1.15)$$

implying perfect resonant behavior at frequency $\omega = \omega_0$, with the approach to resonance following an asymptotic scaling-law $\sim 1/\delta^2$ where $\delta = (\omega - \omega_0)$ is the spectral distance from resonance. The resonator spectrum of Eq. (1.15) is plotted in Fig. 1.2(b).

We now move from consideration of single neurons to the gross behaviors of *large populations* of neurons. Just as a single neuron displays telltale nonlinear increases in responsiveness as it approaches the transition point separating stochastic quiescence from dynamic spiking, we find that the collective behaviors of cooperating neuron populations also exhibit characteristic critical responses as the neural population approaches a change of state. We consider three gross changes of cortical state that are easily detected with a single pair of EEG electrodes: induction of anesthesia; natural sleep cycling from slow-wave sleep into REM sleep; and the nightly transition between wake and sleep.

1.3 The anesthesia state

The ability to render a patient safely and reversibly unconscious via administration of an anesthetic drug is an essential component of modern surgical medicine. Although anesthetic agents have been in use for over 160 years, their mode of action remains poorly understood, and is the focus of ongoing and intensive research.

The state of *general anesthesia* is a controlled and reversible unconsciousness characterized by a lack of pain sensation (analgesia), lack of memory (amnesia), muscle relaxation, and depressed reflex responses. In his classic 1937 textbook for anesthetists [9], Arthur Guedel identified four distinct stages in the induction of general anesthesia using the volatile agent diethyl ether:

1. *Analgesia and amnesia* Patient experiences pain relief and dreamy disorientation, but remains conscious.
2. *Delirium* Patient has lost consciousness, blood pressure rises, breathing can become irregular, pupils dilate. Sometimes there is breath-holding, swallowing, uncontrolled violent movement, vomiting, and uninhibited response to stimuli.
3. *Surgical anesthesia* Return of regular breathing, relaxation of skeletal muscles, eye movements slow, then stop. This is the level at which surgery is safe.
4. *Respiratory paralysis* Anesthetic crisis—respiratory and other vital control centres cease to function, death from circulatory collapse will follow without assisted ventilation and circulatory support.

One might anticipate a roughly linear dose–response in which increasing drug concentration leads to proportionate reductions in brain activity—however, this simple intuition is immediately contradicted by the anomalous patient response reported by Guedel at the stage-2 (delirium) level of anesthesia. A general anesthetic is administered with the aim of quieting or *inhibiting* brain response to noxious stimuli, and yet, on route to the stage-3 fully-inhibited state, the patient transits through a “wild” uncontrolled state of delirium and *uninhibited* response to stimuli. This is a most interesting paradox: the end-state of inhibition is preceded by an intermediate stage of excitation.

1.3.1 Effect of anesthetics on bioluminescence

In the 1970s, researchers reported that the volatile anesthetics ether, halothane, chloroform, and cyclopropane all reversibly reduce the intensity of light emissions from luminescent bacteria [10, 28]. This followed earlier work by Ueda [27] showing that the light emission from the firefly lantern-extract luciferase was reversibly suppressed by both ether and halothane. The anesthetic concentration required to depress bioluminescent intensity by 50% was found to be very similar to the concentrations required for clinical induction in humans. Because of this remarkable scale-invariance (i.e., the light-emitting complex in photo-bacteria and in fireflies,

and the central nervous system in humans, are responsive to similar concentrations of a given anesthetic), and because light intensity can be easily and accurately measured, bioluminescence provided a useful early means for quantifying and comparing anesthetic potency.

Figure 1.5 shows the bioluminescence dose–response for ether reported by Halsey and Smith [10]. At partial pressure $P = 0.026$ atm, the luminous intensity has reduced to 50% of its original value. This partial pressure is similar to the 0.032 atm value quoted in the paper for the abolition of the righting instinct in 50% of mice exposed to ether.³ Of particular interest is their observation that luminescence is *stimulated* by low doses of ether ($P \sim 0.009$ atm), confirming an earlier report by White and Dundas [28]. Halsey and Smith [10] stated that stimulation also occurred at low levels of chloroform, halothane, and nitrous oxide (though for the latter two agents the increase was “not statistically significant”, presumably because the uncertainty bars became very large during this transition phase).

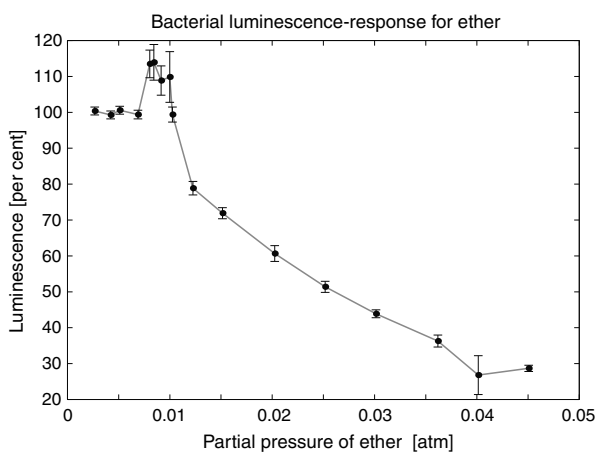


Fig. 1.5 Dose–response curve showing the effect of the volatile anesthetic ether on the luminous intensity of the bacteria *Photobacterium phosphoreum*. Note the anomalous surge, and increased variability, in light output at low ether concentration. (Graph reconstructed from [10].)

Although neither research group offered an explanation for this paradoxical excitation by an inhibitory agent, it seems rather likely that the dilute-ether boost in luminous intensity and variability seen in bacteria could be mapped directly to Guedel’s delirium (stage-2) for ether-induced anesthesia in human patients—though it might be difficult to test this idea quantitatively now, since diethyl-ether is no longer used as an anesthetic agent in hospitals.

³ Prior to the bioluminescence studies, small mammals had been used to calibrate anesthetic potency.

1.3.2 Effect of propofol anesthetic on EEG

Unlike diethyl-ether and the other volatile anesthetic agents (such as those tested in the bioluminescence experiments) that are delivered to the patient by inhalational mask, propofol, a modern and commonly-used general anesthetic, is injected intravenously as a liquid emulsion, so is likely to have a different mode of action. Despite this difference, the onset of propofol anesthesia is also heralded by a *surge* in brain activity that is readily detected as a sudden increase in low-frequency EEG power [14, 15]; this excitation subsides as the patient moves deeper into unconsciousness. Thus propofol, like ether, is “biphasic”, being excitatory at low concentrations, then inhibitory at higher concentrations.

The measurements of Kuizenga *et al.* [15] shown in Fig. 1.6(a) show that a *second* surge in activity occurs as the propofol concentration dissipates, allowing the patient to re-emerge into consciousness. Thus there are *two* biphasic peaks per induction–emergence cycle: the first at or near loss-of-consciousness (LOC), and the second at recovery-of-consciousness (ROC). The onset of the first EEG surge lags ~ 2 min behind the rise in propofol concentration; this delay arises because of the unavoidable mismatch between the site of drug *effect* (the brain) and the site of drug *measurement* (the blood)—it takes about 2 min for the drug to diffuse across the blood–brain barrier. Even after compensating for this delay, there seems to remain a hysteretic separation between that the LOC and ROC biphasic peaks, meaning that the patient awakens at a *lower* drug concentration than that required to put the patient to sleep.

At the individual neuron level, the major effect of propofol is to prolong the duration of inhibitory postsynaptic potential (IPSP) events, thereby increasing the inward flux of chloride ions and thus increasing the hyperpolarizing effectiveness of inhibitory firings by GABAergic interneurons [4, 13].

We developed a model for propofol anesthesia by modifying a set of cortical equations by Liley [16] to include a control parameter λ that lengthens the IPSP decay-constant (by reducing the IPSP rate-constant γ_i) in proportion to drug concentration: $\gamma_i^{-1} \rightarrow \lambda \gamma_i^{-1}$; see Refs [23–25] for full details. For a physiologically plausible set of cortical parameters, we found that, for a given value of anesthetic effect λ , the model cortex could support up to *three* homogeneous steady-states; see Fig. 1.6(b). The upper (active) and lower (quiescent) stable nodes are separated by a saddle-branch that is unstable to small perturbations, suggesting the possibility of a propofol-mediated phase transition between the active (conscious) and quiescent (unconscious) states. A transition from active branch A_1 - A_2 - A_3 to quiescent branch Q_1 - Q_2 - Q_3 becomes increasingly likely as the node-saddle annihilation point A_3 is approached from the left. The abrupt downward transition represents induction of anesthesia (i.e., LOC).

Once unconscious, reductions in λ allow the cortex to move to the left along the bottom branch of Fig. 1.6(b), with the probability of an upward transition (i.e., ROC) rising as the quiescent node–saddle point Q_1 is approached. Thus the model provides a natural explanation for the observed LOC–ROC drug hysteresis.

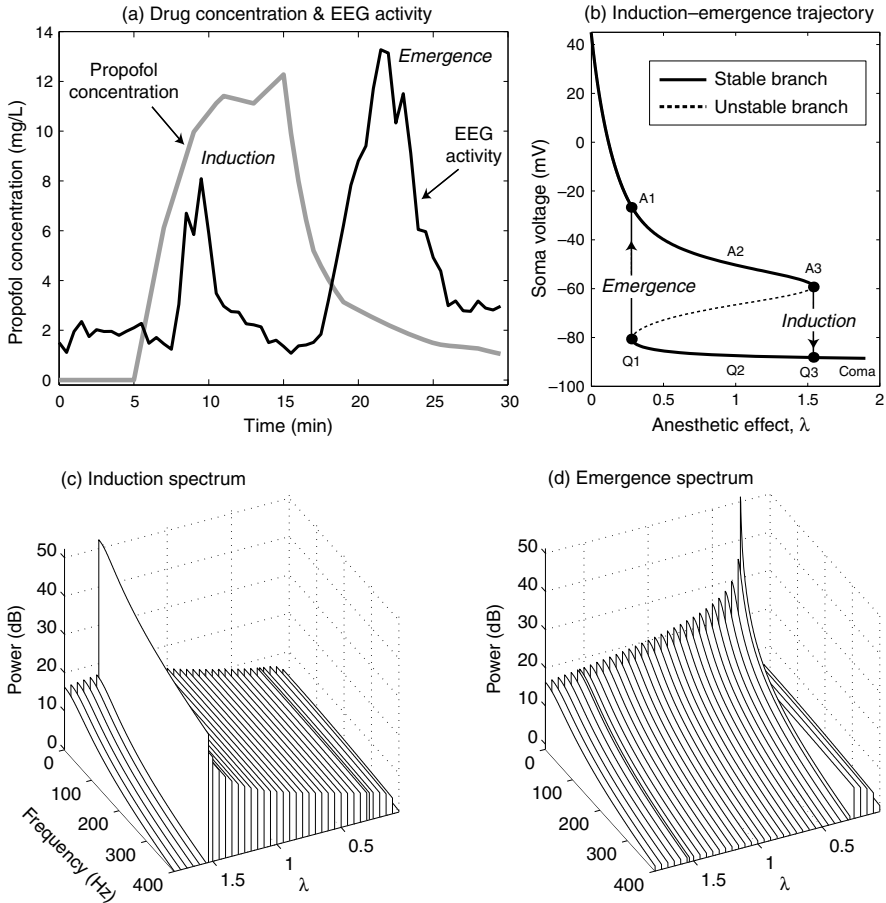


Fig. 1.6 Measured (a) and modeled (b–d) effect of propofol anesthetic agent on EEG activity during induction of, and emergence from, general anesthesia. (a) Time-series showing biphasic surges in 11–15-Hz EEG activity (black curve) in response to increasing and decreasing levels of propofol blood concentration (gray curve) for a patient undergoing a full induction–emergence cycle. [Data provided courtesy of K. Kuizenga.] (b) Trajectory of steady states predicted by a cortical model that assumes propofol acts to prolong IPSP duration by factor λ . Approaches to the saddle-node points A3 (for induction), and Q1 (for emergence), are predicted to show pronounced EEG power surges displayed in (c) and (d) respectively. (Modified from Figs. 3–5 of [23].)

Proximity to either of the node–saddle turning points can be detected by the divergent sensitivity of the cortical model to small disturbances. This increasing susceptibility or “irritability” can be quantified by driving the model with low-level white noise, simulating the biological reality of a continuous background wash of unstructured, nonspecific stimulus entering the cortex from the subcortex. Provided the intensity of the white-noise stimulus is sufficiently small, we can compute exact expressions for the stationary spectrum and correlation properties of the noise-induced fluctuations by applying stochastic theory [8] to the distribution of

eigenvalues obtained from linear stability analysis [23]. The predicted alterations in spectral densities for EEG fluctuation power during anesthetic induction, and during emergence from anesthesia, are plotted in Fig. 1.6(c) and (d) respectively. In both cases, the fluctuation power at zero frequency surges as the node–saddle critical point is approached, providing advance warning of an impending jump in membrane voltage.

According to this model, we can interpret Kuizenga’s observations—of hysteretically separated biphasic surges in EEG activity—as biological evidence supporting the notion that the cortical states of awareness and anesthesia are distinct “phases” of the brain. One could argue that the drug-induced transition into unconsciousness has similarities with physical phase transitions, such as water freezing to ice, with the effect of increasing drug concentration in the brain being analogous to lowering the temperature in the thermodynamic system [25].

1.4 SWS–REM sleep transition

Monitoring the EEG activity of the sleeping human shows natural sleep to consist of two opposed phases: *quiet* slow-wave sleep (SWS) and *active* rapid-eye-movement (REM) sleep. During quiet sleep, the EEG voltage fluctuations are larger, slower—and more coherent across the scalp—than those observed during alert wakefulness. In contrast, during active sleep, the EEG closely resembles wake with its high-frequency, low-amplitude desynchronized patterns. A sleeping adult human cycles between SWS and REM-sleep states at approximately 90-min intervals, for a total of four to six SWS–REM alternations per night.

Figure 1.7 illustrates the cyclic nature of the adult sleep patterns we reported in Ref. [21]. We see four slow surges in EEG power during the six-hour recording, with each surge being terminated by an abrupt decline, signaling the transit from SWS to REM sleep. The increase in fluctuation power is matched by an increase in correlation time⁴ that peaks at the end of each SWS episode, with abruptly lower values in REM sleep. This is consistent with the antiphased changes in low- and high-frequency power fractions of Fig. 1.7(c): SWS is associated with increasing low-frequency activity; REM sleep is associated with diminished low-frequency and enhanced high-frequency EEG fluctuations.

The Fig. 1.8 analysis by Destexhe *et al.* [2] for the sleeping cat shows similar patterns of SWS–REM alternation, albeit with a faster cycling time of ~ 20 min. As was the case for the human sleeper, Fig. 1.8 shows that the sleeping cat exhibits a pronounced increase in low-frequency power prior to transition from SWS into REM. The concomitant increase in “space constant” (correlation *length* for EEG fluctuations) observed for the cat is consistent with an increase in correlation *time* we reported for the human sleeper.

⁴ Correlation time T is the time-lag required for the autocorrelation function for EEG voltage to decay to $1/e$ of its zero-lag peak.

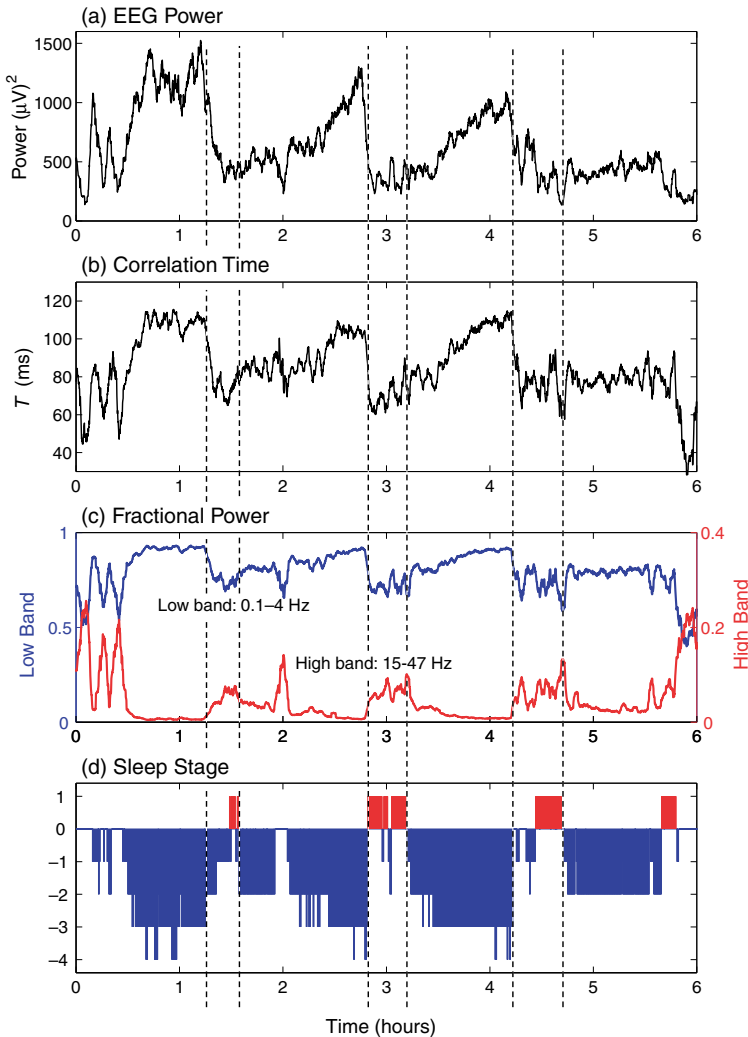


Fig. 1.7 [Color plate] Analysis of an EEG trace recorded from a human sleeper resting overnight in a sleep laboratory. (a) Fluctuation power; (b) correlation time; (c) low- and high-band power fractions; (d) sleep staging as per rulebook of Rechtschaffen and Kale [20]. Key: +1 = REM; 0 = wake; (-1, -2) = light sleep; (-3, -4) = deep sleep (SWS). (Graph reprinted from [21].)

Very similar changes are seen in the ECoG (electrocorticogram) brain activity for a mature fetal sheep. Figure 1.9 shows a 500-s voltage trace, recorded from the cortex of a late-term fetal sheep, that captures the transition between the so-called “high-voltage slow” (i.e., SWS) and “low-voltage fast” (REM sleep) states. As is the case for the human and cat sleepers, the slow-wave state is characterized low-frequency correlated fluctuations that increase in intensity and low-frequency content as the point of transition to active sleep is approached.

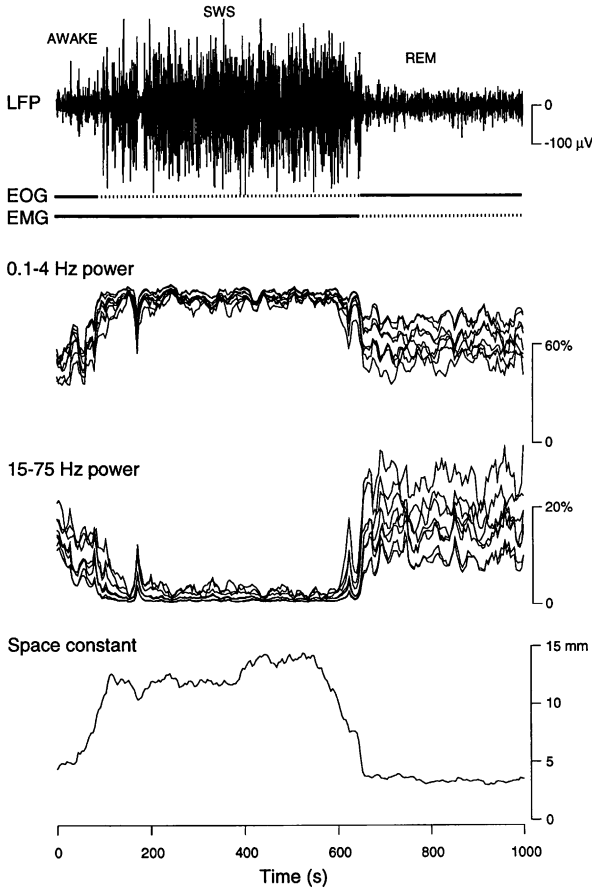


Fig. 1.8 Cortical activity for a cat transitioning from wake to SWS to REM sleep as reported by Destexhe *et al.* [2]. LFP = local field potential (on-cortex EEG); EOG = electrooculogram (eye movement); EMG = electromyogram (muscle tone). REM is identified by reappearance of eye movements (EOG activity) and lack of muscle tone (loss of EMG activity). (Graph reprinted from [2] with permission.)

1.4.1 Modeling the SWS–REM sleep transition

In Ref. [21] we described the construction of a physiologically-based model for the SWS–REM sleep transition that incorporated the two major neuromodulatory influences that are thought to be responsible for the cycles of natural sleep: (a) slow changes in synaptic efficiency ρ_e and resting voltage V_e^{rest} of the population of excitatory neurons caused by the 90-min cycling in acetylcholine (ACh) concentration; and (b) slower changes in resting voltage caused by the gradual elimination during sleep of fatigue agents such as adenosine. The full set of cortical equations are

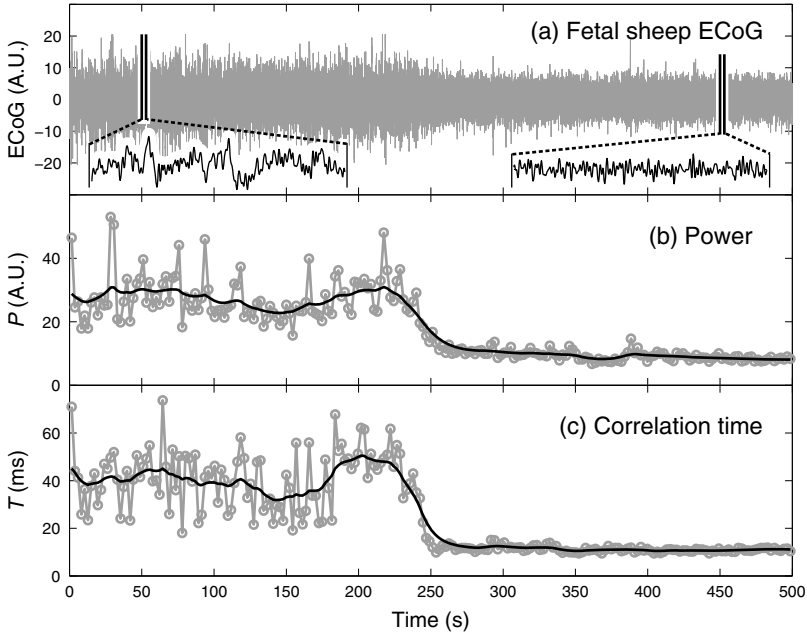


Fig. 1.9 Cortical activity for a full-term fetal sheep (gestational age = 144 days) transitioning from slow sleep (SWS) to active sleep (REM) about 230 s into the recording. (a) ECoG voltage signal (arbitrary units) sampled at 250 s^{-1} . Inserts show 3 s of ECoG signal for the intervals 50–53 s (left), and 450–453 s (right). (b) ECoG power (arbitrary units) computed for 3-s epochs with 25% overlap (gray trace), and smoothed with a Whittaker filter [3] (black trace). (c) ECoG correlation time computed for 3-s epochs with 25% overlap (gray trace), and smoothed with a Whittaker filter (black trace). Note the coincident surge in power and correlation-time prior to transition into REM. (Data provided courtesy of J.J. Wright; analysis by Yanyang Xu.)

described in the chapters by Sleight *et al* and Wilson *et al*.⁵ The model consists of eight differential equations for macrocolumn-averaged soma potentials and synaptic fluxes. Here, we simplify the equations considerably by taking the “slow soma” adiabatic limit in which, relative to the ~ 50 -ms time-scale of the neuron soma, synaptic input events are assumed to be fast and rapidly equilibrating. This simplification reduces the number of state variables from eight to two: V_e and V_i , the average soma potential for the excitatory and inhibitory neural populations. The acetylcholine and adenosine effects are modeled in terms of λ , a multiplicative factor applied to the ρ_e excitatory synaptic efficiency, and ΔV_e^{rest} , an additive adjustment that tends to depolarize (hyperpolarize) the excitatory membrane potential for positive (negative) values of ΔV_e^{rest} .

The λ and ΔV_e^{rest} parameters define a two-dimensional sleep domain for our cortical model. We located the homogeneous equilibrium states (V_e^0, V_i^0) as a function of variations in λ and ΔV_e^{rest} , paying particular attention to those regions of

⁵ See chapters 9 and 10 respectively in the present volume.

the domain that support multiple (up to three) steady states. When plotted in 3-D, the region of multiple steady states appears as a reentrant fold in the sleep-domain manifold of Fig. 1.10. For our reduced adiabatic cortical model, the top and bottom surfaces of the fold contain stable solutions, and only the middle surface (within the overhang outlined in green) contains unstable solutions.⁶

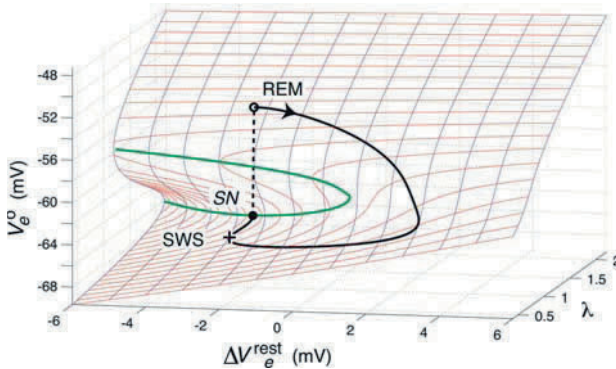


Fig. 1.10 [Color plate] Manifold of homogeneous equilibrium states for SWS–REM sleep-cycling model. Steady-state soma voltage V_e^o is plotted as a function of sleep-domain parameters ΔV_e^{rest} and synaptic efficiency λ . The imposed sleep cycle commences in SWS at (+), encounters the saddle–node critical point SN (•), and jumps vertically into REM sleep (o). (Modified from [21].)

We impose a cyclic tour of the manifold that is proposed to represent a single 90-min SWS-to-REM-to-SWS sleep cycle. This tour, commencing in the quiescent slow-wave sleep state (marked “+” in Fig. 1.10), proceeds clockwise until it encounters the saddle–node annihilation point SN at the lower overhang boundary, whereupon the soma voltage spontaneously makes an upwards jump transition to the activated upper state that we identify as REM sleep.

To visualize the dynamic repertoire available to the sleep-cycling cortex, we perform a numerical simulation of the reduced cortical equations. Voltage fluctuations in soma potential are induced via small-amplitude white-noise stimulations entering the model via the subcortical terms (see [21] for details). The noise-stimulated voltage fluctuations have an amplitude and spectral character that are strongly dependent on the cortical steady-state coordinate. In Fig. 1.11 we have started the numerical simulation very close to the saddle–node critical point SN on the bottom branch of Fig. 1.10. Proximity to the critical point causes the fluctuations to be large and slow; after about 2 s, the fluctuation carries the cortex beyond the basin of attraction of the bottom-branch equilibrium point, and the cortex is promptly drawn to the upper state. Fig. 1.11 shows an abrupt loss of low-frequency activity once the model cortex has transitioned from SWS into the REM (upper) state, similar to the

⁶ Analysis of the full nonadiabatic cortical model shows that, for particular choices of synaptic parameters, the regions of instability can extend beyond the overhang, leading to Hopf and wave instabilities. See chapters 9 (Sleigh *et al*) and 10 (Wilson *et al*) for details.

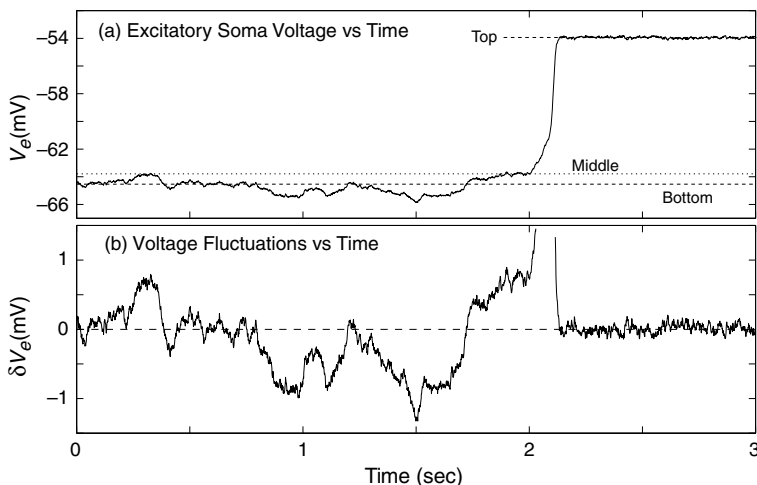


Fig. 1.11 Stochastic simulation of the slow-soma cortical model for SWS-REM sleep cycling. The cortical model is started close to the saddle-node critical point SN on the bottom branch of Fig. 1.10. (a) Soma voltage V_e , and (b) noise-induced fluctuations δV_e versus time. Fluctuations in (b) are measured relative to the bottom (top) steady state for times $t < 2.1$ ($t > 2.1$) s. (Modified from [21].)

spectral changes observed in the EEG for human, cat, and fetal-sheep sleep records (Figs 1.7–1.9 respectively).

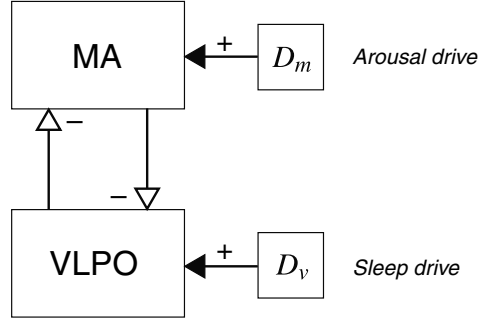
1.5 The hypnic jerk and the wake-sleep transition

The daily cycling of brain state between wake and sleep is a natural phase transition that is synchronized by the diurnal light cycle and regulated by waxing and waning concentrations of neuromodulators such as acetylcholine (ACh) and adenosine. Fulcher, Phillips, and Robinson (FPR) [7, 18] have developed a model for the wake-sleep transition⁷ that we examine briefly here, focusing on the possibility that critical slowing of noise-evoked fluctuations in the wake-sleep control center might provide a natural explanation for the puzzling but common observation of a bodily jerk at the onset of sleep.

The FPR wake-sleep model is expressed in terms of the mutual inhibition between two brainstem neural populations: the *sleep-active* ventrolateral preoptic area (VLPO), and the *wake-active* monoaminergic group (MA). The mutual competition between these populations produces bistable flip-flop behavior that causes the brain state to alternate between wake and sleep states. In the simplest form of the model (Fig. 1.12), the external ACh drive promoting arousal of the MA (and of the cortex) is replaced by a constant excitation voltage $D_m = A = \text{const.}$,

⁷ And see Chap. 8 of this volume.

Fig. 1.12 Schematic for simplified Fulcher–Phillips–Robinson model of the sleep–wake switch. Excitatory (+) and inhibitory (–) interactions are shown with solid and outline arrowheads. Mutual inhibition between VLPO and MA neuron populations results in flip-flop bistability between wake and sleep states.



and the external somnogenic and circadian sleep-promoting drives that activate the VLPO are replaced by a slowly-varying control parameter D_v . For this reduced case, the respective equations of motion for V_v and V_m , the VLPO and MA population voltages (relative to resting voltage) become,

$$\frac{dV_v}{dt} = \frac{1}{\tau} (-V_v + v_{vm}Q_m + D_v), \quad (1.16)$$

$$\frac{dV_m}{dt} = \frac{1}{\tau} (-V_m + v_{mv}Q_v + A), \quad (1.17)$$

where τ is a time-constant, v_{jk} is the coupling strength from population k to j (with $j, k = v$ or m), and Q_k is the sigmoidal mapping from soma voltage V_k to average firing rate [5],

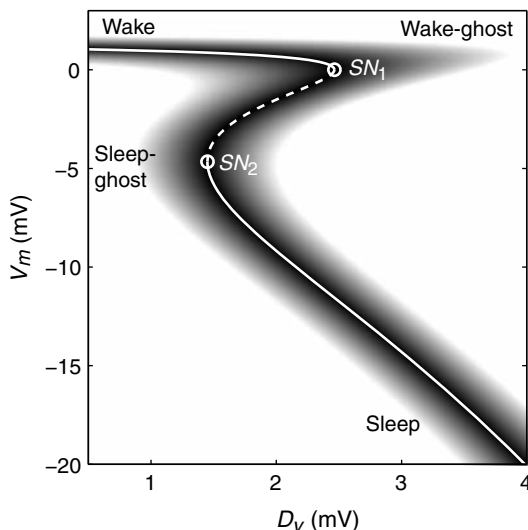
$$Q_k = S(V_k) = Q_{\max} / [1 + \exp(-\{V_k - \theta\}/\sigma')], \quad (1.18)$$

with Q_{\max} being the maximum firing rate, θ the threshold voltage (relative to rest) for firing, and σ' a measure of its spread. (Refer to Table 8.2 for parameter values.)

Setting the time-derivatives in Eqs (1.16–1.17) to zero and solving numerically for the steady states as a function of sleep drive D_v reveals a three-branch locus of equilibria (Fig. 1.13); linear stability analysis indicates that the middle branch is unstable. The top branch has higher V_m values, so is identified with the wake state, while the bottom branch, with lower V_m values, corresponds to sleep. Because the top branch terminates in a saddle–node critical point (SN_1), any noise present in the VLPO–MA flip-flop will produce exaggeratedly enlarged and slowed voltage fluctuations as the awake brain moves to the right along the top branch under the influence of increasing sleep pressure D_v . If D_v^0 is the value of sleep pressure at which the wake state loses stability (i.e., at SN_1), then Eq. (1.12) predicts that the variance of the voltage fluctuations will diverge according to the scaling law

$$\text{var}(V_m) \sim \frac{1}{\sqrt{\varepsilon}}, \quad (1.19)$$

Fig. 1.13 Locus of equilibrium states, saddle-nodes, and ghosts for FPR sleep model as a function of VLPO sleep-drive D_v . White curve shows distribution of stable (solid line) and unstable (dashed line) steady states. Saddle-node bifurcation points SN_1 , SN_2 are marked with open circles (\circ). Regions of slow dynamics are shaded from $\dot{V} = 0$ (black) to $\dot{V} > 0.05$ mV/ms (white). Saddle-node ghosts form in the “shadow zone” that projects beyond the turning points. (Figure modeled on, and modified from, Fulcher *et al* [7].)



where $\varepsilon = |D_v - D_v^0|$ measures the distance to the saddle-node bifurcation point. This theoretical prediction has been verified by numerical simulation—see Fig. 8.9(a) in the chapter by Robinson *et al* (present volume).

It is likely that the brainstem VLPO–MA wake–sleep system projects to other brain areas such as the cerebral cortex and motor cortex. At the point of falling asleep, the VLPO–MA system is close to an instability point, so is highly sensitive to stimulus, thus a sudden impulsive stimulus—either internal (e.g., a spontaneous neural firing) or external (e.g., a sudden noise)—could produce an extravagantly large response. If this disproportionate response were to be transmitted to higher brain areas such as the motor cortex, then we might expect to observe a violent whole-body twitch at or near the transition into sleep. This is a common point-of-sleep experience for many individuals [17], and is known as the *hypnagogic myoclonic twitch* or *hypnic jerk*, but until now has lacked a satisfactory explanation.

The lower-branch turning point SN_2 on Fig. 1.13 marks the position where the sleep state loses stability during the *awakening* phase as the sleep drive D_v reduces in intensity at the end of the diurnal sleep cycle. The fact that this second critical point (for emergence from sleep) occurs at a *lower* value of sleep drive than that required for transition into sleep provides a protective hysteresis that enhances the stability of both states [7, 18]: once asleep, one will tend to remain asleep, and vice versa. Further, if the flip-flop sleep bistability model is correct, we should expect a second nonlinear increase in stimulus sensitivity as the sleep-emerging brain approaches the lower-branch saddle-node critical point. Thus the model predicts that a minor stimulus presented to an almost-awake brain could evoke a disproportionately large response, causing the individual to be startled into wakefulness with a fright. This exaggerated startle response (*hyperekplexia*) can be a common experience at the end of an overnight sleep.

Following Fulcher *et al* [7], we highlight the regions of slow dynamical evolution in the Fig. 1.13 V_m -vs- D_v graph by using a grayscale representation of \dot{V} , the magnitude of the velocity field in V_v - V_m space,

$$\dot{V} = [(dV_v/dt)^2 + (dV_m/dt)^2]^{1/2} \quad (1.20)$$

where dV_v/dt and dV_m/dt are defined respectively by the VLPO and MA equations of motion Eqs (1.16, 1.17). The grayscale shading shows that the regions of slow evolution form a penumbra that brackets the reverse-S locus of fixed points. The low- \dot{V} penumbral region is particularly accentuated in the vicinity of the saddle–node turning points, defining saddle–node remnants or *ghosts* [26]. The wake-ghost (to the right of SN_1) and the sleep-ghost (to the left of SN_2) act as low-velocity bottlenecks, so any trajectory entering a ghostly region will tend to linger there, exhibiting low-frequency enhancement of noise-induced voltage fluctuations as it traverses the bottleneck. This suggests that the region of sleep-onset hypersensitivity to impulsive stimuli could persist beyond the immediate proximity of the wake-branch critical point to include the shadow zone defined by the wake-ghost.

1.6 Discussion

In this chapter we have examined several mathematical models for state transitions in single neurons and in neural populations. The change of state occurs when a control parameter (such as stimulus current, anesthetic concentration, neuromodulator concentration) crosses a critical threshold, causing the initial equilibrium state to lose stability, typically via a saddle–node or a Hopf bifurcation. We can detect proximity to criticality by adding a small white-noise perturbation to induce fluctuations in the observed variable, such as the membrane voltage, allowing the system to explore its nearby state-space. As the system approaches the instability point, the fluctuations grow in amplitude and in spectral coloration with a power-law divergence that depends on the nature of the instability.

For a Hopf bifurcation, the fluctuation power diverges as $1/\varepsilon$ (where ε is the displacement of the control parameter from threshold), and the spectral content becomes “critically tuned” at a nonzero resonance frequency ω_0 with the power spectral density scaling as $1/(\omega - \omega_0)^2$. This is the behavior seen in the subthreshold oscillations of the H.R. Wilson type-II resonator membrane.

For a saddle–node bifurcation, the fluctuation power scales as $1/\sqrt{\varepsilon}$, and the spectral power scales as $1/\omega^2$, implying infinite power at zero frequency. The divergence at dc is the source of the critical slowing seen in the noise-induced fluctuations as a saddle–node annihilation point is approached. This behavior was demonstrated in the type-I integrator neuron, and in the mean-field models for anesthetic induction, SWS-to-REM sleep cycling, and in the FPR model for the wake-to-sleep transition.

We argue that the presence of a neural instability provides a natural demarcation point separating two distinct states, and that a transit across this boundary can be viewed as a phase transition between states. Guedel's observation [9] of an ether-induced delirium phase—separating relaxed consciousness from anesthetic unconsciousness—provided the first historical hint that induction of anesthesia is a brain phase transition. The suppression of bacterial luminescence by volatile agents (ether, chloroform, halothane, nitrous oxide) at clinically relevant concentrations [10, 28] lead to the paradoxical finding of large fluctuations in light intensity at low drug concentrations, consistent with a Guedel-like excited “delirium” phase in bacterial activity. Recent measurements of patient response to the injectable agent propofol show similar “biphasic” (surge followed by decay) brain EEG activity during induction of anesthesia [14, 15], with a second biphasic surge occurring as the patient recovers consciousness. The observation of a hysteresis separation between the induction and emergence biphasic peaks (the recovery biphasic peak occurs at a *lower* drug concentration) suggests that induction of anesthesia can be pictured as a first-order phase transition.

Examination of EEG traces for sleeping mammals (human, cat, fetal sheep) shows broad similarities in sleep patterns, with periodic alternations between a slow, large-amplitude phase (SWS), and a desynchronized lower-amplitude phase (REM sleep). The growth in low-frequency power prior to transition into REM sleep is consistent with the SWS-to-REM sleep phase transition being first-order; the absence of a power surge for REM-to-SWS suggests that this latter transition is continuous.

The FPR phase transition model for the diurnal transition between wake and sleep is based on mutual inhibition of the VLPO and MA brainstem nuclei, resulting in hysteretic flip-flop bistability between wake and sleep states. Each state loses stability via a saddle–node annihilation, so critically-slowed voltage fluctuations, with attendant nonlinear increases in stimulus susceptibility, are predicted in the vicinity of state change. This hypersensitivity to stimulus might provide a natural explanation for the disconcerting hypnic jerk events that are commonly experienced at the moment of sleep onset.

Acknowledgments This research was supported by the Royal Society of New Zealand Marsden Fund, contract UOW-307. We are grateful for assistance from I.P. Gillies, Yanyang Xu, and J.J. Wright.

References

1. Carmichael, H.: *Statistical Methods in Quantum Optics: Master Equations and Fokker–Planck Equations*. Springer, Berlin New York (1999)
2. Destexhe, A., Contreras, D., Steriade, M.: Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states. *J. Neurosci.* **19**(11), 4595–4608 (1999)
3. Eilers, P.H.C.: Smoothing and interpolation with finite differences. In: P.S. Heckbert (ed.), *Graphic Gems IV*, pp. 241–250, Academic Press, San Diego, CA, USA (1994)

4. Franks, N.P., Lieb, W.R.: Anaesthetics set their sites on ion channels. *Nature* **389**, 334–335 (1997)
5. Freeman, W.J.: *Mass Action in the Nervous System*. Academic Press, New York (1975)
6. Freeman, W.J.: Neurodynamics, volume transmission, and self-organization in brain dynamics. *J. Integ. Neurosci.* **4**(4), 407–421 (2005)
7. Fulcher, B.D., Phillips, A.J.K., Robinson, P.A.: Modeling the impact of impulsive stimuli on sleep-wake dynamics. *Phys. Rev. E* **78**(5), 051920 (2008), doi:10.1103/PhysRevE.78.051920
8. Gardiner, C.W.: *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, vol. 13 of Springer Series in Synergetics. Springer-Verlag, Berlin Heidelberg New York, 3 edn. (2004)
9. Guedel, A.E.: *Inhalational Anesthesia: A Fundamental Guide*. Macmillan, New York (1937)
10. Halsey, M.J., Smith, E.B.: Effects of anaesthetics on luminous bacteria. *Nature* **227**, 1363–1365 (1970)
11. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (Lond.)* **117**, 500–544 (1952)
12. Izhikevich, E.M.: Neural excitability, spiking, and bursting. *Internat. J. Bifur. Chaos* **10**, 1171–1266 (2000)
13. Kitamura, A., Marszalec, W., Yeh, J.Z., Narahashi, T.: Effects of halothane and propofol on excitatory and inhibitory synaptic transmission in rat cortical neurons. *J. Pharmacol.* **304**(1), 162–171 (2002)
14. Kuizenga, K., Kalkman, C.J., Hennis, P.J.: Quantitative electroencephalographic analysis of the biphasic concentration–effect relationship of propofol in surgical patients during extradural analgesia. *British Journal of Anaesthesia* **80**, 725–732 (1998)
15. Kuizenga, K., Proost, J.H., Wierda, J.M.K.H., Kalkman, C.J.: Predictability of processed electroencephalography effects on the basis of pharmacokinetic–pharmacodynamic modeling during repeated propofol infusions in patients with extradural analgesia. *Anesthesiology* **95**, 607–615 (2001)
16. Liley, D.T.J., Cadusch, P.J., Wright, J.J.: A continuum theory of electro-cortical activity. *Neurocomputing* **26–27**, 795–800 (1999)
17. Oswald, L.: Sudden bodily jerks on falling asleep. *Brain* **82**, 92–103 (1959)
18. Phillips, A.J.K., Robinson, P.A.: A quantitative model of sleep-wake dynamics based on the physiology of the brainstem ascending arousal system. *J. Biol. Rhythms* **22**(2), 167–179 (2007), doi:10.1177/0748730406297512
19. Phillips, A.J.K., Robinson, P.A.: Sleep deprivation in a quantitative physiologically-based model of the ascending arousal system. *J. Theor. Biol.* **255**, 413–423 (2008), doi:10.1016/j.jtbi.2008.08.022
20. Rechtschaffen, A., Kale, A.: *A Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages of Human Subjects*. U.S. Govt Printing Office, Washington DC (1968)
21. Steyn-Ross, D.A., Steyn-Ross, M.L., Sleigh, J.W., Wilson, M.T., Gillies, I.P., Wright, J.J.: The sleep cycle modelled as a cortical phase transition. *Journal of Biological Physics* **31**, 547–569 (2005)
22. Steyn-Ross, D.A., Steyn-Ross, M.L., Wilson, M.T., Sleigh, J.W.: White-noise susceptibility and critical slowing in neurons near spiking threshold. *Phys. Rev. E* **74**, 051920 (2006)
23. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W.: Modelling general anaesthesia as a first-order phase transition in the cortex. *Progress in Biophysics and Molecular Biology* **85**, 369–385 (2004)
24. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W., Liley, D.T.J.: Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: Evidence for a general anesthetic-induced phase transition. *Phys. Rev. E* **60**, 7299–7311 (1999)
25. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W., Wilcocks, L.C.: Toward a theory of the general anesthetic-induced phase transition of the cerebral cortex: I. A statistical mechanics analogy. *Phys. Rev. E* **64**, 011917 (2001)
26. Strogatz, S.H.: *Nonlinear Dynamics and Chaos*. Westview Press, Cambridge, MA (2000)

27. Ueda, I.: Effects of diethyl ether and halothane on firefly luciferin bioluminescence. *Anesthesiology* **26**, 603–606 (1965)
28. White, D.C., Dundas, C.R.: The effect of anaesthetics on emission of light by luminous bacteria. *Nature* **226**, 456–458 (1970)
29. Wilson, H.R.: Simplified dynamics of human and mammalian neocortical neurons. *J. Theor. Biol.* **200**, 375–388 (1999)
30. Wilson, H.R.: *Spikes, Decisions and Actions: The Dynamical Foundations of Neuroscience*. Oxford University Press, New York (1999)

Chapter 2

Generalized state-space models for modeling nonstationary EEG time-series

A. Galka, K.K.F. Wong, and T. Ozaki

2.1 Introduction

Contemporary neuroscientific research has access to various techniques for recording time-resolved data relating to human brain activity: electroencephalography (EEG) and magnetoencephalography (MEG) record the electromagnetic fields generated by the brain, while other techniques, such as near-infrared spectroscopy (NIRS) and functional magnetic resonance imaging (fMRI) are sensitive to the local metabolic activity of brain tissue.

Time-resolved data contain valuable information on the dynamical processes taking place in brain. EEG and MEG time-series are especially promising, since the electromagnetic fields of the brain are directly reflecting the activation of neural populations; furthermore these time-series can be recorded with high temporal resolution. Extraction of the dynamic changes captured by EEG/MEG recordings is an ideal application for *time-series analysis* [10].

From the multiplicity of concepts and methods for time-series analysis that have been applied to neuroscientific time-series, we focus here on *predictive modeling*, i.e., finding a predictor for future time-series values, based on present and past values. More precisely, we will discuss a particular class of predictive modeling that is attracting considerable attention due to its wide applicability: the *state-space* model [2, 3, 6, 12, 13].

Because nonstationary phenomena—such as sudden phase transitions relating to qualitative changes in dynamical behavior—cannot be modeled well using standard

Andreas Galka

Department of Neurology, University of Kiel, Schittenhelmstrasse 10, 24105 Kiel, Germany.
e-mail: a.galka@neurologie.uni-kiel.de

Kevin Kin Foon Wong

Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA

Tohru Ozaki

Tohoku University, 28 Kawauchi, Aoba-ku, Sendai 980-8576, Japan.

state-space approaches, in this chapter we present a generalization of state-space modeling appropriate for this purpose. This generalized algorithm may also serve as a detector for phase transitions.

2.2 Innovation approach to time-series modeling

Let the data be denoted by $y(t)$, $t = 1, \dots, T$, where T denotes the length of the time-series, i.e., the number of time points at which the data were sampled. In this chapter we will assume the case of univariate (scalar) data, although the modeling algorithms to be presented can also be applied to multivariate (vector) data; techniques like EEG and MEG usually provide multivariate time-series, resulting from a set of up to a few hundred sensors. By confining the analysis to a single channel, we confine our attention to the local brain area for which the chosen sensor is most sensitive.

At a given time point $t - 1$ we intend to predict $y(t)$, employing the data $y(\tau)$, $\tau = t - 1, t - 2, t - 3, \dots$. The optimal predictor is given by the conditional expectation $\mathcal{E}(y(t) \mid y(t - 1), y(t - 2), \dots)$, such that the data model is given by

$$y(t) = \mathcal{E}(y(t) \mid y(t - 1), y(t - 2), \dots) + v(t), \quad (2.1)$$

where $v(t)$ denotes the prediction error or *innovation*. The art of time-series modeling then lies in finding a good approximation to $\mathcal{E}(y(t) \mid y(t - 1), y(t - 2), \dots)$. For an optimal predictor, any correlation structure in the data $y(t)$ is employed for the purpose of prediction, such that, in the time-series of innovations, no correlation of any kind remains, i.e., the innovations are a *white-noise* series. The concept of mapping given data to white innovations represents the core idea of the *innovation approach* to time-series modeling [11].

The theory of innovation approach modeling of Markov processes has been elaborated mainly by Levy [14] and Kailath [12]; one of the main results states that under mild conditions, including continuity of the dynamics, a predictor exists such that the innovations time-series will have a multivariate normal (Gaussian) distribution. We refrain from giving details here; instead the reader is referred to [18].

2.3 Maximum-likelihood estimation of parameters

A parametric function of present and past data, $y(t - 1), y(t - 2), \dots$, may be chosen as an approximation to $\mathcal{E}(y(t) \mid y(t - 1), y(t - 2), \dots)$, i.e., as a predictor; it will typically depend on a set of model parameters, collected in a vector ϑ . Following the concept of maximum-likelihood estimation of statistical parameters, we need to maximize the likelihood defined by the conditional probability distribution

$$L(\vartheta; y(1), \dots, y(T)) = p(y(1), \dots, y(T) \mid \vartheta); \quad (2.2)$$

equivalently, the logarithm of the likelihood, $\log L(\vartheta; y(1), \dots, y(T))$, may be maximized. We will now derive an expression for $\log L(\vartheta; y(1), \dots, y(T))$, to be used in the innovation approach. The joint probability distribution of the data can be expanded as a product

$$p(y(1), \dots, y(T) \mid \vartheta) = p(y(1) \mid \vartheta) p(y(2) \mid y(1), \vartheta) \cdots p(y(T) \mid y(T-1), \dots, y(1), \vartheta), \quad (2.3)$$

where we have used the fact that the data must obey *causality*. The joint probability distribution of the *innovations* has a simpler shape, due to the white-noise property which removes any conditioning on previous values.¹

$$p(v(1), \dots, v(T) \mid \vartheta) = p(v(1) \mid \vartheta) p(v(2) \mid \vartheta) \cdots p(v(T) \mid \vartheta). \quad (2.4)$$

We can employ this simpler expression for deriving the likelihood of the data. The relationship between $p(y(1), \dots, y(T) \mid \vartheta)$ and $p(v(1), \dots, v(T) \mid \vartheta)$ can be found from the function linking these two sets of variables; it is given by Eq. (2.1). According to the standard rules for transforming probability distributions, the Jacobi determinant of this function then arises as a correction to be multiplied with $p(v(1), \dots, v(T) \mid \vartheta)$; however, note that from Eq. (2.1) we have

$$\frac{\partial v(t)}{\partial y(\tau)} = \begin{cases} 1 & \text{for } t = \tau \\ 0 & \text{for } \tau > t, \end{cases} \quad (2.5)$$

where we have used the fact that also the predictor must obey causality. Consequently, the Jacobi determinant is unity, and the joint probability of the given data must be equal to the joint probability of the corresponding innovations,

$$p(v(1), \dots, v(T) \mid \vartheta) = p(y(1), \dots, y(T) \mid \vartheta) \quad (2.6)$$

although the functional form of these two distributions may differ very much. Finally this gives us for the logarithmic likelihood, employing a normal (Gaussian) distribution for the innovations, as argued above,

$$\log L(\vartheta; y(1), \dots, y(T)) = -\frac{1}{2} \left(T \log \sigma_v^2(t) + \sum_{t=1}^T \frac{v^2(t)}{\sigma_v^2(t)} + T \log(2\pi) \right), \quad (2.7)$$

where $\sigma_v^2(t)$ denotes the variance of the innovations.

¹ Here the problem arises that, for the first data value $y(1)$, no previous values exist which could be employed by a predictor. But for sufficiently long time-series, the contribution of the first, or the first few, data values to the likelihood can be neglected.

2.4 State-space modeling

In state-space modeling [2, 3, 6, 12, 13], the data $y(t)$ are modeled by a system of two equations,

$$\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t-1) + \boldsymbol{\eta}(t) \quad (2.8)$$

$$y(t) = \mathbf{C}\mathbf{x}(t) + \varepsilon(t), \quad (2.9)$$

where $\mathbf{x}(t)$ denotes the M -dimensional *state vector*, $\boldsymbol{\eta}(t)$ the *dynamical noise* term and $\varepsilon(t)$ the *observation noise* term; the model parameters are given by the *state transition matrix* \mathbf{A} and the *observation matrix* \mathbf{C} . Furthermore, there are the covariance matrices \mathbf{S}_η and σ_ε^2 of the noise terms (where for univariate data σ_ε^2 is a single variance parameter instead of an actual covariance matrix). Alternatively, the dynamical model, Eq. (2.8), could be chosen as a continuous-time model, i.e., as a stochastic differential equation.

When interpreted as an input–output model, the state-space model of Eqs (2.8, 2.9) produces one output signal $y(t)$ from two input signals $\boldsymbol{\eta}(t)$ and $\varepsilon(t)$. This mapping is not invertible, i.e., the original inputs $\boldsymbol{\eta}(t)$ and $\varepsilon(t)$ cannot be reconstructed from the output $y(t)$. However, it is possible to define a transformed model, such that instead of two input signals just one is present, appearing both in the positions of the dynamical noise and the observation noise; it turns out that this input signal is given by the innovations $\mathbf{v}(t)$ [11]. While the innovations can directly replace observation noise, they need to be multiplied by a problem-specific gain matrix (the *Kalman gain matrix*), before they can replace dynamical noise; in the case of univariate data, this matrix will be an $(M \times 1)$ -dimensional vector.

This transformed model is known as the *innovation representation* or *Kalman filter representation* of the state-space model. It can be shown that the mapping between $y(t)$ and $\mathbf{v}(t)$ is invertible [11]. The existence of this representation provides the justification for practical state-space modeling of time-series.

For given model parameters, the famous Kalman filter algorithm can be applied for the purpose of generating estimates of the state vector [13]; improved estimates can be obtained by additional application of a smoother algorithm [19]. While the Kalman filter performs a pass through the time-series data in forward direction of time, the smoother proceeds in backward direction. Since predictions are only possible in forward direction, it is only the Kalman filter which maps the data to innovations and thereby provides a corresponding value for the likelihood of the data.

2.4.1 State-space representation of ARMA models

A well-established class of predictive models for time-series is given by autoregressive moving-average (ARMA) models [5]. As a simple example for univariate data $y(t)$, we consider the following ARMA(2,1) model:

$$y(t) = a_1 y(t-1) + a_2 y(t-2) + \eta(t) + b_1 \eta(t-1), \quad (2.10)$$

where $\eta(t)$ denotes again a dynamical noise term, with variance σ_η^2 . This model consists of an autoregressive (AR) term of second order, with parameters a_1, a_2 , and a moving-average (MA) term of first order, with parameter b_1 , therefore it is denoted by ARMA(2,1). We can rewrite Eq. (2.10) as

$$\begin{aligned} y(t) &= a_1 y(t-1) + \zeta(t-1) + \eta(t) \\ \zeta(t) &= a_2 y(t-1) + b_1 \eta(t) \end{aligned} \quad (2.11)$$

which is equivalent to

$$\begin{pmatrix} y(t) \\ \zeta(t) \end{pmatrix} = \begin{pmatrix} a_1 & 1 \\ a_2 & 0 \end{pmatrix} \begin{pmatrix} y(t-1) \\ \zeta(t-1) \end{pmatrix} + \begin{pmatrix} 1 \\ b_1 \end{pmatrix} \eta(t), \quad (2.12)$$

where $\zeta(t)$ denotes an auxiliary state variable which can be interpreted as a slightly odd predictor of $y(t+1)$ [2]. We define a state vector as $\mathbf{x}(t) = (y(t), \zeta(t))^\dagger$ (where \dagger denotes matrix transpose) and obtain the state-space model

$$\mathbf{x}(t) = \begin{pmatrix} a_1 & 1 \\ a_2 & 0 \end{pmatrix} \mathbf{x}(t-1) + \begin{pmatrix} 1 \\ b_1 \end{pmatrix} \eta(t) \quad (2.13)$$

$$y(t) = (1, 0) \mathbf{x}(t). \quad (2.14)$$

The dynamical noise term of this model is given by $(1, b_1)^\dagger \eta(t)$; the corresponding covariance matrix follows as

$$S_\eta = \begin{pmatrix} 1 & b_1 \\ b_1 & b_1^2 \end{pmatrix} \sigma_\eta^2. \quad (2.15)$$

In Eq. (2.14) observation noise is absent, $\sigma_\varepsilon^2 = 0$; however, as a generalization we may (and will) allow for nonzero σ_ε^2 .

The specific form of the state transition matrix $\begin{pmatrix} a_1 & 1 \\ a_2 & 0 \end{pmatrix}$ is known as *left companion form*, or (in the language of control theory) *observer canonical form* [12]; it is a characteristic property of the state-space model corresponding to this form that the MA parameter b_1 is accommodated in the covariance matrix of the dynamical noise, while the observation matrix $C = (1, 0)$ keeps a very simple form.

Note that the scaling of the components of the state vector in Eq. (2.13) is directly controlled by the variance σ_η^2 ; since the model is linear, this degree of freedom can be shifted to the observation matrix which then becomes $C = (c_1, 0)$ while the dynamical noise variance can be normalized to $\sigma_\eta^2 = 1$. While in the case of univariate data this is a possible, but not necessary choice, it provides the appropriate generalization for the case of multivariate data; for this reason, we will adopt this choice in this chapter.

The construction leading to the model of Eqs (2.13, 2.14) is easily extended to ARMA($p, p-1$) models with higher order $p > 2$, yielding a state-space model

$$\mathbf{x}(t) = \begin{pmatrix} a_1 & 1 & 0 & \dots & 0 \\ a_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{p-1} & 0 & 0 & \dots & 1 \\ a_p & 0 & 0 & \dots & 0 \end{pmatrix} \mathbf{x}(t-1) + \begin{pmatrix} 1 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{pmatrix} \eta(t) \quad (2.16)$$

$$y(t) = (1, 0, \dots, 0, 0) \mathbf{x}(t). \quad (2.17)$$

The covariance matrix of the dynamical noise term of this model follows as

$$S_\eta = \begin{pmatrix} 1 & b_1 & b_2 & \dots & b_{p-1} \\ b_1 & b_1^2 & b_1 b_2 & \dots & b_1 b_{p-1} \\ b_2 & b_1 b_2 & b_2^2 & \dots & b_2 b_{p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{p-1} & b_1 b_{p-1} & b_2 b_{p-1} & \dots & b_{p-1}^2 \end{pmatrix} \sigma_\eta^2. \quad (2.18)$$

2.4.2 Modal representation of state-space models

The dynamics of any linear state-space model can be characterized by the set of eigenvalues of its state transition matrix A ; the eigenvalues are found by transforming A into a diagonal matrix. If M denotes the dimension of the state-space, there will be M eigenvalues; a certain subset of these eigenvalues will be real, denoted by $a^{(1)}, \dots, a^{(m_1)}$ (where m_1 denotes the number of real eigenvalues), while the remaining eigenvalues will form pairs of complex-conjugated eigenvalues (assuming that all elements of A are real), denoted by $(\psi_{(1)}, \psi_{(1)}^*, \dots, \psi_{(m_2)}, \psi_{(m_2)}^*)$ (where m_2 denotes the number of pairs of complex-conjugated eigenvalues). Then we will have $M = m_1 + 2m_2$.

Real eigenvalues $a^{(k)}$ of A correspond to autoregressive models of first order, AR(1):

$$y(t) = a^{(k)} y(t-1) + \eta(t). \quad (2.19)$$

Each complex-conjugated pair of eigenvalues $\psi_{(k)}, \psi_{(k)}^*$ can be interpreted as an oscillatory eigen-mode of the dynamics, with a resonance frequency $\phi_{(k)}$ (corresponding to the phase of the complex eigenvalues) and an accompanying damping coefficient $\rho_{(k)}$ (corresponding to the modulus of the complex eigenvalues):

$$\psi_{(k)} = \rho_{(k)} \exp i\phi_{(k)}, \quad (2.20)$$

where $i = \sqrt{-1}$.

Consider a complex-conjugated pair of eigenvalues ψ, ψ^* within the diagonalized state transition matrix; it corresponds to a (2×2) -block $\begin{pmatrix} \psi & 0 \\ 0 & \psi^* \end{pmatrix}$ on the diagonal. It is always possible to transform such a block to left companion form $\begin{pmatrix} a_1 & 1 \\ a_2 & 0 \end{pmatrix}$ by a linear transform; therefore each complex-conjugated pair of eigenvalues can be

represented by an ARMA(2,1) model, according to Eq. (2.13). The autoregressive parameters follow from phase and modulus of the complex eigenvalues by

$$a_1^{(k)} = 2\rho_{(k)} \cos \phi_{(k)} \quad , \quad a_2^{(k)} = -\rho_{(k)}^2. \quad (2.21)$$

This transformation has the benefit of removing the complex numbers from the diagonalized state transition matrix.

Finally, the modal representation [23, 24] of the state-space model is given by the transformed state transition matrix:

$$\tilde{\mathbf{A}} = \begin{pmatrix} a^{(1)} & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & a^{(2)} & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & a^{(m_1)} & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & a_1^{(1)} & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & a_2^{(1)} & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & a_1^{(2)} & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & a_2^{(2)} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & a_1^{(m_2)} & 1 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & a_2^{(m_2)} & 0 \end{pmatrix} \quad (2.22)$$

where we have ordered the dimensions of the transformed state-space, such that dimensions corresponding to real eigenvalues come first, followed by dimensions corresponding to complex eigenvalues.²

Note that this matrix is block-diagonal, such that no dynamical interactions between blocks, and therefore between the corresponding AR(1) and ARMA(2,1) components, will occur; however, it has to be kept in mind that in general the dynamical noise covariance matrix \mathbf{S}_η of the state-space model will not be block-diagonal, thereby creating instantaneous correlations between components.

2.4.3 The dynamics of AR(1) and ARMA(2,1) processes

We shall briefly discuss some dynamical properties of the components defined in the previous section. For an ARMA(2,1) process, as defined by Eq. (2.10) or in state-space representation by Eq. (2.13), the corresponding pair of eigenvalues should lie inside the unit circle of the complex plane, otherwise the dynamics would be unstable, i.e., there is a stability condition for the modulus of the eigenvalues,

² In the case of repeated eigenvalues, the transformation to the modal representation will not be possible, but this case is unlikely to arise for real-world data.

$0.0 < \rho < 1.0$. The closer ρ approaches the unit circle, the sharper the resonance will become; a sine wave corresponds to the limit case of $\rho = 1.0$.

The frequency-domain transfer function of an ARMA(2,1) process with AR parameters a_1, a_2 and MA parameter b_1 is given by

$$h(f) = \frac{1 + b_1 \exp(-2\pi i f)}{1 - a_1 \exp(-2\pi i f) - a_2 \exp(-4\pi i f)}, \quad (2.23)$$

where $i = \sqrt{-1}$ and $0 \leq f \leq 0.5$. The behavior of the real part of this function is shown in Fig. 2.1 for a fixed value of ϕ and a set of values for ρ . It can be seen that only for values of ρ close to 1.0 a sharp resonance peak appear. The first-order moving average term $b_1 \eta(t-1)$ produces a distortion of the curves; for the case $b_1 = 1.0$ this distortion is most pronounced, since the numerator of Eq. (2.23) becomes zero at $f = 0.5$. We remark that for ARMA(p, q) models with MA model order $q > 1$ the MA component may impose more complicated changes on the transfer function, since then zeros of the numerator may occur at any frequency.

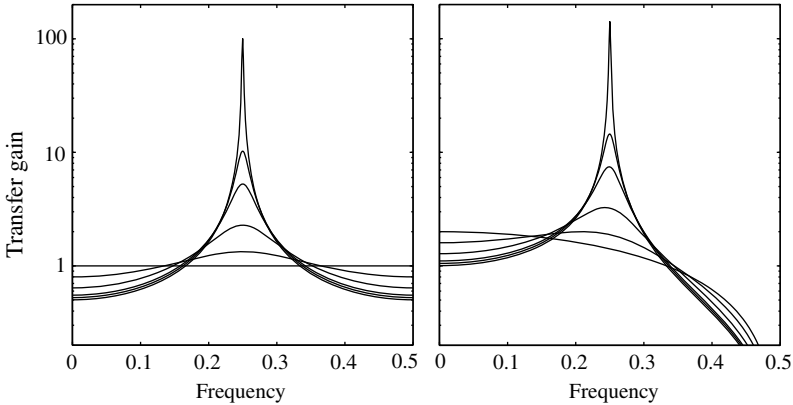


Fig. 2.1 Real part of the transfer function of an ARMA(2,1) process for resonant frequency $\phi = 0.25$, damping coefficients $\rho = 0.0, 0.5, 0.75, 0.9, 0.95, 0.995$ (curves from bottom to top at frequency 0.25) and moving average parameters $b_1 = 0.0$ (left figure) and $b_1 = 1.0$ (right figure). Note the logarithmic scale of the vertical axis.

For AR(1) processes, according to Eq. (2.19), there is only a single real eigenvalue of the transition matrix, which is equal to the first-order autoregressive parameter itself; here we denote this parameter simply by a , for ease of notation. It is obvious that also a real eigenvalue should lie inside of the unit circle, i.e., it should fulfill the stability condition $|a| < 1.0$. The case $a = 1.0$ corresponds to a random walk. AR(1) components cannot have resonant frequencies,³ but they can serve the

³ A somewhat pathological exception is the case $a < 0.0$ which corresponds to an oscillation with precisely the Nyquist frequency; however, this oscillation will not produce an actual resonance peak.

purpose of describing random-walk-like behavior, such as slow drifts and trends in the data, especially if a is close to 1.0.

2.4.4 State-space models with component structure

The modal representation of state-space models corresponds to a model that is organized into sets of AR(1) and ARMA(2,1) components; we can generalize this structure by allowing for higher-order components, i.e., ARMA($p, p-1$) components with $p > 2$, as given by Eqs (2.16), (2.17) and (2.18). For each p , up to some maximum model order, we may choose a set of n_p ARMA($p, p-1$) components and arrange their individual ($p \times p$)-dimensional state transition matrices on the diagonal of the state transition matrix of the state-space model with component structure. The new state dimension will then result as $M = \sum n_p p$, and the state transition matrix will again have a block-diagonal structure, with all remaining elements vanishing.⁴ ARMA($p, p-1$) components with $p > 2$ can be regarded as summarizing a subset of the eigenvalues of the state transition matrix within one ($p \times p$)-dimensional block in left companion form.

If we intend to design a state-space model consisting of mutually independent components, we should choose for the covariance matrix of the dynamical noise S_η the same block-diagonal structure as for the state transition matrix. The corresponding blocks are then given simply by nonzero values for the variances of the AR(1) components, and by ($p \times p$)-dimensional block matrices, as shown in Eq. (2.18), for the ARMA($p, p-1$) components, while again all elements outside of these blocks vanish. For this model structure, there exist no ways by which correlations, instantaneous or delayed, could arise between components, except for coincidental correlations due to limited data-set size.

Finally, the ($1 \times M$)-dimensional observation matrix of the state-space model with component structure is given by

$$C = \left(c_1^{(1)}, c_2^{(1)}, \dots, c_{n_1}^{(1)}, c_1^{(2)}, 0, c_2^{(2)}, 0, \dots, c_{n_2}^{(2)}, 0, \dots, c_1^{(3)}, 0, 0, c_2^{(3)}, 0, 0, \dots, c_{n_3}^{(3)}, 0, 0, \dots \right), \quad (2.24)$$

where the $c_i^{(p)}$ are model parameters, if the corresponding dynamical noise variances σ_η^2 have been normalized to unity.

⁴ As a generalization, it would be possible to use some of the elements outside the blocks for introducing coupling between components.

2.5 State-space GARCH modeling

The state-space model, as presented above, is sufficient for modeling a given time-series under the assumption of *stationarity*. In the case of *nonstationarity*,⁵ the dynamical properties of the time-series data change with time; in this case, some of the model parameters would have to change their values as well, in order to adapt to these changing properties. This additional freedom may either be given to the deterministic part of the model (the first term on the rhs of Eq. (2.13), i.e., the AR term) or to the stochastic part (the second term, i.e., the MA term).

Here we choose the second option, i.e., we allow the dynamical noise covariance to change with time. By this step we approach the concept of *stochastic volatility* modeling [21], which consists of defining the dynamical noise (co)variance itself as a set of new state variables, obeying a separate stochastic dynamical model. For this additional dynamical model a new dynamical noise term is required, which renders this model estimation problem considerably more complicated; however, there exists a famous approximation to full stochastic volatility modeling, known as *generalized autoregressive conditional heteroscedastic*⁶ (GARCH) modeling [4, 7]. GARCH modeling was introduced in the field of financial data analysis.

Originally, GARCH modeling was developed only for the direct modeling of data through AR/ARMA models; its core idea is to use the innovation at the previous time point, $v(t-1)$, as an estimate of the noise input to the additional volatility model. Recently, the method has been generalized to the situation of state-space modeling [8, 25]. The main problem in this generalization is given by the fact that, in the case of state-space models, we would need to employ *state prediction errors* as an estimate of the noise input, but all that is available is the set of *data prediction errors*, i.e., innovations.

2.5.1 State prediction error estimate

In order to derive a state-space version of GARCH modeling, it is necessary to derive a suitable estimator $\hat{v}_x(t)$ of the state prediction error. The first choice for a simple estimator is given by $\hat{v}_x(t) = K(t)v(t)$, where $K(t)$ denotes the $(M \times 1)$ -dimensional Kalman gain matrix of a Kalman filter, used for estimating states from given time-series data; $K(t)$ can be regarded as a regularized pseudo-inverse of the observation matrix C . However, in practical applications this simple estimator displays poor performance, whence we will use a refined estimator, derived in [25]:

$$\hat{v}_x^2(t) = S_\eta(t) - S_\eta(t)C^\dagger \sigma_v^{-2}(t)CS_\eta(t) + K(t)v^2(t)K^\dagger(t) \quad (2.25)$$

⁵ We note that, within the framework of linear modeling, *nonlinearity* may be indistinguishable from *nonstationarity*.

⁶ The term *heteroscedasticity* refers to the situation in which, within a set of stochastic variables, different variables have different variances. Here, the term *scedasticity*, from Greek *skedasis* for “dispersion”, is yet another word for “variance”.

which is, strictly speaking, an estimator of the square of the state prediction error; the square is inherited from the square in the definition of (co)variances. In Eq. (2.25), $\sigma_v^2(t)$ denotes the innovation variance, provided by the Kalman filter. From Eq. (2.25), $\hat{v}_x^2(t)$ is a square matrix; in order to obtain the noise estimates for the individual state components, we pick out the diagonal values from this matrix. While this uniquely defines the noise terms for AR(1) components, ARMA($p, p-1$) components pose the problem that there are p diagonal elements; here we have chosen to simply average over these elements, but other choices would be possible. The resulting average is denoted by $\hat{v}_x^2(k, t)$ for the k th component of a state-space model with component structure.

2.5.2 State-space GARCH dynamical equation

The design of a state-space GARCH model contains various details of implementation which need to be chosen, and in several cases it is not obvious which choice would be best; instead, practical experience is employed.⁷ We found useful the particular implementation which we now describe.

In our implementation, the new time-dependent GARCH state variables correspond roughly to standard deviations, rather than variances; however, in contrast to standard deviations, these variables may also become negative. The state-space GARCH model itself is given by another ARMA(r, s) model,

$$\sigma(k, t) = \sigma(k, 0) + \sum_{\tau=1}^r \alpha(k, \tau) \sigma(k, t - \tau) + \sum_{\tau=1}^s \beta(k, \tau) \hat{v}_x^2(k, t - \tau), \quad (2.26)$$

such that for each component there is an additional set of state-space GARCH parameters $\sigma(k, 0), \alpha(k, 1), \dots, \alpha(k, r), \beta(k, 1), \dots, \beta(k, s)$; these parameters become an additional part of the vector of model parameters ϑ . However, in practice we do not need a state-space GARCH model for each component of a given state-space model, but only for the particular component which actually contains the nonstationary phenomena to be modeled. For the other components we set

$$\sigma(k, 0) = 1, \alpha(k, 1) = \dots = \alpha(k, r) = 0, \beta(k, 1) = \dots = \beta(k, s) = 0.$$

The choice of the GARCH model orders r, s forms again part of the model design. In the application examples to be presented in this chapter, we have decided to use $r = 1, s = 10$; experience has shown that sometimes it is advantageous to include a longer history of previous noise estimates into the model. However, in other cases also the choice $r = 1, s = 1$ has yielded good results [25]. In order to simplify the parameter estimation step, we define a constraint $\beta(k, 1) = \beta(k, 2) = \dots = \beta(k, 10)$,

⁷ This situation is not unusual in statistical modeling of data, since it will rarely be possible to set up a model which faithfully reproduces the structure of the underlying natural processes; rather, models have always to be regarded as approximations. At least, this is the situation we are facing in the study of systems of enormous complexity, such as the human brain.

such that in effect we are using an average of the last 10 noise estimates and just one MA parameter. In Refs. [8] and [25], state-space GARCH models were introduced in which the logarithm of the variance, $2 \log \sigma(k, t)$, was used as GARCH state variable, but in this chapter we have decided to formulate the model directly in the standard deviations $\sigma(k, t)$.

2.5.3 Interface to Kalman filtering

At each time point t , the current value of the GARCH state variable, $\sigma(k, t)$, is passed through a “nonlinear” observation function by taking the square, thereby becoming a genuine non-negative variance $\sigma^2(k, t)$; this variance then replaces, for component k , the term σ_η^2 which appears in Eqs (2.15, 2.18) of the stationary state-space model. The corresponding dynamical noise covariance matrix of component k then enters the block-diagonal covariance matrix of the state-space model at the appropriate block position of component k , such that this matrix itself becomes time-dependent.

This step represents a major modification of the usual Kalman filter iteration, since the continuous changes of one of the main matrices of the model prevent the filter from reaching its steady state.

2.5.4 Some remarks on practical model fitting

The generalized state-space models discussed in this chapter are parametric models, consisting of a model structure and a parameter vector ϑ . The following table lists the parameter sets contained in ϑ , also giving the dimension of each set:

Description	Symbol	Dimension ⁸
state transition matrix parameters ⁹	$a^{(k)}, \phi^{(k)}, \rho^{(k)}$	$m_1 + 2m_2$
moving-average parameters	b_i	m_2
observation matrix parameters	c_i	$m_1 + m_2$
observation noise variance	σ_ε^2	1
GARCH parameters	$\sigma(k, 0), \alpha(k, \tau), \beta(k, \tau)$	$r + s + 1$ or ¹⁰ $r + 2$
initial state vector	$\mathbf{x}(0)$	$m_1 + 2m_2$

⁸ In the table, m_1 and m_2 denote the number of real eigenvalues and of pairs of complex eigenvalues, respectively, regardless of how these eigenvalues are distributed over the ARMA($p, p-1$) components of the state-space model.

⁹ Optimizing $\phi^{(k)}, \rho^{(k)}$ instead of the corresponding AR parameters $a_1^{(k)}, a_2^{(k)}$ has the advantage that the stability constraint $\rho^{(k)} < 1.0$ can be directly imposed; furthermore, prior knowledge about the frequencies $\phi^{(k)}$ can be conveniently incorporated into the model, or particular frequencies can be excluded from the optimization process.

¹⁰ if the constraint $\beta(k, 1) = \beta(k, 2) = \dots = \beta(k, s)$ is applied

For a given choice of ϑ , the Kalman filter provides the corresponding value of the likelihood. Model fitting consists of maximizing the likelihood, or, more conveniently, the logarithmic likelihood, with respect to ϑ by numerical optimization [9]. For this purpose, we are employing standard optimization algorithms, namely the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton algorithm and the Nelder-Mead simplex algorithm. Sometimes the simplex algorithm can be employed in situations where the quasi-Newton algorithm fails due to numerical problems.

Several optimization steps should be iterated, such that in some steps the optimization is limited to subsets of parameters. With the sole exception of σ_ε^2 , all parameters can be uniquely assigned to one of the components; we recommend performing a series of optimization steps such that each step is confined to one component. Some optimization steps may also be confined to state transition matrix parameters, or to observation matrix parameters, etc.

A good initial model is of crucial importance for successful modeling. We recommend fitting an autoregressive model, $AR(p)$, of sufficiently high model order, say $p = 30$, to the given data; fitting of pure AR models, without MA terms, can be done very efficiently by standard least-squares regression [5]. This model is then converted into a state-space model, as discussed above, and the resulting state-space model is transformed into its modal representation; thereby a model consisting of a set of AR(1) and ARMA(2,1) components is obtained. Later, higher-order ARMA components can be created by merging pairs of these AR(1) and/or ARMA(2,1) components. The dynamical noise covariance matrix is constrained to the same block-diagonal structure as the state transition matrix by setting all other elements to zero.

At this point there is a need for subjective interference with the modeling process: usually a subset of the initial components will capture the most important features of the data and of the underlying dynamics, such as frequencies known to play an important role, or prominent time-domain patterns, while other components will describe rather unspecific activity. Only this subset of important components should be selected as initial model, while the remaining components should be discarded. Also, the decision as to which components, if any, are to be merged later to form higher-order ARMA components depends on subjective assessment of the dynamics represented by the components.

Keeping all components from the modal representation would also be possible, but it would result in a very large model with many redundant components; such a model could be employed as an alternative initial model, and later the model could gradually be “pruned” during the optimization process, but this procedure would be very demanding in terms of computational time consumption.

For the observation noise variance σ_ε^2 and the state-space GARCH parameters, no initial values can be obtained by this approach. For σ_ε^2 , a small initial value should be chosen, maybe about 10^{-3} times smaller than the variance of the data, unless we have reason to assume that there was considerably more observation noise in the data. Larger initial values for σ_ε^2 may create the risk that the Kalman filter would incorrectly allocate a large fraction of the variance of the data to the

observation noise term. The procedure for the state-space GARCH parameters is described below.

For the application examples which will be discussed below, the dimension of ϑ is about 35; about 10 of these parameters form the initial state vector $\mathbf{x}(0)$. We recommended to keep $\mathbf{x}(0)$ initially at zero and to optimize only the remaining parameters, except for the state-space GARCH parameters, while omitting the contributions of the first (approximately) 20 data points to the likelihood, in order to allow a transient of the Kalman filter to die out. Once a first set of optimization steps has been applied, such that approximate estimates of the main parameter sets of the model have been obtained, the full likelihood is evaluated and the initial state vector is included into the remaining optimization steps.

During the first part of the model fitting procedure, there should not yet be any state-space GARCH models, i.e., the state-space GARCH parameters should be fixed as $\sigma(k, 0) = 1$, $\alpha(k, 1) = 0$, $\beta(k, 1) = 0$, $\beta(k, 2) = 0$, \dots . After the estimates of the other parameter sets have converged to stable values, it can be decided which component should be given a state-space GARCH model. Usually, the nonstationary behavior to be modeled is represented only by one or possibly two components, and only these components should be given state-space GARCH models. Experience has shown that if state-space GARCH models are given to all components of a state-space model, components tend to become blurry and featureless, since too much freedom is available to each component. After estimates of the state-space GARCH parameters have been obtained, also all other model parameters need to be refitted, since the introduction of state-space GARCH models may considerably change the dynamics of the complete model.

In many cases, we probably cannot expect to reliably find global maxima in a 25-dimensional, highly heterogeneous parameter space. After the optimization procedure, the Hessian matrix at the obtained solution should routinely be computed, in order to check for the possibility of saddle points; nevertheless we may find only local maxima. Refined studies of the geometry of these parameter spaces would be needed, in order to obtain additional insight into this problem. However, we expect that for practical purposes a good solution will be almost as useful as the perfect solution. In the end, the properties of the innovations will always allow an assessment of the quality of the obtained model; major problems during the optimization step will usually also be reflected in the innovations.

2.6 Application examples

In the remaining part of this chapter we will discuss the application of state-space modeling, with component structure and state-space GARCH components, to three examples of EEG time-series; all contain nonstationary phenomena: in the first example, due to the transition from the conscious state to anesthesia; in the second, due to the transition from one sleep stage into another; and in the third, due to the occurrence of an epileptic seizure.

2.6.1 Transition to anesthesia

As the first example we choose an EEG time-series recorded from a patient being anesthetized (with propofol) prior to surgery. Sampling rate was $f_s = 100$ Hz. From the full clinical data set we select the time-series recorded at the T4 electrode versus average reference; we select $T = 2000$ sample points, starting with the moment when induction of anesthesia was begun. The data are shown in Fig. 2.2A; the same data were also analyzed in [25]. It can be seen that the qualitative appearance of the trace changes within the 20 seconds covered by this time-series, i.e., the data contain pronounced nonstationarity: high-amplitude oscillations in the delta frequency range gradually become stronger, corresponding to the loss of consciousness. The transition from the conscious state to anesthesia may be regarded as a phase transition in brain dynamics [22].

We model the data by a state-space model consisting of $m_2 = 5$ mutually independent ARMA(2,1) components; the model is fitted by maximizing the log-likelihood until convergence. It is found that one of the components represents the gradually increasing delta range oscillation; in a second modeling step, a state-space GARCH model is added to this component, but not to the remaining four components. The state-space GARCH model orders are $r = 1$, $s = 10$, but we apply to the MA parameters the constraint introduced above, $\beta(k, 1) = \beta(k, 2) = \dots = \beta(k, 10)$. The three additional parameters of the state-space GARCH model are also fitted by maximizing the log-likelihood; then the other sets of model parameters are refitted, starting at their previous non-GARCH values, in order to allow the model to adapt to the presence of the state-space GARCH model. Joint and alternate optimization of state-space GARCH model parameters and other parameters are iterated a few times, again until convergence.

The resulting five components are shown in Fig. 2.2B; together they represent a decomposition of the data of Fig. 2.2A. The figure shows smoothed state estimates, as obtained by a standard Rauch-Tung-Striebel smoother [19] which performs a backward pass through the time-series; during optimization only the forward pass of the Kalman filter is performed, since it is this pass which transforms the data into innovations and thereby produces a value for the likelihood.

Note that in Fig. 2.2B all components are displayed with the same variance, such that their dynamical properties can be compared; their actual variances in state-space will differ considerably, since we have chosen to normalize the variances of the dynamical noises to 1, instead of the variances of the estimated states¹¹.

In Fig. 2.2B components are ordered according to increasing frequency; this is possible since all components are modeled by ARMA(2,1) processes, such that there is a single resonance frequency for each component. At the top we find the nonstationary delta range component (labeled c1), with frequency¹² $f = 0.422$ Hz and

¹¹ The effective variances of the time-series of estimated state components do not represent model parameters, therefore they would be inaccessible for the purpose of normalization.

¹² The physical frequency f is related to the phase ϕ of the corresponding pair of complex eigenvalues, as defined by Eq. (2.20), by $\phi = 2\pi f/f_s$, where f_s denotes the sampling frequency of the data.

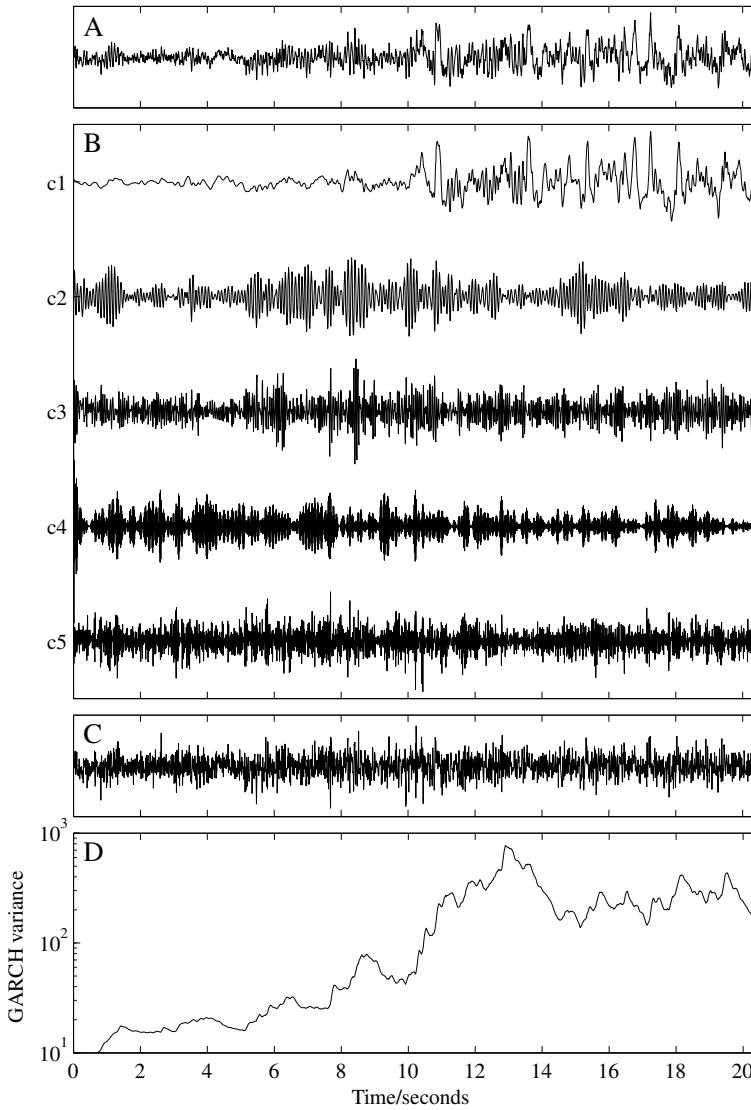


Fig. 2.2 EEG time-series with transition to anesthesia: Data (A); state-space decomposition (B); innovations (C); and state-space GARCH variance of component c1 (D). Vertical axes for all graphs in subfigures A, B and C have been rescaled individually for convenience of graphical display. Resonance frequencies f and damping coefficients ρ of components: c1: $f = 0.422$ Hz, $\rho = 0.690$; c2: $f = 10.495$ Hz, $\rho = 0.946$; c3: $f = 17.463$ Hz, $\rho = 0.772$; c4: $f = 45.34$ Hz, $\rho = 0.910$; c5: $f = 48.649$ Hz, $\rho = 0.292$.

damping coefficient $\rho = 0.690$; the gradual increase of delta amplitude is clearly visible. The next two components (c2 and c3) represent alpha and beta range components, with $f = 10.495$ Hz, $\rho = 0.946$ for c2; and $f = 17.463$ Hz, $\rho = 0.772$ for c3. The remaining two components (c4 and c5) represent high-frequency noise components, with $f = 45.34$ Hz, $\rho = 0.910$ for c4; and $f = 48.649$ Hz, $\rho = 0.292$ for c5. Note that for this data set the Nyquist frequency lies at $f_s/2 = 50$ Hz.

In Fig. 2.2C the weighted innovations are shown, confirming that little, if any, structure has remained in the innovations. The raw innovations (prediction errors) of the state-space model have been weighted by being divided at each time point by the square root of the corresponding innovation variance, as provided by the Kalman filter; remember that in presence of a state-space GARCH model the Kalman filter will not reach its steady state, such that also the innovation variance (or, more generally, covariance) will not converge to a constant value.

Finally, in Fig. 2.2D the time-dependent variance of the delta range component is shown, as described by the state-space GARCH model. Note that the vertical axis of this figure is logarithmic. This graph should be studied together with the delta range component itself, the top graph in Fig. 2.2B. It can be seen that the variance increases from values around 20 in the first few seconds to values around 200–300 at the end of the time-series; this increase may be interpreted as a data-derived quantitative representation of the phase transition process.

At the beginning of the time-series, the dynamics of the variance of the delta range component was initialized at an arbitrary value of 1.0, from which it has to rise to more realistic values during a short transient which is not explicitly resolved in the figure. The maximum-likelihood estimates of the state-space GARCH parameters are $\sigma(k, 0) = 0.0837$, $\alpha(k, 1) = 0.975$, $\beta(k, 1 \dots 10) = 4.111 \times 10^{-5}$.

2.6.2 Sleep stage transition

The second example is given by an EEG time-series recorded from the surface of a fetal sheep brain (144 days gestation age). The original sampling rate was 250 Hz, but we decide to subsample the data to $f_s = 125$ Hz. A single electrode is selected. Out of a longer data set, a subset of $T = 50000$ sample points (at 125 Hz) is selected, covering a transition between slow-wave sleep (SWS) to REM sleep. The data are shown in Fig. 2.3A. The transition is discernible by a decrease of signal amplitude with concomitant fading of the characteristic slow-wave activity.

For modeling the sleep data we choose the same model structure as used for the anesthesia study, i.e., we choose a state-space model consisting of $m_2 = 5$ mutually independent ARMA(2,1) components; the model is fitted by maximizing the log-likelihood until convergence. Again it is found that only one of the components captures the nonstationary behavior representing the sleep stage transition; in a second modeling step, a state-space GARCH model is added to this component, but not to the remaining four components. State-space GARCH model orders are the same

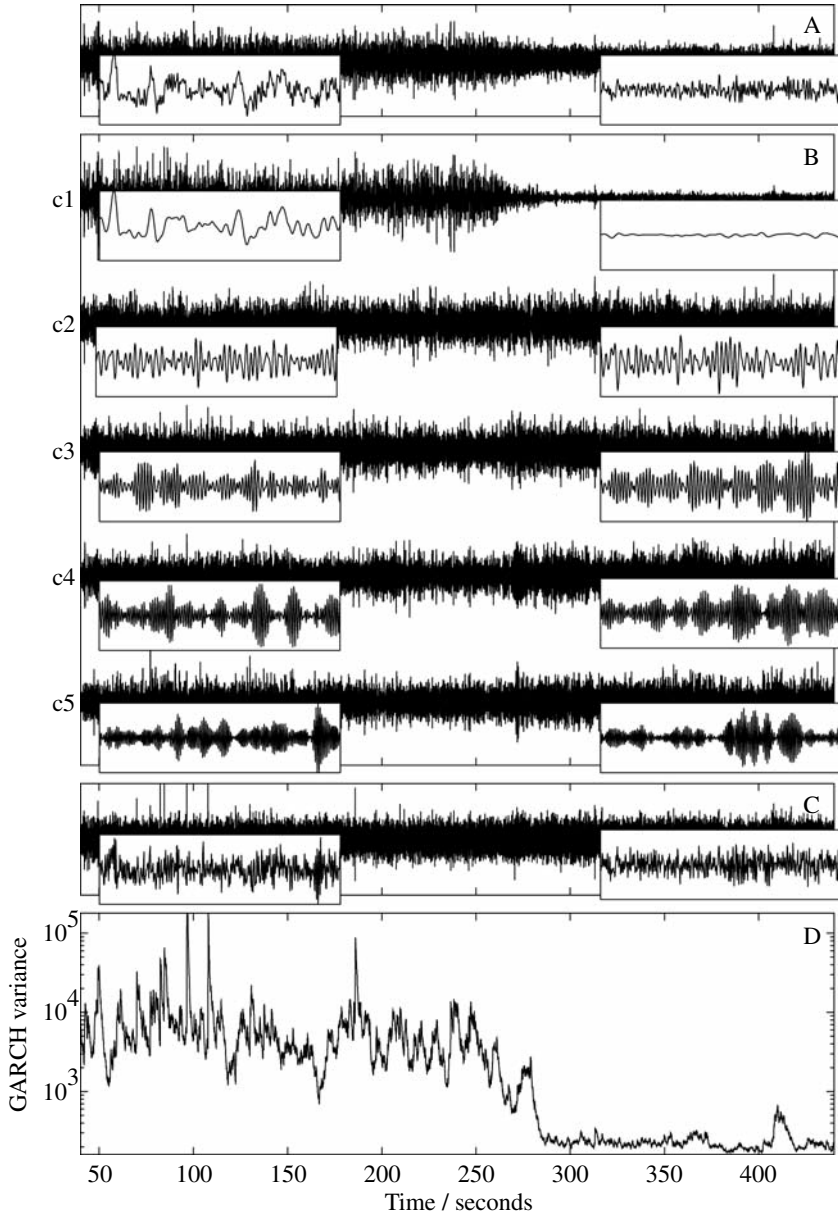


Fig. 2.3 EEG time-series from fetal sheep brain with transition from slow-wave sleep to REM sleep: Data (A); state-space decomposition (B); innovations (C); and state-space GARCH variance of component c1 (D). Vertical axes for all graphs in subfigures A, B and C have been rescaled individually for convenience of graphical display. Resonance frequencies f and damping coefficients ρ of components: c1: $f = 3.811$ Hz, $\rho = 0.910$; c2: $f = 11.465$ Hz, $\rho = 0.882$; c3: $f = 18.796$ Hz, $\rho = 0.926$; c4: $f = 24.896$ Hz, $\rho = 0.951$; c5: $f = 30.801$ Hz, $\rho = 0.945$. Insets show enlarged parts of data, state-space components and innovations: 100–104 s (left) and 400–404 s (right).

as for the anesthesia example: $r = 1$, $s = 10$, and the same constraint for the MA parameters is employed. Model parameters are optimized until convergence.

The resulting five components are shown in Fig. 2.3B, ordered according to increasing frequency; again, smoothed state estimates are shown, rescaled to the same variance. The first component, labeled c1, is the nonstationary component; its frequency and damping coefficient is $f = 3.811$ Hz, $\rho = 0.910$. For the remaining components, frequencies and damping coefficients are $f = 11.465$ Hz, $\rho = 0.882$ for c2; $f = 18.796$ Hz, $\rho = 0.926$ for c3; $f = 24.896$ Hz, $\rho = 0.951$ for c4; and $f = 30.801$ Hz, $\rho = 0.945$ for c5. The Nyquist frequency lies at $f_s/2 = 62.5$ Hz. Note that damping coefficients for all components are fairly close to 1.0, indicating pronounced oscillatory behavior.

The weighted innovations and the time-dependent variance of the nonstationary component are shown in Figs. 2.3C and 2.3D, respectively. It can be seen that the variance decreases from values around 5000 in the first part of the time-series (representing slow-wave sleep) to values around 200 in the latter part (representing REM sleep). If the variance is used as a quantitative measure for the transition between the two sleep stages, the time point at which the transition occurs can be identified to within a time-interval of no more than 5 s; however, note that also within each of the two sleep stages there are slow changes of the variance which may reflect changes of the underlying physiological state.

Also for this model, the dynamics of the variance of the nonstationary component was initialized at a value of 1.0, from which it rises to appropriate values around 5000 during a short transient. The maximum-likelihood estimates of the state-space GARCH parameters are $\sigma(k, 0) = 0.176$, $\alpha(k, 1) = 0.985$, $\beta(k, 1 \dots 10) = 2.68 \times 10^{-6}$.

In this time-series, we have the example of a nonstationarity where a state with large variance passes to a state with smaller variance; we remark that we were also able to model data displaying the opposite situation, i.e., the transition from REM sleep to slow-wave sleep, from the same experiment (the same fetus) with the same model class.

2.6.3 Temporal-lobe epilepsy

As the third example we choose an EEG time-series recorded from a patient suffering from temporal-lobe epilepsy, during awake resting state with open eyes. Sampling rate was $f_s = 200$ Hz. From the full clinical data set we select the time-series recorded at the Fz electrode versus linked earlobes; out of a longer data set, we select $T = 2048$ sample points, covering one short generalized epileptic seizure of a type characteristic for temporal-lobe epilepsy. The data are shown in Fig. 2.4A. In the figure, it can be seen that at time near 7 s the qualitative appearance of the trace changes abruptly, with a series of periodic high-amplitude spike-wave patterns emerging; these patterns are typical of the ictal regime (containing the seizure), while the earlier part of the trace represents the preictal regime. The transition from

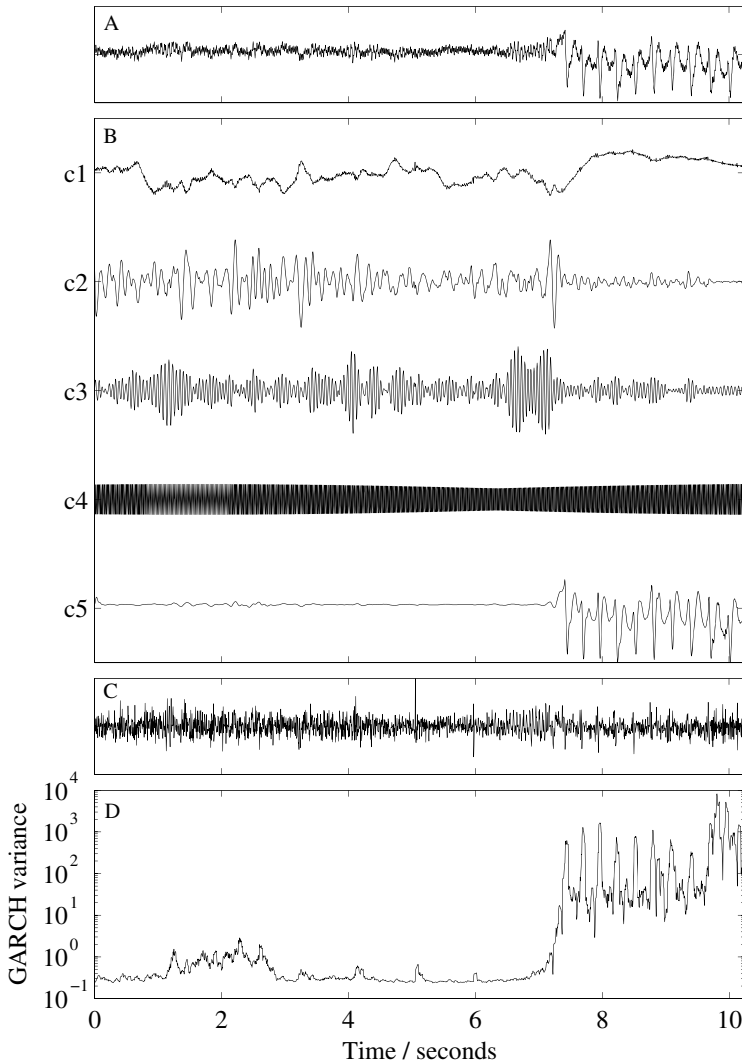


Fig. 2.4 EEG time-series with epileptic seizure: Data (A); state space decomposition (B); innovations (C); and state-space GARCH variance of component c5 (D). Vertical axes for all graphs in subfigures A–C have been rescaled individually for convenience of graphical display. AR-parameter of component c1 is $a_1 = 0.985$. Resonance frequencies f and damping coefficients ρ of other components: c2: $f = 7.978\text{Hz}$, $\rho = 0.876$; c3: $f = 17.288\text{Hz}$, $\rho = 0.976$; c4: $f = 50.025\text{Hz}$, $\rho = 1.0$; component c5 has frequencies $f = 3.274\text{Hz}$, $f = 18.171\text{Hz}$ and corresponding damping coefficients $\rho = 0.870$, $\rho = 0.883$.

the preictal to the ictal regime has recently been discussed by Milton *et al.* [15] in analogy with phase transitions in physics.

We model the data by a state-space model consisting of one AR(1), three ARMA(2,1) and one ARMA(4,3) components (corresponding to $m_1 = 1$ real eigenvalues and $m_2 = 5$ complex-conjugated pairs of eigenvalues); this structure is chosen according to the transformation of an initial AR model into modal representation which reveals at least two components representing the epileptic seizure; these two components are merged into a single fourth-order component. Again, the initial state-space model is fitted by maximizing the log-likelihood until convergence. The ARMA(4,3) component, representing the seizure activity, is then provided with a state-space GARCH model, while the remaining four components are not. Again, the state-space GARCH model orders are $r = 1$, $s = 10$, and the same constraint for the MA parameter as before is employed. Fitting of the three additional parameters and refitting of the other sets of model parameters proceeds in the same way as for the earlier anesthesia and fetal sleep examples.

The resulting five components are shown in Fig. 2.4B, ordered according to increasing frequency, with the ARMA(4,3) component at the bottom of the figure; together these components represent a decomposition of the data of Fig. 2.4A. Also this figure shows smoothed state estimates; again, for convenience of graphical display, variances of components have been normalized to the same value.

At the top of the figure, the single AR(1) component is shown, labeled c1; its state transition parameter a_1 is 0.985, and thereby well suited for describing slow drifts and trends in the data. In this time-series, there seems to be a slow shift of potential during the seizure; we see that the AR(1) component captures this shift well, thereby facilitating the modeling of the oscillatory pattern during the seizure by another component. In the preictal regime, the first-order component also captures some unspecific low-frequency activity.

Below the AR(1) component, we see in Fig. 2.4B the three ARMA(2,1) components, with frequencies and damping coefficients $f = 7.978$ Hz, $\rho = 0.876$ for c2, $f = 17.288$ Hz, $\rho = 0.976$ for c3, and $f = 50.025$ Hz, $\rho = 1.0$ for c4; the Nyquist frequency lies at $f_s/2 = 100$ Hz. Components c2 and c3 represent alpha- and beta-range components, respectively; the beta activity is clearly visible in the data. Component c4 represents the frequency of the electrical power supply, i.e., an artifact of technical origin; the damping coefficient of $\rho = 1.0$ clearly reveals an undamped oscillation.

At the bottom of Fig. 2.4B, the ARMA(4,3) component representing the epileptic seizure is shown; it can be seen that this component displays only weak activity until the seizure commences. The seizure itself is well extracted, without leakage into the other components. The two frequencies of this component are $f = 3.274$ Hz and $f = 18.171$ Hz; the corresponding damping coefficients are $\rho = 0.870$ and $\rho = 0.883$. It is obvious that the first of these frequencies describes the main periodicity of the ictal spike-wave patterns.

In Fig. 2.4C the weighted innovations are shown; again they are weighted by being divided at each time point by the square root of the corresponding innovation variance. While it can be seen that most of the structure has been removed, there are

still some remains of seizure-related structure in the innovations. This can be seen most clearly from the series of sharp spikes in the innovations which correspond well with the epileptic spikes in the data. The last 40 samples of the innovations are probably dominated by muscle artifact effects.

Finally, in Fig. 2.4D the time-dependent variance of the epileptic seizure component is shown, as described by the state space GARCH model. Note that again the vertical axis of this figure is logarithmic. This graph should be studied together with the epileptic seizure component itself, the bottom graph in Fig. 2.4B. Again, at the beginning of the time-series, the dynamics of the variance was initialized at an arbitrary value of 1.0; the variance then drops to a somewhat smaller value and mostly stays close to this value for several seconds, until the seizure commences. The maximum-likelihood estimates of the state space GARCH parameters are $\sigma(k, 0) = 0.465$, $\alpha(k, 1) = 5.044 \times 10^{-3}$, $\beta(k, 1 \dots 10) = 3.941 \times 10^{-3}$; from these values it is not surprising that the variance stays close to the constant term $\sigma(k, 0) = 0.465$, as long as the innovations remain small.

However, as soon as the seizure starts, the variance rises to values of almost 10^3 ; then, the variance oscillates roughly between 10 and 10^3 , thereby following the spike-wave oscillation of the seizure. We thus have two regimes of different behavior of variance, preictal and ictal; if the transition between these two regimes is regarded as a phase transition, the concurrent rise of the variance may again be interpreted as a data-derived quantitative representation of this phase transition process. We emphasize that no prior information—relating to either the components in the data or the timing of seizure onset—was given to the algorithm.

Also in the preictal regime, the time-dependent variance shows some structure, such as a transient increase of variance between 1.0 and 3.5 s into the time-series; whether this structure actually reveals relevant information on the epileptic seizure component cannot be decided on the basis of the analysis of just a single seizure.

2.7 Discussion and summary

For centuries, the ability to make quantitative predictions has been regarded as one of the ultimate goals of science. Our present work, which aims to construct predictive models for particular brain phenomena that are accessible to direct observation, is motivated by the same goal.

Much is now known about the elementary constituents of the human brain: the neurons, synapses, neurotransmitters and ion channels. In principle, it should be possible to use this knowledge to set up a detailed model of the dynamics of brain; such a model would allow reliable predictions of the observable phenomena generated by the brain. However, due to the enormous numbers of these constituents and the complexity of their interconnections, this is not (yet) a practicable task.

Alternatively, a predictive model may be set up predominantly or exclusively based on the available data, and this is the path we have followed in this chapter. More specifically, we have studied how such a model can be set up for the

purpose of efficient description of transitions between qualitatively different dynamical states, i.e., of *nonstationary* behavior. The resulting models summarize various useful statistics about the data, mainly encoded in the properties of the state-space components, into which the data are decomposed. Each component is characterized by one or several resonant frequencies, but also by the corresponding damping coefficients; furthermore the total power in the data is distributed over the components in a specific way. Nonstationary components are modeled by additional state-space GARCH models, and the time-dependent variance information provided by such models offers additional information on the processes underlying the data; it may be used also for purposes of automatic classification and event detection. In particular, phase transitions represent an example of nonstationary processes; thus the time-dependent variance may serve as a data-derived quantitative representation of the underlying phase transition processes.

In this chapter, we have sketched a systematic approach to building state-space models for univariate time-series data; the generalization to multivariate data is straightforward. State-space models are predictive models, mapping the time-series data to a time-series of prediction errors, denoted as *innovations*. The innovation approach to data modeling aims at whitening the data, i.e., at removing all correlations from the innovations; this is the condition for the validity of the expression for the logarithmic likelihood of the data given by Eq. (2.7).

The innovations are also a source of information for further improvement of models; a good example is given by the third application example of this chapter. Epileptic spike-wave patterns are known to be difficult to model by autoregressive models [16]; the strongly anharmonic waveforms, in combination with poor stability of the main frequency, pose considerable challenges. An improved model for the epileptic seizure component, possibly incorporating also nonlinear elements, should be able to reduce the amount of seizure-related residual structure in the innovations which is visible in Fig. 2.4C; alternatively, or additionally, the design of the state-space GARCH model may be further improved.

The choice of the model order of certain components represents a question of model design, i.e., the choice of model structure; a related problem is that of model comparison. This is a much more difficult problem than estimating model parameters within a fixed model structure, and a full discussion would go beyond the scope of this chapter. For the purpose of time-series decomposition and characterization of nonstationarities, we have found the approach of fitting a set of mutually independent $\text{ARMA}(p, p-1)$ components useful; the choice of the number of components and their model orders will, to some extent, remain a subjective decision. However, such subjective decisions may be partly based on prior knowledge about the properties of physiologically meaningful components, or of well-known artifacts.

Fitting larger models with larger numbers of model parameters will usually improve the likelihood, when compared with smaller models. It is well known that this effect invites the risk of overfitting, against which the maximum-likelihood method itself has no protection. Information criteria like the Akaike Information Criterion (AIC) [1] or the Bayesian Information Criterion (BIC) [20] have been introduced, for replacing the likelihood $L(\vartheta; y(1), \dots, y(T))$, or, more precisely,

replacing $-2\log L(\vartheta; y(1), \dots, y(T))$; these criteria contain a penalty term for the number of model parameters, such that it can be decided whether the improvement of likelihood resulting from extending a model is worth the price of additional model complexity. Recently, *logarithmic evidence* has been proposed as an alternative for AIC and BIC [17].

For the application examples presented in this chapter we have not reported detailed values of log-likelihood, AIC or BIC; but we remark that the comparison of both AIC and BIC for the best non-GARCH models with the final models including state-space GARCH modeling has consistently favored the latter models.

Information criteria like AIC or BIC are best known as tools for estimating optimal model orders for model classes like $AR(p)$ models; but in fact these measures permit the comparison of the performance of models in a much wider setting, such as non-nested models, or even, with respect to their structure, mutually unrelated models. Then, in principle, the process of model design could be based completely on comparison of such criteria, instead on subjective decisions; the problem here is that, for each competing model, a time-consuming numerical optimization procedure would have to take place before the values of the criteria would become available; this would make such an approach very time-consuming. But the power of information criteria for quantitative model comparison should be kept in mind.

Also, other design choices of the modeling algorithm discussed in this chapter could be investigated in the light of information criteria. As an example, we again mention details of the implementation of state-space GARCH modeling, such as model orders, or the choice of the estimator for the state prediction errors, Eq. (2.25); the quadratic estimator which we have employed, following [25], draws its main justification from its superior performance in practical applications, also in terms of information criteria, as compared to other estimators.

Use of state-space GARCH modeling to describe nonstationary structure in time-series—in the absence of prior information on the timing of the nonstationary changes—represents a comparatively new approach that will require more study, both in simulations and applications, in order to become an established tool for time-series analysis. In this chapter we have demonstrated its rich potential for modeling phase transitions and other nonstationary behavior in electroencephalographic time-series data.

Acknowledgments The work of A. Galka was supported by Deutsche Forschungsgemeinschaft (DFG) through SFB 654 “Plasticity and Sleep”. The anesthesia EEG data set was kindly provided by W.J. Kox and S. Wolter, Department of Anesthesiology and Intensive Care Medicine, Charité-University Medicine, Berlin, Germany, and by E.R. John, Brain Research Laboratories, New York University School of Medicine, New York, USA. The fetal sleep data set was kindly provided by A. Steyn-Ross, Department of Engineering, University of Waikato, New Zealand. The epilepsy EEG data set was kindly provided by K. Lehnertz and C. Elger, Clinic for Epileptology, University of Bonn, Germany.

References

1. Akaike, H.: A new look at the statistical model identification. *IEEE Trans. Autom. Contr.* **19**, 716–723 (1974)
2. Akaike, H., Nakagawa, T.: *Statistical Analysis and Control of Dynamic Systems*. Kluwer, Dordrecht (1988)
3. Åström, K.J.: Maximum likelihood and prediction error methods. *Automatica* **16**, 551–574 (1980)
4. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327 (1986), doi:10.1016/0304-4076(86)90063-1
5. Box, G.E.P., Jenkins, G.M.: *Time Series Analysis, Forecasting and Control*, 2. edn. Holden-Day, San Francisco (1976)
6. Durbin, J., Koopman, S.J.: *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, New York (2001)
7. Engle, R.F.: Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**, 987–1008 (1982), doi:10.2307/1912773
8. Galka, A., Yamashita, O., Ozaki, T.: GARCH modelling of covariance in dynamical estimation of inverse solutions. *Physics Letters A* **333**, 261–268 (2004), doi:10.1016/j.physleta.2004.10.045
9. Gupta, N., Mehra, R.: Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculations. *IEEE Trans. Autom. Contr.* **19**, 774–783 (1974), doi:10.1109/TAC.1974.1100714
10. Hamilton, J.D.: *Time Series Analysis*. Princeton University Press, Princeton, New Jersey (1994)
11. Kailath, T.: An innovations approach to least-squares estimation – Part I: Linear filtering in additive white noise. *IEEE Trans. Autom. Control* **13**, 646–655 (1968), doi:10.1109/TAC.1968.1099025
12. Kailath, T.: *Linear Systems*. Information and System Sciences Series. Prentice-Hall, Englewood Cliffs (1980)
13. Kalman, R.E.: A new approach to linear filtering and prediction problems. *J. Basic Engin.* **82**, 35–45 (1960)
14. Lévy, P.: Sur une classe de courbes de l'espace de Hilbert et sur une équation intégrale non linéaire. *Ann. Sci. École Norm. Sup.* **73**, 121–156 (1956)
15. Milton, J.G., Chkhenkeli, S.A., Towle, V.L.: Brain connectivity and the spread of epileptic seizures. In: V.K. Jirsa, A.R. McIntosh (eds.) *Handbook of Brain Connectivity*, pp. 477–503. Springer-Verlag, Berlin, Heidelberg, New York (2007)
16. Ozaki, T., Valdes, P., Haggan-Ozaki, V.: Reconstructing the nonlinear dynamics of epilepsy data using nonlinear time-series analysis. *J. Signal Proc.* **3**, 153–162 (1999)
17. Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J.: Comparing dynamic causal models. *NeuroImage* **22**, 1157–1172 (2004), doi:10.1016/j.neuroimage.2004.03.026
18. Protter, P.: *Stochastic Integration and Differential Equations*. Springer-Verlag, Berlin, Heidelberg, New York (1990)
19. Rauch, H.E., Tung, G., Striebel, C.T.: Maximum likelihood estimates of linear dynamic systems. *American Inst. Aeronautics Astronautics (AIAA) Journal* **3**, 1445–1450 (1965)
20. Schwarz, G.: Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978)
21. Shephard, N.: Statistical aspects of ARCH and stochastic volatility. In: D.R. Cox, D.V. Hinkley, O.E. Barndorff-Nielsen (eds.) *Time Series Models in Econometrics, Finance and Other Fields*, pp. 1–67. Chapman & Hall, London (1996)
22. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleight, J.W., Wilcocks, L.C.: Toward a theory of the general-anesthetic-induced phase transition of the cerebral cortex. I. A thermodynamics analogy. *Phys. Rev. E* **64**, 011917 (2001), doi:10.1103/PhysRevE.64.011917
23. Su, G., Morf, M.: Modal decomposition signal subspace algorithms. *IEEE Trans. Acoust. Speech Signal Proc.* **34**, 585–602 (1986)

24. West, M.: Time series decomposition. *Biometrika* **84**, 489–494 (1997)
25. Wong, K.F.K., Galka, A., Yamashita, O., Ozaki, T.: Modelling nonstationary variance in EEG time-series by state space GARCH model. *Computers Biol. Med.* **36**, 1327–1335 (2006), doi:10.1016/j.combiomed.2005.10.001

Chapter 3

Spatiotemporal instabilities in neural fields and the effects of additive noise

Axel Hutt

3.1 Introduction

The spatiotemporal activity of neural populations may be measured by various experimental techniques, such as optical sensitive-dye imaging [2], multi-unit local field potentials [20] or electroencephalography [41]. Most of this experimentally observed activity results from the interaction of a large number of neurons [56]. Consequently, to describe theoretically such experimental data, the best model choice is a mesoscopic description of populations with a typical spatial scale of few millimeters [55]. Moreover, to understand the underlying dynamics of observed activity, it is important to investigate neural population models that are extended in space. A well-studied mesoscopic population model is the neural field, which assumes a continuous space and may involve various spatial axonal interactions, axonal temporal and spatiotemporal delays, various synaptic time-scales and external inputs. This chapter presents an analysis of such a neural-field model, with the aim of allowing deeper insight into the activity of neural populations.

The chapter derives a basic neural population model (Sect. 3.1.1) and subsequently introduces an extended model involving local and nonlocal spatial interactions subject to different axonal transmission delays. Since the understanding of spatiotemporal activity in neural populations necessitates the knowledge of its basic dynamical properties, such as the linear response to external stimuli, the following sections study the linear stability of the system at the absence of noise (Sect. 3.2). In this context, time-independent and time-dependent phase transitions subject to axonal conduction delays are discussed analytically and numerically. In a subsequent analysis step, the stability and the linear response theory in the presence of noisy inputs is discussed (Sect. 3.3). Finally, Sect. 3.4 considers the nonlinear behavior

Axel Hutt
LORIA, Campus Scientifique-BP 239, 54506 Vandoeuvre-Les-Nancy Cedex, France.
e-mail: axel.hutt@loria.fr

of the system close to an instability, both in the absence of and in the presence of noise.

3.1.1 The basic model

The neural model under discussion may be seen as a *recurrent network* consisting of several functional building elements and assuming a population of densely packed neurons. A recurrent network is one which allows connections from each node to every other node. This topology contrasts to *feed-forward networks* showing sequential information transmission from one sub-network to another, which we do not consider here.

One important element of the model under study is the branching system of axonal fibers which connect neurons to each other. The connection elements between neurons are chemical synapses located on dendritic branches. The postsynaptic potentials generated by the synapses propagate to the neuron body (which is an additional model element), where they are summed to produce the effective membrane potential. When the effective membrane potential exceeds a threshold value, the neuron fires, and action potentials propagate along the axonal branches to downstream chemical synapses adjoining other neurons, and the model circle is closed.

In more detail, chemical synapses bind to dendrites which exhibit passive spread of current. According to this approach, Freeman [24] was one of the first to show experimentally that the action potentials arriving at the synapses convolve mathematically with an impulse response function $h_e(t)$ and $h_i(t)$ at excitatory and inhibitory synapses, respectively, to generate excitatory and inhibitory postsynaptic potentials (PSP), $V^e(t)$ and $V^i(t)$, respectively. Since a dendritic branch of a single neuron typically contacts many synapses (~ 8000 [50]), many PSPs occur in a short time on the dendritic branch and sum up to generate an effective current. If the incoming action potentials are uncorrelated, it is reasonable to interpret the incoming action potentials as an effective population firing activity of high rate. Consequently, considering a short time window Δt , the single action potentials in this time window can be replaced by the mean presynaptic firing rate [34], and the PSPs obey

$$\bar{V}^{e,i}(t) = \int_{-\infty}^t h_{e,i}(t - \tau) \bar{P}_{e,i}(\tau) d\tau, \quad (3.1)$$

with $h_{e,i}(t) = \bar{g}_{e,i}h(t)$. Here, $\bar{V}^{e,i}$ is the time-averaged PSP and \bar{g}_e and \bar{g}_i represent the average synaptic gain of excitatory and inhibitory synapses, respectively. Further, $\bar{P}_e(x,t)$ and $\bar{P}_i(x,t)$ represent the corresponding presynaptic population pulse rate, which terminates at excitatory and inhibitory synapses, respectively. We point out that the time window Δt is longer than the duration of a single action potential but shorter than the typical decay time of chemical synapses, i.e., $\Delta t \approx 2-5$ ms.

The synaptic response behavior is defined by the corresponding synaptic response functions $h(t)$. In general mathematical terms, $h(t)$ is the Green's function

for the temporal operator $\hat{T} = \hat{T}(\partial/\partial t)$ with $\hat{T}h(t) = \delta(t)$, and $\delta(t)$ being the Dirac delta-distribution. Here, $\hat{T}(\partial/\partial t)$ denotes the operator \hat{T} being a function of the operator $\partial/\partial t$, e.g., $\hat{T} = (\partial/\partial t)^2 + (\partial/\partial t) + 1$. Subsequently, the integral equation (3.1) may be formulated as the ordinary differential equation

$$\hat{T}\bar{V}^{e,i}(t) = \bar{g}_{e,i}\bar{P}_{e,i}(t). \quad (3.2)$$

We point out that the PSPs are temporal averages over a short time-scale of few milliseconds and thus the model is coarse-grained in time. Moreover, the model assumes spatial averaging over small patches at the millimeter scale, which correspond to the macrocolumns found experimentally in cortical [30, 54] and sub-cortical [62] areas. Such macrocolumns represent ensembles of interacting neurons, which evolve coherently. They are also called neuronal ensembles or neuronal pools [21, 26].

Further, the model considers variations of synaptic properties in the neuronal population [43] and thus PSPs $V^{e,i}(t)$ at single neurons are random variables with the corresponding probability distributions $p_s^e(V^e - \bar{V}^e)$ and $p_s^i(V^i - \bar{V}^i)$. Since excitatory and inhibitory PSPs sum up at the soma [25], the probability density function $p_S(V - \bar{V})$ of the effective membrane potential $V = V^e - V^i$ is a function of the probability density functions of excitatory and inhibitory PSPs. Here, \bar{V} represents the average effective membrane potential in the population.

When the effective membrane potential exceeds a threshold V_{th} at time t , the neuron generates an action potential. Thus, the probability of a single neuron firing is $\Theta(V(t) - V_{\text{th}}(t))$, where Θ denotes the Heaviside function. For an ensemble of neurons, there is a distribution of firing thresholds $D(V_{\text{th}} - \bar{V}_{\text{th}}, t)$ at time t . Here \bar{V}_{th} denotes the mean firing threshold and $D(V_{\text{th}} - \bar{V}_{\text{th}}, t)dV_{\text{th}}$ is the number of neurons at time t in the interval $[V_{\text{th}}, V_{\text{th}} + dV_{\text{th}}]$, which are not in a refractory period and thus can fire. Consequently D may change in time. Hence the expected number of firing neurons is

$$\begin{aligned} n(t) &= \int_{V_{\min}}^{V_{\max}} dV p_S(V - \bar{V}(t)) \int_{V_\ell}^{V_h} dV_{\text{th}} \Theta(V - V_{\text{th}}) D(V_{\text{th}} - \bar{V}_{\text{th}}, t) \\ &= \int_{V_{\min} - \bar{V}}^{V_{\max} - \bar{V}} dw \int_{V_\ell - \bar{V}_{\text{th}}}^{V_h - \bar{V}_{\text{th}}} du \Theta(w + \bar{V}(t) - \bar{V}_{\text{th}}) p_S(w) D(u, t), \end{aligned}$$

where V_{\min} and V_{\max} are the minimum and maximum values of the membrane potentials, respectively, and V_ℓ and V_h denote the lowest and highest firing thresholds, respectively. Subsequently, the time-averaged pulse activity in the time window Δt at time t reads

$$\begin{aligned} f(t) &= \frac{1}{\Delta t} \int_t^{t+\Delta t} n(\tau) d\tau \\ &\approx \int_{V_{\min} - \bar{V}}^{V_{\max} - \bar{V}} dw p_S(w) \int_{V_\ell - \bar{V}_{\text{th}}}^{V_h - \bar{V}_{\text{th}}} du \Theta(w + \bar{V}(t) - \bar{V}_{\text{th}} - u) \bar{D}(u, t) \quad (3.3) \end{aligned}$$

with the distribution of firing thresholds $\bar{D}(u, t)$ per unit time Δt at time t . Here, $f(t)$ represents the average number of firing neurons per unit time. In other words, $f(t)$ is the population firing rate at time t , and thus Eq. (3.3) is the general definition of the so-called transfer function [1, 34]. Moreover, it can be shown that time-independent Gaussian distributions $p_S = \mathcal{N}(0, \sigma_S^2)$, $\bar{D} = S_m \mathcal{N}(0, \sigma^2)$ and the conditions $V_{\max}, V_h \rightarrow \infty$ and $V_{\min}, V_\ell \rightarrow -\infty$ lead to the well-known transfer function

$$f(t) = S_m \underbrace{\frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\bar{V}(x, t) - \bar{V}_{\text{th}}}{\sqrt{2}\eta_k} \right) \right)}_{S(V(x, t))}. \quad (3.4)$$

Here, $\eta^2 = \sigma_S^2 + \sigma^2$, $\operatorname{erf}(x)$ represents the Gaussian error function, and S_m denotes the maximum firing rate. By virtue of the unimodality of the distribution functions p_S and \bar{D} , $f(t)$ has a sigmoidal shape and the maximum steepness is $S_m/(\sqrt{2\pi}\eta)$. We abbreviate the sigmoidal function by $S(V)$ to indicate its shape. Equation (3.4) shows accordance to the results of previous studies [1].

Typically, the sigmoid function is formulated as $S(V) = S_0/(1 + \exp(-C(V - V_0)))$. This formulation and Eq. (3.4) share their maximum and the locations at the $1/e$ height for the choice $S_0 = -4S_m/\sqrt{\pi} \ln(a)$, $V_0 = \bar{V}_{\text{th}}$ and $C = -\ln(a)/\sqrt{2}\eta$ with $a = 2e - 1 \pm \sqrt{(2e - 1)^2 - 1}$. Hence, the more similar the firing thresholds in the neural populations (i.e., the smaller σ and η), the larger C and thus the steeper the sigmoidal function.

To close the circle of model elements, we consider the axonal connections between neurons which link the somata to dendritic structures of terminal neurons distant in space. By virtue of these spatial connections, the corresponding spatial interactions are nonlocal and yield temporal propagation delays due to the finite axonal propagation speed c . To consider such spatial interactions, we conceive a field of densely packed spatial patches, which are connected according to probability distributions. Hence, the presynaptic pulse activities at excitatory and inhibitory synapses taken from Eq. (3.1) read

$$\bar{P}_e(x, t) = \int_{\Omega} dy K_e(x, y) f(y, t - \frac{|x - y|}{c}) \quad (3.5)$$

$$\bar{P}_i(x, t) = \int_{\Omega} dy K_i(x, y) f(y, t - \frac{|x - y|}{c}), \quad (3.6)$$

in which Ω denotes the spatial domain of the field. Here we assume a single neuron type; the kernels $K_{e,i}$ denote the probability densities of synaptic connections from neurons to synapses of type e or i .

After inserting Eqs (3.5), (3.6) into Eq. (3.2), the effective membrane potential $\bar{V} = \bar{V}^e - \bar{V}^i$ obeys

$$\hat{T}V(x, t) = \int_{-\infty}^{\infty} dy [a_e K_e(x - y) - a_i K_i(x - y)] S \left(V(y, t - \frac{|x - y|}{v_i}) \right) + I(x, t) \quad (3.7)$$

with $a_{e,i} = \bar{g}_{e,i} S_{\max}$. Moreover, the kernels are chosen symmetric and homogeneous, i.e., $K_{e,i}(x, y) = K_{e,i}(x - y) = K_{e,i}(|x - y|)$, and we choose the spatial domain Ω to be the real line if not stated otherwise. Equation (3.7) also considers an external input $I(x, t)$ which may originate from other neural populations or from external stimulation.

3.1.2 Model properties and the extended model

In previous studies, different variations and extensions of the basic model (3.7) have been investigated [4, 10, 12, 36, 68, 70]. The average membrane potential in our model obeys an integral-differential equation (IDE), while other neural population models are formulated as partial differential equations (PDE) [8, 59, 65, 72, 73]. In recent years several studies have shown the relationship between the two model types, concluding that the IDE-model generalizes the PDE-models [13, 31, 33, 34].

To illustrate this relation, assume a one-dimensional spatial field $u(x)$ with spatial kernel function $K(x)$. Then the relation

$$\int_{-\infty}^{\infty} K(x-y)u(y)dy = \int_{-\infty}^{\infty} \tilde{K}(k)\tilde{u}(k)e^{ikx}dk \quad (3.8)$$

holds with the Fourier transforms \tilde{K} and \tilde{u} of K and u , respectively. Specifying $K(x) = \exp(-|x|/\sqrt{D})/2\sqrt{D}$, one finds $\tilde{K}(k) = 1/(1 + Dk^2) = 1 - Dk^2 + D^2k^4 + \dots$ and Eq. (3.8) reads

$$\int_{-\infty}^{\infty} K(x-y)u(y)dy \approx u(x, t) + D\Delta u(x, t) + D^2\Delta^2 u(x, t) + \dots$$

with $\Delta = \partial^2/\partial x^2$. This means that fast-decaying integral kernels $K(x)$ with $\sqrt{D} \ll 1$ represent a local diffusion process with diffusion constant D , while slowly-decaying integral kernels represent higher orders of spatial derivatives, and thus reflect nonlocal interactions of a long range. Following this approach, extended types of integral-differential equations generalize four types of PDE-models, namely wave equations [49], reaction-diffusion models [13, 31], the Swift–Hohenberg equation and the Kuramoto–Sivashinsky equation [31]. These well-studied PDE-models describe pattern propagation in physical and chemical systems [14].

To capture several aspects of spatially extended neural populations in one model, we consider an extended scalar IDE-model which allows various types of spatial interactions and two propagation speeds:

$$\begin{aligned} \hat{T}V(x, t) = & g[V(x, t)] + I(x, t) + \int_{\mathfrak{R}} a_K K(x-y)S \left[V(y, t - \frac{|x-y|}{c_K}) \right] \\ & + a_L L(x-y)S \left[V(y, t - \frac{|x-y|}{c_L}) \right] dy. \end{aligned} \quad (3.9)$$

Here the functional $g[V(x,t)]$ defines the local dynamics at spatial location x , and the kernel functions K, L represent the probability density functions of spatial connections in two networks, i.e., $K(x)$ and $L(x)$ are normalized to unity in the spatial domain. Moreover, a_K and a_L are the total amount of activity in the corresponding networks. The functional $S[V]$ represents the firing-rate function derived in Sect. 1.1. Further, the model considers propagation delays between spatial locations x and y at distance $|x - y|$ due to the finite transmission speed c_K and c_L corresponding to the two networks.

This model allows the study of a wide range of combined spatial interaction. For instance, spatial systems involving both mean-field and local interactions may be modeled by $K(x) = \text{const.}$ and $L(x) = \delta(x)$, respectively, while the case $a_K > 0$, $a_L < 0$ represents excitatory and inhibitory interactions.

In a recent work, Venkov and Coombes introduced an even more general spatiotemporal scalar model which captures several different integral-differential models [68]. Moreover, other previous studies investigated scalar models in two dimensions [47, 57] and models involving two IDEs [7, 29, 40, 48].

The following sections aim to illustrate the analysis of IDEs, and thus focus on the more specific model Eq. (3.9). Further, we study the interaction of excitation and inhibition, i.e., $a_K > 0$ and $a_L \rightarrow -a_L$, $a_L > 0$.

3.2 Linear stability in the deterministic system

First let us determine the stationary state of Eq. (3.9) constant in space and time. Applying a constant external input $I(x,t) = I_0$, we find the implicit equation $V_0 = g(V_0) + (a_K - a_L)S(V_0) + I_0$. Then small deviations $u(x,t)$ about the stationary state obey a linear evolution equation and evolve in time according to $u(x,t) \sim \exp(\lambda t)$ with the complex Lyapunov exponent $\lambda \in \mathcal{C}$. If the exponent's real part is negative, i.e., $\text{Re}(\lambda) < 0$, then the small deviations relax to the stationary state, while $\text{Re}(\lambda) > 0$ yields divergent activity for large times.

The subsequent decomposition of $u(x,t)$ into the continuous Fourier basis $\{e^{ikz}\}$ yields the implicit condition

$$T(\lambda) = g_0 + S' \int_{\mathfrak{R}} \left(a_K K(y) e^{-\lambda|y|/c_K} - a_L L(y) e^{-\lambda|y|/c_L} \right) e^{iky} dy, \quad (3.10)$$

with $g_0 = \partial g / \partial V$, $S' = \partial S / \partial V$ computed at $V = V_0$. Here, the function $T(\lambda)$ is gained by replacing in \hat{T} the operators $\partial / \partial t$ by λ . Since g_0 and S' depend on the stationary state, which in turn depends on the external stimulus, the Lyapunov exponents depend strongly on the external stimulus and consequently I_0 is a reasonable control parameter.

Equation (3.10) is difficult to solve exactly for λ due to the finite propagation speeds. To gain some insights into the stability of the system, we consider large but finite propagation speeds [3] and find $\exp(-\lambda|z|/c) \approx 1 - \lambda|z|/c$ for $\lambda|z|/c \ll 1$. Since $\exp(-\lambda|z|/c)$ occurs in the integral in Eq. (3.10), $\tau = |z|/c$ is interpreted as

the relation of the characteristic spatial scale of the kernel and the maximum speed of activity, i.e., the characteristic propagation delay. Hence this approximation is valid if $1/\lambda \gg \tau$, i.e., the time-scale of the evolving system is much larger than the propagation delay. To get more insight to this approximation, we recall that most sensory and cognitive processes evolve in the frequency band 5–80 Hz, i.e., on time scales 12.5–200 ms. Since $1/\lambda$ represents the time-scale of the system, we find $12.5 < 1/\lambda < 200$ ms. Moreover, assuming the lateral interactions between macro-columns have a spatial range of 2 mm with typical intracortical axonal propagation speed of 0.5 m/s, the characteristic propagation delay is $\tau = (2/0.5)$ ms = 4 ms, much shorter than the system time-scale, so the approximation condition is valid.

Considering the latter approximation, Eq. (3.10) reads

$$T(\lambda) + \lambda S' \tilde{M}_1(k) \approx g_0 + S' \tilde{M}_0(k) \quad (3.11)$$

with

$$\tilde{M}_0(k) = a_K \tilde{K}^{(0)}(k) - a_L \tilde{L}^{(0)}(k), \quad \tilde{M}_1(k) = \frac{a_K}{c_K} \tilde{K}^{(1)}(k) - \frac{a_L}{c_L} \tilde{L}^{(1)}(k)$$

and the kernel Fourier moments $\tilde{M}^{(n)}(k) = \int_{-\infty}^{\infty} M(z) |z|^n \exp(-ikz) dz$ [3]. The term $\tilde{M}_0(k)$ represents the Fourier transform and $\tilde{M}_1(k)$ is the first kernel Fourier moment of the effective kernel function $M(x) = a_K K(x) - a_L L(x)$.

Now let us specify the temporal operator \hat{T} . Typically the synaptic response function $h(t)$ has two time scales, namely the fast rise-time τ_2 and the slower decay-time τ_1 . We choose $h(t) = (\exp(-t/\tau_1) - \exp(-t/\tau_2))/(\tau_1 - \tau_2)$, and scale the time by $t \rightarrow \sqrt{\tau_1 \tau_2}$ to gain

$$\hat{T} = \frac{\partial^2}{\partial t^2} + \gamma \frac{\partial}{\partial t} + 1$$

with $\gamma = \sqrt{\tau_2/\tau_1} + \sqrt{\tau_1/\tau_2}$, and we find $\gamma \geq 2$.

Now let us discuss the occurrence of instabilities. For vanishing input $I_0 = 0$, we assume that the system is stable, i.e., $\text{Re}(\lambda) < 0$. Then, increasing I_0 , the Lyapunov exponent may change its sign at $\text{Re}(\lambda) = 0$ and thus the stationary state becomes unstable and the system approaches a new state. In this section we describe the different types of stability loss, but do not compute the new state approached by the system in case of an instability. This is discussed in Sect. 3.4.

Consequently, it is sufficient to study two cases: $\lambda = 0$, which is the stability threshold for *static* (or *Turing*) instabilities; and $\lambda = i\omega$, defining the stability threshold for *nonstationary* instabilities oscillating with frequency ω . Specifically, the corresponding conditions are obtained from Eq. (3.11) to be

$$\frac{1 - g_{0,c}}{S'_c} = \tilde{M}_0(k_c) \quad (\text{stationary instability}), \quad (3.12)$$

$$\frac{\gamma}{S'_c} = -\tilde{M}_1(k_c) \quad (\text{nonstationary instability}). \quad (3.13)$$

In other words, the global maximum of $\tilde{M}_0(k_c)$ and $-\tilde{M}_1(k_c)$ define the critical wavenumber k_c and the critical control parameter via S'_c and $g_{0,c}$ of stationary and nonstationary instabilities, respectively. Figure 3.1 illustrates this result for nonstationary instabilities. Consequently, Eqs (3.12) and (3.13) give the rule for the emergence of instabilities: increasing the control parameter from low values (where the system is stable), the left-hand sides of Eqs (3.12), (3.13) change, and the condition that holds first determines the resulting instability.

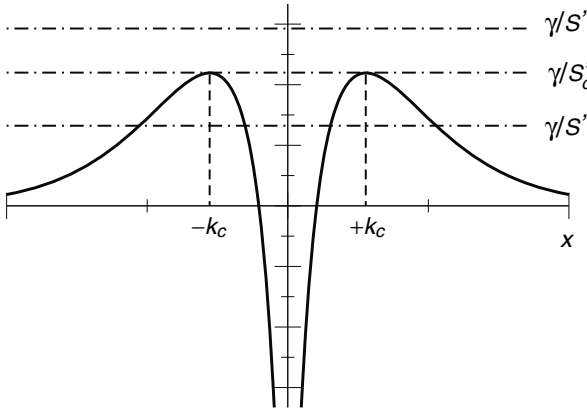


Fig. 3.1 Sketch to illustrate the mechanism of nonstationary instabilities. The plot shows the right-hand side of Eq. (3.13) (solid line) and its left-hand side (dot-dashed line). If the stability threshold is not reached, i.e., $S' < S'_c$, the system is stable, while $S' = S'_c$ ($S' > S'_c$) reflects a marginally stable (unstable) system.

3.2.1 Specific model

To gain more insight, we neglect local interactions, i.e., $g = 0$, and the spatial connection kernels are chosen to be,

$$K(x-y) = \frac{1}{2r_e^p \Gamma(p)} |x-y|^{p-1} e^{-|x-y|/r_e}, \quad L(x,y) = \frac{1}{2r_i} e^{-|x-y|/r_i}, \quad (3.14)$$

where $p > 0$ is a parameter, $\Gamma(p)$ denotes the gamma function and r_e, r_i are the spatial ranges of excitatory and inhibitory connections, respectively. This choice of kernels allows for studies of manifold interactions. For instance, $p = 1$ represents axonal connections which are maximal at zero distance and monotonically decreasing for increasing distance. This combination of excitatory and inhibitory axonal interaction may yield four different spatial interactions, namely pure excitation, pure inhibition, local excitation–lateral inhibition (Mexican hat) and local inhibition–lateral excitation (inverse Mexican hat). Moreover, $p > 1$ represents the

case of zero local excitation but strong lateral excitation which has been found in theoretical [50] and experimental studies [56]. The decreasing spatial connectivity of inhibitory connections is motivated by successful theoretical studies [64, 69] and the strong anatomical evidence for inhibitory self-connections in cat visual cortex [66]. In the case of $p < 1$, the probability of local excitations is infinite. Though infinite probability densities exist in some statistical systems [22], this type of interaction has not been found yet in experiments. However its study yields interesting theoretical effects as discussed in Sect. 3.2.3.

After scaling the space by $x \rightarrow x/r_e$, we obtain the kernel functions

$$K(x) = \frac{1}{2\Gamma(p)} |x|^{p-1} e^{-|x|}, \quad L(x) = \frac{1}{2\xi} e^{-|x|/\xi} \quad (3.15)$$

with $\xi = r_i/r_e$. Hence the spatial interaction is governed by ξ which represents the relation of inhibitory and excitatory spatial range, and by p reflecting the degree of excitatory self-interaction. Figure 3.2 shows the kernel functions and four major resulting types of spatial interaction.

3.2.2 Stationary (Turing) instability

For the specific choice of kernels (3.15), the condition for the Turing instability (3.12) is $S'_c = 1/\tilde{M}_0(k_c)$. Since $\tilde{M}_0(k) \rightarrow 0$ as $k \rightarrow \infty$, $\tilde{M}_0(k_c)$ exhibits a positive local maximum for finite k_c [34]. Figure 3.3(a) shows the parameter space of the condition for a local maximum, and Fig. 3.3(b) presents some example kernel functions and the corresponding Fourier transforms. We observe that $p = 1$ yields local excitation–lateral inhibition interaction for $\xi > 1$, i.e., for $r_i > r_e$ (open triangles in Fig. 3.3(b)), and local inhibition–lateral excitation for $\xi < 1$, i.e., $r_i < r_e$ (filled triangles in Fig. 3.3(b)). In addition, $\xi > 1$ yields a maximum of the Fourier transform $\tilde{M}_0(k)$ at $k_c \neq 0$ which allows for a Turing instability. This finding is consistent with the well-known idea that small excitation ranges yield an enhancement of the local activity, while the larger range of inhibition diminishes the lateral spread of activity. Consequently one observes local bumps. However, this idea no longer holds for $p = 2$. Figure 3.3(b) shows that both values of $\xi > 1$ (open and filled diamonds) represent local inhibition–lateral excitation interaction, which however are expected to reflect local excitation–lateral inhibition according to the standard idea of spatial interactions. Further, the Fourier transform $\tilde{M}_0(k)$ exhibits a local maximum at $k_c \neq 0$ for the larger value of ξ but not for the lower value though both parameters reflect similar spatial interaction. In other words, we find Turing patterns for local inhibition–lateral excitation interaction [34]. Hence the value of p and thus the shape of the excitatory interaction plays an important role here.

Fig. 3.4 presents the simulated Turing instability in the neural population for both $p = 1$ and $p = 2$.

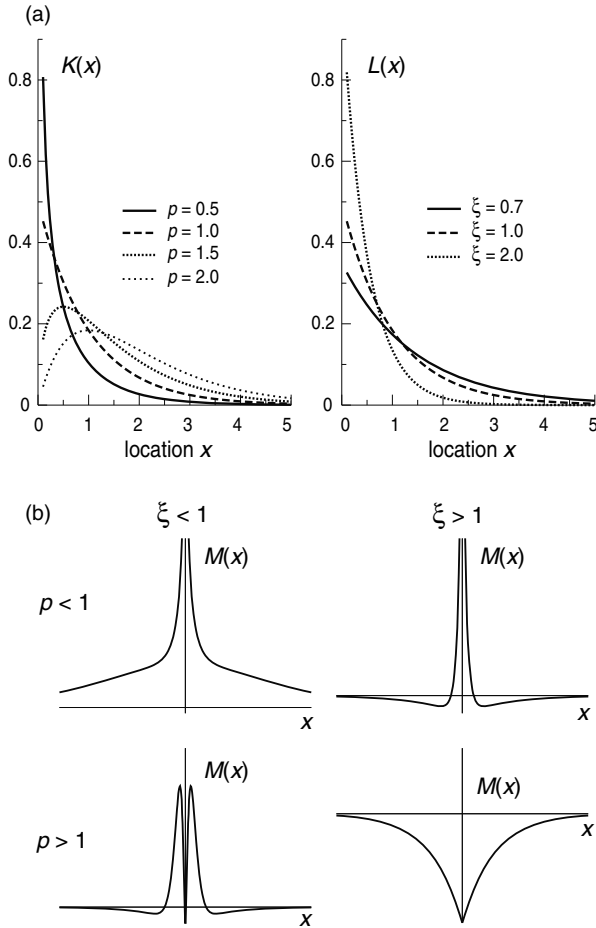


Fig. 3.2 Illustration of spatial interaction types for different parameters p and $\xi = r_i/r_e$. Panel (a) shows the kernel functions $K(x)$ and $L(x)$ from Eq. (3.15) while (b) presents four major interaction types with the effective kernel function $M(x) = K(x) - L(x)$. More specific types are discussed in Sect. 3.2.3.

Eventually the question arises whether Turing bifurcations are possible in the brain. To this end, we recall experiments which indicate intracortical inhibitory connections of a spatial range $r_i \approx 1$ mm and cortico-cortical connections of a range of $r_e = 20$ mm, i.e., $\xi = 0.05$, $p = 3$ and $a_K > a_L$ [56]. Consequently Turing patterns do not occur for such cortico-cortical interactions according to Fig. 3.3. However different spatial scales and connectivity distribution functions may yield Turing pattern formation in the brain and we mention previous studies which argue that Turing phase transitions may be responsible for the pattern formation in populations with intra-cortical connectivity, such as the periodic anatomical structure of the visual cortex [42, 67, 71].

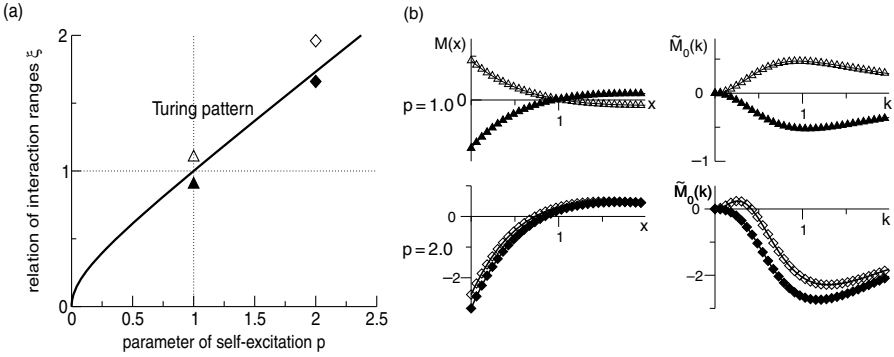


Fig. 3.3 Turing bifurcations for various parameter sets. Panel (a) shows the sufficient parameter regimes of Turing patterns; (b) presents the effective kernel function $M(x)$ for the four parameter sets given in (a) and shows the corresponding Fourier transforms $\tilde{M}_0(k)$. The values of the four parameter sets are $p = 1.0$, $\xi = 1.1$ (open triangle), $p = 1.0$, $\xi = 0.9$ (filled triangle), $p = 2.0$, $\xi = 1.96$ (open diamond) and $p = 2.0$, $\xi = 1.66$ (filled triangle). Further parameters are $a_K = a_L = 5$.

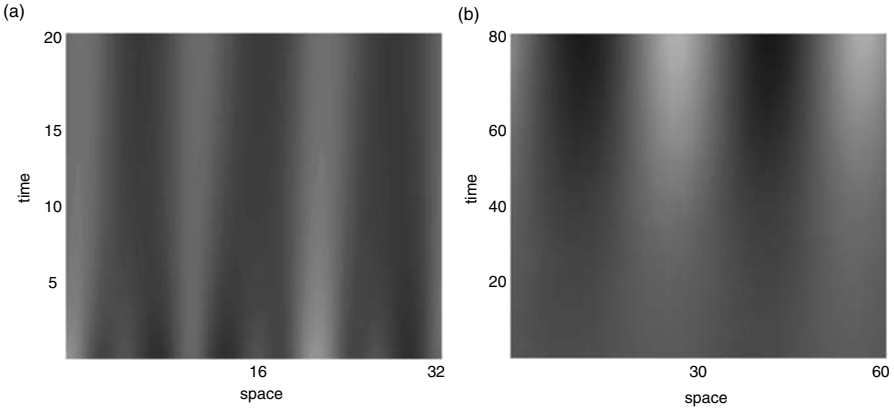


Fig. 3.4 Space-time plots of Turing instabilities. Panel (a) shows the case of local excitation–lateral inhibition interaction, and (b) for local inhibition–lateral excitation interaction. Parameters are (a) $p = 1.0$, $\xi = 2.0$, $a_K = 6.0$, $a_L = 5.0$, $c_K = c_L = 10.0$, $I_0 = 2.36$; (b) $p = 2.0$, $\xi = 1.92$, $a_K = 131.0$, $a_L = 130.0$, $c_K = c_L = 10.0$, $I_0 = 2.2$. The grayscale encodes the deviation from the stationary state, and the theoretical values k_c show good accordance with the k_c observed.

3.2.3 Oscillatory instability

Now we focus on time-dependent instabilities which occur if Eq. (3.13) holds. In order to obtain the critical control parameter and the critical wavenumber and oscillatory frequency, we compute the first kernel Fourier moment in Eq. (3.13). Since $\tilde{M}_1(k) \rightarrow 0$ for $|k| \rightarrow \infty$, there is a maximum of $-\tilde{M}_1(k)$ at some $|k| = |k_c|$ with $-\tilde{M}_1(k_c) > 0$ if $-\tilde{M}_1(0) > 0$, $-d^2\tilde{M}_1(0)/dk^2 > 0$ [33]. This estimation yields necessary conditions for wave instabilities shown in Fig. 3.5.

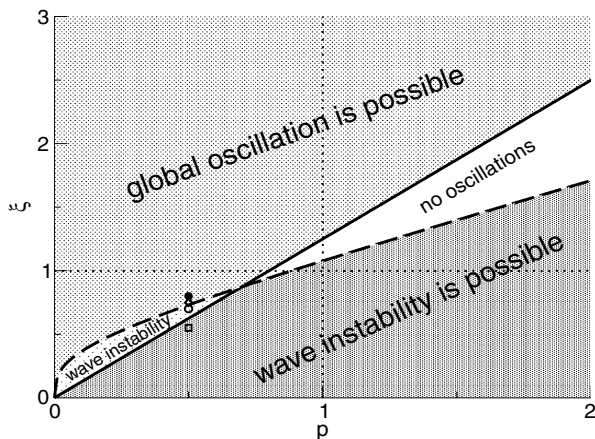


Fig. 3.5 The occurrence of oscillatory instabilities subject to the parameters ξ and p .

For $p = 1$, according to the standard idea of spatial interactions we expect traveling waves for local inhibition–lateral excitation interaction, i.e., $\xi = r_i/r_e < 1$, and global oscillatory activity for local excitation and lateral inhibition, i.e., $\xi > 1$. Figure 3.6 shows the corresponding spatiotemporal activity and thus confirms this idea.

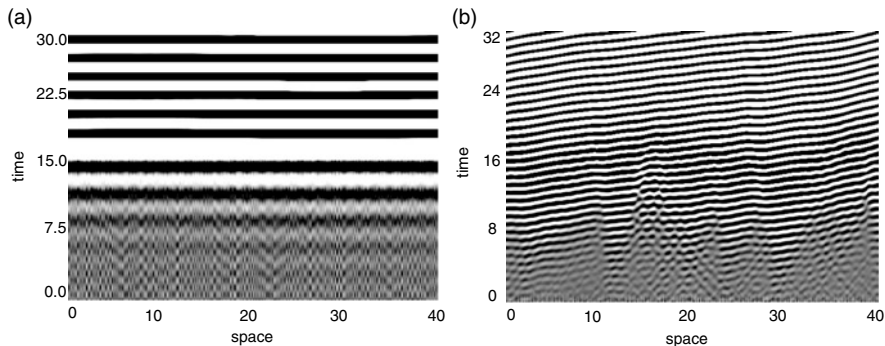


Fig. 3.6 Space-time plots of (a) global oscillations, and (b) traveling waves for $p = 1$. The panels show the deviations from the stationary state for (a) $\xi = 5$, (b) $\xi = 0.33$. Modified from [3].

We have seen in the previous section that the sign of the kernel function itself does not reflect all properties of the system. Though we have found that the Fourier transform determines the occurrence of Turing instabilities, we may still ask which feature of the kernel function might indicate the corresponding instability. To investigate this question for the case of oscillatory instabilities, we consider the number of roots of the kernel function as a possible criterion. To illustrate the relation of the spatial interaction type and the resulting nonstationary instability, Fig. 3.7 plots the

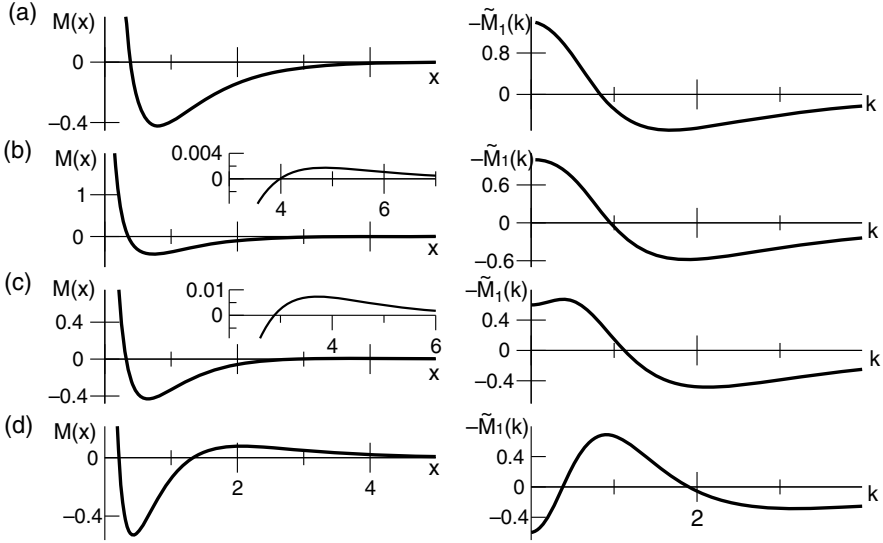


Fig. 3.7 The kernel functions $M(x)$ and the negative first kernel Fourier moment $-\tilde{M}_1(k)$ for different values of ξ and $p = 0.5$. (a) $\xi = 0.8$, (b) $\xi = 0.75$, (c) $\xi = 0.7$ and (d) $\xi = 0.55$. The insets in the left column represent foci of the corresponding plot to illustrate the additional lateral excitation. Modified from [33].

kernel functions $M(z)$ and the corresponding negative first kernel Fourier moments $-\tilde{M}_1(k)$ for four different cases of spatial interactions with $p = 0.5$. We find local excitation–lateral inhibition for $\xi = 0.8$ yielding global oscillations with $k_c = 0$. For $\xi = 0.75$, the kernel function exhibits two roots, i.e., local and long-range excitation and mid-range inhibition, which still leads to global oscillations. In case of $\xi = 0.70$, the sufficient condition for wave instabilities is fulfilled, as $-\tilde{K}_1(k)$ reveals a critical wavenumber $k_c \neq 0$ (Fig. 3.7(c)). For $\xi = 0.55$ there is also a wave instability showing local excitation, mid-range inhibition and lateral excitation. Figure 3.8 shows the spatiotemporal evolution of the oscillatory instability for the parameters of Fig. 3.7(c).

Concluding, the number of roots and the sign of the spatial interaction function between the roots do *not* indicate the type of oscillatory instability, i.e., constant oscillation of traveling wave, as assumed in many previous studies. In contrast, first kernel Fourier moment contains all necessary information, and its study elucidates the system phase transitions.

Now the question arises whether one could understand the previous result in physical terms. Following the idea of the interplay between excitation and inhibition, the spatial kernel for $\xi = 0.8$ enhances strongly local activity and diminishes the activity at distant spatial locations, see Fig. 3.7. Hence we might expect stationary activity. Now decreasing ξ , the lateral inhibition increases while the local excitation decreases and a long-range excitation occurs. Consequently, the major

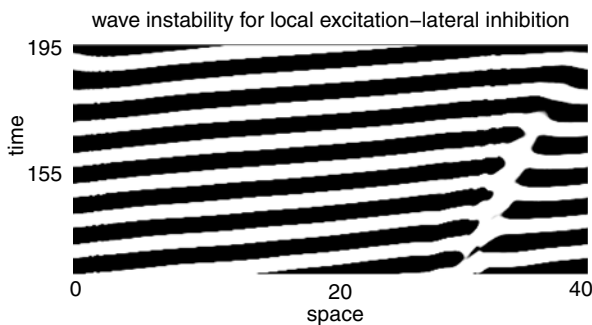


Fig. 3.8 Space-time plot showing a wave instability for local excitation–lateral inhibition interactions. The white kink represents an irregularity originating from the initial conditions and dying out for large times. Here parameters are $p = 0.5$, $\xi = 0.7$. Modified from [33].

spatial interactions in the system are the mid-range inhibition and the long-range excitation while the local excitation is negligible. This interplay of excitation and inhibition yields traveling waves. However, this interpretation may fail if the propagation speed is large enough and condition (3.13) does not hold anymore. In addition we point out that the spatial scale of the instability is given by the maximum of $-\tilde{K}_1$ and thus depends on the specific shape of the interaction kernels in a non-trivial way.

3.3 External noise

Now we discuss the effect of random fluctuations on the dynamics of neural populations. Several previous studies showed that randomness may be beneficial to the information processing in the brain [11, 15–17, 51–53]. The origin of spontaneous fluctuations in the living brain, i.e., fluctuations unrelated to external stimuli, is poorly understood though there is evidence for stochastic ion channel dynamics in the cell membrane and spontaneous activity failure in synapses [45]. The latter two examples of fluctuations represent random changes of neuron properties over space and time. Since most of these properties are formulated mathematically as variables multiplied to the activity variable in the model, the spontaneous random fluctuations yield so-called multiplicative noise.

Besides these spontaneous, or internal, fluctuations one may consider external fluctuations, which may originate from other neural populations or from external stimuli [16, 17, 53]. In the following we study this type of fluctuation for stable systems close to the deterministic instabilities discussed in the previous section. For simplicity we neglect propagation delay effects, i.e., $c_K, c_L \rightarrow \infty$ in Eq. (3.9), and refer the reader to previous work [37] for the discussion of propagation delay.

In this and in the following sections, we assume a finite spatial domain Ω with periodic boundary conditions and a temporal operator $\hat{T} = \partial/\partial t + 1$. Then the model Eq. (3.9) reads

$$\frac{\partial V(x,t)}{\partial t} = H[V](x,t) + I(x,t) \quad (3.16)$$

with

$$H[V](x,t) = -V(x,t) + g[V(x,t)] + \int_{\Omega} M(x-y)S[V(y,t)] dy$$

and with the spatial connectivity function $M(x) = a_K K(x) + a_L L(x)$. In contrast to the previous section, now the external input $I(x,t)$ is the sum of a constant input I_0 and random Gaussian fluctuations $\xi(x,t)$, i.e., $I(x,t) = I_0 + \xi(x,t)$ with $\langle \xi(x,t) \rangle = 0$ where $\langle \cdot \rangle$ denotes the ensemble average.

Since the spatial domain is a finite circle, we introduce a discrete Fourier expansion with the basis $\{\exp(-ik_n x)/\sqrt{|\Omega|}\}$, $k_n = n2\pi/|\Omega|$ and the orthogonality condition

$$\frac{1}{|\Omega|} \int_{\Omega} e^{i(k_m - k_n)x} dx = \delta_{nm}.$$

Then $V(x,t)$ may be written as

$$V(x,t) = \frac{1}{\sqrt{|\Omega|}} \sum_{n=-\infty}^{\infty} u_n(t) e^{ik_n x} \quad (3.17)$$

with the corresponding Fourier projections $u_n(t) \in \mathcal{C}$, $u_n = u_{-n}$.

Inserting the Fourier projection (3.17) into Eq. (3.16) yields the stochastic differential equations

$$du_n(t) = \tilde{H}_n[\{u_j\}]dt + \frac{1}{\sqrt{|\Omega|}} \int_{\Omega} dx d\Gamma(x,t) e^{-ik_n x}, \quad -\infty < n < \infty, \quad (3.18)$$

where $\tilde{H}_n[\cdot]$ denotes the Fourier transform of $H[\cdot]$ and $d\Gamma(x,t)$ are the general random fluctuations. Equation (3.18) represents a system of infinitely many stochastic differential equations. Now the random fluctuations $d\Gamma(x,t)$ are assumed to represent a superposition of independent fluctuating sources with

$$d\Gamma(x,t) = \int_{\Omega} dy c(x,y) dW(y,t), \quad (3.19)$$

where $c(x,y)$ is a weight kernel function [39]. The terms $dW(y,t)$ represent the differentials of independent Wiener processes satisfying

$$\langle dW(y,t) \rangle = 0, \quad \langle dW(x,t) dW(y,t') \rangle = 2\delta(x-y)\delta(t-t') dt dt'.$$

In Sects. 3.3.2 and onwards, we shall specify $c(x, y) = \eta' \delta(y)$ leading to $d\Gamma(x, t) = \eta' dW(0, t) = dW(t)$. In other words, at time t the random fluctuations are the same at each spatial location and we refer to this fluctuation type as global fluctuations [32, 38]. Consequently Eq. (3.18) reads

$$du_n(t) = \tilde{H}_n[\{u_j\}]dt, \quad \forall n \neq 0 \quad (3.20)$$

$$du_0(t) = \tilde{H}_0[\{u_j\}]dt + \eta dW(t) \quad (3.21)$$

with $\eta = \eta' \sqrt{|\Omega|}$. Hence all spatial modes $u_{n \neq 0}(t)$ obey the deterministic ordinary differential equation (3.20) and the dynamics of $u_0(t)$ is governed by the stochastic differential equation (3.21) subjected to the Wiener process $dW(t)$. The corresponding Fokker–Planck equation of the system (3.20), (3.21) reads

$$\frac{\partial P(\{u_n\}, t)}{\partial t} = - \sum_j \frac{\partial}{\partial u_j} \tilde{f}[\{u_n\}] P(\{u_n\}, t) + \eta^2 \frac{\partial^2}{\partial u_0^2} P(\{u_n\}, t). \quad (3.22)$$

3.3.1 Stochastic stability

In Sect. 3.2, we determined the stationary state constant in space and time and studied the Lyapunov exponent of the linearized problem about the stationary state. At a first glance this approach does not apply here since the system is time-dependent for all times due to the external random fluctuations. Hence the question arises: How can one define a stationary state in this case? At first let us recall the definition of stability. Several definitions of stability exist [46, 61], such as asymptotic stability considered in the previous section, or the mean-square stability: if the deviation about a system state is $u(x, t)$ and $\langle |u(x, t)|^2 \rangle < \delta$ holds for $\delta > 0$, then this system state is called mean-square stable. If in addition $\lim_{t \rightarrow \infty} \langle |u(x, t)|^2 \rangle \rightarrow 0$, then the state is called asymptotically mean-square stable. In other words, a system state may be called stable if the system evolves in a small neighborhood of this state. Hence we argue that the stable system state might be equivalent to the deterministic stable stationary state V_0 and the system evolves in a small neighborhood about it due to the external random fluctuations. To quantify this, we obtain the (presumed) stationary state for $I(x, t) = I_0$ and gain from Eq. (3.16)

$$V_0 = g[V_0] + (a_k + a_L)S(V_0) + I_0.$$

Up to this point, the stability of V_0 has yet to be confirmed. The subsequent paragraphs study the dependence of the system dynamics about this state on Gaussian-distributed random fluctuations

$$\langle \xi(x, t) \xi(x', t') \rangle = Q \delta(x - x') \delta(t - t') \quad (3.23)$$

and answer the question of stability. From Eq. (3.23) the random fluctuations have variance Q and are uncorrelated in space and time.

Further presuming the fluctuation variance Q being small may yield small deviations from the stationary state $u(x,t) = V(x,t) - V_0 \ll V_0$ and we obtain

$$\hat{T}(\partial/\partial t)u(x,t) = g_0u(x,t) + S' \int_{\Omega} M(x-y)u(y,t) dy + \xi(x,t) \quad (3.24)$$

with $g_0 = \delta g/\delta V$, $S' = \delta S/\delta V$ computed at $V = V_0$. In the following paragraphs, we shall use S' as the control parameter. Now let us study Eq. (3.24) in the Fourier space similar to the deterministic case. With the Fourier transformation Eq. (3.17) we find for each spatial mode the stochastic differential equation

$$\hat{T}\tilde{u}(k,t) = \left(g_0 + \sqrt{\Omega}S'\tilde{M}(k)\right)\tilde{u}(k,t) + \tilde{\xi}(k,t). \quad (3.25)$$

The random fluctuations $\tilde{\xi}(k,t)$ of the spatial mode with wave number k obey a Gaussian distribution with

$$\langle \tilde{\xi}(k,t) \rangle = 0 \quad , \quad \langle \tilde{\xi}^*(k,t)\tilde{\xi}(k',t') \rangle = Q\delta(k-k')\delta(t-t'), \quad (3.26)$$

i.e., the $\tilde{\xi}(k,t)$ are uncorrelated in k -space and in time.

Now interpreting $\tilde{\xi}(k,t)$ as an external perturbation, linear response theory gives the general solution of Eq. (3.25) by

$$\tilde{u}(k,t) = \tilde{u}_h(k,t) + \int_{-\infty}^{\infty} dt' G(k,t-t')\tilde{\xi}(k,t'). \quad (3.27)$$

Here, $\tilde{u}_h(k,t)$ represents the homogeneous solution of Eq. (3.25), i.e., for $\tilde{\xi}(k,t) = 0$, and $G(k,t-t')$ is the Green's function of the spatial mode with wave number k . Applying standard techniques in linear response theory, the Green's function reads

$$G(k,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \frac{e^{-i\omega t}}{T(-i\omega) - \left(g_0 + \sqrt{\Omega}S'\tilde{M}(k)\right)}. \quad (3.28)$$

After applying the residue theorem, we obtain essentially

$$u(x,t) = \tilde{u}_h(x,t) + \frac{1}{\sqrt{\Omega}} \sum_{l=1}^m \int_0^t \int_{-\infty}^{\infty} e^{\lambda_l(k)(t-t')} \tilde{s}(k,t') r_l(k) e^{ikx} dk dt', \quad (3.29)$$

assuming the initial time $t = 0$ and $r_l(k)$ are constants. In addition, we introduce $\lambda_l = i\Omega_l$ with Ω_l as the roots of $T(-i\Omega) - \left(g_0 + \sqrt{\Omega}S'\tilde{M}(k)\right) = 0$ in Eq. (3.28) and m is the number of roots. Further, we point out the important fact that λ_l are the Lyapunov exponents of the homogeneous system and thus define its (deterministic) stability.

Recalling the definition of mean-square stability and utilizing Eq. (3.26), we find

$$\begin{aligned}
\langle |u^2(x,t)| \rangle &< |\tilde{u}_h(x,t)|^2 + \frac{Q}{2\pi} \sum_{l,n} \int_{-\infty}^{\infty} r_l^*(k) r_n^*(k) \frac{1 - e^{(\lambda_l(k) + \lambda_n(k))t}}{\lambda_l(k) + \lambda_n(k)} dk \\
&< |\tilde{u}_h(x,t)|^2 + \frac{Q\eta}{2\pi} \sum_{l,n} \int_{-\infty}^{\infty} \frac{|1 - e^{(\lambda_l(k) + \lambda_n(k))t}|}{|\lambda_l(k) + \lambda_n(k)|} dk \\
&< |\tilde{u}_h(x,t)|^2 + \frac{Q\eta}{2\pi} \sum_{l,n} \int_{-\infty}^{\infty} \frac{1}{|\lambda_l(k) + \lambda_n(k)|} dk
\end{aligned}$$

for bounded values $|r_l^*(k) r_n^*(k)| < \eta$, $\eta > 0$ and $\text{Re}(\lambda_l) < 0$. Since $\text{Re}(\lambda_l) < 0$ implies stability and thus the boundedness of the homogeneous solution, we find finally

$$\langle |u^2(x,t)| \rangle < \delta,$$

i.e., mean-square stability and thus bounded areas of the stochastic system. We point out that the system is not asymptotically mean-square stable.

We conclude that the mean-square stability of the (deterministic) stationary state V_0 in the presence of external additive fluctuations is given by the Lyapunov exponents obtained from the deterministic case.

3.3.2 Noise-induced critical fluctuations

Now let us study the statistical properties of the activity about the stationary state V_0 for global fluctuations. Equation (3.16) yields the implicit equation $H[V_0] = 0$ and small deviations $z(x,t) = V(x,t) - V_0$ obey

$$dz(x,t) = \left((-1 + g_0)z(x,t) + S' \int_{\Omega} M(x-y)z(y,t) dy \right) dt + \eta dW(t). \quad (3.30)$$

Applying the Fourier transformation to Eq. (3.30), we obtain

$$du_n(t) = \alpha_n u_n(t) dt + \eta \delta_{n,0} dW_n(t) \quad (3.31)$$

with $\alpha_n = -1 + g_0 + S' \sqrt{\Omega} \tilde{M}(k)$, $\alpha_n = \alpha_{-n}$ and the Fourier transform $u_n(t)$ of $z(x,t)$. To study the stability of Eq. (3.31), we use the results gained in the previous section and find the Lyapunov exponent $\lambda = \alpha_n$. Since the kernel functions and the functional $g[V]$ take real values, λ is real-valued and no oscillatory activity is present. Further the system is mean-square stable if $\alpha_n < 0$.

The system stability is lost if some spatial modes with wavenumber k_n satisfy $\alpha_n = 0$. In this case, the system becomes unstable and exhibits a non-oscillatory instability (or phase transition in physical terms) with the critical wavenumber $k_c = \arg \max_{k_n} (-1 + g_0 + S' \sqrt{\Omega} \tilde{M}(k_n))$, see Eq. (3.12). To gain further information on the stationary stochastic process near the stability threshold, we examine the joint probability density $P_s(\{u_j\})$ of Eq. (3.31) by the corresponding Fokker-Planck

equation

$$\frac{\partial P(\{u_n\})}{\partial t} = - \sum_j \frac{\partial}{\partial u_n} \alpha_n P(\{u_n\}) + \eta^2 \frac{\partial^2}{\partial u_0^2} P(\{u_n\}).$$

Since the spatial modes $u_n(t)$ are independent from each other, we find $P(\{u_n\}, t) = \prod_n^N P(u_n, t)$ with the probability density function $P(u_n, t)$ of the spatial mode n . For $\alpha_n < 0$, the stationary distribution $P_s(u_n)$ is [60]

$$P_s(\{u_j\}) = \frac{\sqrt{|\alpha_0|}}{(\sqrt{2\pi}\eta)} e^{-|\alpha_0|u_0^2/2\eta^2} \prod_{n=1}^{N-1} \delta(u_n) \delta(u_{-n}). \quad (3.32)$$

In the absence of random fluctuations, i.e., $\eta \rightarrow 0$, we have $P_s \rightarrow \prod_n \delta(u_n)$, while P_s exhibits the variance $\sigma = \eta^2 / \sqrt{|\alpha_0|}$ for $\eta > 0$. In addition, just below the stability threshold some spatial modes u_c exhibit $\alpha_c \approx 0$. If $u_c = u_0$ then the variance σ is very large. This phenomena has been observed experimentally in many spatial systems, see e.g., [27, 28], and the system exhibits so-called *critical fluctuations*. In the case of $\alpha_n > 0$, no stationary probability density exists.

The latter analysis considered the stability criteria and the behavior close to instabilities characterized in the linear regime. When the system becomes unstable, the system's activity may still be bounded for larger deviations from the stationary state for suitable nonlinearities. The subsequent section considers this nonlinear saturation in some detail for the deterministic system and for a system subjected to global fluctuations.

3.4 Nonlinear analysis of the Turing instability

In order to illustrate the nonlinear behavior near the bifurcation point, this section focuses on the nonlinear amplitude equation of the Turing instability. First, we discuss the deterministic case, which gives some insight into the method. Subsequently, we consider the stochastic case showing the effect of global fluctuations on the system stability.

3.4.1 Deterministic analysis

Expanding Eq. (3.16) to cubic nonlinear order in V about V_0 for $\xi(x, t) = 0$, we obtain

$$\frac{\partial V(x, t)}{\partial t} = \int_{\Omega} dy K_1(x-y) V(y, t) + K_2(x-y) V^2(y, t) + K_3(x-y) V^3(y, t) \quad (3.33)$$

with

$$K_1(x) = \frac{\delta S}{\delta V} M(x) + \left(\frac{\delta g}{\delta V} - 1 \right) \delta(x), \quad K_2(x) = \frac{1}{2} \frac{\delta S^2}{\delta V^2} M(x) + \frac{1}{2} \frac{\delta^2 g}{\delta V^2} \delta(x),$$

$$K_3(x) = \frac{1}{6} \frac{\delta^3 S}{\delta V^3} M(x) + \frac{1}{6} \frac{\delta^3 g}{\delta V^3} \delta(x),$$

where the functional derivatives are computed at the stationary state $V = V_0$. Now inserting the Fourier expansion (3.17) into Eq. (3.33), multiplying both sides by $\exp(-ik_n x)$ and integrating over space, we obtain the infinite set of ODEs in the Fourier picture

$$du_n(t) = \alpha_n u_n(t) + \beta_n \sum_l u_l(t) u_{n-l}(t) + \gamma_n \sum_{l,m} u_l(t) u_m(t) u_{n-l-m}(t) \quad (3.34)$$

with the constants

$$\alpha_n = \tilde{K}_1(k_n), \quad \beta_n = \frac{\tilde{K}_2(k_n)}{\sqrt{|\Omega|}}, \quad \gamma_n = \frac{\tilde{K}_3(k_n)}{|\Omega|}, \quad (3.35)$$

where $\alpha_n = \alpha_{-n}$, $\beta_n = \beta_{-n}$, $\gamma_n = \gamma_{-n}$. Recall that $u_n = u_{-n}$ take real values according to the previous sections, i.e., $u_n = u_{-n} \in \mathcal{R}$, which will be taken into account in the following. Figure 3.9 illustrates the typical dependence of the parameters α_n on the discrete wavenumber k_n and shows the critical wavenumber $k_{\pm c}$ with $\alpha_c = 0$. Moreover, the illustration shows values $\alpha_2 \ll \alpha_0, \alpha_c$, i.e., a separation of values α_c , α_0 and α_2 .

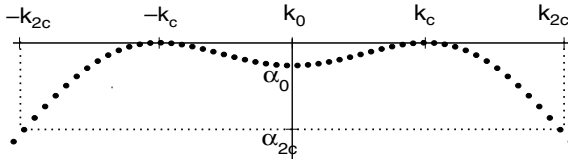


Fig. 3.9 Illustration of $\alpha_n = \alpha(k_n)$. The stability threshold is given by $\alpha_c = 0$ at k_c and k_{-c} .

Taking a closer look at the system of modes (3.34), it splits into critical modes u_c and stable modes $u_{i \neq c}$

$$\dot{u}_c = \alpha_c u_c + \beta_c \sum_n u_n u_{c-n} + 2\gamma_c u_c^3 + \gamma_c \sum_{\substack{n,m \\ n+m \neq \pm c}} u_n u_m u_{c-n-m} \quad (3.36)$$

$$\dot{u}_i = \alpha_i u_i + 2\beta_i (u_c u_{i-c} + u_{-c} u_{i+c}) + \beta_i \sum_{n \neq \pm c} u_n u_{i-n} + \gamma_i \sum_{n,m} u_n u_m u_{i-n-m}, \quad \forall i \neq \pm c. \quad (3.37)$$

If the system evolves on the threshold, i.e., $\alpha_c = 0$ and $\alpha_i < 0$, $i \neq c$, the deterministic center manifold theorem applies [58] and the stable modes $u_{i \neq c}$ depend on u_c , i.e., $u_i = u_i(u_c)$. Hence inserting the polynomial ansatz

$$u_i = a_i u_c^2 + b_i u_c^3, \quad i \neq \pm c \quad (3.38)$$

into Eqs (3.36), (3.37) and comparing the coefficients of orders u_c^2 and u_c^3 yields

$$a_i = -\frac{4\beta_i(\delta_{i,0} + \delta_{i,2c})}{\alpha_i}, \quad b_i = 0 \quad \rightarrow \quad u_0, u_{2c} \sim u_c^2. \quad (3.39)$$

Consequently there are different classes of modes defined corresponding to their scaling behavior with respect to u_c : the critical mode u_c with wave number k_c , and the subset of stable modes u_n with $\{k_n\}$, $n = 0, 2c, 3c, \dots$. Since the distinction of stable and critical modes in Eqs (3.36), (3.37) implies small time-scales $1/|\alpha_k|$, $|k| \geq 2c$, the corresponding modes decrease rapidly in time, i.e., $u_k \approx 0$ for $|k| \geq 2c$ at large times. This assumption is valid if $\alpha_n \ll \alpha_0, \alpha_c$ with $|n| \geq 2c$, cf. Fig. 3.9.

To examine the time-scales of the critical and stable modes, let us introduce the scaling factor $\varepsilon > 0$ with

$$\alpha_c \sim O(\varepsilon), \quad \alpha_{i \neq c}, \beta_i, \gamma_i \sim O(1).$$

In other words, we assume the system to be very close to the threshold and $\varepsilon \ll 1$ is proportional to the magnitude of α_c and thus quantifies the distance from the threshold $\alpha_c = 0$. Hence, the larger the distance from the stability threshold, i.e., the larger ε , the larger the deviations from the stationary state and thus the stronger the nonlinear effects. Then according to Eq. (3.39), the mode amplitudes may be written as $u_c = x\varepsilon^m$, $u_0 = y\varepsilon^{2m}$, $u_{2c} = z\varepsilon^{2m}$ for some constant $m \in \mathcal{R}$ and x, y, z are independent of ε . Inserting these expressions into Eqs (3.36), (3.37) we find $m = 1/2$,

$$u_c \sim O(\varepsilon^{1/2}), \quad u_{0,2c} \sim O(\varepsilon) \quad (3.40)$$

and $dx/dt \sim O(\varepsilon)$ and $dy/dt, dz/dt \sim O(1)$. Hence the critical mode evolves on the time scale of order $O(\varepsilon)$ which is slow compared to the stable mode time scale of order $O(1)$.

Summarizing the latter discussion, the stable modes obey the dynamics of the critical modes on the center manifold just around the stability threshold, while they evolve faster than the critical modes. In physical terms, the slow critical modes enslave the fast stable modes, which in turn obey the dynamics of the critical modes. This dependence is also called the slaving principle and the circular dependence is known as circular causality [27].

Applying the previous result, considering the lowest order $O(\varepsilon^{3/2})$ and taking into account the modes u_0 , u_c and u_{2c} only, Eqs. (3.36), (3.37) read

$$\dot{u}_c = \alpha_c u_c + 2\beta_c u_c (u_0 + u_{2c}) + 2\gamma_c u_c^3, \quad (3.41)$$

$$\dot{u}_0 = \alpha_0 u_0 + 4\beta_0 u_c^2. \quad (3.42)$$

Then the center manifolds (3.38) read

$$u_0(u_c, t) = -\frac{4\beta_0}{\alpha_0}u_c^2. \quad (3.43)$$

Inserting these results into the evolution equation of u_c we obtain the final equation

$$\dot{u}_c = \alpha_c u_c - a u_c^3 \quad (3.44)$$

with $a = 8\beta_0\beta_c/\alpha_0 - 2\gamma_c$.

The latter reduction to a single variable can also be understood from a more physical point of view. Since $|\alpha_0| \gg |\alpha_c|$, the mode u_0 relaxes fast to its stationary state determined by $\dot{u}_0 = 0$ for large times while choosing $\dot{u}_0 = 0$ yields directly Eq. (3.43). This procedure is called adiabatic approximation [27].

A detailed analysis of Eq. (3.44) reveals a stable solution $u_c = 0$ for $\alpha_c < 0$, $a > 0$ and a stable solution $u_c = \pm\sqrt{\alpha_c/a}$ for $\alpha_c > 0$, $a > 0$. In other words, the linearly unstable solutions for $\alpha_c > 0$ remain bounded if $a > 0$. If $a < 0$ there does not exist a stable state for $\alpha_c > 0$. This instability type is called a *pitchfork bifurcation*.

Summarizing, near the instability threshold the system evolves according to Eq. (3.44) and thus may exhibit a Turing instability with wave number k_c if $\alpha_c > 0, a > 0$.

3.4.2 Stochastic analysis at order $O(\varepsilon^{3/2})$

Now we study the nonlinear behavior close to the stability threshold at the presence of global fluctuations. At order $O(\varepsilon^{3/2})$, then Eqs (3.41), (3.42) read

$$du_c = (\alpha_c u_c + 2\beta_c u_c(u_0 + u_{2c}) + 2\gamma_c u_c^3) dt \quad (3.45)$$

$$du_0 = (\alpha_0 u_0 + 4\beta_0 u_c^2) dt + \eta dW(t) \quad (3.46)$$

Here we have neglected the mode u_{2c} for simplicity, which however does not affect the major result [39].

In recent years several methods have been developed to analyse the nonlinear behavior of systems subject to additive random fluctuations [5, 6, 9, 19, 32, 38, 39, 44, 63]. In the following we apply the stochastic center-manifold analysis [9, 74]. Similar to its deterministic version, this analysis assumes the stochastic center manifold of u_0 and u_{2c} is of the form

$$u_0(u_c, t) = h_0(u_c, t) = \sum_{n=2}^{\infty} h_0^{(n)}(u_c, t) \quad (3.47)$$

with $h_0^{(n)}, h_{2c}^{(n)} \sim O(\varepsilon^{n/2})$. We point out that the manifold is dependent on time due to the random fluctuations, which contrasts to the deterministic center manifold analysis. We obtain the evolution equations of the center manifolds to [39]

$$\left(\frac{\partial}{\partial t}dt - \alpha_0 dt\right) h_0 = \left[4\beta_0 u_c^2 - \frac{\partial h_0}{\partial u_c} (\alpha_c u_c + 2\beta_c u_c (h_0 + h_c) + 2\gamma_c u_c^3)\right] dt + \eta dW(t).$$

Now the stochastic center manifold contributions $h_0^{(n)}, h_{2c}^{(n)}$ are calculated at subsequent orders $O(\varepsilon^n)$ and we find at order $O(\varepsilon)$

$$h_0^{(2)}(u_c, t) = -\frac{4\beta_0}{\alpha_0} u_c^2 + \eta Z(t), \quad Z(t) = \int_{-\infty}^t e^{\alpha_0(t-\tau)} dW_0(\tau), \quad (3.48)$$

and at order $O(\varepsilon^{3/2})$ we find $h_0^{(3)}(u_c, t) = 0$. Inserting these expressions into Eq. (3.45), the stochastic differential equation for the critical mode reads at order $O(\varepsilon^{3/2})$

$$du_c = (\alpha_c u_c + a u_c^3 + b \eta u_c Z(t)) dt \quad (3.49)$$

with $b = 2\beta_c$. From Eq. (3.49) we learn that the mode u_c is subjected to multiplicative noise $Z(t)$. Since $Z(t)$ also represents an Ornstein–Uhlenbeck process, Eq. (3.49) can be extended by the corresponding SDEs $dZ(t) = \alpha_0 Z dt + dW(t)$. In this formulation, u_c depends on the additional variable Z .

To obtain a single equation (the so-called *order-parameter equation*) similar to the deterministic case a different approach is necessary, since the re-application of the stochastic center-manifold approach yields the same two equations. The chosen method involves an adiabatic elimination procedure based on the corresponding Fokker–Planck equation to gain a single order-parameter equation [18, 19]. After rescaling the variables $\bar{u}_c = u_c/\sqrt{\varepsilon}$, $\bar{\eta} = \eta/\varepsilon$, $\bar{\alpha}_c = \alpha_c/\varepsilon$ with Z , α_0 , a , $b \sim O(1)$ according to our previous results (3.40), the Fokker–Planck equation for $u_c(t)$ reads

$$\frac{\partial P(\bar{u}_c, t)}{\partial t} = -\frac{\partial}{\partial \bar{u}_c} [(\bar{\alpha}_c \bar{u}_c + a \bar{u}_c^3) + b \bar{\eta} \bar{u}_c \langle Z | \bar{u}_c \rangle] P(\bar{u}_c, t) \varepsilon$$

with $\langle Z | \bar{u}_c \rangle = \int_{-\infty}^{\infty} Z P(Z | \bar{u}_c, t) dZ$ and the conditional probability density $P(Z | \bar{u}_c, t)$. We learn that the probability density $P(\bar{u}_c, t)$ evolves on the slow time-scale of order $O(\varepsilon)$ and depends on Z via $\langle Z | \bar{u}_c \rangle$. Further computing the Fokker–Planck equation of the joint probability density function $P(Z, t)$, it turns out that $P(Z, t)$ evolves on a fast time-scale of order $O(1)$ and is independent of u_c . Then we obtain

$$\frac{\partial P(Z | \bar{u}_c, t)}{\partial t} = -\frac{\partial}{\partial Z} (\alpha_0 Z) P(Z | \bar{u}_c, t) + \frac{\partial^2}{\partial Z^2} P(Z | \bar{u}_c, t). \quad (3.50)$$

Here we have approximated $P(\bar{u}_c, t)$ by a constant on the time scale $O(1)$ which reflects the idea of an adiabatic behavior. In other words, the dynamics of $P(\bar{u}_c, t)$ is much slower than the dynamics on the time-scale $O(1)$ and thus may be treated as stationary. Then the stationary solution of Eq. (3.50) on the time scale $O(1)$ reads $P_s(Z | u_c) = \sqrt{|\alpha_0|} \exp(-|\alpha_0|Z^2/2)/\sqrt{2\pi}$ and we gain $\langle Z | \bar{u}_c \rangle = 0$. Hence the probability density function of the order parameter obeys

$$\frac{\partial P(u_c, t)}{\partial t} = -\frac{\partial}{\partial u_c} (\alpha_c u_c + a u_c^3) P(u_c, t), \quad (3.51)$$

whose stationary solution is $P_s(u_c) = \delta(x)$ for $\alpha_c < 0$ and $P_s(u_c) = \delta(x - x_0)/2 + \delta(x + x_m)/2$ for $\alpha_c \geq 0$ with the roots x_0 of $\alpha_c x_0 + a x_0^3 = 0$, i.e., we find again the pitchfork bifurcation gained in section 3.4.1 [28]. Consequently, additive global fluctuations do not affect the stability of the system at this order $O(\varepsilon^{3/2})$.

3.4.3 Stochastic analysis at order $O(\varepsilon^{5/2})$

The stochastic analysis of the next higher order $O(\varepsilon^2)$ leads to the same result as for $O(\varepsilon^{3/2})$. Consequently we focus on the higher order $O(\varepsilon^{5/2})$ with the evolution equations

$$du_c = (\alpha_c + b u_0 u_c + 2\gamma_c u_c^3 + 3\gamma_c u_c u_0^2) dt, \quad (3.52)$$

$$du_0 = (\alpha_0 u_0 + 4\beta_0 u_c^2 + \beta_0 u_0^2 + 2\gamma_0 u_0 u_c^2) dt + \eta dW(t). \quad (3.53)$$

The subsequent application of the stochastic center manifold analysis retains the lower order terms $h_0^{(2)}$ and $h_0^{(3)}$ and yields the additional term

$$h_0^{(4)}(u_c) = \beta_0 \eta^2 Z_5 - 8 \frac{\beta_0 \alpha_c}{\alpha_0^2} u_c^2 + 4 \frac{B}{b} \eta Z_4 u_c^2 + A u_c^4$$

$$h_0^{(5)}(u_c, t) = 0$$

with the constants A , B and the colored random fluctuations

$$Z_4(t) = \int_{-\infty}^t e^{\alpha_0(t-\tau)} Z_0(\tau) d\tau, \quad Z_5(t) = \int_{-\infty}^t e^{\alpha_0(t-\tau)} Z_0^2(\tau) d\tau.$$

Applying some approximations to $Z_4(t)$ and $Z_5(t)$ [39], the final Fokker–Planck equation for the order parameter reads at order $O(\varepsilon^2)$

$$\frac{\partial P(u_c)}{\partial t} = -\frac{\partial}{\partial u_c} [(\alpha_c - \alpha_{\text{th}}(\eta)) u_c + C u_c^3 + D u_c^5] P(u_c) \quad (3.54)$$

with

$$\alpha_{\text{th}}(\eta) = \eta^2 \left(\frac{\beta_0 b}{\alpha_0^2} - 3 \frac{\gamma_c}{|\alpha_0|} \right) \quad (3.55)$$

and constants C, D .

We observe that the order parameter u_c obeys the deterministic equation

$$\dot{u}_c = (\alpha_c - \alpha_{c, \text{th}}(\eta)) u_c + C u_c^3 + D u_c^5, \quad (3.56)$$

where α_{th} defines the new stability threshold, which now depends on the fluctuation strength η and thus reflects noise-induced transitions. Hence for $\alpha_{\text{th}} > 0$ the noise retards the emergence of the instability with increasing α_c and thus stabilizes the neural field [39]. From a physical point of view, the global fluctuations represent an external stimulus which forces the system to obey the stimulus dynamics. The stronger the stimulus is, i.e., the larger the fluctuation strength η , the stronger the spatial mode $k = 0$ is and thus the smaller the contribution of the unstable mode $k = k_c$.

Figure 3.10 shows the space-time activity of a stochastic Turing instability in the absence and presence of global fluctuations. We observe the onset of a spatial pattern in the absence of random fluctuations (left panel), while global fluctuations stabilize the system and no spatial pattern occurs (right panel).

Concluding, additive global fluctuations change the stability of the neural population and thus affect the neural processing.

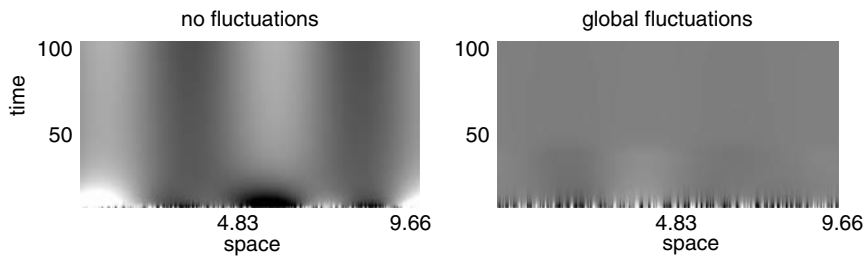


Fig. 3.10 Effects of global fluctuations on Turing instabilities. The spatiotemporal activity is shown subtracted from its spatial average at each time for the fluctuation strengths $\eta' = 0$ (no fluctuations) and $\eta' = 0.03$ (global fluctuations). Modified from [39].

3.5 Conclusion

The previous sections have shown the analytical treatment of nonlocally interacting neural populations in the linear and nonlinear regime. We have found several types of spatiotemporal instabilities and have studied briefly the statistical properties of the system close to such instabilities. Further, the last section illustrates one way to treat the nonlinear behavior of neural populations.

Such investigations are necessary in order to understand the information processing in neural populations, as they reflect the linear response to incoming stimuli. This response has been examined further in some detail in recent studies [23, 35, 37, 59] and promises deep insights to neural processing.

References

1. Amit, D.J.: Modeling brain function: The world of attractor neural networks. Cambridge University Press, Cambridge (1989)
2. Arieli, A., Shoham, D., Hildesheim, R., Grinvald, A.: Coherent spatiotemporal pattern of ongoing activity revealed by real-time optical imaging coupled with single unit recording in the cat visual cortex. *J. Neurophysiol.* **73**, 2072–2093 (1995)
3. Atay, F.M., Hutt, A.: Stability and bifurcations in neural fields with finite propagation speed and general connectivity. *SIAM J. Appl. Math.* **65**(2), 644–666 (2005), doi:10.1137/S0036139903430884
4. Atay, F.M., Hutt, A.: Neural fields with distributed transmission speeds and constant feedback delays. *SIAM J. Appl. Dyn. Syst.* **5**(4), 670–698 (2006), doi:10.1137/050629367
5. Berglund, N., Gentz, B.: Geometric singular perturbation theory for stochastic differential equations. *J. Diff. Eq.* **191**, 1–54 (2003), doi:10.1016/S0022-0396(03)00020-2
6. Berglund, N., Gentz, B.: Noise-Induced Phenomena in Slow-Fast Dynamical Systems: A Sample-Paths Approach. Springer, Berlin (2006)
7. Blomquist, P., Wyller, J., Einevoll, G.T.: Localized activity patterns in two-population neuronal networks. *Physica D* **206**, 180–212 (2005), doi:10.1016/j.physd.2005.05.004
8. Bojak, I., Liley, D.: Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* **71**, 041902 (2005), doi:10.1103/PhysRevE.71.041902
9. Boxler, P.: A stochastic version of the center manifold theorem. *Probab. Theory. Rel.* **83**, 509–545 (1989), doi:10.1007/BF01845701
10. Bressloff, P.C.: Synaptically generated wave propagation in excitable neural media. *Phys. Rev. Lett.* **82**(14), 2979–2982 (1999), doi:10.1103/PhysRevLett.82.2979
11. Chacron, M.J., Longtin, A., Maler, L.: The effects of spontaneous activity, background noise and the stimulus ensemble on information transfer in neurons. *Network-Comp. Neural* **14**, 803–824 (2003), doi:10.1088/0954-898X/14/4/010
12. Coombes, S., Lord, G., Owen, M.: Waves and bumps in neuronal networks with axo-dendritic synaptic interactions. *Physica D* **178**, 219–241 (2003), doi:10.1016/S0167-2789(03)00002-2
13. Coombes, S., Venkov, N., Shiau, L., Bojak, I., Liley, D., Laing, C.: Modeling electrocortical activity through improved local approximations of integral neural field equations. *Phys. Rev. E* **76**, 051901–8 (2007), doi:10.1103/PhysRevE.76.051901
14. Cross, M.C., Hohenberg, P.C.: Pattern formation outside of equilibrium. *Rev. Mod. Phys.* **65**(3), 851–1114 (1993), doi:10.1103/RevModPhys.65.851
15. Destexhe, A., Contreras, D.: Neuronal computations with stochastic network states. *Science* **314**, 85–90 (2006), doi:10.1126/science.1127241
16. Doiron, B., Chacron, M., L.Maler, Longtin, A., Bastian, J.: Inhibitory feedback required for network burst responses to communication but not to prey stimuli. *Nature* **421**, 539–543 (2003), doi:10.1038/nature01360
17. Doiron, B., Lindner, B., Longtin, A., L.Maler, Bastian, J.: Oscillatory activity in electrosensory neurons increases with the spatial correlation of the stochastic input stimulus. *Phys. Rev. Lett.* **93**, 048101 (2004), doi:10.1103/PhysRevLett.93.048101
18. Drolet, F., Vinals, J.: Adiabatic reduction near a bifurcation in stochastically modulated systems. *Phys. Rev. E* **57**(5), 5036–5043 (1998), doi:10.1103/PhysRevE.57.5036
19. Drolet, F., Vinals, J.: Adiabatic elimination and reduced probability distribution functions in spatially extended systems with a fluctuating control parameter. *Phys. Rev. E* **64**, 026120 (2001), doi:10.1103/PhysRevE.64.026120
20. Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., Reitboeck, H.: Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biol. Cybern.* **60**, 121–130 (1988), doi:10.1007/BF00202899
21. Eggert, J., van Hemmen, J.L.: Modeling neuronal assemblies: Theory and implementation. *Neural Comput.* **13**(9), 1923–1974 (2001)
22. Feller, W.: An introduction to probability theory and its applications. Wiley, New York (1966)
23. Folias, S., Bressloff, P.: Breathers in two-dimensional excitable neural media. *Phys. Rev. Lett.* **95**, 208107 (2005), doi:10.1103/PhysRevLett.95.208107

24. Freeman, W.J.: *Mass Action in the Nervous System*. Academic Press, New York (1975)
25. Freeman, W.: Tutorial on neurobiology: from single neurons to brain chaos. *Int. J. Bifurcat. Chaos* **2**(3), 451–482 (1992), doi:10.1142/S0218127492000653
26. Gerstner, W.: Time structure of the activity in neural network models. *Phys. Rev. E* **51**(1), 738–758 (1995), doi:10.1103/PhysRevE.51.738
27. Haken, H.: *Synergetics*. Springer, Berlin (2004)
28. Horsthemke, W., Lefever, R.: *Noise-induced transitions*. Springer, Berlin (1984)
29. Huang, X., Troy, W., Schiff, S., Yang, Q., Ma, H., Laing, C., Wu, J.: Spiral waves in disinhibited mammalian neocortex. *J. Neurosci.* **24**(44), 9897–9902 (2004), doi:10.1523/JNEUROSCI.2705-04.2004
30. Hubel, D.H., Wiesel, T.N.: Receptive fields of cells in striate cortex of very young, visually unexperienced kittens. *J. Physiol* **26**, 994–1002 (1963)
31. Hutt, A.: Generalization of the reaction-diffusion, Swift-Hohenberg, and Kuramoto-Sivashinsky equations and effects of finite propagation speeds. *Phys. Rev. E* **75**, 026214 (2007), doi:10.1103/PhysRevE.75.026214
32. Hutt, A.: Additive noise may change the stability of nonlinear systems. *Europhys. Lett.* **84**, 34003 (2008), doi:10.1209/0295-5075/84/34003
33. Hutt, A.: Local excitation-lateral inhibition interaction yields oscillatory instabilities in non-locally interacting systems involving finite propagation delay. *Phys. Lett. A* **372**, 541–546 (2008), doi:10.1016/j.physleta.2007.08.018
34. Hutt, A., Atay, F.M.: Analysis of nonlocal neural fields for both general and gamma-distributed connectivities. *Physica D* **203**, 30–54 (2005), doi:10.1016/j.physd.2005.03.002
35. Hutt, A., Atay, F.M.: Spontaneous and evoked activity in extended neural populations with gamma-distributed spatial interactions and transmission delay. *Chaos Solitons Fract* **32**, 547–560 (2007), doi:10.1016/j.chaos.2005.10.091
36. Hutt, A., Bestehorn, M., Wennekers, T.: Pattern formation in intracortical neuronal fields. *Network-Comp. Neural* **14**, 351–368 (2003), doi:10.1088/0954-898X/14/2/310
37. Hutt, A., Frank, T.D.: Critical fluctuations and $1/f$ -activity of neural fields involving transmission delays. *Acta Phys. Pol. A* **108**(6), 1021 (2005)
38. Hutt, A., Longtin, A., Schimansky-Geier, L.: Additive global noise delays Turing bifurcations. *Phys. Rev. Lett.* **98**, 230601 (2007), doi:10.1103/PhysRevLett.98.230601
39. Hutt, A., Longtin, A., Schimansky-Geier, L.: Additive noise-induced Turing transitions in spatial systems with application to neural fields and the Swift-Hohenberg equation. *Physica D* **237**, 755–773 (2008), doi:10.1016/j.physd.2007.10.013
40. Hutt, A., Schimansky-Geier, L.: Anesthetic-induced transitions by propofol modeled by non-local neural populations involving two neuron types. *J. Biol. Phys.* **34**(3-4), 433–440 (2008), doi:10.1007/s10867-008-9065-4
41. Jirsa, V., Jantzen, K., Fuchs, A., Kelso, J.: Spatiotemporal forward solution of the EEG and MEG using network modelling. *IEEE Trans. Med. Imag.* **21**(5), 493–504 (2002), doi:10.1109/TMI.2002.1009385
42. Kaschube, M., Schnabel, M., Wolf, F.: Self-organization and the selection of pinwheel density in visual cortical development. *New J. Phys.* **10**, 015009 (2008), doi:10.1088/1367-2630/10/1/015009
43. Katz, B. (ed.): *Nerve, Muscle and Synapse*. McGraw-Hill, New York (1966)
44. Knobloch, E., Wiesenfeld, K.: Bifurcations in fluctuating systems: The center-manifold approach. *J. Stat. Phys.* **33**(3), 611–637 (1983), doi:10.1007/BF01018837
45. Koch, C.: *Biophysics of Computation*. Oxford University Press, Oxford (1999)
46. Kozin, F.: A survey of stability of stochastic systems. *Automatica* **5**, 95–112 (1969)
47. Laing, C.: Spiral waves in nonlocal equations. *SIAM J. Appl. Dyn. Syst.* **4**(3), 588–606 (2005), doi:10.1137/040612890
48. Laing, C., Coombes, S.: The importance of different timings of excitatory and inhibitory models. *Network. Comput. Neur. Syst.* **17**(2), 151–172 (2006), doi:10.1080/09548980500533461
49. Laing, C., Troy, W.: PDE methods for nonlocal models. *SIAM J. Appl. Dyn. Syst.* **2**(3), 487–516 (2003), doi:10.1137/030600040

50. Liley, D., Wright, J.: Intracortical connectivity of pyramidal and stellate cells: estimates of synaptic densities and coupling symmetry. *Network-Comp. Neural* **5**, 175–189 (1994), doi:10.1088/0954-898X/5/2/004
51. Lindner, B., Schimansky-Geier, L.: Transmission of noise coded versus additive signals through a neuronal ensemble. *Phys. Rev. Lett.* **86**, 2934–2937 (2001), doi:10.1103/PhysRevLett.86.2934
52. Longtin, A., Moss, F., Bulsara, A.: Time interval sequences in bistable systems and noise induced transmission of neural information. *Phys. Rev. Lett.* **67**, 656–659 (1991), doi:10.1103/PhysRevLett.67.656
53. Masuda, N., Okada, M., Aihara, K.: Filtering of spatial bias and noise inputs by spatially structured neural networks. *Neural Comp.* **19**, 1854–1870 (2007)
54. Mountcastle, V.B.: Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Neurophysiol.* **20**, 408–434 (1957)
55. Nunez, P.: The brain wave equation: A model for the EEG. *Math. Biosc.* **21**, 279–291 (1974)
56. Nunez, P.: *Neocortical dynamics and human EEG rhythms*. Oxford University Press, New York - Oxford (1995)
57. Owen, M.R., Laing, C.R., Coombes, S.: Bumps and rings in a two-dimensional neural field: splitting and rotational instabilities. *New J. Phys.* **9**, 378 (2007), doi:10.1088/1367-2630/9/10/378
58. Perko, L.: *Differential Equations and Dynamical Systems*. Springer, Berlin (1998)
59. Rennie, C., Robinson, P., Wright, J.: Unified neurophysical model of EEG spectra and evoked potentials. *Biol. Cybern.* **86**, 457–471 (2002), doi:10.1007/s00422-002-0310-9
60. Risken, H.: *The Fokker-Planck equation — Methods of solution and applications*. Springer, Berlin (1989)
61. R.Z.Khasminskij: *Stochastic stability of differential equations*. Alphen aan den Rijn (1980)
62. Sanderson, K.: The projection of the visual field to the lateral geniculate and medial interlaminar nuclei in the cat. *J. Comp. Neurol.* **143**, 101–118 (1971)
63. Schimansky-Geier, L., Tolstopjatenko, A., Ebeling, W.: Noise-induced transitions due to external additive noise. *Phys. Lett. A* **108**(7), 329–332 (1985), doi:10.1016/0375-9601(85)90107-0
64. Somers, D., Nelson, S., Sur, M.: An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.* **15**(8), 5448–5465 (1995)
65. Steyn-Ross, M., Steyn-Ross, D., Wilson, M., Sleight, J.: Gap junctions mediate large-scale Turing structures in a mean-field cortex driven by subcortical noise. *Phys. Rev. E* **76**, 011916 (2007), doi:10.1103/PhysRevE.76.011916
66. Tamas, G., Buhl, E.H., Somogyi, P.: Massive autaptic self-innervation of GABAergic neurons in cat visual cortex. *J. Neurosci.* **17**(16), 6352–6364 (1997)
67. Thomson, J.R., Zhang, Z., Cowan, W., Grant, M., Hertz, J.A., Zuckermann, M.J.: A simple model for pattern formation in primate visual cortex for the case of monocular deprivation. *Phys. Scr.* **T33**, 102–109 (1990)
68. Venkov, N.A., Coombes, S., Matthews, P.C.: Dynamic instabilities in scalar neural field equations with space-dependent delays. *Physica D* **232**, 1–15 (2007), doi:10.1016/j.physd.2007.04.011
69. Wennekers, T.: Orientation tuning properties of simple cells in area V1 derived from an approximate analysis of nonlinear neural field models. *Neural Comput.* **13**, 1721–1747 (2001)
70. Wilson, H., Cowan, J.: Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–24 (1972)
71. Wolf, F.: Symmetry, multistability, and long-range interactions in brain development. *Phys. Rev. Lett.* **95**, 208701 (2005), doi:10.1103/PhysRevLett.95.208701
72. Wright, J.J.: Simulation of EEG: dynamic changes in synaptic efficacy, cerebral rhythms, and dissipative and generative activity in cortex. *Biol. Cybern.* **81**, 131–147 (1999)
73. Wright, J., Liley, D.: A millimetric-scale simulation of electrocortical wave dynamics based on anatomical estimates of cortical synaptic density. *Network-Comp. Neural* **5**(2), 191–202 (1994), doi:10.1088/0954-898X/5/2/005
74. Xu, C., Roberts, A.: On the low-dimensional modelling of Stratonovich stochastic differential equations. *Physica A* **225**, 62–80 (1996), doi:10.1016/0378-4371(95)00387-8

Chapter 4

Spontaneous brain dynamics emerges at the edge of instability

V.K. Jirsa and A. Ghosh

4.1 Introduction

Neuroscience has prominently attracted mathematicians and physicists to understand complex dynamics of the brain. The mathematical framework has benefited neuroscience in explaining observed neuronal behavior in both quantitative and qualitative manners. Mathematical models of neuronal communication and synaptic plasticity, nonlinear dynamical systems theory, and use of probability theory to quantify anatomical observations, all serve to illustrate the extensive application of mathematical tools and physical laws to explain the complexity of brain. On the other hand, mathematics has also benefitted from the rich dynamical repertoire of neurodynamics by motivating studies of bifurcation theory to explore various theoretical concepts.

The dynamics of individual neurons—often described by the Hodgkin–Huxley model—is well studied. Simplified versions of that complex model are also well studied and used in different contexts. However, the brain is a collection of billions of such units, and collectively exhibits a range of dynamics like synchronization, self-organization, etc.

A first question that can be asked is: How are these neurons spatially connected? The *anatomical connectivity* can be links between individual neurons and synapses, and also connections between neuronal populations and pathways. Unlike lattice models frequently studied in physics, neural networks not only contain short-range connections but also long-range connections. For large-scale cortical network

Viktor K. Jirsa

Theoretical Neuroscience Group, Institut des Sciences du Mouvement, Etienne-Jules Marey UMR 6233, Université de la Méditerranée, 163 Avenue de Luminy, CP 910 13288 Marseille cedex 9 France. e-mail: viktor.jirsa@univmed.fr

Anandamohan Ghosh

Theoretical Neuroscience Group, Institut des Sciences du Mouvement, Etienne-Jules Marey UMR 6233, Université de la Méditerranée, 163 Avenue de Luminy, CP 910 13288 Marseille cedex 9 France. e-mail: anandamohan@gmail.com

studies, one can make use of the information on anatomical connectivity available for the primate brain.

A second question that follows is: How do functions emerge, and how are they interrelated? The *functional connectivity* of a neuronal system consists of statistical temporal correlations due to the dynamics of different functional units. Two principles leading to organization of functional behavior are: *segregation*—neuronal units are capable of performing specific functions and are segregated from each other; and *integration*—coherent brain processes like perception and cognition emerge due to functional integration. The competition between segregation and integration leads to functional connectivity [15].

Questions that naturally follow are: How are anatomical and functional connectivity related? And: How do we explain various states emerging from the complex dynamics of the brain? The *resting-state* of the brain provides an ideal setup to test hypotheses and to develop a comprehensive mathematical framework to explain criticality in neuronal dynamics. We have developed a large-scale neuronal model, and conjecture that brain dynamics evolves close to the edge of instability such that intrinsic fluctuations allow the brain to explore its dynamic repertoire [6]. Our proposed model successfully identifies the correlated and anticorrelated resting-state network observed in experimental studies [3].

In this chapter we discuss the role of noise in exploring the neighborhood of instability, leading to the emergence of spontaneous neural activity. In Sect. 2 we lay the theoretical foundation; in Sect. 3 we will provide details of our previous modeling study and its validation with experimental findings; in Sect. 4 we illustrate how dynamical features underlying these concepts can be extracted from experimental EEG signals. This is followed by some concluding remarks.

4.2 Concept of instability, noise, and dynamic repertoire

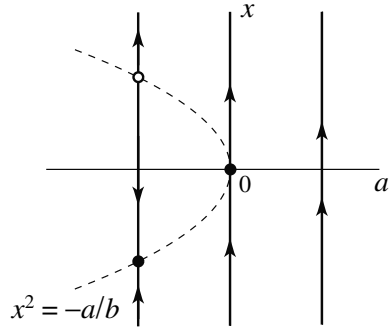
The mechanism of neuronal firing due to synaptic currents can be well understood within the framework of a dynamical systems perspective. Simple geometrical tools can qualitatively and quantitatively explain the evolution of complex behavior of neurons arising from competition between a stable and an unstable equilibrium in phase space. However, the model of a single neuron is highly nonlinear [8], so mathematical simplifications are often useful to gain insight into the underlying mechanism in a qualitative manner. This intrinsic nonlinearity leads to an abrupt behavior due to a change of system parameters, and is often studied by bifurcation analysis.

Here we illustrate a bifurcation in a simple system given by the ordinary differential equation,

$$\frac{dx}{dt} = a + bx^2, \quad (4.1)$$

due to a variation in the parameter a . For $a < 0$, the system has two equilibrium points, one stable and the other unstable; it has a metastable equilibrium for $a = 0$; and no equilibrium for $a > 0$. This scenario is known as a *saddle-node bifurcation* where two fixed points collide and disappear (Fig. 4.1).

Fig. 4.1 Bifurcation diagram for a saddle-node bifurcation. The dashed curve shows the distribution of equilibria as a function of control parameter a . The arrows on the vertical lines indicate flow toward a stable equilibrium, and away from an unstable equilibrium.



Close to the bifurcation point, noisy input results in fluctuations around the equilibrium point, which are irregularly occurring departures from, and returns to, the equilibrium. These transient processes reflect the properties of the noise, as well as the flow in the neighborhood of the equilibrium point. In conjunction, they define the *dynamic repertoire* of the system. Figure 4.2 shows the time-series corresponding to three different values of the Eq. (4.1) control parameter $a < 0$ close to its critical value of zero; we observe that noise-induced fluctuations become more pronounced as the distance to the critical point decreases.

Since neuronal dynamics is oscillatory in nature, it is interesting to look into systems that show oscillations due to parametric variations. One such bifurcation scenario is captured by the Hopf bifurcation. In this mechanism, a limit cycle is born as the equilibrium point changes stability giving rise to pair of purely imaginary eigenvalues. Depending on the stability property of the limit cycle, a Hopf bifurcation

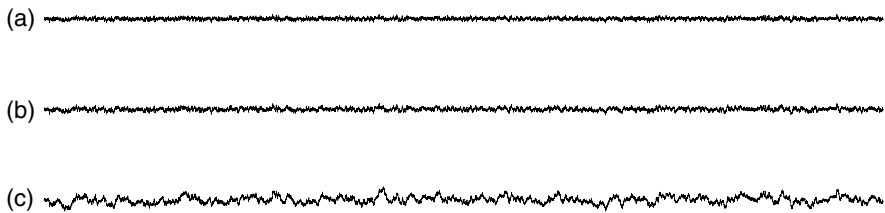


Fig. 4.2 Time-series corresponding to three different values of $a < 0$ for the saddle-node bifurcation in Fig. 4.1: (a) $a = -4$; (b) $a = -1.21$; (c) $a = -0.04$. Traces show the increase in noise-induced fluctuation activity about equilibrium as the control parameter a approaches its critical value $a = 0$. With $b = 1$, the equilibrium values are (a) $x^* = -2$; (b) $x^* = -1.1$; (c) $x^* = -0.2$ respectively. The noise-source used for each simulation run was identical.

can be subcritical (unstable) or supercritical (stable). For a two-dimensional system, such supercritical behavior can be observed in the system represented by,

$$\frac{dx_1}{dt} = ax_1 - x_2 - x_1(x_1^2 + x_2^2), \quad (4.2a)$$

$$\frac{dx_2}{dt} = x_1 + ax_2 - x_2(x_1^2 + x_2^2). \quad (4.2b)$$

It is easy to see that for $a < 0$ there is one stable equilibrium point at $(0, 0)$; but for $a > 0$, a stable attracting limit cycle is born [12]. Characteristic trajectories in phase space for various values of the control parameter a are shown in Fig. 4.3 with no noise.

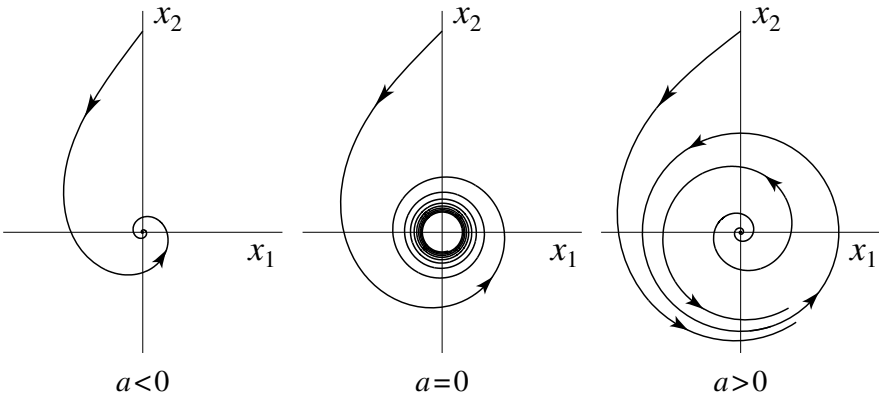


Fig. 4.3 Supercritical Hopf bifurcation. The trajectories are plotted in phase-space for a -values below ($a < 0$), equal to ($a = 0$), and above ($a > 0$) the critical value.

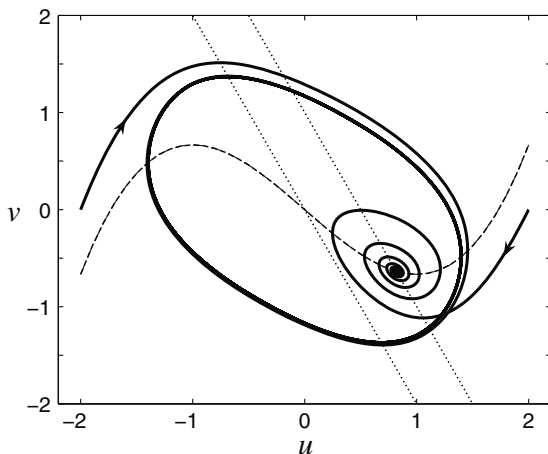
Simple models undergoing a Hopf bifurcation can explain neuronal spiking and are very common in the literature. A well-established model is the FitzHugh–Nagumo model [2, 13]. The FitzHugh–Nagumo model is a simplification of the more biologically realistic Hodgkin–Huxley model for a single neuron. Its trajectories below and above the Hopf bifurcation point are illustrated in Fig. 4.4, and its dynamics is represented by

$$\frac{du}{dt} = u - \frac{u^3}{3} + v + I, \quad (4.3a)$$

$$\frac{dv}{dt} = a - u - bv, \quad (4.3b)$$

where u is voltage-like variable having a cubic nonlinearity, and v is a recovery-like variable. By increasing input current I in Eq. (4.3a), or, equivalently, by reducing parameter a in Eq. (4.3b), the neuronal dynamics can be made to undergo a bifurca-

Fig. 4.4 Phase space trajectories of the FitzHugh–Nagumo neuron. Dashed line identifies the cubic nullcline for $du/dt = 0$; dotted lines identify the nullclines for $dv/dt = 0$ for two sets of parameter values: $I = 0$, $b = 0.5$, and $a = 0$ or $a = 0.5$. Bold lines trace out representative trajectories: a limit cycle ($a = 0$), and a stable spiral ($a = 0.5$). The intersection of the two nullclines determines the fixed point. It is the fixed point near $u = 0.82$ that undergoes a stability change through a Hopf bifurcation.



tion from quiescent subthreshold behavior to periodic-spiking behavior. Figure 4.5 shows the response of a near-threshold Hopf oscillator when perturbed with noise of increasing intensity.

Here we wish to postulate that the working point of the brain during rest is often in the neighborhood of the critical boundary separating stable and unstable regions. More specifically, the expression “working point” refers to all the values of a set of parameters characterizing the brain—such as excitability, synaptic strength, etc. We have studied in detail the resting-state of the brain [6], and show that, by using a realistic primate connectivity matrix in the presence of noise and time-delay, it is possible to explore the dynamic repertoire of the brain’s resting-state.

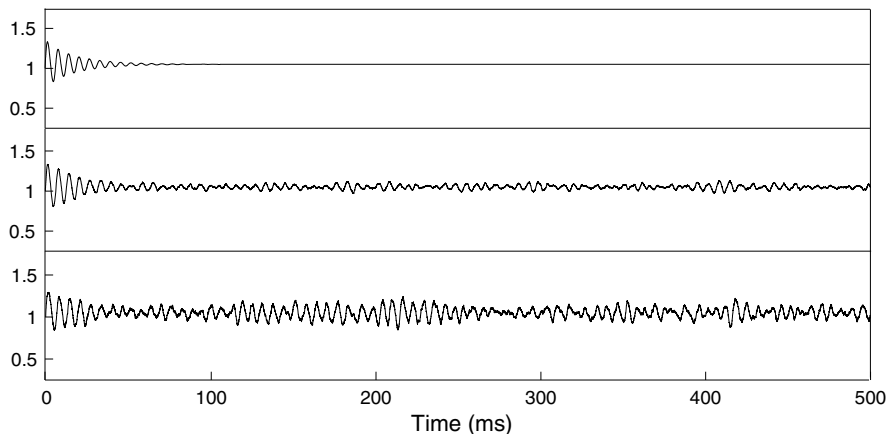


Fig. 4.5 Hopf oscillator with noise. Spontaneous oscillations emerge when the intensity of the noise is increased for the Hopf oscillator operating in the neighborhood of its critical value $a = 0$.

4.3 Exploration of the brain's instabilities during rest

When subjects are not actively engaged in goal-directed mental activity, spontaneous brain activity has been suggested not to represent simply “noise”, but rather to implicate spontaneous and transient processes involved in task-unrelated imagery and thought. The resting-state networks that are not associated with sensory or motor regions have been thought of as a “default-mode” network specific to the human, and include medial prefrontal, parietal, and posterior and anterior cingulate cortices. The dynamics of these spontaneous fluctuations evolves on a slow time-scale of multiple seconds in the Blood Oxygen Level Dependent (BOLD) signal. However, the computational models to explain the generating mechanisms are few, and do not satisfactorily explain how the default network relates to the complex spatiotemporal dynamics.

To shed light on the emergence of the resting-state networks and their dynamics on various temporal scales, we performed a network simulation study in which the major ingredients were biologically realistic primate connectivity of brain areas, time-delays via signal propagation between areas, and noise. We initially consider only the spatial aspect of the couplings. The connectivity matrix collated from macaque tracing studies comprises 38 nodes with weights ranging from 0 to 3. The corresponding “regional map” gives the translation between macaque and human neuroanatomy [10, 11]. It is to be noted that some connections between some areas are not known. When computing various graph-theoretical measures for the weighted graphs we do not observe any of the areas to have significant features to emerge as a central hub. Thus anatomical connectivity of the large scale network does not suffice to reliably identify the network constituents during rest.

To study the rest-state dynamics, we place oscillatory neuronal populations at each network node, and couple these via time-delayed interaction terms. Each population is characterized by a degree of excitability in which the increase of excitation parameterizes the onset of oscillations emerging from a quiescent state. When the populations are embedded in a network, the network's dynamic repertoire will be shaped by the space–time structure of the couplings. To quantify the total connectivity strength, we introduce a parameter, c , which scales all connection strengths without altering the connection topology of the weight distribution of the matrix, nor affecting the associated time-delays $\Delta t = d/v$. The network model is implemented as,

$$\frac{du_i}{dt} = g(u_i, v_i) - c \sum_{j=1}^N f_{ij} u_j(t - \Delta t_{ij}) + n_u(t), \quad (4.4a)$$

$$\frac{dv_i}{dt} = h(u_i, v_i) + n_v(t). \quad (4.4b)$$

where u_i, v_i are the state variables of the i -th neural population, and f_{ij} is the connectivity matrix. White Gaussian noises $n_u(t), n_v(t)$ are introduced additively. The functions $g(\cdot)$ and $h(\cdot)$ are based on FitzHugh–Nagumo systems with

$$g(u_i, v_i) = \tau \left[\gamma u_i - \frac{u_i^3}{3} + v_i \right], \quad (4.5a)$$

$$h(u_i, v_i) = \frac{1}{\tau} [\alpha - u_i - \beta v_i]. \quad (4.5b)$$

We apply a linear stability analysis to the system described by Eqs (4.4), and identify the critical boundary which separates the stable, quiescent state from the unstable regions in the parameter space of c and v (see Fig. 4.6(a)). In its immediate proximity (but still in the stable region), the effect of noise driving the network transiently out of its equilibrium state will be most prominent and hence easiest to identify.

To perform a spatiotemporal analysis of the network dynamics, we identify the dominating subnetworks involved in the ongoing transient oscillatory dynamics. During the transition, we use a sliding temporal-window analysis, and perform a PCA (principal components analysis) to identify the dominant network modes shown in Fig. 4.6(b).

We find that prefrontal, parietal, and cingulate cortices rank highest in this ordering scheme, and hence contribute most to the two network patterns present during the transient of the instability. We confirm our findings by performing a complete computational network simulation with noise just below the critical boundary, and verify that these subnetworks are most commonly present during the transient oscillations of rest-state activity (see Fig. 4.7).

Another important aspect that we addressed in the present study is by generating the ultra-slow oscillating BOLD signals [5] and identifying the correlated and anticorrelated networks [3]. Fox and colleagues chose six predefined seed regions, and computed the correlations against all other regions. These seed regions included three so-called “task-positive” regions, routinely exhibiting activity increases during task performance; and three “task-negative” regions, routinely exhibiting activity decreases during task performance.

Task-positive regions were centered in the intraparietal sulcus (IPS; in our notation: PCIP (intraparietal sulcus cortex)); the frontal eye field (FEF) region (same in our notation); and the middle temporal region (MT; in our notation this area is part of VACD (dorsal anterior visual cortex)).¹

Task-negative regions were centered in the medial prefrontal cortex (MPF; in our notation this area corresponds mostly to PFCM (medial prefrontal cortex) and to a lesser extent to PFCPOL (prefrontal polar cortex)); posterior cingulate precuneus (PCC; in our notation CCP (posterior cingulate cortex), but note that the precuneus comprises also our medial parietal cortex PCM); and lateral parietal cortex (LP; in our notation PCI (inferior parietal cortex)). We compute the cross-correlations of the seed regions from our simulated data set, and find excellent agreement with experimental observations. Fig. 4.8 shows a summary of our findings (obtained from Ghosh *et al.* (2008) [6]).

¹ Brain-region abbreviations are listed at the end of this chapter.

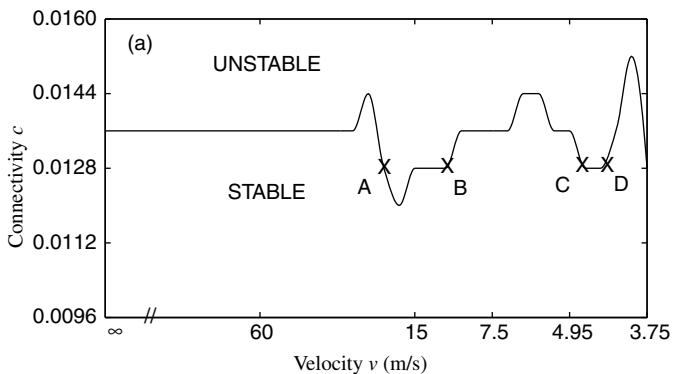


Fig. 4.6 (a) The critical line separating unstable and stable regions is shown as a function of velocity v and connectivity c .

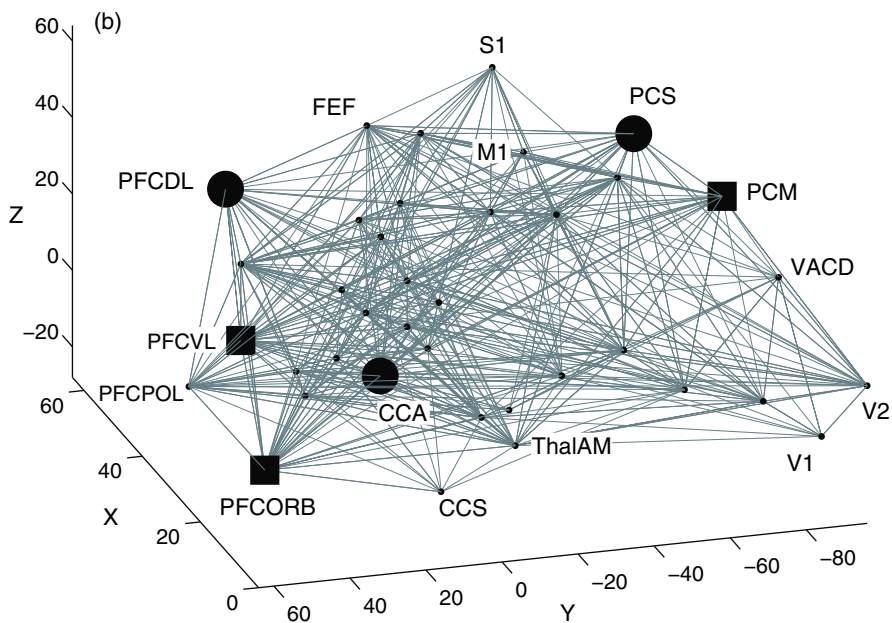


Fig. 4.6 (b) The emergent resting-state networks are indicated in 3-D physical space. Position coordinates carry units of mm. (Node acronyms are listed at the end of this chapter.)

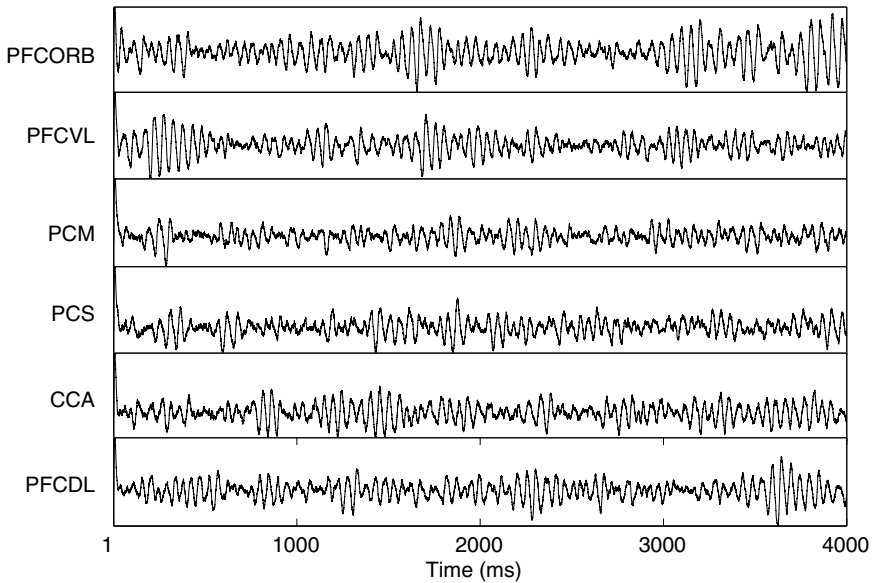


Fig. 4.7 Characteristic simulated time-series are shown for nodes of the resting-state network.

4.4 Dynamical invariants of the human resting-state EEG

Cortical dynamics of resting brain exhibits spontaneous neuronal activity. With state-of-the-art experimental techniques (EEG), it is possible to record cortical activity simultaneously at multiple sensor locations. However, EEG recordings are often corrupted by various artefacts (eye blinks, eye wanderings, etc) and make meaningful analysis difficult. We systematically eliminate artefacts by a combination of regression and wavelet transform, and subject the data to time-series analysis. We demonstrate that it is possible to identify features of standing, traveling, and rotating waves from the EEG data. These features remain invariant for EEG recordings performed on several subjects, and can be considered to be a signature of resting brain.

The EEG recordings are performed on 16 healthy subjects under eyes-open condition. Rest-state activity was recorded while the subjects altered between eyes-open and eyes-closed in a block design. Blocks consist of periods of about 40 s duration (to avoid eye-blink artefacts). EEG was recorded using a 128-channel Neuroscan system (Compumedics USA, Inc., El Paso, TX) that provides high-density, full-head coverage. The exact location of each EEG electrode was determined with respect to standard fiducials (bilateral preauricular points and the nasion) using a Polhemus Fastrack system.

The artefact-removed data reveal characteristic spindle structures, and its Fourier transform in Fig. 4.9 shows a power spectrum with a slope of approximately -2 in

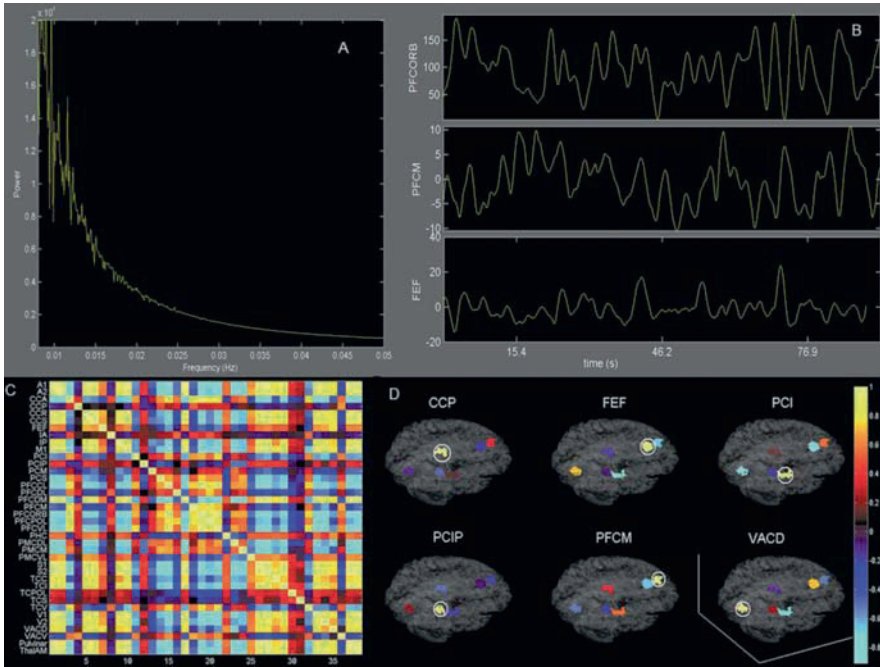


Fig. 4.8 [Color plate] Analysis of BOLD signal activity. (A) Fourier power spectrum of the BOLD signal corresponding to PFCORB node. (B) BOLD signal time-series shown for PFCORB, PFCM, FEF. (C) 38×38 correlation matrix computed from the simulated BOLD signals. (D) BOLD signal activity for the six regions corresponding to the report of Fox *et al.* (2005) [3]. (Reprinted with permission from [6]).

the log-log plot, with a dominant frequency component around 10 Hz, corresponding to the alpha band.

4.4.1 Time-series analysis

Previous attempts at demonstrating low-dimensional nonlinear structure in EEG data have had only limited success [14, 17]. Here we will compute various measures typically used in analyzing complex time-series data.

The time-delayed *mutual information* takes into account nonlinear correlations, and is a useful tool to determine the delay-time for embedding time-series data [4], and is defined as

$$M(\tau) = - \sum_{ij} p_{ij}(\tau) \ln \frac{p_{ij}(\tau)}{p_i p_j}, \quad (4.6)$$

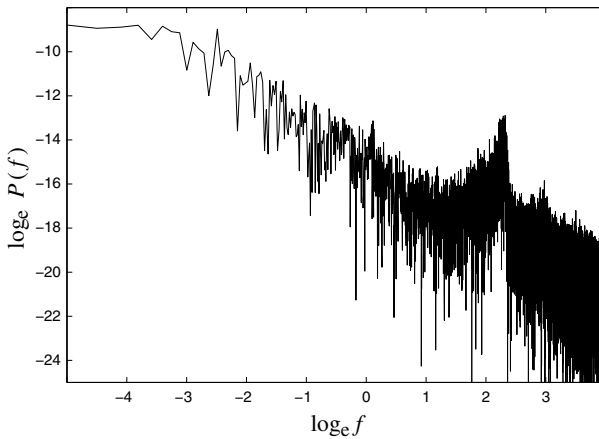


Fig. 4.9 Characteristic power spectrum of the resting-state EEG shows a $1/f$ -like decay with an approximate slope -2 and a peak at ~ 10 Hz (alpha waves).

where $p_{ij}(\tau)$ is the joint probability that an observation lies in i^{th} interval, given that observation after time τ lies in j^{th} interval. The first minimum in $M(\tau)$ vs τ plot is a good estimate of time-delay. Our estimated time-delay is $\tau = 40$ ms (see Fig. 4.10(a)). Using the delay information, we reconstruct the phase space by time-delayed embedding of the time-series data. For low-dimensional chaotic signals, the phase space is expected to show a strange attractor whose dimension can be estimated by computing the correlation sum [7],

$$C(m, \varepsilon) = \frac{1}{N} \sum_{j=m}^N \sum_{k < j} \Theta(\varepsilon - |x_j - x_k|). \quad (4.7)$$

For sufficiently small length-scale ε , and embedding dimension m exceeding the box dimension of the attractor, the correlation sum scales as $C(m, \varepsilon) = \varepsilon^D$, where D is the *correlation dimension*. In Fig. 4.10(b), we plot $\log C(m, \varepsilon)$ vs $\log \varepsilon$ as a function of embedding dimension, m , and do not observe convergence with increasing m . Hence we conclude that the underlying attractor of the rest dynamics lies in a high-dimensional space.

Moreover, it needs to be seen if the rest-state EEG has any characteristic chaotic signature. Chaotic time-series are characterized by their spectrum of Lyapunov exponents that describe the exponential growth-rate of infinitesimal perturbations [9]. The maximal Lyapunov exponent is estimated from the linear slope of

$$S(\varepsilon, m, t) = \left\langle \ln \left(\frac{1}{|U|} \sum_{x_m \in U} |x_{n+t} - x_{m+t}| \right) \right\rangle_n \quad (4.8)$$

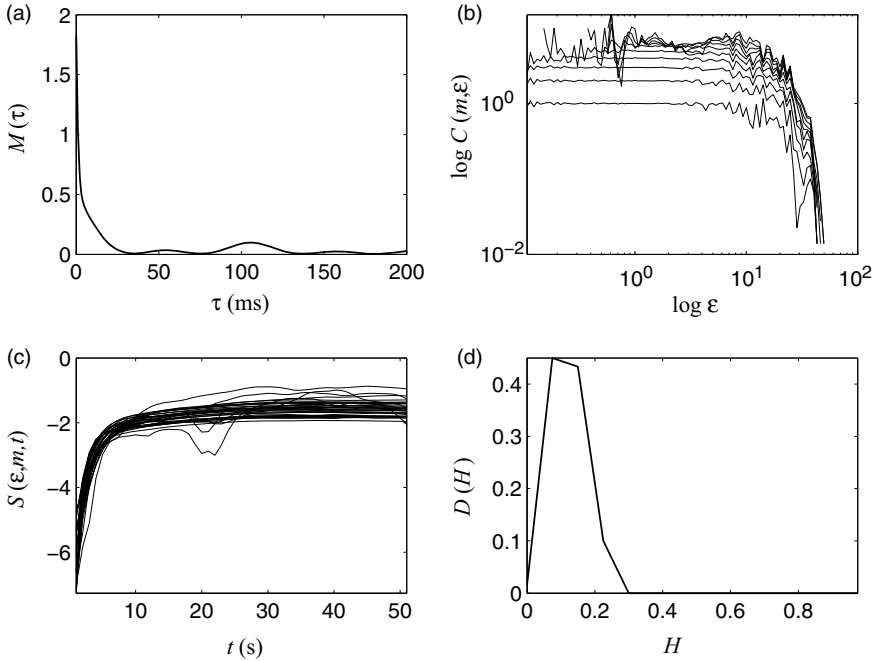


Fig. 4.10 (a) Mutual information indicates time-delay is $\tau = 40$ ms; (b) $\log C(m, \epsilon)$ vs $\log \epsilon$ plot, does not show convergence as a function of embedding dimension, m ; (c) $S(\epsilon, m, t)$ does not show a significant linear regime for calculating Lyapunov exponents; (d) Distribution of Hurst exponents shows significant peak at $H > 0$.

plotted as a function of t for all embedding dimension m for reasonable neighborhood size ϵ . However, we observe in Fig. 4.10(c) that $S(\epsilon, m, t)$ does not show a significant linear regime, reflecting lack of exponential divergence of nearby trajectories. Our findings indicate that rest-state EEG dynamics does not conform to low-dimensional chaos, but instead indicate high-dimensional complexity. At this point one may also suspect that the rest-state EEG signal is predominantly noise.

The next measure that we propose to distinguish the rest-state activity from noise is the Hurst exponent [1]. The Hurst exponent is a quantification of the degree of independence or the relative tendency of observables to cluster towards certain values, and can be estimated from the structure function:

$$S_q = \langle |x(t + \tau) - x(t)|^q \rangle_T \approx \tau^{qH(q)}, \quad (4.9)$$

where $q > 0$, τ is the time-delay, and averaging is done over time-window T . For Gaussian white noise $H(q) = 0$ indicating statistical independence, while for correlated observations $H(q) \rightarrow 1$. The distribution of Hurst exponents estimated for all EEG channel recordings during rest-state dynamics has a nonzero mean, indicating that the data have significant correlations and are not white noise (see Fig. 4.10(d)). However, from time-series analysis we do not gain sufficient insight into the

dynamics of the rest-state, so resort to spatiotemporal analysis to reveal the dynamical features.

4.4.2 Spatiotemporal analysis

An important feature of EEG is that the data are recorded at a number of electrode locations simultaneously. Thus the aim of a spatiotemporal analysis is to reveal if there are any inter-area correlations. A first measure in this direction is the degree of *phase synchronization*. This can be characterized by quantifying phase-locking in terms of a synchronization index,

$$\sigma = (S_{\max} - S)/S_{\max}, \quad \text{where} \quad S = -\sum p_k \ln p_k \quad (4.10)$$

is the entropy of the distribution

$$\phi_{nm} = n\phi_i - m\phi_j. \quad (4.11)$$

The phase ϕ_i of electrode location i is computed by the Hilbert transform of the corresponding EEG signal. The normalized index is in the range of $0 \leq \sigma \leq 1$, where $\sigma = 0$ indicates no synchronization, and $\sigma = 1$ shows complete synchronization [16].

Here we observe that while time-series analyses of individual EEG channel recordings exhibit no signature of low-dimensional chaos, their distribution of synchronization indices ($n:m = 1:1$) computed for all channels and subjects exhibit a high degree of synchronization (see Fig. 4.11). Mutual correlation of EEG data

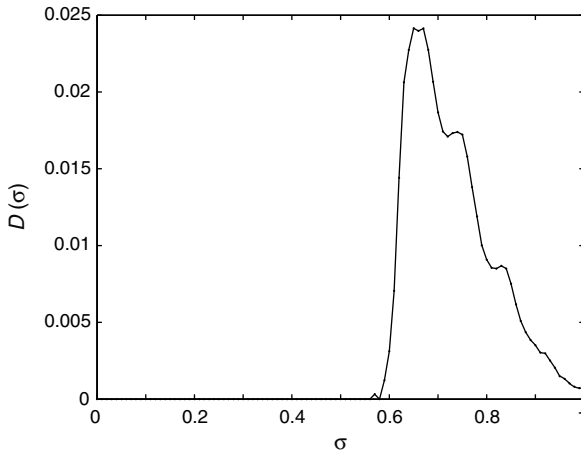


Fig. 4.11 Distribution of the synchronization index indicates a high degree of synchronization across electrodes during the resting state.

indicates that the data can be subjected to principal component analysis (PCA) to identify the spatial structures.

Let the spatiotemporal data be decomposed by PCA as,

$$x(i, t) = \sum_k \xi_k(i) \psi_k(t), \quad (4.12)$$

where $\xi_k(i)$ are the spatial modes spanning the space of electrode locations, and $\psi_k(t)$ are the corresponding temporal coefficients. Now the spatiotemporal resting-state data are subjected to PCA. The cumulative sum of the normalized PCA eigenvalues indicates that the first four principal modes are sufficient to capture $\sim 95\%$ of the data (see upper asterisked-curve in Fig. 4.12(a)). For comparison we also show the cumulative sum of PCA normalized eigenvalues computed from random variables. The spatial modes are plotted in interpolated scalp surface (viewed from above; see Fig. 4.12(b)). The first four principal modes are similar to the first lowest-order spherical harmonics.

The Fourier power spectrum of temporal coefficients, $\psi(t)$, retains power-law fluctuations and the dominant alpha oscillations (see Fig. 4.13(a)). Now we bandpass the temporal coefficients and select the signal in the alpha-band, i.e., 8–13 Hz. Convolution of bandpassed $\psi_k^\alpha(t)$ and $\xi_k(i)$ exhibits alpha waves. Alpha waves can be standing, traveling, or rotating in nature, and can be quantified in the following way: we compute the phase variables $\phi(t)$ from $\psi(t)$ by Hilbert transform, and the phase differences Φ between different PCA modes. The distribution of $\Phi_{12} = \phi_1 - \phi_2$ (i.e., phase difference between the first two modes) shows a peak around $\Phi_{12} \approx \pi/2$ (see Fig. 4.13(b)), implying the presence of *longitudinal* waves traveling from anterior to posterior. *Transverse* traveling waves are less frequent as the distribution of $\Phi_{13} = \phi_1 - \phi_3$ has no significant peak. Moreover, the $\Phi_{23} = \phi_2 - \phi_3$ distribution also shows a peak at around $\pi/2$, explaining occurrences of mostly counterclockwise rotating waves.

4.5 Final remarks

When the brain is at rest, its resting activity is *not* zero—in fact, it displays a rich dynamics. The concept of rest of the brain is difficult on various levels of consideration. What does it actually mean when the brain “does nothing”? One of the most often posed questions at conferences is, “How do you actually know that the brain is at rest?” We do know. Why? By definition: when a human subject is properly instructed to close the eyes and attempt to neither move nor think of anything, the associated brain activity *is* the resting-state of the brain.

Another argument often raised is that it is impossible to think of nothing, since thoughts will appear involuntarily, even though typically only for a brief time. These occurrences of transient thoughts do not violate the constraint of “doing nothing”; in fact, these transient explorations of the brain’s dynamic repertoire is what we seek to understand. When the brain activates certain network configurations (but not in

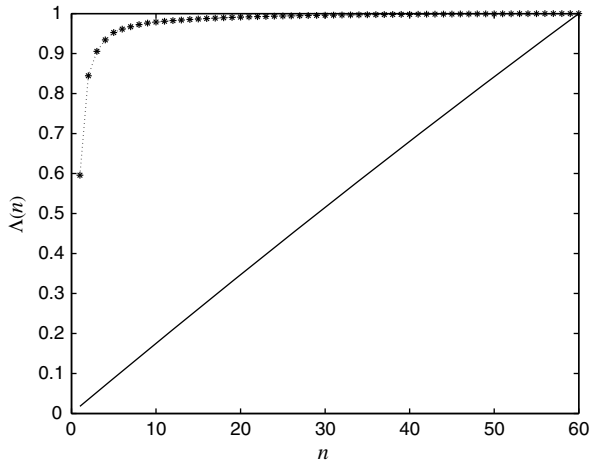


Fig. 4.12 (a) Cumulative sum of PCA eigenvalues for resting state (*-*), and for random variables (-). The first four modes, capturing 95% of the variation, are plotted below in the Fig. 4.12(b) color-map.

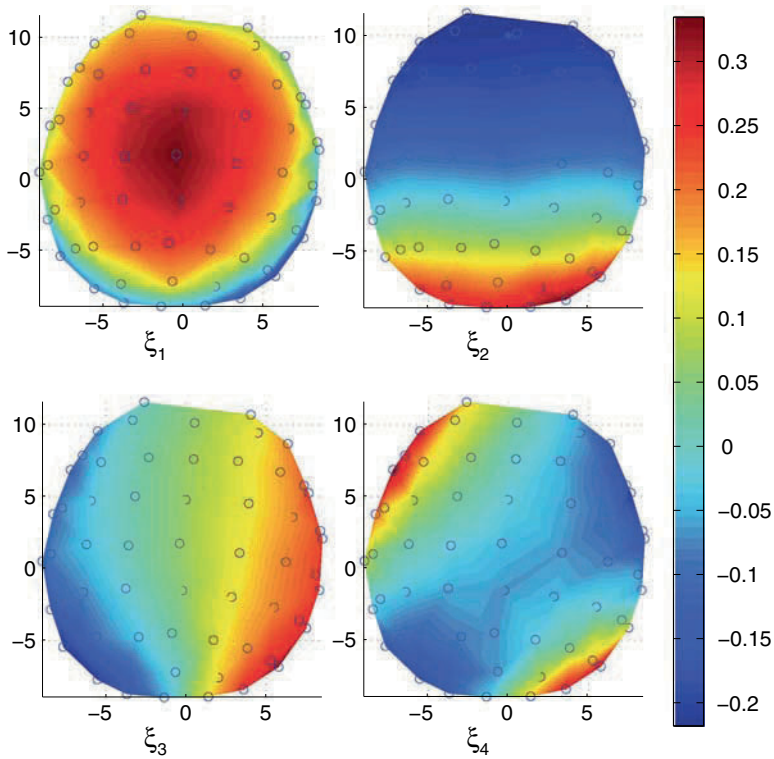


Fig. 4.12(b) [Color plate] First four spatial modes obtained from PCA for the resting-state EEG. Modes are plotted in interpolated scalp surface, viewed from above.

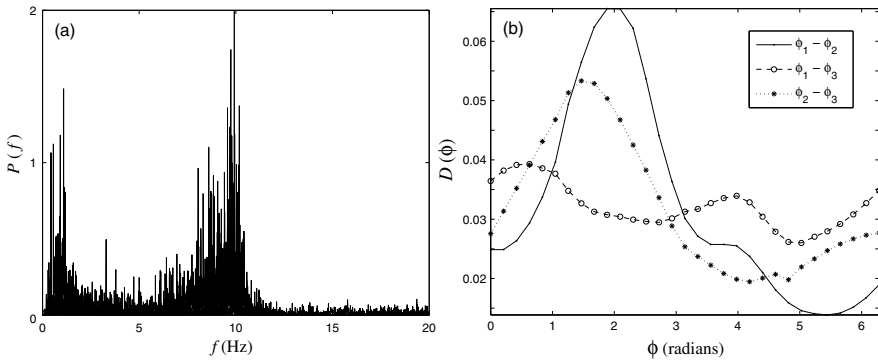


Fig. 4.13 (a) Fourier power spectrum of all four temporal modes of PCA shows that the alpha band has the most significant contribution. (b) Distribution of phase differences Φ_{ij} with $i, j = 1, 2, 3$.

the sense of a “first mover”) during resting activity, such transient thoughts should even be expected.

The characteristic resting-state networks include as a subset the default network of the brain, which is *ex negativo* associated with cognitive processes. Philosophically, this imposes only mild constraints upon the reader’s favorite cognitive theory: an identity of thought and brain activations is not required, a mild constructivism regarding the emergence of thought processes being fully sufficient. However, in this chapter we have identified several of the key ingredients necessary for the emergence of undirected thought, i.e., for the emergence of the resting-state of the brain. Without the noise in physiological systems such as the brain, the resting-state would be truly a state and truly at rest; in other words, the brain’s dynamics would relax to its stable equilibrium point and remain there, until a new stimulation occurs or a new action is required. No undirected thought, no transient emergence of thoughts would be possible (within the framework developed in this chapter). It is the presence of noise that initiates these processes, but also determines their irregularity. It is the deterministic skeleton of the brain, though, which prescribes the coherence and consistency of these transient thought processes. This deterministic skeleton consists of the anatomical connectivity, its time-delay structure, and the response properties of individual brain areas (their intrinsic dynamics). In conjunction, these elements define a deterministic set of behaviors (the brain’s dynamic repertoire) open to exploration through the noise. However, some degree of tuning is required. In order to allow the noise to perform such exploration, the brain’s deterministic skeleton must be close to an instability (or bifurcation or phase transition), else the effect of the noise will be negligible.

We have listed here only the architectural elements necessary to achieve a network dynamics as observed in noninvasive brain imaging during rest. Such has been the intention of this chapter. We have also discussed its dynamic implications and consequences. What we have not commented upon, is any implicated function or purpose of the resting-state.

Acknowledgments V.K.J. acknowledges support by the ATIP and Neuroinformatique programs of the CNRS, as well as support by the J.S. McDonnell Foundation and Codebox Research.

List of abbreviations

A1	primary auditory cortex	PFCVL	ventrolateral prefrontal cortex
A2	secondary auditory cortex	PHC	parahippocampal cortex
CCA	anterior cingulate cortex	PMCDL	dorsolateral premotor cortex
CCP	posterior cingulate cortex	PMCM	medial (supplementary) premotor cortex
CCR	retrosplenial cingulate cortex	PMCVL	ventrolateral premotor cortex
CCS	subgenual cingulate cortex	Pulvinar	pulvinar thalamic nucleus
FEF	frontal eye field	S1	primary somatosensory cortex
IA	anterior insula	S2	secondary somatosensory cortex
IP	posterior insula	TCC	central temporal cortex
M1	primary motor cortex	TCI	inferior temporal cortex
PCI	inferior parietal cortex	TCPOL	polar temporal cortex
PCIP	intraparietal sulcus cortex	TCS	superior temporal cortex
PCM	medial parietal cortex	TCV	ventral temporal cortex
PCS	superior parietal cortex	ThalAM	thalamus
PFCCL	centrolateral prefrontal cortex	V1	primary visual cortex
PFCDL	dorsolateral prefrontal cortex	V2	secondary visual cortex
PFCDM	dorsomedial prefrontal cortex	VACD	dorsal anterior visual cortex
PFCM	medial prefrontal cortex	VACV	ventral anterior visual cortex
PFCORB	orbital prefrontal cortex		
PFCPOL	polar prefrontal cortex		

References

1. Di Matteo, T.: Multi-scaling in finance. *Quant. Financ.* **7**, 21–26 (2007), doi:10.1080/14697680600969727
2. FitzHugh, R.: Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.* **1**, 445–466 (1961), doi:10.1016/S0006-3495(61)86902-6
3. Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., van Essen, D.C., Raichle, M.E.: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. USA* **102**(27), 9673–9678 (2005), doi:10.1073/pnas.0504136102
4. Fraser, A.M., Swinney, H.L.: Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* **33**, 1134–1140 (1986), doi:10.1103/PhysRevA.33.1134
5. Friston, K.J., Ashburner, J., Frith, C.D., Pline, J.B., Heather, J.D., Frackowiak, R.S.J.: Spatial registration and normalization of images. *Hum. Brain Mapp.* **2**, 165–189 (1995), doi:10.1002/hbm.460030303
6. Ghosh, A., Rho, Y., McIntosh, A.R., Kötter, R., Jirsa, V.K.: Noise during rest enables exploration of the brain’s dynamic repertoire. *PLoS Comput. Biol.* **4**(10), e1000196 (2008), doi:10.1371/journal.pcbi.1000196
7. Grassberger, P., Procaccia, I.: Characterization of strange attractors. *Phys. Rev. Lett.* **50**, 346–349 (1983), doi:10.1103/PhysRevLett.50.346
8. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**, 500–544 (1952)

9. Kantz, H., Schreiber, T.: *Nonlinear time series analysis*. Cambridge University Press, Cambridge (1997)
10. Kötter, R.: Online retrieval, processing, and visualization of primate connectivity data from the CoCoMac database. *Neuroinformatics* **2**, 127–144 (2004), doi:10.1385/NI:2:2:127
11. Kötter, R., Wanke, E.: Mapping brains without coordinates. *Phil. Trans. R. Soc. Lond. B* **360**, 751–766 (2005), doi:10.1098/rstb.2005.1625
12. Kunetsov, Y.A. *Elements of Applied Bifurcation Theory*. Applied Mathematical Sciences **112**. Springer-Verlag, New York (1995)
13. Nagumo, J., Arimoto, S., Yoshizawa, S.: An active pulse transmission line simulating nerve axon. *Proc. IRE* **50**, 2061–2070 (1962), doi:10.1109/jrproc.1962.288235
14. Palus, M.: Nonlinearity in normal human EEG: cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biol. Cybern.* **75**, 389–396 (1996), doi:10.1007/s004220050304
15. Sporns, O., Tononi, G.: Structural determinants of functional brain dynamics. In: V.K. Jirsa, A.R. McIntosh (eds.), *Handbook of Brain Connectivity*, chap. 4, pp. 117–148, Springer, Heidelberg (2007), doi:10.1007/978-3-540-71512-2_4
16. Tass, P., Rosenblum, M.G., Weule, J., Kurths, J., Pikovsky, A., Volkman, J., Schnitzler, A., Freund, H.J.: Detection of n:m phase locking from noisy data: Application to magnetoencephalography. *Phys. Rev. Lett.* **81**, 3291–3294 (1998), doi:10.1103/PhysRevLett.81.3291
17. Theiler, J., Rapp, P.: Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram. *Electroen. Clin. Neuro.* **98**, 213–222 (1996), doi:10.1016/0013-4694(95)00240-5

Chapter 5

Limited spreading: How hierarchical networks prevent the transition to the epileptic state

M. Kaiser and J. Simonotto

5.1 Introduction

An essential requirement for the representation of functional patterns in complex neural networks, such as the mammalian cerebral cortex, is the existence of stable network activations within a limited critical range. In this range, the activity of neural populations in the network persists between the extremes of quickly dying out, or activating the whole network. The latter case of large-scale activation is visible in the transition to the epileptic state. It is known in neuroanatomy that the neuronal network of the mammalian cerebral cortex possesses a modular organization across several levels of organization—from cortical clusters such as the visual cortex at the highest level, to individual columns at the lowest level. Using a basic spreading model of a network without inhibitory units, we investigate how functional activations of nodes propagate through such a hierarchically clustered network. Simulations demonstrate that persistent and scalable activation can be produced in clustered networks, but not in random networks of the same size. Moreover, the parameter range yielding critical activations is substantially larger in hierarchical cluster networks than in same-sized small-world networks. These findings indicate that a hierarchical cluster architecture may provide the structural backbone for the stable and diverse functional patterns observed in cortical networks additional to the known role of inhibitory neurons. Such *topological inhibition* might help to maintain healthy levels of neural activity. For readers who are unfamiliar with the emerging area of network science, we provide a glossary of key terms at the end of the chapter.

Natural systems operate within a critical functional range, sustaining diverse dynamical states [5, 41]. For instance, in neural systems, such as the cerebral cortical

Marcus Kaiser · Jennifer Simonotto

School of Computing Science, Newcastle University, Newcastle-upon-Tyne NE1 7RU, U.K.

Institute of Neuroscience, Newcastle University, Newcastle-upon-Tyne NE2 4HH, U.K.

e-mail: m.kaiser@newcastle.ac.uk jennifer.simonotto@newcastle.ac.uk

<http://www.biological-networks.org/>

networks of the mammalian brain, this critical range is indicated by the fact that initial activations result in various neuronal activity patterns that are neither dying out too quickly, nor spreading across the entire network too often as large-scale activation is infrequent [7]. What are the essential structural and functional parameters that allow complex neural networks to maintain such a dynamic balance? In particular, which factors limit the spreading of neural activity through the whole brain, thus preventing a pathological state resembling epilepsy? Preventing the spreading is important as there are few processing steps in the brain as indicated by the analysis of cortical connectivity [19, 28] and of cortical latencies [54].

Most current models of neural network dynamics focus on maintaining the right balance of activation through functional interactions among populations of inhibitory and excitatory nodes [7, 18]. However, the topology of the networks may also make a significant contribution toward critical network dynamics, even in the absence of inhibitory nodes. Earlier studies at a single level of neural organization had shown that a small-world organization of a network of excitatory neurons was related to patterns of synchrony [37] and epilepsy spreading [11, 40]. In our model, we will observe how hierarchies, in addition to properties of small-world networks, influence network dynamics.

5.1.1 Self-organized criticality and avalanches

Nonlinear dynamics and criticality arise in natural systems through the interplay of many variables and degrees of freedom. In theoretical and computational models, these systems can be represented by differential or difference equations, and typically have at least three variables, or degrees of freedom (the logistic map being a notable exception). The nonlinear aspect of these interactions causes systems to have varying responses to stimuli and input, based on the “state” of the system as a whole. For example, some input at one point in time may have a certain output, but an identical input at some later time can result in a very different output of the system, due to different initial conditions. Taken’s theory of embedding [51] and Sauer’s extension to time-delay embedding [47] allow one to recreate these state spaces, allowing one to visualize attractors. Thus, one may understand both temporally local and temporally global dynamics: locally, a linear approximation to translate output to input is possible; globally, if one watches the dynamics long enough, one may reconstruct the entire attractor. However, prediction of intermediate-term behavior is not currently possible.

Examination of how these attractors change when variables are changed allows one to identify critical points within a system; these critical points are phase states in which very different types of behavior result from mildly different initial conditions. In some systems, critical points are the attractors of a system; in this case, the variables themselves are less important and it is from the inputs that one sees critical-point transitions in behavior. Such systems are referred to as self-organized critical systems [4, 5]; earthquakes, sandpile avalanches, and large ensembles of neurons are examples. Self-organized critical systems are typically slow-driven

nonequilibrium systems, with a large degree of freedom and high nonlinearity, but there is no set of characteristics that guarantees that a given system will display self-organized criticality [52].

Another characteristic of self-organized critical systems is scale invariance, in which fluctuations have no characteristic time or spatial scale. Spatially extended critical systems which exhibit scale invariance [14] are of increasing interest in many natural systems, including the brain. Variability also exists in the underlying network topology of systems. For scale-free networks, connections between nodes of a network are not uniformly randomly distributed, but follow a power-law of distribution, with certain nodes acting as highly-connected hubs [6]. This type of connectedness gives robustness of operation even with loss of random connections (“damage”) between nodes, so long as the hubs are not completely disconnected from the network [2]. A similar response to structural damage as for scale-free networks was observed for cortical networks [30].

5.1.2 Epilepsy as large-scale critical synchronized event

Epilepsy affects 3–5% of the population worldwide. Seizures are the clinical manifestation of an abnormal and excessive excitation and synchronization of a population of cortical neurons. These seizures can spread along network connections to other parts of the brain (depending on the type and severity of the seizure), and can be quite debilitating in terms of quality of life, cognitive function, and development. In the vast majority of cases, seizures arise from medial temporal structures that have been damaged (due to injury or illness) months to years before the onset of seizures [12]. Over this “latent period”, cellular and network changes are thought to occur which precipitate the onset of seizures.

It is not understood exactly how these seizures come about, but is thought to be due to structural changes in the brain, as in the loss of inhibitory neurons, the strengthening of excitatory networks, or the suppression of GABA receptors [12, 31]. Cranstoun et al. (2002) reported self-organized criticality in EEG (electroencephalogram) recordings from human epileptic hippocampus; thus applying network analysis to this system may reveal useful information about the development (and possible prevention) of seizures. As the networks that support the spread of seizure activity are the very same networks that also support normal cognitive activity, it is important to understand how this type of activity arises in networks in general [16]. The question of how seizures are initiated (ictogenesis) is also of great interest, as further elucidation of either epileptogenesis or ictogenesis may have considerable impact on the treatment (and possible cure) of epilepsy [24].

5.1.3 Hierarchical cluster organization of neural systems

It is known from the anatomy of the brain that cortical architecture and connections are organized in a hierarchical and modular way, from cellular microcircuits in

cortical columns [8] at the lowest level, via cortical areas at the intermediate level, to clusters of highly connected brain areas at the global systems level [19, 20, 50]. At each level, clusters arise, with denser connectivity within than between modules. This means that neurons within a column, area, or area cluster are more frequently linked with each other than with neurons in the rest of the network.

Cluster organization at the global level is, for example, visible in the pattern of corticocortical connectivity between brain areas in the cat [48, 49]. Based on the structural connectivity of anatomical fiber tracts it is possible to distinguish four clusters which closely resemble different functional tasks (Fig. 5.1). Cluster organization is also visible at the level of cortical areas, for example, about 30–40% of synapses within visual areas come from distant cortical areas or thalamic nuclei [54], thus the majority of connections runs within an area. Within cortical columns of area 17 of the cat, two-thirds of synapses within layers come from external neurons in different layers [8]. Nonetheless, a neuron is more likely to connect to a neuron in the same layer than to a neuron in a different layer. After discussing the transition to the epileptic state in the next section, we will show how the cluster organisation of neural systems can prevent this transition in the normal brain, and we will identify which changes could lead to seizures in epileptic patients.

5.2 Phase transition to the epileptic state

The phase transition to the epileptic (“ictal”) state is abrupt from a behavioral point of view (seizures start suddenly), but from an electrical/network point of view, there are subtle connectivity and synchronization-related changes in network activity that can indicate that a seizure will occur soon (with a prediction window ranging from minutes to hours). The existence of this so-called “pre-ictal” period—in which one is neither “inter-ictal” (between seizure states), nor currently having a seizure—has been the subject of intense debate in the literature, but more and more evidence points to its existence [24]. Epileptogenesis typically has a longer timescale of development (months to years) than ictogenesis (weeks to days), but understanding the changes of epileptogenesis and how seizures become more easily generated is also of considerable interest, as characterization of network changes may allow one to treat epilepsy in a more precise manner (i.e., with no systemic drug application or removal of whole brain areas in order to eliminate malfunctioning pathways).

5.2.1 Information flow model for brain/hippocampus

The hippocampus is a well-studied part of the brain, and is an especially important part of the limbic system, as it has to do with memory formation and information processing. Limbic epilepsy, in particular temporal lobe epilepsy, is a particularly debilitating form as it can be difficult to treat surgically without adverse quality-of-life effects [12]. Avoli et al. (2002) examined information flow within

the hippocampus and reported changes of this structure in animal models of limbic epilepsy. They reported a change within the information flow of the hippocampus, involving the loss of connectivity from the CA3 area to the rest of the hippocampus, as well as increased connectivity from the entorhinal cortex, an area normally made quiescent by a 0.5–1-Hz signal from CA3. The time-course of these changes, and the nature of the structure–function alterations as the animal behavior alters (from normal to epileptogenic) are difficult to characterize. This is because these changes occur over an extended period of time, so require large-scale storage and computing facilities in order to contain and analyze data of sufficiently high spatial and temporal resolution, captured over the entire critical period.

5.2.2 *Change during epileptogenesis*

The Chronic Limbic Epilepsy [35, 36] model is a rodent model of limbic epilepsy in which the animal is kindled into status epilepticus for one hour. Following a recovery period of 12–24 hours, spontaneous seizures occur within 2–8 weeks, which are recurrent and chronic.¹ A total of 32 tungsten microwire electrodes were implanted in the CA1 and dentate gyrus subfields of the hippocampus bilaterally, with ~ 8 microwires implanted into each field. The electrodes were implanted in two rows spaced 420 μm apart, with each electrode in the row spaced at 210- μm intervals. Electrode voltages were digitized at 16 bits, and recorded continuously at 12 kHz using custom-written acquisition software and a Tucker-Davis Pentusa DSP, which employed a hardware bandpass filter set from 0.5 Hz to 6 kHz.

Two weeks of baseline data were recorded after the animal had had sufficient time to recover from electrode implantation. The animal was then kindled in the manner prescribed for the Chronic Limbic Epilepsy Animal model [35]. Continuous recording began within a day of kindling, and continued until after the spontaneous electrographic and behavioral seizures had ended. A control animal was recorded using the same protocol. All animals were continuously video-recorded to monitor for seizures.

Coherence, defined as

$$C_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)},$$

is a measure used to determine the degree of linear similarity between two signals [23, 43]. Coherence has been applied to the human EEG in order to determine the relationship between signals for determining seizure propagation delay [15, 17]. A significant increase (or decrease) in coherence would indicate whether two time-series have quantifiably similar frequency properties (or have more dissimilar frequency properties) over that time period. During the latent period before the onset of a seizure, one might predict that an increase in coherence might occur across

¹ The work described here was undertaken at the University of Florida as part of the “Evolution into Epilepsy” NIH/NHS joint research project (grant no. 1R01EB004752).

the epileptic brain compared to normal animals, or that changes might occur preferentially in different frequency bands.

Averaged coherence of inter-hemispherical activity in high gamma to ripple frequencies (40 to 200 Hz, called “low band” in analysis and subsequent figures) showed a significant suppression of coherence between hemispheres ($p < 0.0015$) in stimulated animals compared to nonstimulated animals (see Fig. 5.2). Medvedev reported findings that coherence decreased in the hippocampus at frequencies from 20–100 Hz, suggesting an “anti-binding” mechanism; our findings indicate that this decrease in coherence is evident for spontaneous epilepsy [38].

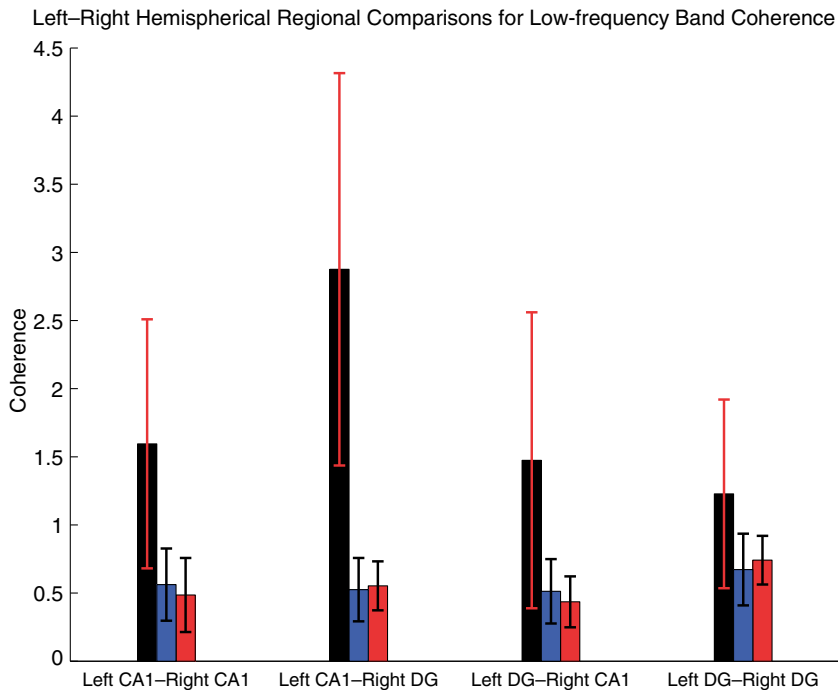


Fig. 5.2 [Color plate] Coherence comparisons of stimulated vs. nonstimulated animals. Mean and standard deviation of coherence in the 40–200-Hz band for two stimulated animals (blue and red bars), and one nonstimulated animal (black bar) are shown. Note that the inter-hemispherical coherence is suppressed in the stimulated animals.

5.3 Spreading in hierarchical cluster networks

5.3.1 Model of hierarchical cluster networks

How can the topology of neuronal networks reduce or enhance the probability of a transition to the epileptic state? We used a basic spreading model to explore the

role played by different network topologies in producing persistent, yet contained, activations. Spreading analysis has also been applied to cortical networks at the global level [33], and to other complex networks with a nonrandom organization [10, 21, 44].

The present model operates without inhibitory units such as cortical inhibitory interneurons, as we were specifically interested in the contribution of network topology. This lack of inhibition is also reflective of structural attributes of cortical networks [34], and other complex networks such as social networks [44].

In our model, individual network vertices represent cortical columns whose connectivity follows the levels of hierarchical organization (Fig. 5.3(a)). Networks were undirected graphs with $N = 1000$ vertices and $E = 12000$ edges. To create the hierarchical cluster network, 1000 vertices were divided up into 10 disjoint sets (“clusters”), each consisting of 100 vertices. Each cluster was further split into 10 “sub-clusters” containing 10 vertices each. The network was wired randomly, such that 4000 edges (one third of the total 12000 connections) connected vertices within the same subclusters, 4000 edges connected vertices within the same clusters, and 4000 were randomly distributed over all nodes of the network (Fig. 5.3(b)). The edge density in these networks was 0.025 whereas the clustering coefficient was 0.15. The characteristic path length (2.6), however, was similar to that of random networks (2.5), indicating properties of small-world networks [53].

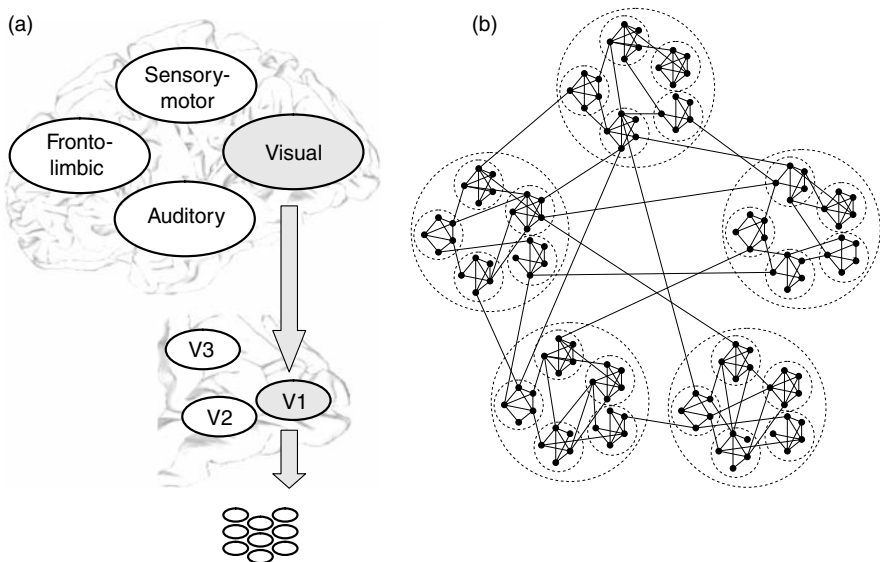


Fig. 5.3 (a) The hierarchical network organization ranges from cluster (e.g., visual cortex), to subcluster (e.g., V1), to individual nodes (cortical columns). (b) Schematic view of a hierarchical cluster network with five clusters, each containing five subclusters.

We compared spreading (i.e., propagation of activation) in hierarchical networks with spreading in random and small-world benchmark networks with the same number of vertices and edges [25]. The small-world networks with a rewiring

probability of $p = 0.5$ had similar clustering coefficient values (0.11) and characteristic path lengths (2.6) to that of the hierarchical networks, but lacked the characteristic cluster architecture. We also generated Erdős-Rényi random networks [13].

5.3.2 Model of activity spreading

We used a simple threshold model for activity-spreading in which a number i of randomly selected nodes were activated in the first step. An additional component was the extent of localization of the initial activation i_0 . For initialization, i ($i \leq i_0$) nodes among the nodes 1 to i_0 were randomly selected and activated in the first time-step. The network nodes were numbered consecutively. For example, by setting i_0 to 10, 20 or 100, only nodes in the first subcluster, the first two subclusters, or the first cluster, respectively, were activated during initialization. Thus, i determined the number of initially activated nodes while i_0 controlled the localization of initial activations, with smaller values resulting in more localized initial activity. At each time-step, inactive nodes become activated if at least k neighbors were activated (neighbors of a node are nodes to which direct connections exist). Activated nodes could become inactive with probability v . As default we used $k = 6$ and $v = 0.3$. The state of the network was determined after 200 steps of the simulation as activity was either dying out (zero activation), spreading through the whole network (more than 50% of the nodes were active), or balanced for an intermediate activation level.

5.3.3 Spreading simulation outcomes

Across different simulation conditions, hierarchical cluster networks show a larger variety of behaviors than do random or small-world networks, and produce persistent yet balanced network activity for a wider range of initial conditions.

Examples exhibiting the behaviors of the different networks are shown in Fig. 5.4. The figure shows the result of 20 simulations in the three network types when 10%

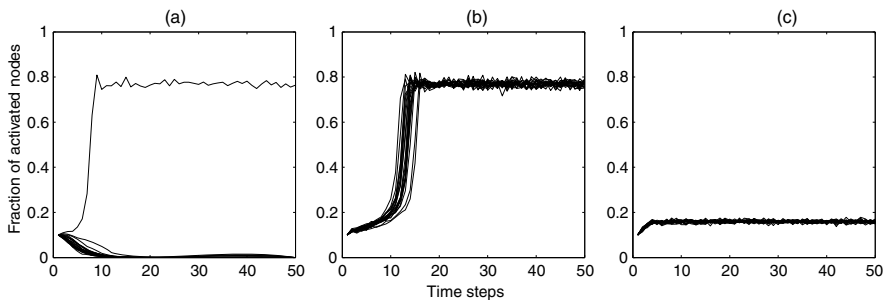


Fig. 5.4 Examples for spread of activity in (a) random, (b) small-world and (c) hierarchical cluster networks ($i = 100$, $i_0 = 150$), based on 20 simulations for each network.

of the nodes were randomly selected for initial activation. In the random network, activity dies out in most cases. In the small-world network, spread of activity results in almost complete activation (NB: 100% activation cannot be achieved due to the deactivation probability for active nodes at each step). In contrast, the hierarchical cluster network produces cases in which spreading is limited. Such persistent activation can be sustained with different patterns and varying extent of involved nodes (see Fig. 5.5).

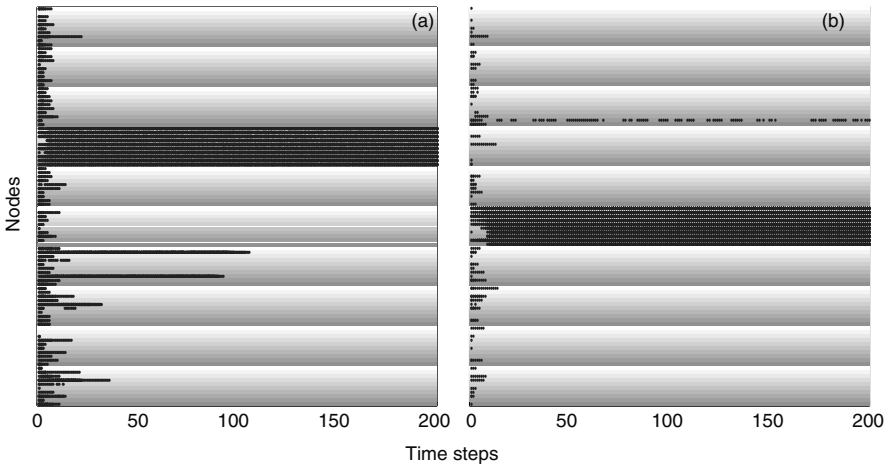


Fig. 5.5 Examples for different sustained activity patterns in hierarchical cluster networks ($i = 90, i_0 = 1000$). Graded gray background shading indicates the 10 subclusters within each of the 10 clusters. Black dots represent nodes active at the respective time-step. (a) One cluster showing sustained activity. (b) One cluster remaining active with frequent co-activation of one external subcluster.

5.3.3.1 Delay until large-scale activation

Does the “speed” with which the whole network can become activated depend on the network topology? For those cases where large-scale activation was observed, we looked at the number of time-steps required to reach this state. For the random network, if activity spread at all, it did so rapidly, typically in less than 10 time-steps. Even if the initial activity was in the borderline range for all-or-none network activation, not more than 15 time-steps were required in any of the cases. This was in contrast to the small-world and hierarchically clustered networks, for which a wide range of delay times was observed. For the small-world network, delayed spreading depended on whether initial activity was strictly localized ($i_0 = i$). Setting $i_0 = i = 90$ typically resulted in about 40 time-steps for spreading, whereas for $i_0 = 190, i = 90$, spreading in the small-world network appeared similar to that in the random network. By contrast, for the hierarchically clustered network, spreading to the global

level did not arise when the initial activation was too strictly localized. A maximum delay for spreading was achieved by localizing the initial activity within two or three clusters (e.g., delay around 40 steps for $i_0 = 200, i = 90$). Thus neighborhood clustering in the small-world and hierarchical networks slows down the spreading of activation. Note that the increase in delay compared to the random network is larger than would be expected from the increase of the characteristic path length. These results indicate that limiting the number of short-cuts or connections between clusters acts as a bottleneck for the spreading of activation. We will come back to this point later.

5.3.3.2 Robustness of sustained-activity cases

The higher likelihood of sustained activation in hierarchical networks is largely independent of our choice of model parameters. We systematically explored the network activation behaviors resulting from different settings of the initial node activation and localization parameters. Both the number of initially activated nodes and their localization had a critical influence on the resulting spreading patterns [25]. Since at any given time only a fraction of neurons in a neural system will be in the activated state, we limited the maximum number of initially active nodes to 250, that is, one-quarter of all network nodes. Persistent contained activity in hierarchical networks was robust for a wide range of initial localization and activation parameters (indicated by the gray parameter domain in Fig. 5.6). For small-world networks, however, parameters needed to be finely tuned in order to yield sustained activity. Thus, hierarchical networks showed sustained activity for a wider range of initial activation conditions.

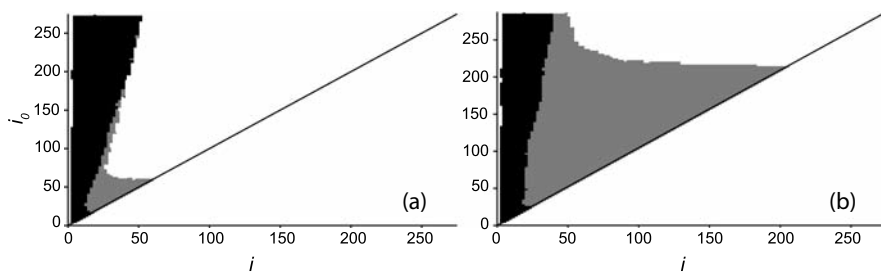


Fig. 5.6 Parameter space exploration of the critical range for all combinations of initial activation parameter i and localization parameter i_0 , based on 1000 test cases. Simulation outcomes are indicated by gray level (black: activity died out; gray: limited spreading; white: complete spreading). (a) Small-world network; (b) hierarchical cluster network.

The results were also robust in terms of the spreading parameters k and ν . Using a Monte Carlo approach, for each pair of k and ν , we generated 20 small-world and 20 hierarchical networks. For each network, the dynamics for 1000 randomly chosen parameters i and i_0 were tested (see Fig. 5.7). A trial was considered to show

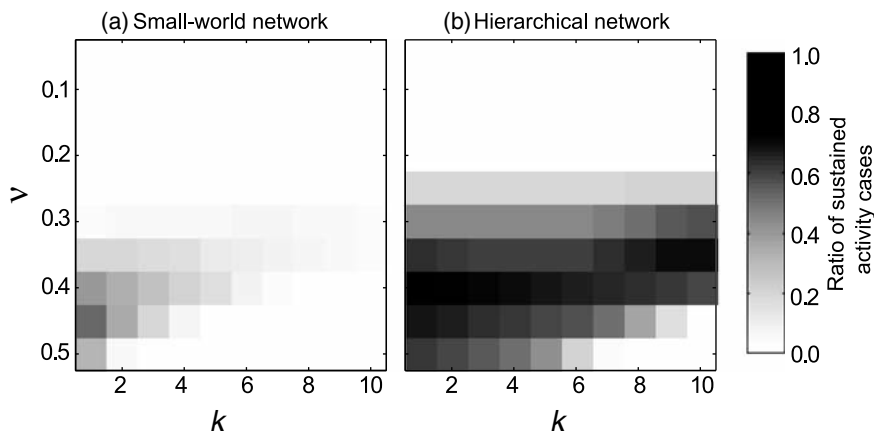


Fig. 5.7 Ratio of sustained activity cases depending on the spreading parameters k (activation threshold) and v (deactivation probability) for (a) small-world, and (b) hierarchical cluster networks.

sustained activity if at least one, but no more than 50%, of all nodes were activated at the end of the simulation. For each pair of spreading parameters k and v , the average ratio of cases for which sustained activity occurred, related to the ratio of the gray space in Fig. 5.6, was larger for hierarchical cluster networks than for small-world networks. The maximum ratio was 67% of the cases for hierarchical cluster networks compared to 30% for small-world networks.

Sustained spreading in hierarchical cluster networks still occurred for different ratios of connectivity within and between clusters and subclusters. However, results differed for large changes in the proportion of connections between modules (clusters or subclusters; see [25] for details): *Reducing the proportion of connections between modules* led to a higher proportion of cases with sustained activity. While the total number of edges was kept constant, the number of connections between cluster and subclusters was reduced. Now, three or more clusters could be persistently activated without a subsequent spread through the whole network (Fig. 5.8a). In these cases, the limited number of inter-cluster connections formed a bottleneck for activation of the remaining clusters. *Increasing the proportion of connections between modules* blurred the boundaries of local network modules and reduced the proportion of cases with sustained activity, but the proportion was still larger than that for small-world networks. However, for these networks, initially contained activation was able to spread through the network at later stages of the simulation (Fig. 5.8b). These results for spreading dynamics are in line with earlier studies on the important role of inter-cluster connections for structural network integrity [26].

For the above model, activated nodes might stay active for a long time, potentially until the end of the simulation. However, energy resources for sustaining neural network activations are limited in real neural systems. For instance, exhaustion occurs during epileptic seizures, reducing the duration of large-scale cortical activation. Therefore, we also tested the effect of restricting the number of

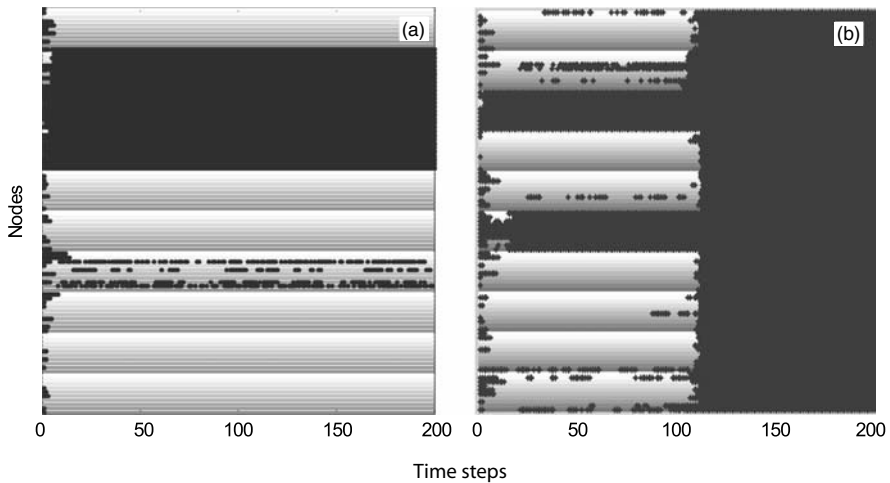


Fig. 5.8 (a) Sustained activity in three clusters, without subsequent spreading through the rest of the network, was possible when the number of connections between clusters was reduced. These few inter-cluster connections created a bottleneck for further activity spreading. (b) When the number of inter-cluster connections was increased, activity was more likely to spread through the entire network. The figure shows an activation that is initially limited to two clusters and subsequently spreads through the whole network.

time-steps that nodes could be consecutively active from seven steps to a single step. Sustained network activation could still occur in the hierarchical cluster network, despite different degrees of limiting node exhaustion: sustained activity was largely independent of the exhaustion threshold parameter. The range of parameters for which sustained activity occurred remained similar to that in the previous analyses, with no clear correlation to the number of steps (average ratio of sustained-activity cases over all pairs of the spreading parameters was 0.272 ± 0.068).

We also tested whether these findings were specific to the threshold activation model described here. Simulations with integrate-and-fire (IF) neurons [32] as network nodes led to similar results. In comparison to random networks, hierarchical cluster network simulations showed easier activation, and exhibited intermediate states of activation [T. Jucikas, private communication]. Thus, our results do not appear to depend on the specific activation model, but are general properties of the topology of the network.

5.4 Discussion

Our simulations demonstrate the strong influence of network topology on spreading behavior. Clustered networks are more easily activated than random networks of the same size. This is due to the higher density of connections within the clusters,

facilitating local activation. At the same time, the sparser connectivity between clusters prevents the spreading of activity across the whole network. The prevalence of persistent yet contained activity in hierarchical cluster networks is robust over a large range of model parameters and initial conditions. In contrast, small-world networks *without* hierarchical modules frequently show a transition to the putative epileptic state of large-scale activation.

The present hierarchical cluster model, which reflects the distributed multilevel modularity found in biological neural networks, is different from previously studied “centralistic” hierarchical modular networks in which most nodes are linked to network hubs [45]. While developmental algorithms have been suggested for the latter type of network, there are currently no algorithms for producing the hierarchical cluster networks presented here. However, single-level clustered network architectures can be produced by models for developmental spatial growth [27, 29, 42] or dynamic self-organization of neural networks [22]; such models may serve as a starting point for exploring the biological mechanisms for developing multilevel clustered neural architectures.

Present results provide a proof of concept for three points. First, persistent but contained network activation can occur in the absence of inhibitory nodes. This might explain why cortical activity does not normally spread to the whole brain, even though top-level links between cortical areas are exclusively formed by excitatory fibers [34]. While the involvement of inhibitory neurons and other dynamic control mechanisms may further extend the critical range, the present results indicate that the hierarchical cluster architecture of complex neural networks, such as the mammalian cortex, may provide the principal structural basis for their stable and scalable functional patterns. Second, in hierarchical clustered networks, activity can be sustained without the need for random input or noise as an external driving force. Third, multiple clusters in a network influence activity spreading in two ways: bottleneck connections between clusters limit global spreading, whereas a higher connection density within clusters sustains recurrent local activity.

5.5 Outlook

It will be important to see how the *topological inhibition* based on the cluster architecture relates to *neuronal inhibition* from inhibitory interneurons. For topological inhibition, an increase in the number of axons between clusters will enhance the likelihood for activity spreading. At the cortical level, this could be visualized as changes in white matter volume that could be detected by tract tracing or diffusion tensor imaging. An alternative way to increase the probability of activity spreading to other clusters would be a larger connection strength of existing inter-cluster connections. For neuronal inhibition, the most effective way for inhibitory neurons to limit large-scale activity spreading from its own cluster to another cluster would be to reduce the activity in excitatory neurons that project to the other cluster. This would be the network analogue to the frequent positioning of inhibitory synapses

close to the axon hillock to prevent the spreading of activation at the individual neuron level. If activity of another cluster—independent of the activity level of an inhibitory neuron’s own cluster—is to be reduced, a direct long-range inhibitory projection to that cluster is needed.

The model of topological inhibition may have practical implications and may guide future research. For instance, it might be worthwhile to test whether epileptic patients show a higher degree of connectivity between cortical network clusters or other changes in structural connectivity which would facilitate spreading. Such changes might be reflected in certain aspects of functional connectivity [1, 46], or might be demonstrated more directly by observing structural changes in brain connectivity (using, for example, diffusion tensor imaging).

Acknowledgments We thank Claus Hilgetag, Matthias Görner and Bernhard Kramer for helpful comments on this chapter. We also thank Tadas Jucikas for performing control simulations with integrate-and-fire networks. Financial support from the German National Merit Foundation, EPSRC (EP/E002331/1), and Royal Society (RG/2006/R2) is gratefully acknowledged.

Glossary: Graph theory and network science

Adjacency (connection) matrix The adjacency matrix of a *graph* is an $n \times n$ matrix with entries $a_{ij} = 1$ if node j connects to node i , and $a_{ij} = 0$ if there is no connection from node j to node i .

Characteristic path length The characteristic path length L (also called “path length” or “average shortest path”) is the global mean of the finite entries of the *distance matrix*. In some cases, the median or the harmonic mean may provide a better estimate.

Clustering coefficient The clustering coefficient C_i of node i is the number of existing connections between the node’s neighbors divided by all their possible connections. The clustering coefficient ranges between 0 and 1 and is typically averaged over all nodes of a *graph* to yield the graph’s clustering coefficient C .

Cycle A *path* which links a node to itself.

Degree The degree of a node is the sum of its incoming (afferent) and outgoing (efferent) connections. The number of afferent and efferent connections is also called the in-degree and out-degree, respectively.

Distance The distance between a source node i and a target node j is equal to the length of the shortest path.

Distance matrix The entries d_{ij} of the distance matrix correspond to the *distance* between node j and i . If no path exists, $d_{ij} = \infty$.

Graph Graphs are a set of n nodes (vertices, points, units) and k edges (connections, arcs). Graphs may be undirected (all connections are symmetrical)

or directed. Because of the polarized nature of most neural connections, we focus on directed graphs, also called digraphs.

Path A path is an ordered sequence of distinct connections and nodes, linking a source node i to a target node j . No connection or node is visited twice in a given path. The length of a path is equal to the number of distinct connections.

Random graph A graph with uniform connection probabilities and a binomial degree distribution. All node degrees are close to the average degree (“single-scale”).

Scale-free graph Graph with a power-law degree distribution. “Scale-free” means that degrees are not grouped around one characteristic average degree (scale), but can spread over a very wide range of values, often spanning several orders of magnitude.

Small-world graph A graph in which the *clustering coefficient* is much higher than in a comparable random network, but the *characteristic path length* remains about the same. The term “small-world” arose from the observation that any two persons can be linked over few intermediate acquaintances [39].

References

1. Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E.: A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* **26**, 63–72 (2006), doi:10.1523/jneurosci.3874-05.2006
2. Albert, R., Jeong, H., Barabási, A.L.: Error and attack tolerance of complex networks. *Nature* **406**, 378–382 (2000), doi:10.1038/35019019
3. Avoli, M., D’Antuono, M., Louvel, J., Köhling, R.: Network and pharmacological mechanisms leading to epileptiform synchronization in the limbic system. *Prog. Neurobiol.* **68**, 167–207 (2002), doi:10.1016/S0301-0082(02)00077-1
4. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality. *Phys. Rev. A* **38**, 364–374 (1988), doi:10.1103/PhysRevA.38.364
5. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality: an explanation of the $1/f$ noise. *Phys. Rev. Lett.* **59**, 381–384 (1987), doi:10.1103/PhysRevLett.59.381
6. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. *Science* **286**, 509–512 (1999), doi:10.1126/science.286.5439.509
7. Beggs, J.M., Plenz, D.: Neuronal avalanches in neocortical circuits. *J. Neurosci.* **23**, 11167–11177 (2003)
8. Binzegger, T., Douglas, R.J., Martin, K.A.C.: A quantitative map of the circuit of cat primary visual cortex. *J. Neurosci.* **24**, 8441–8453 (2004), doi:10.1523/jneurosci.1400-04.2004
9. Cranstoun, S., Worrell, G., Echauz, J., Litt, B.: Self-organized criticality in the epileptic brain. *Proc. Joint EMBS/BMES Conf. 2002* **1**, 232–233 (2002)
10. Dezso, Z., Barabási, A.L.: Halting viruses in scale-free networks. *Phys. Rev. E* **65**, 055103 (2002), doi:10.1103/PhysRevE.65.055103

11. Dyhrfeld-Johnsen, J., Santhakumar, V., Morgan, R.J., Huerta, R., Tsimring, L., Soltesz, I.: Topological determinants of epileptogenesis in large-scale structural and functional models of the dentate gyrus derived from experimental data. *J. Neurophysiol.* **97**, 1566–1587 (2007), doi:10.1152/jn.00950.2006
12. Engel, J.: *Surgical Treatment of the Epilepsies*. Lippincott Williams & Wilkins (1993)
13. Erdős, P., Rényi, A.: On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–61 (1960)
14. Erice workshop on Complexity, Metastability and Nonextensivity: Networks as Renormalized Models for Emergent Behavior in Physical Systems (2004), doi:10.1142/9789812701558_0042
15. Gevins, A., Rémond, A.: *Methods of Analysis of Brain Electrical and Magnetic Signals*. Elsevier (1987)
16. Gómez-Gardeñes, J., Moreno, Y., Arenas, A.: Synchronizability determined by coupling strengths and topology on complex networks. *Phys. Rev. E* **75**, 066106 (2007), doi:10.1103/PhysRevE.75.066106
17. Gotman, J.: Measurement of small time differences between EEG channels: Method and application to epileptic seizure propagation. *Electroenceph. Clin. Neurophysiol.* **56**(5), 501–14 (1983), doi:10.1016/0013-4694(83)90235-3
18. Haider, B., Duque, A., Hasenstaub, A.R., McCormick, D.A.: Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J. Neurosci.* **26**(17), 4535–4545 (2006), doi:10.1523/jneurosci.5297-05.2006
19. Hilgetag, C.C., Burns, G.A.P.C., O’Neill, M.A., Scannell, J.W., Young, M.P.: Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat. *Phil. Trans. R. Soc. Lond. B* **355**, 91–110 (2000), doi:10.1098/rstb.2000.0551
20. Hilgetag, C.C., Kaiser, M.: Clustered organisation of cortical connectivity. *Neuroinf.* **2**, 353–360 (2004), doi:10.1385/NI:2:3:353
21. Hufnagel, L., Brockmann, D., Geisel, T.: Forecast and control of epidemics in a globalized world. *Proc. Natl. Acad. Sci. USA* **101**, 15124–15129 (2004), doi:10.1073/pnas.0308344101
22. Izhikevich, E.M., Gally, J.A., Edelman, G.M.: Spike-timing dynamics of neuronal groups. *Cereb. Cortex* **14**, 933–944 (2004), doi:10.1093/cercor/bhh053
23. Jenkins, G.M., Watts, D.G.: *Spectral Analysis and Its Applications*. Holden-Day (1968)
24. Jung, P., Milton, J.: *Epilepsy as a Dynamic Disease*. Biological and Medical Physics Series, Springer (2003)
25. Kaiser, M., Goerner, M., Hilgetag, C.C.: Criticality of spreading dynamics in hierarchical cluster networks without inhibition. *New J. Phys.* **9**, 110 (2007), doi:10.1088/1367-2630/9/5/110
26. Kaiser, M., Hilgetag, C.C.: Edge vulnerability in neural and metabolic networks. *Biol. Cybern.* **90**, 311–317 (2004), doi:10.1007/s00422-004-0479-1
27. Kaiser, M., Hilgetag, C.C.: Spatial growth of real-world networks. *Phys. Rev. E* **69**, 036103 (2004), doi:10.1103/PhysRevE.69.036103
28. Kaiser, M., Hilgetag, C.C.: Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Comput. Biol.* e95 (2006), doi:10.1371/journal.pcbi.0020095
29. Kaiser, M., Hilgetag, C.C.: Development of multi-cluster cortical networks by time windows for spatial growth. *Neurocomputing* **70**(10–12), 1829–1832 (2007), doi:10.1016/j.neucom.2006.10.060
30. Kaiser, M., Martin, R., Andras, P., Young, M.P.: Simulation of robustness against lesions of cortical networks. *European J. Neurosci.* **25**, 3185–3192 (2007), doi:10.1111/j.1460-9568.2007.05574.x
31. Khalilov, I., Quyen, M.L.V., Gozlan, H., Ben-Ari, Y.: Epileptogenic actions of GABA and fast oscillations in the developing hippocampus. *Neuron* **48**, 787–796 (2005), doi:10.1016/j.neuron.2005.09.026
32. Koch, C., Laurent, G.: Complexity and the nervous system. *Science* **284**, 96–98 (1999), doi:10.1126/science.284.5411.96

33. Kötter, R., Sommer, F.T.: Global relationship between anatomical connectivity and activity propagation in the cerebral cortex. *Philos. Trans. R. Soc. Lond. B* **355**, 127–134 (2000), doi:10.1098/rstb.2000.0553
34. Latham, P.E., Nirenberg, S.: Computing and stability in cortical networks. *Neural Comput.* **16**, 1385–1412 (2004), doi:10.1162/089976604323057434
35. Lothman, E.W., Bertram, E.H., Bekenstein, J.W., Perlin, J.B.: Self-sustaining limbic status epilepticus induced by ‘continuous’ hippocampal stimulation: Electrographic and behavioral characteristics. *Epilepsy Res.* **3**(2), 107–19 (1989)
36. Lothman, E.W., Bertram, E.H., Kapur, J., Stringer, J.L.: Recurrent spontaneous hippocampal seizures in the rat as a chronic sequela to limbic status epilepticus. *Epilepsy Res.* **6**(2), 110–8 (1990), doi:10.1016/0920-1211(90)90085-A
37. Masuda, N., Aihara, K.: Global and local synchrony of coupled neurons in small-world networks. *Biol. Cybern.* **90**, 302–309 (2004), doi:10.1007/s00422-004-0471-9
38. Medvedev, A.V.: Epileptiform spikes desynchronize and diminish fast (γ) activity of the brain: An ‘anti-binding’ mechanism? *Brain Res. Bull.* **58**(1), 115–28 (2002), doi:10.1016/S0361-9230(02)00768-2
39. Milgram, S.: The small-world problem. *Psychol. Today* **1**, 60–67 (1967)
40. Netoff, T.I., Clewley, R., Arno, S., Keck, T., White, J.A.: Epilepsy in small-world networks. *J. Neurosci.* **24**, 8075–8083 (2004), doi:10.1523/jneurosci.1509-04.2004
41. Newman, M.E.J.: Power laws, pareto distributions and Zipf’s law. *Contemp. Phys.* **46**, 323–351 (2005), doi:10.1080/00107510500052444
42. Nisbach, F., Kaiser, M.: Developmental time windows for spatial growth generate multiple-cluster small-world networks. *European Phys. J. B* **58**, 185–191 (2007), doi:10.1140/epjb/e2007-00214-4
43. Otnes, R.K., Enochson, L.: *Digital Time Series Analysis*. John Wiley and Sons (1972)
44. Pastor-Satorras, R., Vespignani, A.: Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200 (2001), doi:10.1103/PhysRevLett.86.3200
45. Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N., Barabási, A.L.: Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551–1555 (2002), doi:10.1126/science.1073374
46. Salvador, R., Suckling, J., Coleman, M.R., Pickard, J.D., Menon, D., Bullmore, E.: Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb. Cortex* **15**(9), 1332–1342 (2005), doi:10.1093/cercor/bhi016
47. Sauer, T., Yorke, J., Casdagli, M.: Embedology. *J. Stat. Phys.* **65**, 579–616 (1991), doi:10.1007/BF01053745
48. Scannell, J.W., Burns, G.A., Hilgetag, C.C., O’Neil, M.A., Young, M.P.: The connectional organization of the cortico-thalamic system of the cat. *Cereb. Cortex* **9**(3), 277–299 (1999), doi:10.1093/cercor/9.3.277
49. Scannell, J., Blakemore, C., Young, M.: Analysis of connectivity in the cat cerebral cortex. *J. Neurosci.* **15**(2), 1463–1483 (1995)
50. Sporns, O., Chialvo, D.R., Kaiser, M., Hilgetag, C.C.: Organization, development and function of complex brain networks. *Trends Cogn. Sci.* **8**, 418–425 (2004), doi:10.1016/j.tics.2004.07.008
51. Taken, F.: Detecting strange attractors in turbulence. *Lecture Notes in Mathematics* **898**, 366–381 (1981), doi:10.1007/BFb0091924
52. Turcotte, D.: *Fractals and Chaos in Geology and Geophysics*. Cambridge University Press (1997)
53. Watts, D.J., Strogatz, S.H.: Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440–442 (1998), doi:10.1038/30918
54. Young, M.P.: The architecture of visual cortex and inferential processes in vision. *Spat. Vis.* **13**(2–3), 137–146 (2000), doi:10.1163/156856800741162

Chapter 6

Bifurcations and state changes in the human alpha rhythm: Theory and experiment

D.T.J. Liley, I. Bojak, M.P. Dafilis, L. van Veen, F. Frascoli,
and B.L. Foster

6.1 Introduction

The alpha rhythm is arguably the most ubiquitous rhythm seen in scalp-recorded electroencephalogram (EEG). First discovered by Hans Berger in the 1920s [27] and later confirmed by Adrian and Mathews in the early 1930s [1], it has played a central role in phenomenological descriptions of brain electrical activity in cognition and behavior ever since. While the definition of classical alpha is restricted to that 8–13-Hz oscillatory activity recorded over the occiput, which is reactive to eyes opening and closing, it is now widely acknowledged that activity in the same frequency range can be recorded from multiple cortical areas. However, despite decades of detailed empirical research involving the relationship of this rhythm to cognition, we remain essentially ignorant regarding the mechanisms underlying its genesis and its relevance to brain information processing and function [74].

Broadly speaking we are certain of only two essential facts: first, alpha activity can be recorded from scalp; and second, it bears some relationship to brain function. However a raft of recent modeling work suggests that alpha may be conceived as a marginally stable rhythm in the Lyapunov sense, and hence represents a brain state which can be sensitively perturbed by a range of factors predicted to also include

David T.J. Liley · Mathew P. Dafilis · Federico Frascoli · Brett L. Foster
Brain Sciences Institute (BSI), Swinburne University of Technology, P.O. Box 218, Victoria 3122,
Australia. e-mail: DLiley@groupwise.swin.edu.au

Lennaert van Veen
Department of Mathematics and Statistics, Faculty of Arts and Sciences, Concordia University,
1455 de Maisonneuve Blvd. W., H3G 1M8 Montreal, Quebec, Canada.
e-mail: lvanveen@mathstat.concordia.ca

Ingo Bojak
Department of Cognitive Neuroscience (126), Donders Institute for Neuroscience, Radboud University
Nijmegen Medical Centre, Postbus 9101, 6500 HB Nijmegen, The Netherlands.
e-mail: i.bojak@science.ru.nl

afferent sensory stimuli. In this view, which we will elaborate on in some detail, the alpha rhythm is best conceived as a readiness rhythm. It is not a resting or idling rhythm, as originally suggested by Adrian and Mathews [1], but instead represents a physiologically meaningful state from which transitions can be made from or to. This perspective echoes that of EEG pioneer Hans Berger [27]:

I also continue to believe that the *alpha waves* are a concomitant phenomenon of the continuous automatic physiological activity of the cortex.

This chapter is divided into three main sections: The first section gives a succinct overview of the alpha rhythm in terms of phenomenology, cerebral extent and mechanisms postulated for its genesis. It concludes by arguing that its complex features and patterns of activity, its unresolved status in cognition, and the considerable uncertainty still surrounding its genesis, all necessitate developing a more mathematical approach to its study. The second section provides an overview of our mean-field approach to modeling alpha activity in the EEG. Here we outline the constitutive equations and discuss a number of important features of their numerical solutions. In particular we illustrate how model dynamics can switch between different, but electroencephalographically meaningful, states. The third and final section outlines some preliminary evidence that such switching dynamics can be identified in scalp recordings using a range of nonlinear time-series analysis methods.

6.2 An overview of alpha activity

Between 1926 and 1929 Hans Berger laid the empirical foundations for the development of electroencephalography in humans. In the first of a number of identically titled reports [27], Berger described the alpha rhythm, its occipital dominance, and its attenuation with mental effort or opened eyes. This, and the subsequent reports, evinced virtually no interest from the neurophysiological community until Edgar Douglas Adrian (later Lord Adrian) and his colleague Bryan Mathews reproduced these results in a public demonstration that in addition revealed how easy the alpha rhythm was to record.

Following its demonstration by Adrian and Mathews [27], interest in the alpha rhythm and electroencephalography in general accelerated, to the point that considerable funding was devoted to its investigation. However, by the 1950s much of the early promise—that EEG research would elucidate basic principles of higher brain function—had dissipated. Instead, a much more pragmatic assessment of its utility as a clinical tool for the diagnosis of epilepsy prevailed. By the 1970s, rhythmicity in the EEG had been effectively labeled an epiphenomenon, assumed to only coarsely relate to brain function. However, the temporal limitations of functional magnetic resonance imaging and positron emission tomography have in the last decades renewed interest in its genesis and functional role.

6.2.1 Basic phenomenology of alpha activity

Classically, the term “alpha rhythm” is restricted to EEG activity that fulfills a number of specific criteria proposed by the International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFSECN) [32]. The most important of these are: the EEG time-series reveal a clear 8–13-Hz oscillation; this oscillation is principally located over posterior regions of the head, with higher voltages over occipital areas; it is best observed in patients in a state of wakeful restfulness with closed eyes; and it is blocked or attenuated by attentional activity that is principally of a visual or mental nature. However, alpha-band activity is ubiquitously recorded from the scalp with topographically variable patterns of reactivity. A slew of studies have revealed that the complex distribution of oscillations at alpha frequency have different sources and patterns of reactivity, suggesting that they subserve a range of different functional roles. Indeed W. Grey Walter, the pioneering British electroencephalographer, conjectured early on that “there are many alpha rhythms”, see [70]. Because the original IFSECN definition of alpha rhythm does not extend to these oscillations, they are typically referred to as *alpha activity* [15].

To date, two types of nonclassical alpha have been unequivocally identified. The first is the *Rolandic (central) mu rhythm*, first described in detail by Gastaut [25]. It is reported as being restricted to the pre- and post-central cortical regions, based on its pattern of blocking subsequent to contralateral limb movement and/or sensory activity. Like alpha activity in general, the mu rhythm does not appear to be a unitary phenomenon. For example, Pfurtscheller *et al.* [61] have observed that the mu rhythm is comprised of a great variety of separate alpha activities. The other well-known nonclassical alpha activity is the *third rhythm* (also *independent temporal alphoid rhythm* or *tau rhythm*). It is hard to detect in scalp EEG unless there is a bone defect [28], but is easily seen in magnetoencephalogram (MEG) recordings [77]. While no consensus exists regarding its reactivity or function, it appears related to the auditory cortex, as auditory stimuli are most consistently reported to block it [50, 70]. There have also been other demonstrations of topographically distinct alpha activity, whose status is much less certain and controversial. These include the alphoid *kappa rhythm* arising from the anterior temporal fossae, which has been reported to be non-specifically associated with mentation [36], and a 7–9-Hz MEG rhythm arising from second somatosensory cortex in response to median nerve stimulation [49].

Because historically the most common method of assessing the existence of alpha activity has been counting alpha waves on a chart, incorrect impressions regarding the distribution and neuroanatomical substrates of the various alpha rhythms are likely [55]. Thus the current nomenclature has to be viewed as somewhat provisional. Nevertheless, the global ubiquity of alpha activity and its clear associations with cognition suggest that understanding its physiological genesis will contribute greatly to understanding the functional significance of the EEG. This possibility was recognised by the Dutch EEG pioneer Willem Storm van Leeuwen who is cited in [4] as commenting:

If one understands the alpha rhythm, he will probably understand the other EEG phenomena.

6.2.2 Genesis of alpha activity

To date, two broad approaches have emerged for explaining the origin of the alpha rhythm and alpha activity. The first approach conceives of alpha as arising from cortical neurons being paced or driven at alpha frequencies: either through the intrinsic oscillatory properties of other cortical neurons [44, 71], or through the oscillatory activity of a feed-forward subcortical structure such as the thalamus [30, 31]. In contrast, the second approach assumes that alpha emerges through the reverberant activity generated by reciprocal interactions of synaptically connected neuronal populations in cortex, and/or through such reciprocal interactions between cortex and thalamus.

While Berger was the first to implicate the role of the thalamus in the generation of the alpha rhythm [27], it was the work of Andersen and Andersson [2] that popularised the notion that intrinsic thalamic oscillations, communicated to cortical neurons, are the source of the scalp-recorded alpha rhythm. Their essential assumption was that barbiturate-induced spindle oscillations recorded in the thalamus of the cat were the equivalent of the alpha oscillations recorded in humans. However, the notion that spindle oscillations are the source of alpha activity has not survived subsequent experimental scrutiny [74]. Spindle oscillations only occur during anaesthesia and the retreat into sleep, whereas alpha oscillations occur most prominently during a state of wakeful restfulness. Further, while the frequency of spindle oscillations and alpha activity overlap, spindles occur as groups of rhythmic waves lasting 1–2 s recurring at a rate of 0.1–0.2 Hz, whereas alpha activity appears as long trains of waves of randomly varying amplitude. A range of other thalamic local field oscillations with frequencies of approximately 10 Hz have been recorded in cats and dogs [13, 14, 30, 31], and have been considered as putative cellular substrates for human alpha activity. Nevertheless, there remains considerable controversy regarding the extent and mode of thalamic control of human alpha activity [70].

Indeed, there are good reasons to be suspicious of the idea that the thalamus is the principal source of scalp-recorded alpha oscillations. First, thalamocortical synapses are surprisingly sparse in cortex. Thalamocortical neurons project predominantly to layer IV of cerebral cortex, where they are believed to synapse mainly on the dendrites of excitatory spiny stellate cells. A range of studies [6, 12, 57, 58] have revealed that only between 5–25% of all synapses terminating on spiny stellate cells are of thalamic origin. Averaged over the whole of cortex, less than 2–3% of all synapses can be attributed to thalamocortical projections [10]. Second, recent experimental measurements reveal that the amplitude of the unitary thalamocortical excitatory postsynaptic potential is relatively small, of the order of 0.5 mV, on its own insufficient to cause a postsynaptic neuron to fire [12]. This raises the question whether weak thalamocortical inputs can establish a regular cortical rhythm even

in the spiny stellate cells, which would then require transmission to the pyramidal cells, whose apical dendrites align to form the dipole layer dominating the macroscopic EEG signal. Third, coherent activity is typically stronger between cortical areas than between cortical and thalamic areas [47, 48], suggesting cortical dominance [74]. Fourth, isolated cerebral cortex is capable of generating bulk oscillatory activity at alpha, beta and gamma frequencies [19, 37, 76]. Finally, pharmacological modulation of alpha oscillatory activity yields different results in thalamus and cortex. In particular, low doses of benzodiazepines diminish alpha-band activity but promotes beta-band activity in EEG recorded from humans, but in cat thalamus instead appear to promote lower frequency local-field potential activity by enhancing total theta power [30, 31].

For these, and a variety of other reasons [55], it has been contended that alpha activity in the EEG instead reflects the dynamics of activity in distributed reciprocally-connected populations of cortical and thalamic neurons. Two principal lines of evidence have arisen in support of this view. First, empirical evidence from multichannel MEG [16, 83] and high density EEG [55] has revealed that scalp-recorded alpha activity arises from a large number or continuum of equivalent current dipoles in cortex. Secondly, a raft of physiologically plausible computational [38] and theoretical models [40, 54, 66, 80], developed to varying levels of detail, reveal that electroencephalographically realistic oscillatory activity can arise from the synaptic interactions between distributed populations of excitatory and inhibitory neurons.

6.2.3 *Modeling alpha activity*

The staggering diversity of often contradictory empirical phenomena associated with alpha activity speaks against the notion of finding a *simple* unifying biological cause. This complexity necessitates the use of mathematical models and computer simulations in order to understand the underlying processes. Such a quantitative approach may help address three essential, probably interrelated, questions regarding the alpha rhythm and alpha activity. First, can a dynamical perspective shed light on the functional roles of alpha and its attenuation (or blocking)? While over the years a variety of theories and hypotheses have been advanced, all are independent of any physiological mechanism accounting for its genesis. The most widespread belief has been that the alpha rhythm has a clocking or co-ordinating role in the regulation of cortical neuronal population dynamics, see for example Chapter 11 of [70]. This simple hypothesis is probably the reason that the idea of a subcortical alpha pacemaker has survived despite a great deal of contradictory empirical evidence. The received view on alpha blocking and event-related desynchronisation (ERD), is that they represent the electrophysiological correlates of an activated, and hence more excitable, cortex [59]. However, this view must be regarded as, at best, speculative due to the numerous reports of *increased* alpha activity [70] in tasks requiring levels of attention and mental resource above a baseline that already exhibits strong alpha activity.

Second, what is the relationship between alpha and the other forms of scalp-recordable electrical activity? Activity in the beta band (13–30 Hz) is consistently linked to alpha-band activity. For instance, blocking of occipital alpha is almost always associated with corollary reductions in the amplitude of beta activity [60]. Further, peak occipital beta activity is, on the basis of large cross-sectional studies involving healthy subjects, almost exactly twice the frequency of peak occipital alpha, in addition to exhibiting significant phase coherence [52]. Significant phase correlation between alpha and gamma (> 30 Hz) activity has also been reported in EEG recorded from cats and monkeys [67]. Less is known about the connection to the low-frequency delta and theta rhythms.

Finally, what is the link between activity at the single neuronal level and the corresponding large-scale population dynamics? Can knowledge of the latter enable us to make inferences regarding the former, and can macroscopic predictions be deduced from known microscopic or cellular level perturbations? This becomes particularly pertinent for attempts to understand the mesoscopic link between cell (membrane) pharmacology and physiology, and co-existing large-scale alpha activity [20].

6.3 Mean-field models of brain activity

Broadly speaking, models and theories of the electroencephalogram can be divided into two complementary kinds. The first kind uses spatially discrete network models of neurons with a range of voltage- and ligand-dependent ionic conductances. While these models can be extremely valuable, and are capable of giving rise to alpha-like activity [38], they are limited since the EEG is a bulk property of populations of cortical neurons [45]. Further, while a successful application of this approach may suggest physiological and anatomical prerequisites for electrorhythmogenesis, it cannot provide explicit mechanistic insight due to its own essential complexity. In particular, a failure to produce reasonable EEG/electrocorticogram (ECoG) does not *per se* suggest which additional empirical detail must be incorporated. A more preferable approach exists in the continuum or mean-field method [33, 40, 54, 66, 80]. Here it is the bulk or population activity of a region of cortex that is modeled, more optimally matching the scale and uncertainties of the underlying physiology. Typically the neural activity over roughly the extent of a cortical macrocolumn is averaged.

However three general points need to be noted regarding the continuum mean-field approach and its application to modeling the EEG. First, in general, all approaches dynamically model the mean states of cortical neuronal populations, but only in an effective sense. Implicitly modeled are the intrinsic effects of non-neuronal parts of cortex upon neuronal behavior, e.g., glia activity or the extracellular diffusion of neurotransmitters. In order to treat the resulting equations as closed, non-neuronal contributions must either project statically into neuronal ones (e.g., by changing the value of some neuronal parameter) or be negligible in the chosen

observables (e.g., because their time-scale is slower than the neuronal dynamics of interest). Where this cannot be assumed, one must “open” the model equations by modifying the neuronal parameters dynamically.

Second, intrinsic parts or features of the brain that are not modeled (e.g., the thalamus or the laminarity of cortex) or extrinsic influences (e.g., drugs or sensory driving) likewise must be mapped onto the neuronal parameters. One may well question whether any modeling success achieved by freely changing parameters merely indicates that a complicated enough function can fit anything. There is no general answer to this criticism, but the following Ockhamian guidelines prove useful: the changes should be limited to few parameters, there should be some reason other than numerical expediency for choosing which parameters to modify, the introduced variations should either be well understood or of small size relative to the standard values, and the observed effect of the chosen parameter changes should show some stability against modifications of other parameters. If systematic tuning of the neuronal parameters cannot accommodate intrinsic or extrinsic contributions, then the neuronal model itself needs to be changed.

Third, the neuronal mean-fields modeled generally match the limited spatial resolution of functional neuroimaging, since they average over a region C surrounding a point \mathbf{x}_{cort} on cortex¹: $\underline{f} \equiv f(\mathbf{x}_{\text{cort}}, t) = 1/C \times \int_C d\mathbf{x}' f(\mathbf{x}', t)$. In the foreseeable future images of brain activity will not have spatial resolutions better than 1–2 mm², about the size of a cortical macrocolumn containing $T = 10^6$ neurons. Temporal coherence dominates quickly for signals from that many neurons. A signal from N coherent neurons is enhanced linearly $\sim N = p \times T$ over that of a single neuron, whereas for M incoherent neurons enhancement is stochastic $\sim \sqrt{M} = \sqrt{(1-p) \times T}$. $p = 1\%$ coherent neurons thus produce a *10 times stronger signal* than the 99% incoherent neurons. If p is too low, then the coherent signal will be masked by incoherent noise. In the analysis of experimental data, such time-series are typically discarded.

A mean-field prediction hence need not match all neuronal activity. It is sufficient if it effectively describes the *coherent* neurons actually causing the observed signal. Neurons in strong temporal coherence are likely of similar kind and in a similar state, thus approximating them by equations for a single “effective” neuron makes sense. Other neurons or cortical matter influence the coherent dynamics only *incoherently*, making it more likely that disturbances on average only result in static parameter changes. A crucial modeling choice is hence the number of coherent groups within C , since coherent groups will not “average out” in like manner. Every coherent neural group is modeled by equations describing its separate characteristic dynamics, which are then coupled to the equations of other such groups according to the assumed connectivity. For example, the Liley model in Fig. 6.1 shows two different C as two columns drawn side by side. We hence see that, per C , it requires equations for one excitatory group and one inhibitory group, respectively, which will then be coupled in six ways (four of which are local).

¹ Underlined symbols denote functions spatially averaged in the following manner.

The construction of a mean-field model requires the specification of three essential structural determinants: (i) the number of coherent neuronal populations modeled; (ii) the degree of physiological complexity modeled for each population; and (iii) the connectivity between these populations. While the majority of mean-field theories of EEG model the dynamics of at least two cortical neuronal populations (excitatory and inhibitory), details of the topology of connectivity can vary substantially. Figure 6.1 illustrates the connectivity of a number of competing modeling approaches.

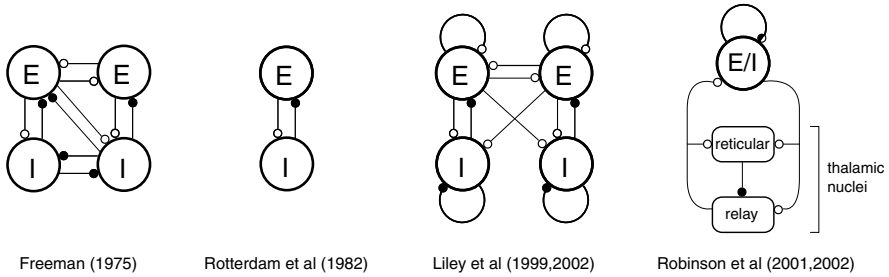


Fig. 6.1 Schematic outline of the connection topologies of a number of mean-field approaches. “E” stands for excitatory, “I” for inhibitory neuronal populations. Open circles represent excitatory connections, filled circles inhibitory ones.

6.3.1 Outline of the extended Liley model

The theory of Liley *et al* [18, 40, 43] is a relatively comprehensive model of the alpha rhythm, in that it is capable of reproducing the main spectral features of spontaneous EEG in addition to being able to account for a number of qualitative and quantitative EEG effects induced by a range of pharmacological agents, such as benzodiazepines and a range of general anesthetic agents [7, 39, 42, 75].

Like many other models, the Liley model considers two (coherent) neuronal populations within C , an excitatory one and an inhibitory one. These two populations are always indicated below by setting the subscript $k = e$ and $k = i$, respectively. In the absence of postsynaptic potential (PSP) inputs \underline{I} , the mean soma membrane potentials \underline{h} are assumed to decay exponentially to their resting value h^r with a time constant τ :

$$\tau_k \frac{\partial}{\partial t} \underline{h}_k = h_k^r - \underline{h}_k + \frac{h_{ek}^{\text{eq}} - h_k}{|h_{ek}^{\text{eq}} - h_k^r|} \times \underline{I}_{ek} + \frac{h_{ik}^{\text{eq}} - h_k}{|h_{ik}^{\text{eq}} - h_k^r|} \times \underline{I}_{ik}. \quad (6.1)$$

Double subscripts indicate first source and then target, thus for example \underline{I}_{ei} indicates PSP inputs from an excitatory to an inhibitory population. Note that PSP inputs, which correspond to transmitter activated postsynaptic channel conductance, are weighted by the respective ionic driving forces $h_{jk}^{\text{eq}} - \underline{h}_k$, where $h_{ek,ik}^{\text{eq}}$ are the

respective *reversal potentials*. All these weights are normed to one at the relevant soma membrane resting potentials.

Next consider four types ($k = e, i$) of PSP inputs:

$$\left(\frac{\partial}{\partial t} + \gamma_{ek}\right) \left(\frac{\partial}{\partial t} + \tilde{\gamma}_{ek}\right) \underline{I}_{ek} = \Gamma_{ek} \underbrace{\gamma_{ek} \exp(\tilde{\gamma}_{ek} \delta_{ek})}_{=\tilde{\gamma}_{ek} \exp(\gamma_{ek} \delta_{ek})} \times \left[N_{ek}^\beta \underline{S}_e + \underline{p}_{ek} + \underline{\Phi}_{ek} \right], \quad (6.2)$$

$$\left(\frac{\partial}{\partial t} + \gamma_{ik}\right) \left(\frac{\partial}{\partial t} + \tilde{\gamma}_{ik}\right) \underline{I}_{ik} = \Gamma_{ik} \underbrace{\gamma_{ik} \exp(\tilde{\gamma}_{ik} \delta_{ik})}_{=\tilde{\gamma}_{ik} \exp(\gamma_{ik} \delta_{ik})} \times \left[N_{ik}^\beta \underline{S}_i + \underline{p}_{ik} \right]. \quad (6.3)$$

The terms in the square brackets correspond to different classes of sources for incoming action potentials: local \underline{S} , extra-cortical \underline{p} , and cortico-cortical $\underline{\Phi}$. Only excitatory neurons project over long distances, thus there is no $\underline{\Phi}_{ik}$ in Eq. (6.3). However, long-range inhibition can still occur, namely by an excitation of an inhibitory populations via $\underline{\Phi}_{ei}$.

For a single incoming Dirac impulse $\delta(t)$, the above equations respond with

$$R(t) = \Gamma \gamma \exp(\tilde{\gamma} \delta) \times \frac{\exp(-\gamma t) - \exp(-\tilde{\gamma} t)}{\tilde{\gamma} - \gamma} \Theta(t), \quad (6.4)$$

where Θ is the Heaviside function. Here, δ is the rise-time to the maximal PSP response:

$$\delta = \frac{\ln \tilde{\gamma} - \ln \gamma}{\tilde{\gamma} - \gamma} \implies R(t = \delta) = \Gamma. \quad (6.5)$$

$R(t)$ describes PSPs from the ‘‘fast’’ neurotransmitters AMPA/kainate and GABA_A, respectively. Sometimes instead the simpler ‘‘alpha form’’² is used:

$$R_0(t) = \Gamma \gamma \exp(1) \times t \exp(-\gamma t) \Theta(t) \implies R_0(t = \delta_0 = 1/\gamma) = \Gamma. \quad (6.6)$$

Note that as $\tilde{\gamma} \rightarrow \gamma$: $R \rightarrow R_0$. Equation (6.4) must be *invariant* against exchanging $\tilde{\gamma} \leftrightarrow \gamma$, see Eqs (6.2) and (6.3), since the change induced by $\tilde{\gamma} \neq \gamma$ cannot depend on naming the decay constant values. In the ‘‘alpha form’’, the time at which the response decays again to Γ/e is coupled to the rise-time: $\zeta_0 = -W_{-1}\left(-\frac{1}{e^2}\right) \times \delta_0 \simeq 3.1462/\gamma$, with the Lambert W function. Anaesthetic agents can change the decay time of PSPs independently and hence require the biexponential form [7]:

$$\tilde{\gamma} \simeq \gamma: \zeta = -W_{-1}\left(-\frac{1}{e^2}\right) \times \delta + \mathcal{O}\left(|\tilde{\gamma} - \gamma|^2\right), \quad \tilde{\gamma} \gg \gamma: \zeta \simeq \frac{\tilde{\gamma}/\gamma}{\tilde{\gamma} - \gamma}. \quad (6.7)$$

In [7] further results were derived for the specific parametrisation $\tilde{\gamma} = \exp(\varepsilon)\gamma$.

² In this context, ‘‘alpha’’ refers to a particular single-parameter function, the so-called *alpha function*, often used in dendritic cable theory to model the time-course of a single postsynaptic potential.

If time delays for local connections are negligible, then the number of incoming action potentials will be the number of local connections N^β times the *current* local firing rate \underline{S} , see Eqs (6.2) and (6.3). Assume that threshold potentials in the neural mass are normally distributed with mean μ and standard deviation σ . Then the fraction of neurons reaching their firing threshold h^{th} is

$$\int_{-\infty}^{\underline{h}} dh^{\text{th}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(h^{\text{th}} - \mu)^2}{2\sigma^2}\right) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{\underline{h} - \mu}{\sqrt{2}\sigma}\right)\right]. \quad (6.8)$$

We approximate $(1 + \operatorname{erf}x)/2 \simeq [1 + \exp(-2x)]^{-1}$ and associate the theoretical limit of infinite \underline{h} not with excitation block but with the maximal mean firing rate S^{max} :

$$\underline{S}_k = S_k^{\text{max}} \left[1 + \exp\left(-\sqrt{2}\frac{h_k - \mu_k}{\sigma_k}\right)\right]^{-1}. \quad (6.9)$$

By construction, this is a good approximation for regular \underline{h} but will fail for unusually high mean potentials. Note that Eq. (6.9) reduces to $\underline{S}_k = S_k^{\text{max}} \Theta(\underline{h} - \mu)$ for $\sigma \rightarrow 0$.

Next we consider extra-cortical sources \underline{p} . Unless some of these inputs are strongly *coherent* (e.g., for sensory input), their average over a region will be noise-like even if the inputs themselves are not. Our ansatz is hence

$$\underline{p}_{ek} = \mathcal{L} \left[\operatorname{rand}_n(\bar{p}_{ek}, \Delta p_{ek}) \right] + \underline{p}_{ek}^{\text{coh}}, \quad (6.10)$$

$$\underline{p}_{ik} = \mathcal{L} \left[\operatorname{rand}_n(\bar{p}_{ik}, \Delta p_{ik}) \right] + \underline{p}_{ik}^{\text{coh}}, \quad (6.11)$$

with spatiotemporal “background noise” potentially overlaid by coherent signals. The noise is normally distributed with mean \bar{p} and standard deviation Δp , and shaped by some filter function \mathcal{L} . Since neurons cannot produce arbitrarily high firing frequencies, \mathcal{L} should include a lowpass filter. In practice, we often set $\underline{p}_{ik} \equiv 0$, since likely extracortical projections are predominantly excitatory. Further, for stochastic driving noise in \underline{p}_{ee} alone is sufficient. We take $\underline{p}^{\text{coh}} \equiv 0$ unless known otherwise. In particular we do not assume coherent thalamic pacemaking. However, the $\underline{p}^{\text{coh}}$ provide natural ports for future extensions, e.g., an explicit model of the thalamus could be interfaced here.

An “ideal” ansatz for cortico-cortical transmission is given by

$$\mathcal{G}_{ek}(r, t) = \frac{N_{ek}^\alpha}{2\pi} \tilde{\Lambda}_{ek}^2 \times \exp(-\tilde{\Lambda}_{ek} r) \times \delta\left(t - \frac{r}{\tilde{v}_{ek}}\right), \quad (6.12)$$

where r measures distances along cortex. With this Green’s function, impulses would propagate distortion-free and isotropically at velocity \tilde{v} . The metrics of connectivity are seen to be

$$n_{ek}^\alpha(r) = \int_0^\infty dt \mathcal{G}_{ek}(r, t) = \frac{N_{ek}^\alpha}{2\pi} \tilde{\Lambda}_{ek}^2 \times \exp(-\tilde{\Lambda}_{ek} r), \quad \int_0^\infty dr 2\pi r n_{ek}^\alpha(r) = N_{ek}^\alpha, \quad (6.13)$$

and thus the N_{ek}^α long-range connections per cortical neuron are distributed exponentially with a characteristic distance $\tilde{\Lambda}_{ek}$. One can Fourier transform Eq. (6.12)

$$\mathcal{G}_{ek}(k, \omega) = \frac{N_{ek}^\alpha \tilde{v}_{ek}^2 \tilde{\Lambda}_{ek}^2 (\mathbf{i}\omega + \tilde{v}_{ek} \tilde{\Lambda}_{ek})}{\left[(\mathbf{i}\omega + \tilde{v}_{ek} \tilde{\Lambda}_{ek})^2 + \tilde{v}_{ek}^2 k^2 \right]^{3/2}} \equiv \frac{\mathcal{N}(k, \omega)}{\mathcal{D}(k, \omega)}, \quad (6.14)$$

and write $\mathcal{N}\phi = \mathcal{D}\underline{S}$ with $\mathbf{i}\omega \rightarrow \partial/\partial t$ and $k^2 \rightarrow -\nabla^2$ to obtain an equivalent PDE. Unfortunately this \mathcal{D} is non-local (i.e., evaluating this operator with a finite difference scheme at one discretization point would require values from all points over the domain of integration). By expanding for large wavelengths $2\pi/k$: $\mathcal{D}(k, \omega) \simeq (\mathbf{i}\omega + \tilde{v}_{ek} \tilde{\Lambda}_{ek}) [(\mathbf{i}\omega + \tilde{v}_{ek} \tilde{\Lambda}_{ek})^2 + \frac{3}{2} \tilde{v}_{ek}^2 k^2]$, and with $v \equiv \sqrt{3/2} \tilde{v}$, $\Lambda \equiv \sqrt{2/3} \tilde{\Lambda}$, we obtain an inhomogeneous two-dimensional telegraph (or: transmission line) equation [40, 66]:

$$\left[\frac{1}{v_{ek}^2} \frac{\partial^2}{\partial t^2} + \frac{2\Lambda_{ek}}{v_{ek}} \frac{\partial}{\partial t} - \nabla^2 + \Lambda_{ek}^2 \right] \underline{\Phi}_{ek} = N_{ek}^\alpha \Lambda_{ek}^2 \underline{S}_e, \quad (6.15)$$

where the forcing term is simply the firing \underline{S} of the sources.

Note that Eq. (6.15) is a special case. If we substitute

$$\underline{\Phi}_{ek} = e^{-\Lambda v t} \underline{\varphi}_{ek} \implies \left[\frac{1}{v_{ek}^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right] \underline{\varphi}_{ek} = e^{\Lambda v t} N_{ek}^\alpha \Lambda_{ek}^2 \underline{S}_e, \quad (6.16)$$

then $\underline{\varphi}$ obeys an inhomogeneous wave equation. (Equation (6.16) corrects a sign error in Eq. (61) of Ref. [66], which is likely to have influenced their numerical results.) The impulse response is hence that of the 2-D wave equation multiplied by an exponential decay:

$$\underline{\mathcal{G}}_{ek}(r, t) = \frac{N_{ek}^\alpha}{2\pi} \Lambda_{ek}^2 \times \exp(-\Lambda_{ek} v_{ek} t) \times \frac{\Theta(t - r/v_{ek})}{\sqrt{t^2 - r^2/v_{ek}^2}}. \quad (6.17)$$

We can compare with (6.12) to see the effects of the approximation: Impulse propagation is now faster $v = \sqrt{3/2} \tilde{v}$ and distorted by a brief ‘‘afterglow’’ $\sim 1/\sqrt{t - r/v}$. Connectivity $n_{ek}^\alpha(r) = N_{ek}^\alpha \Lambda_{ek}^2 / (2\pi) \times K_0(\Lambda_{ek} r)$ follows now a zeroth-order modified Bessel function of the second kind. Compared to Eq. (6.13), it is now radially weaker for $1.0 \lesssim r \tilde{\Lambda} \lesssim 4.9$, and stronger otherwise.

This completes our description of the extended Liley model: Eqs (6.1), (6.2), (6.3), and (6.15) determine its spatiotemporal dynamics, (6.9) computes local firing rates, whereas (6.10) and (6.11) define the external inputs. An important feature of this model is that there are no ‘‘toy parameters’’ in the constitutive equations, i.e., every parameter has a biological meaning and its range can be constrained by physiological and anatomical data. All model parameters could depend on the position on cortex or even become additional state variables, e.g., $\mu \rightarrow \mu(\mathbf{x}_{\text{cort}}) \rightarrow \underline{\mu}$. The only exceptions are the parameters of Eq. (6.15), since the equation is derived

assuming globally constant parameters. However, this mathematical restriction can be loosened somewhat [17, 65].

6.3.2 Linearization and numerical solutions

Linearization investigates small disturbances around fixed points of the system, i.e., around state variables $\underline{\mathbf{Z}} = \underline{\mathbf{Z}}^*$ which are *spatiotemporally constant* solutions of the PDEs. For hyperbolic fixed points (i.e., all eigenvalues have nonzero real part), the Hartman–Grobman theorem states that a linear expansion in $\underline{\mathbf{z}}$ with $\underline{\mathbf{Z}} = \underline{\mathbf{Z}}^* + \underline{\mathbf{z}}$ will capture the essential local dynamics. Thus we define a state vector

$$\underline{\mathbf{Z}} \equiv (h_e, h_i, I_{ee}, I_{ei}, I_{ie}, I_{ii}, \Phi_{ee}, \Phi_{ei})^T, \quad (6.18)$$

and rewrite Eqs (6.10) and (6.11) with $\underline{p} \equiv \bar{p} + \underline{P}$, setting $\underline{P} \equiv 0$ for now. Then the fixed points are determined by

$$h_k^* = h_k^r + \frac{h_{ek}^{\text{eq}} - h_k^*}{|h_{ek}^{\text{eq}} - h_k^r|} I_{ek}^* + \frac{h_{ik}^{\text{eq}} - h_k^*}{|h_{ik}^{\text{eq}} - h_k^r|} I_{ik}^*, \quad \Phi_{ek}^* = N_{ek}^\alpha S_k^* = \frac{N_{ek}^\alpha S_k^{\text{max}}}{1 + \exp\left(-\sqrt{2} \frac{h_k^* - \mu_k}{\sigma_k}\right)},$$

$$I_{ek}^* = \Gamma_{ek} \frac{e^{\tilde{\gamma}_{ek} \delta_{ek}}}{\tilde{\gamma}_{ek}} \left[N_{ek}^\beta S_e^* + \bar{p}_{ek} + \Phi_{ek}^* \right], \quad I_{ik}^* = \Gamma_{ik} \frac{e^{\tilde{\gamma}_{ik} \delta_{ik}}}{\tilde{\gamma}_{ik}} \left[N_{ik}^\beta S_i^* + \bar{p}_{ik} \right], \quad (6.19)$$

which immediately reduces to just two equations in h_e^* and h_i^* . If multiple solutions exist, we define a “default” fixed point $\underline{\mathbf{Z}}^{*,r}$ by choosing the h_e^* closest to rest h_e^r .

We use the following ansatz for the perturbations

$$\underline{\mathbf{z}} \equiv \mathbf{a} \times \exp(\lambda t) \times \exp(\mathbf{ik} \cdot \mathbf{x}_{\text{cort}}), \quad (6.20)$$

and expand linearly in components $[\mathbf{a}]_m$. For example, the equation for Φ_{ee} becomes

$$\left(\frac{1}{v_{ee}^2} \lambda^2 + \frac{2\Lambda_{ee}}{v_{ee}} \lambda + k^2 + \Lambda_{ee}^2 \right) [\mathbf{a}]_7 = N_{ee}^\alpha \Lambda_{ee}^2 \frac{S_e^{\text{max}} \sqrt{2} v}{\sigma_e (1 + v)^2} [\mathbf{a}]_1, \quad (6.21)$$

with $v \equiv \exp[-\sqrt{2}(h_e^* - \mu_e)/\sigma_e]$. Treating all PDEs in a similar fashion, we end up with an equation set

$$\sum_j B_{ij}(\lambda, k) [\mathbf{a}]_j = 0, \quad \text{with } i, j = 1, \dots, 8. \quad (6.22)$$

In matrix notation $\mathbf{B}(\lambda, k) \mathbf{a} = 0$. Nontrivial solutions exist only for

$$\mathcal{E}(\lambda, k) \equiv \det \mathbf{B}(\lambda, k) = 0. \quad (6.23)$$

However, searching for roots $\lambda(k)$ of Eq. (6.23) is efficient only in special cases. Instead, introduce auxiliary variables $\underline{\mathcal{Z}}_{9, \dots, 14} = \partial \underline{\mathcal{Z}}_{3, \dots, 8} / \partial t$, with $\underline{\mathcal{Z}}_{9, \dots, 14}^* = 0$, to eliminate second-order time derivatives. Our example (6.21) becomes

$$[\mathbf{a}]_{15} = \lambda [\mathbf{a}]_7, \quad \left(\frac{1}{v_{ee}^2} \lambda + \frac{2\Lambda_{ee}}{v_{ee}} \right) [\mathbf{a}]_{15} + (k^2 + \Lambda_{ee}^2) [\mathbf{a}]_7 = N_{ee}^\alpha \Lambda_{ee}^2 \frac{S_e^{\max} \sqrt{2} v}{\sigma_e (1+v)^2} [\mathbf{a}]_1. \quad (6.24)$$

Treating all PDEs likewise, we can write a new but equivalent form

$$\sum_j B_{ij}(\lambda, k) [\mathbf{a}]_j = \sum_j [A_{ij}(k) - \lambda \delta_{ij}] [\mathbf{a}]_j = 0, \quad \text{with } i, j = 1, \dots, 14, \quad (6.25)$$

with the Kronecker δ_{ij} . In matrix notation $A(k)\mathbf{a} = \lambda\mathbf{a}$, hence $\lambda(k)$ solutions are eigenvalues. Powerful algorithms are readily available to solve (6.25) as

$$\sum_l A_{il} R_{lj} = \lambda_j R_{ij}, \quad \sum_l L_{il} A_{lj} = \lambda_i L_{ij}, \quad \lambda_j \sum_l L_{il} R_{lj} = \lambda_i \sum_l L_{il} R_{lj}, \quad (6.26)$$

with $i, j, l = 1, \dots, 14$, and all quantities are functions of k . The λ_j denote 14 eigenvalues with corresponding right $[\mathbf{r}_j]_i = R_{ij}$ (columns of R) and left $[\mathbf{l}_j]_i = L_{ji}$ (rows of L) eigenvectors. The third equation in (6.26) implies orthogonality for non-degenerate eigenvalues $\sum_l L_{il} R_{lj} = \delta_{ijn}$. In this case one can orthonormalize $LR = RL = \mathbf{1}$. For spatial distributions of perturbations, different k -modes will generally mix quickly with time.

For numerical simulations one can model the cortical sheet as square, connected at the edges to form a torus, and discretize it $N \times N$ with sample length d_s [7, 9]. Time then is also discretized $t = nt_s$ with $n = 0, 1, \dots$. We substitute Euler forward-time derivatives and five-point Laplacian formulae, and solve the resulting algebraic equations for the next time-step. The five-point Laplacian is particularly convenient for parallelization [7], since only one-point-deep edges of the parcellated torus need to be communicated between nodes. The Euler-forward formulae will converge slowly $\mathcal{O}(t_s^2, d_s^2)$ but *robustly*, which is important since the system dynamics can change drastically for different parameter sets. The Courant–Friedrichs–Lewy condition for a wave equation, cf Eq. (6.16), is simply $t_s < d_s/(\sqrt{2}v)$. If we consider a maximum speed of $v = 10$ m/s, and a spatial spacing of $d_s = 1$ mm for Eq. (6.15), then $t_s < 7.1 \times 10^{-5}$ s. In practice, we choose $t_s = 5 \times 10^{-5}$ s. We initialize the entire cortex to its (default) fixed point value $\mathbf{Z}(\mathbf{x}_{\text{cort}}) = \mathbf{Z}^*$ at $t = 0$. For parameter sets that have no fixed point in physiological range, we instead set $h_e(\mathbf{x}_{\text{cort}}) = h_e^r$ and $h_i(\mathbf{x}_{\text{cort}}) = h_i^r$, and other state variables to zero. Sometimes it is advantageous to have no external inputs: then any observed dynamics must be self-sustained. In this case some added spatial variation in $h_e(\mathbf{x}_{\text{cort}})$ helps to excite $k \neq 0$ modes quickly.

6.3.3 Obtaining physiologically plausible dynamics

For “physiological” parameters, a wide range of model dynamics can be encountered. However, proper parameterisations should produce electroencephalographi-

cally plausible dynamics. In general, two approaches can be employed to generate such parameter sets. The first is to fit the model to real electroencephalographic data. However, there is still considerable uncertainty regarding the reliability, applicability and significance of using experimentally obtained data for fitting or estimating sets of ordinary differential equations [79]. Alternatively one can explore the physiologically admissible multi-dimensional parameter space in order to identify parameter sets that give rise to “suitable” dynamics, e.g., those showing a dominant alpha rhythm.

With regard to the extended Liley model outlined in the previous section, one could stochastically or heuristically explore the parameter space by solving the full set of spatiotemporal equations. However, the computational costs of this approach are forbidding at this point in time. Alternatively, the parameter space of a simplified model, e.g., spatially homogeneous without the Laplacian in Eq. (6.15), can be searched. This can provide sufficient simulation speed gains to allow iterative parameter optimization. Finally, *if* the defining system can be approximated by linearization, then one can estimate the spatiotemporal dynamics merely from the resulting eigensystem. Such an analysis is exceedingly rapid compared with the direct solution of the equations. One can then simply test parameter sets randomly sampled from the physiologically admissible parameter space. Thus, for example [7] shows how one can model plausible EEG recorded from a single electrode: the power spectrum, $S(\omega)$, can be estimated for subcortical noise input $\hat{\mathbf{p}}$ by

$$S(\omega) = \frac{1}{2\pi} \int dk k \left| \Psi(k) \left\{ \mathbf{R} \cdot \text{diag} \left[\frac{1}{i\omega - \lambda_n(k)} \right] \cdot \mathbf{L} \cdot \hat{\mathbf{p}} \right\} \right|_1^2, \quad (6.27)$$

and then evaluated for physiological veracity. The left and right eigen-matrices, \mathbf{L} and \mathbf{R} , are defined in Eq. (6.26), here $\mathbf{L}\mathbf{R} = \mathbf{1}$ and $\Psi(k)$ is the electrode point-spread function. The obvious drawback is that nonlinear solutions of potential physiological relevance will be missed. However, as will be illustrated in the next section, “linear” parameter sets can be continued in one- and two-dimensions to reveal a plethora of electroencephalographically plausible nonlinear dynamical behavior.

6.3.4 Characteristics of the model dynamics

Numerical solutions to Eqs (6.1–6.15) for a range of physiologically admissible parameter values reveal a large array of deterministic and noise-driven dynamics, as well as bifurcations, at alpha-band frequencies [8, 18, 40, 43]. In particular, alpha-band activity appears in three distinct dynamical scenarios: as linear noise-driven, limit-cycle, or chaotic oscillations. Thus this model offers the possibility of characterizing the complex changes in dynamics that have been inferred to occur during cognition [79] and in a range of central nervous system diseases, such as epilepsy [46]. Further, our theory predicts that reverberant activity between inhibitory neuronal populations is causally central to the alpha rhythm, and hence the strength

and form of inhibitory \rightarrow inhibitory synaptic interactions will be the most sensitive determinants of the frequency and damping of emergent alpha-band activity. If resting eyes-closed alpha is indistinguishable from a filtered random linear process, as some time-series analyses seem to suggest [72, 73], then our model implies that electroencephalographically plausible “high quality” alpha ($Q > 5$) can be obtained only in a system with a conjugate pair of weakly damped (marginally stable) poles at alpha frequency [40].

Numerical analysis has revealed regions of parameter space where abrupt changes in alpha dynamics occur. Mathematically these abrupt changes correspond to bifurcations, whereas physically they resemble phase transition phenomena in ordinary matter. Figure 6.2 displays such a region of parameter space for the 10-dimensional local reduction of our model ($\Phi_{ek} = 0$). Variations in $\langle p_{ee} \rangle$ (excitatory input to excitatory neurons) and $\langle p_{ei} \rangle$ (excitatory to inhibitory) result in the system producing a range of dynamically differentiated alpha activities. If $\langle p_{ei} \rangle$ is much larger than $\langle p_{ee} \rangle$, a stable equilibrium is the unique state of the EEG model. Driving the model in this state with white noise typically produces sharp alpha resonances [40]. If one increases $\langle p_{ee} \rangle$, this equilibrium loses stability in a Hopf bifurcation and periodic motion sets in with a frequency of about 11 Hz. For still larger $\langle p_{ee} \rangle$ the fluctuations can become irregular and the limiting behavior of the model is governed by a chaotic attractor. The different dynamical states can be distinguished by computing the largest Lyapunov exponent (LLE), which is negative for equilibria, zero for (quasi)-periodic fluctuations, and positive for chaos. Bifurcation analysis [81] indicates that the boundary of the chaotic parameter set is formed by infinitely many saddle–node and period-doubling bifurcations, as shown in Fig. 6.2(a). All these bifurcations converge to a narrow wedge for negative, and hence unphysiological, values of $\langle p_{ee} \rangle$ and $\langle p_{ei} \rangle$, literally pointing to the crucial part of the diagram where a *Shilnikov saddle–node homoclinic bifurcation* takes place.

Figure 6.2(b) shows a sketch of the bifurcation diagram at the tip of the wedge: the blue line with the cusp point **c** separates regions with one and three equilibria, and the line of Hopf bifurcations terminates on this line at the Bogdanov–Takens point **bt**. The point **gh** is a generalised Hopf point, where the Hopf bifurcation changes from sub- to super-critical. The green line which emanates from **bt** represents a homoclinic bifurcation, which coincides with the blue line of saddle–node bifurcations on an open interval, where it denotes an orbit homoclinic to a saddle node. In the normal form, this interval is bounded by the points n_1 and n_2 , at which points the homoclinic orbit does not lie in the local center manifold. While the normal form is two-dimensional and only allows for a single orbit homoclinic to the saddle–node equilibrium, the high dimension of the macrocolumnar EEG model ($\Phi_{ek} = 0$) allows for several orbits homoclinic to the saddle–node. If we consider the numerical continuation of the homoclinic along the saddle–node curve, starting from n_1 as shown in Figure 6.2(b), it actually overshoots n_2 and folds back at t_1 , where the center-stable and center-unstable manifolds of the saddle node have a tangency. In fact, the curve of homoclinic orbits folds several times before it terminates at n_2 . This creates an interval, bounded by t_1 and t_2 , in which up to four homoclinic orbits coexist—signaling the existence of infinitely many periodic orbits, which is the hallmark of chaos.

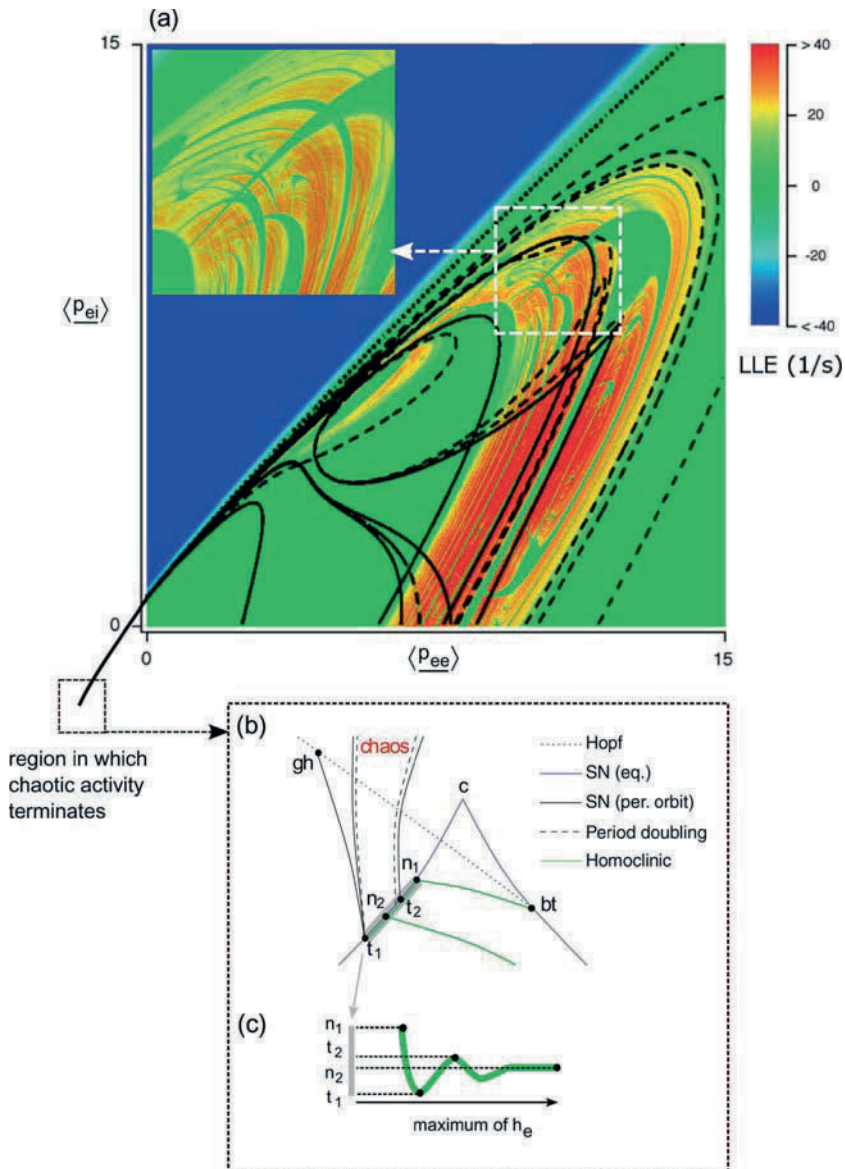


Fig. 6.2 [Color plate] (a) The largest Lyapunov exponent (LLE) of the dynamics of a simplified local model ($\Phi_{ek} = 0$) for a physiologically plausible parameter set exhibiting robust (fat-fractal) chaos [18]. Superimposed is a two parameter continuation of saddle–node and period-doubling bifurcations. The leftmost wedge of chaos terminates for negative values of the exterior forcings, $\langle P_{ee} \rangle$ and $\langle P_{ei} \rangle$. (b) Schematic bifurcation diagram at the tip of the chaotic wedge. bt = Bogdanov–Takens bifurcation, gh = generalized Hopf bifurcation, and SN = saddle node. Between t_1 and t_2 multiple homoclinic orbits coexist and Shilnikov’s saddle–node bifurcation takes place. (c) Schematic illustration of the continuation of the homoclinic orbit between points n_1 and t_1 . (Figure adapted from [81] and [18].)

It is important to understand that in contrast to the homoclinic bifurcation of a saddle focus, commonly referred to as the Shilnikov bifurcation, this route to chaos has not been reported before in the analysis of any mathematical model of a physical system. While the Shilnikov saddle node bifurcation occurs at negative, and thus unphysiological, values of $\langle p_{ee} \rangle$ and $\langle p_{ei} \rangle$, it nevertheless organizes the qualitative behavior of the EEG model in the biologically meaningful parameter space. Further, it is important to remark that this type of organization persists in a large part of the parameter space: if a third parameter is varied, the codimension-two points **c**, **bt** and **gh** collapse onto a degenerate Bogdanov–Takens point of codimension three, which represents an organizing center controlling the qualitative dynamics of an even larger part of the parameter space.

Parameter sets that have been chosen to give rise to physiologically realistic behavior in one domain, can produce a range of unexpected, but physiologically plausible, activity in another. For example, parameters were chosen to accurately model eyes-closed alpha and the surge in total EEG power during anesthetic induction [7]. Among other conditions, parameter sets were required to have a sharp alpha resonance ($Q > 5$) and moderate mean excitatory and inhibitory neuronal firing rates $< 20/s$. Surprisingly, a large fraction of these sets also produced limit cycle (non-linear) gamma band activity under mild parameter perturbations [8]. Gamma band (> 30 Hz) oscillations are thought to be the *sine qua non* of cognitive functioning. This suggests that the existence of weakly damped, noise-driven, linear alpha activity can be associated with limit cycle 40-Hz activity, and that transitions between these two dynamical states can occur. Figure 6.3 illustrates a bifurcation diagram for one such set (column 11 of Table V in [7], see also Table 1 in [8]) for the spatially homogeneous reduction $\nabla^2 \rightarrow 0$ of Eq. (6.15). The choice of bifurcation parameters is motivated by two observations: (i) differential increases in $\Gamma_{ii,ie}$ have been shown to reproduce a shift from alpha to beta band activity, similar to what is seen in the presence of low levels of GABA_A agonists such as benzodiazepines [42]; and (ii) the dynamics of linearized solutions for the case when $\nabla^2 \neq 0$ are particularly sensitive to variations of parameters affecting inhibitory→inhibitory neurotransmission [40], such as N_{ii}^β and $\langle p_{ii} \rangle$.

Specifically, Fig. 6.3 illustrates the results of a two-parameter bifurcation analysis for changes in the inhibitory PSP amplitudes via $\Gamma_{ie,ii} \rightarrow r\Gamma_{ie,ii}$ and changes in the total number of inhibitory→inhibitory connections via $N_{ii}^\beta \rightarrow kN_{ii}^\beta$. The parameter space has physiological meaning only for positive values of r and k . The saddle–node bifurcations of equilibria have the same structure as for the 10-dimensional homogeneous reduction discussed previously, in that there are two branches joined at a cusp point. Furthermore, we have two branches of Hopf bifurcations, the one at the top being associated with the birth of alpha limit cycles and the other with gamma limit cycles. This former line of Hopf points enters the wedge-shaped curve of saddle–nodes of equilibria close to the cusp point and has two successive tangencies in fold–Hopf points (fh). The fold–Hopf points are connected by a line of tori. The same curve of Hopf points ends in a Bogdanov–Takens (bt) point, from which a line of homoclinics emanate. Contrary to the previous example, this line of homoclinics does not give rise to a Shilnikov saddle–node bifurcation. Instead it gives

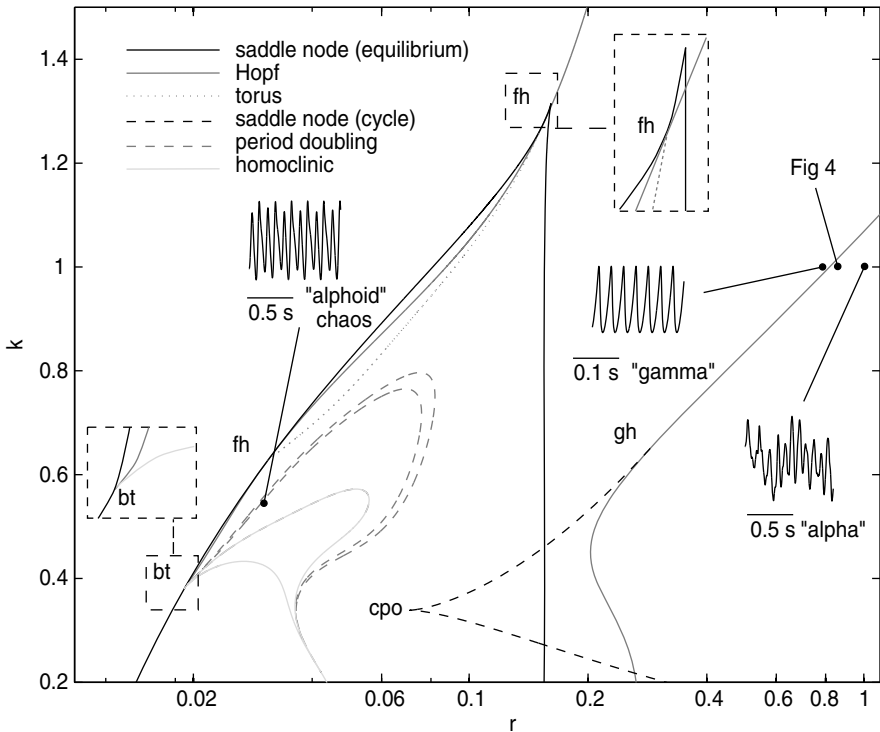


Fig. 6.3 Partial bifurcation diagram for the spatially homogeneous model, $\nabla^2 \rightarrow 0$ in Eq. (6.15), as a function of scaling parameters k and r , defined by $\Gamma_{ie,ii} \rightarrow r\Gamma_{ie,ii}$ and $N_{ii}^\beta \rightarrow kN_{ii}^\beta$, respectively. Codimension-two points have been labeled fh for fold-Hopf, gh for generalized Hopf, and bt for Bogdanov-Takens. The right-most branch of Hopf points corresponds to emergence of gamma frequency (≈ 37 Hz) limit-cycle activity via subcritical Hopf bifurcation above the point labeled gh. A homoclinic doubling cascade takes place along the line of homoclinics emanating from bt. Insets on the left show schematic blowups of the fh and bt points. Additional insets show time-series of deterministic (limit-cycle and chaos) and noise-driven dynamics for a range of indicated parameter values.

rise to a different scenario leading to complex behavior (including chaos), called the *homoclinic doubling cascade*.

In this scenario, a cascade of period-doubling bifurcations collides with a line of homoclinics. As a consequence, not only are infinitely many periodic orbits created, but so are infinitely many homoclinic connections [56]. All these periodic and homoclinic orbits coexist with a stable equilibrium. The second line of Hopf bifurcations in the gamma frequency range (> 30 Hz) does not interact with the lines of saddle nodes in the relevant portion of the parameter space. Both branches of Hopf points change from super- to subcritical at gh around $r^* = 0.27$, so that bifurcations are “hard” for $r > r^*$ in either case. These points are also the end points of folds for the periodic orbits, and the gamma frequency ones form a cusp (cpo) inside the wedge of saddle-nodes of equilibria.

Because the partial bifurcation diagram of Fig. 6.3 has necessarily been determined for the spatially homogeneous model equations, it will not accurately reflect the stability properties of particular spatial modes (nonzero wavenumbers) in the full set of model 2-D PDEs. For example, at the spot marked “Fig 4” in Fig. 6.3, for wavenumbers around $0.6325/\text{cm}$ the eigenvalues of the corresponding alpha-rhythm are already unstable, implying that these modes have undergone transition to the subcritical gamma-rhythm limit cycle. If one starts a corresponding numerical simulation with random initial h_e , but without noise driving, one finds that there is, at first, a transient organization into alpha-rhythm regions of a size corresponding to the unstable wavenumber (graph labeled “0 ms” in Fig. 6.4). The amplitude of these alpha-oscillations grows, and is then rapidly replaced by “gamma hotspots”, which are phase synchronous with each other (graphs up to “480 ms” in Fig. 6.4). It may be speculated from a physiological perspective that the normal organization of the brain consists of regions capable of producing stable weakly-

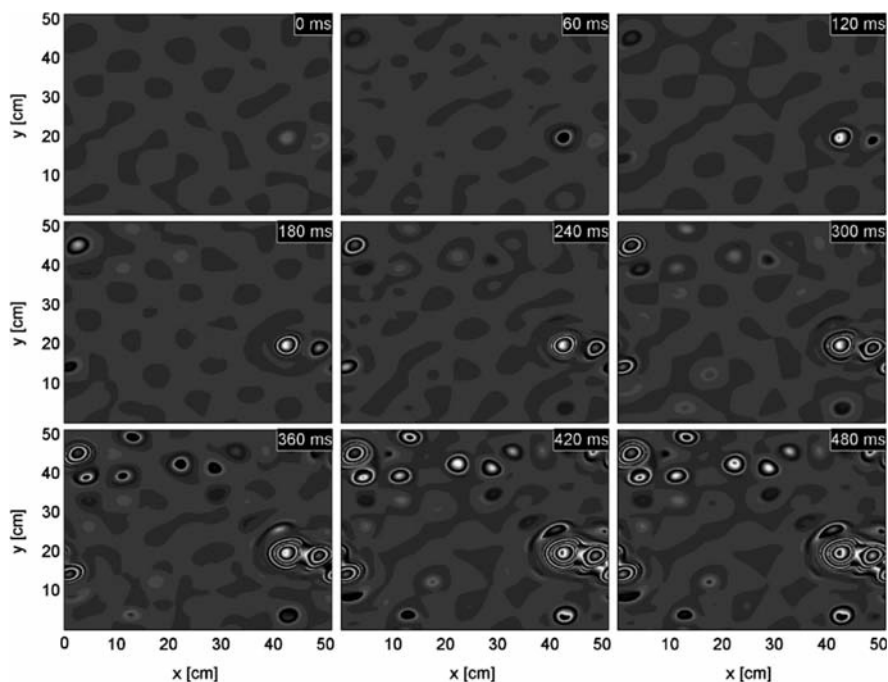


Fig. 6.4 Numerical solutions of 2-D model equations (Sect. 3.1) for a human-sized cortical torus with $k = 1$ and $r = 0.875$ (see Fig. 6.3). Here, h_e is mapped every 60 ms (grayscale: -76.9 mV black to -21.2 mV white). For $r = 0.875$, linearization becomes unstable for a range of wavenumbers around $0.6325/\text{cm}$. Starting from random h_e , one initially sees transient spatially-organized alpha oscillations ($t = 0$, starting transient removed) from which synchronized gamma activity emerges. Gamma-frequency spatial patterns, with a high degree of phase correlation (“gamma hotspots”) form with a frequency consistent with the predicted subcritical Hopf bifurcations of the spatially homogeneous equations, compare Fig. 6.3. (Figure reproduced from [8].)

damped alpha-oscillations for all wavenumbers, but, due to variations in one or more bifurcation parameters, is able to become critical at a particular wavenumber, thereby determining, in some fashion, the spatial organization of the subsequently generated coherent gamma-oscillations. However, the influences of noisy inputs (and environments), inhomogeneous neuronal populations, and anisotropic connectivity, are likely to be significant for actual transitions, and require further study.

6.4 Determination of state transitions in experimental EEG

Our theoretical analysis so far suggests that cortex may be conceived as being in a state of marginal linear stability with respect to alpha activity, which can be lost by a range of perturbations and replaced by a rapidly emerging (≈ 150 ms) spatially synchronized, nonlinear, oscillatory state. It is therefore necessary to examine real EEG for evidence of transitions between noise-driven linear and nonlinear states. While the theory of nonlinear, deterministic dynamical systems has provided a number of powerful methods to characterize the dynamical properties of time-series, they have to be applied carefully to the dynamical characterization of EEG, where any deterministic dynamics are expected to be partly obscured by the effects of noise, non-stationarity and finite sampling. For such weakly nonlinear systems, the preferred approach to characterizing the existence of any underlying deterministic dynamics has been the surrogate data method [34].

In this approach a statistic, λ , which assigns a real number to a time-series and is sensitive to deterministic structure, is computed for the original time-series, λ_0 , and compared to the distribution of values, $\{\lambda_i\}$, obtained for a number of suitably constructed “linear” surrogate data sets. Then one estimates how likely it is to draw λ_0 from the distribution of values obtained for the surrogates $\{\lambda_i\}$. For example, if we have reason to believe that λ is normally distributed, we estimate its mean $\bar{\lambda}$ and variance σ_λ^2 . Then if $|\lambda_0 - \bar{\lambda}| < 2\sigma_\lambda$, we would not be able to reject the null hypothesis H_0 that λ_0 was drawn from $\{\lambda_i\}$ at the $p = 0.05$ level for a two-tailed test. Typically though, there is no *a priori* information regarding the distribution of $\{\lambda_i\}$, and hence a rank-based test is generally used.

However, rejection of the null hypothesis in itself does not provide unequivocal statistical evidence for the existence of deterministic dynamics. In particular, non-stationarity is a well-known source of false rejections of the linear stochastic null hypothesis. To deal with this, two general strategies are employed. First, the null hypothesis is evaluated on time-series segments short enough to be assumed stationary, but long enough to allow the meaningful evaluation of the nonlinear statistic. Second, if some measure of stationarity can be shown to be equivalent in the original and surrogate data time-series, then it may be assumed that nonstationarity is an insignificant source of false positives.

6.4.1 Surrogate data generation and nonlinear statistics

The features that the “linear” surrogate data sets must have depend on the null hypothesis that is to be tested. The most common null hypothesis is that the data comes from a stationary, linear, stochastic process with Gaussian inputs. Therefore almost all surrogate data generation schemes aim to conserve linear properties of the original signal, such as the auto- and cross-spectra. The simplest method of achieving this is by phase randomisation of the Fourier components of the original time-series. However, such a simple approach results in an unacceptably high level of false positives, because the spectrum and amplitude distribution of the surrogates has not been adequately preserved. For this reason a range of improvements to the basic phase-randomised surrogate have been developed [69]. Of these, the iterated amplitude-adjusted FFT surrogate (IAFFT) seems to provide the best protection against spurious false rejections of the linear stochastic null hypothesis [34].

A large number of nonlinear test statistics are available to evaluate time-series for evidence of deterministic/nonlinear structure using the surrogate data methodology. The majority of these quantify the predictability of the time-series in some way. While there is no systematic way to choose one statistic over another, at least in the analysis of EEG the zeroth-order nonlinear prediction error (0-NLPE) seems to be favored. Indeed, Schrieber and Schmitz [68], by determining the performance of a number of commonly used nonlinear test statistics, concluded that the one-step-ahead 0-NLPE gave consistently good discrimination power even against weak nonlinearities. The idea behind the NLPE is relatively simple: delay-embed a time-series x_n to obtain the vectors $\mathbf{x}_n = (x_{n-(m-1)\tau}, x_{n-(m-2)\tau}, \dots, x_{n-\tau}, x_n)$ in \mathbb{R}^m , and use the points closer than ε to each \mathbf{x}_N , i.e., $\mathbf{x}_m \in \mathcal{U}_\varepsilon(\mathbf{x}_N)$, to predict \mathbf{x}_{N+1} as the average of the $\{\mathbf{x}_{m+1}\}$. Formally [34]

$$\hat{x}_{N+1} = \frac{1}{|\mathcal{U}_\varepsilon(\mathbf{x}_N)|} \sum_{\mathbf{x}_m \in \mathcal{U}_\varepsilon(\mathbf{x}_N)} x_{m+1}, \quad (6.28)$$

where $|\mathcal{U}_\varepsilon(\mathbf{x}_N)|$ is the number of elements in the neighborhood $\mathcal{U}_\varepsilon(\mathbf{x}_N)$. The one-step-ahead 0-NLPE is then defined as the root-mean-square prediction error over all points in the time-series, i.e., $\lambda^{\text{NLPE}} = \sqrt{\langle (\hat{x}_{N+1} - x_{N+1})^2 \rangle}$.

Other nonlinear statistics include the correlation sum, the maximum likelihood estimator of the Grassberger–Procaccia correlation dimension D_2 , and a variety of higher-order autocorrelations and autocovariances. Of the latter, two are of particular note due to their computational simplicity and their applicability to short time series. These are the third-order autocovariance, $\lambda^{\text{C3}}(\tau) = \langle x_n x_{n-\tau} x_{n-2\tau} \rangle$, and time-reversal asymmetry, $\lambda^{\text{TREV}}(\tau) = \langle (x_n - x_{n-\tau})^3 \rangle / \langle (x_n - x_{n-\tau})^2 \rangle$.

6.4.2 Nonlinear time-series analysis of real EEG

The surrogate data method has produced uncertain and equivocal results for EEG [72]. An early report, using a modified nonlinear prediction error [73], suggested

that resting EEG contained infrequent episodes of deterministic activity. However, a later report [26], using third-order autocovariance and time-reversal asymmetry, revealed that in a significant fraction (up to 19.2%) of examined EEG segments, the null hypothesis of linearity could not be rejected. Therefore, depending on the nonlinear statistic used, a quite different picture regarding the existence of dynamics of deterministic origin in the EEG may emerge. Thus attempts to identify transitions between putatively identified linear and nonlinear states using surrogate data methods will need to use a range of nonlinear discriminators.

Figure 6.5 shows a subset of the results obtained from a multivariate surrogate data based test of nonlinearity for eyes-closed resting EEG recorded from a healthy male subject. The important points to note are: (i) the fraction of epochs tentatively identified as nonlinear is small for all nonlinear statistics; (ii) temporal patterns of putatively identified nonlinear segments differ depending on the nonlinear statistic used; and (iii) the power spectra of nonlinear segments are associated with a visible sharpening of the alpha resonance for all nonlinear statistical discriminators. It is this latter feature that is of particular interest to us. It suggests, in the context of our theory, that the linear stochastic system underlying the generation of the alpha activity has become more weakly damped and is thus more prone to being “excited” into a nonlinear or deterministic state.

Because we theoretically envision a system intermittently switching between linear and deterministic (nonlinear) states there is a reduced need to identify the extent to which nonstationarity acts as a source of false positives in our surrogate data nonlinear time-series analysis. For if our system switches between linear and nonlinear states on a time-scale less than the length of the interval over which nonlinearity is characterized, deterministic dynamics and nonstationarity necessarily co-exist.

Thus this preliminary experimental evidence, involving the detection of weak nonlinearity in resting EEG using an extension of the well-known surrogate data method, suggests that nonlinear (deterministic) dynamics are more likely to be associated with weakly damped alpha activity and that either a dynamical bifurcation has occurred or is more likely to occur.

6.5 Discussion

We have outlined a biologically plausible mean-field approximation of the dynamics of cortical neural activity which is able to capture the chief properties of mammalian EEG. Central to this endeavor has been the modeling of human alpha activity, which is conceived as the central organizing rhythm of spontaneous EEG.

A great deal of modern thinking regarding alpha activity in general, and the alpha rhythm in particular, has focused on its variation during task performance and/or stimulus presentation, and therefore attempts to describe its function in the context of behavioral action or perception. These attempts to characterize alpha activity in terms of its psychological correlates, together with its inevitable appearance in scalp-recorded EEG has meant that specific research aimed at understanding this

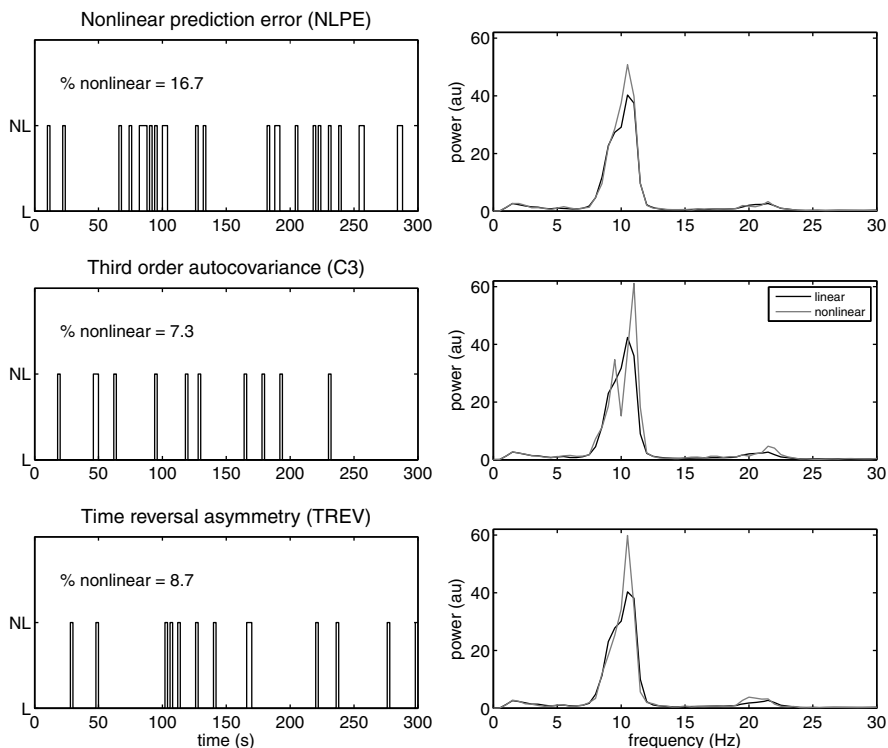


Fig. 6.5 Nonlinear surrogate data time-series analysis of parieto-occipally recorded EEG from a healthy male subject. Left-hand panels show the temporal sequence of putatively identified nonlinear 2-s EEG segments for channel P4 for three nonlinear discriminators: λ^{NLPE} , λ^{C3} and λ^{TREV} . Right-hand panels show the corresponding averaged power spectra for segments identified as nonlinear, compared with the remaining segments. Three-hundred seconds of artifact-free 64-channel (modified-expanded 10–20 system of electrode placement; linked mastoids) resting eyes-closed EEG was recorded, bandpass filtered between 1 and 40 Hz and sampled at 500 Hz. EEG was then segmented into contiguous multichannel epochs of 2-s length from which multivariate surrogates were created. H_0 (data results from a Gaussian linear stochastic process) was then tested for each channel at the $p = 0.05$ level using a nonparametric rank-order method together with a step-down procedure to control for familywise type-I error rates. Power spectra were calculated using Hamming-windowed segments of length 1000.

oscillatory phenomenon is more the exception than the rule. In a prescient review regarding electrical activity in the brain, W. Grey Walter in 1949 [82], whilst talking about spontaneous activity, remarked:

The prototype in this category, the alpha rhythm, has been seen by every electroencephalographer but studied specifically by surprisingly few.

While we have proposed a theory for the dynamical genesis of alpha activity, and via large-scale parameter searches established plausible physiological domains that can produce alpha activity, we do not understand the basis for the parameterizations so found. Our theory suggests that the reason human alpha activity shows complex

and sensitive transient behavior is because it is readily perturbed from a dynamical state of marginal linear stability. It is therefore not inconceivable that the system producing alpha activity has, through as yet unknown mechanisms, a tendency to organize itself into a state of marginal stability. This line of thinking relates in a general form to the ideas of self-organized criticality [5], especially in the context of the near $1/f$ distribution of low-frequency power reported in EEG [62] ECoG [22] and MEG recordings [53].

Our view of electrorhythmogenesis and brain functioning emphasizes the self-organized structure of spontaneous neural dynamics as active (or ready) and chiefly determined by the bulk physiological properties of cortical tissue, which is perturbed or modulated by a variety of afferent influences arising from external sources and/or generated by other parts of the brain. Alpha activity is hypothesised to be the source of this self-organizing process, providing the background dynamical state from which transitions to emergent, and thus information creating, nonlinear states are made. In a general sense then, alpha activity provides ongoing dynamical predicates for subsequently evoked activity. Such an approach is not uncommon among neurophysiologists who have emphasized the importance of ongoing neural dynamics in the production of evoked responses, see for example [3]. Indeed, this point was highlighted early on by Donald O. Hebb [29]:

Electrophysiology of the central nervous system indicates in brief that the brain is continuously active, in all its parts, and an afferent excitation must be superimposed on an already existent excitation. It is therefore impossible that the consequence of a sensory event should often be uninfluenced by the pre-existent activity.

6.5.1 Metastability and brain dynamics

Although early attempts to dynamically describe brain function sought to prescribe explicit attractor dynamics to neural activity, more recent thinking focuses on transient nonequilibrium behavior [63]. In the context of the mesoscopic theory of alpha activity presented here, it is suggested that these transient states correspond to coherent mesoscopic gamma oscillations arising from the bifurcation of noise-driven marginally stable alpha activity.

From a Hebbian perspective, such a bifurcation may represent the regenerative activation of a cell assembly through the mutual excitation of its component neurons. However, Hebb's original notion of a cell assembly did not incorporate any clear mechanism for the initiation or termination of activity in cell assemblies. As originally formulated, Hebbian cell assemblies could only generate run-away excitation due to the purely excitatory connections among the assembly neurons. In the theory presented here, the possibility arises that the initiation and termination of cell assembly activity (assuming it corresponds to synchronized gamma band activity) might occur as a consequence of modulating local reverberant inhibitory neuronal activity through either disinhibition (variations in $\langle p_{ii} \rangle$) or transient modifications in inhibitory→inhibitory synaptic efficacy (N_{ii}^{β} , Γ_{ii}) [8]. Because local inhibition has

been shown to be a sensitive determinant of the dynamics of emergent model alpha activity [40], it may be hypothesized that it is readily influenced by the relatively sparse thalamocortical projections.

Given that neuronal population dynamics have been conceived as evolving transiently, rarely reaching stability, a number of authors have opted to describe this type of dynamical regime as metastability [11, 21, 23, 35, 64]. Common to many of these descriptions is an ongoing occurrence of transitory neural events, or state transitions, which define the flexibility of cognitive and sensori-motor function. Some dynamical examples include the chaotic itinerancy of Tsuda [78], in which neural dynamics transit in a chaotic motion through unique attractors (Milnor), or the liquid-state machine of Rabinovich *et al* [63], where a more global stable heteroclinic channel is comprised of successive local saddle states. More specific neurodynamical approaches include the work of Kelso [35], Freeman [21] and Friston [24].

In developing mathematical descriptions of metastable neural dynamics, many of the models are often sufficiently general to allow for a standard dynamical analysis and treatment. For this reason, much of the dynamical analysis of EEG has focused on the identification of explicit dynamical states. However attempts to explore the attractor dynamics of EEG have produced at best equivocal results, suggesting that such simplistic dynamical metaphors have no real neurophysiological currency. Modern surrogate data methods have revealed that normal spontaneous EEG is only weakly nonlinear [72], and thus more subtle dynamical methods and interpretations, motivated by physiologically meaningful theories of electrorhythmogenesis, need to be developed.

References

1. Adrian, E.D., Matthews, B.H.C.: The Berger rhythm, potential changes from the occipital lobe in man. *Brain* **57**, 355–385 (1934), doi:10.1093/brain/57.4.355
2. Andersen, P., Andersson, S.A.: *Physiological basis of the alpha rhythm*. Appelton-Century-Crofts, New York (1968)
3. Arieli, A., Sterkin, A., Grinvald, A., Aertsen, A.: Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses. *Science* **273**, 1868–1871 (1996), doi:10.1126/science.273.5283.1868
4. Başar, E., Schürmann, M., Başar-Eroglu, C., Karakaş, S.: Alpha oscillations in brain functioning: An integrative theory. *Int. J. Psychophysiol.* **26**, 5–29 (1997), doi:10.1016/S0167-8760(97)00753-8
5. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality: An explanation of the $1/f$ noise. *Phys. Rev. Lett.* **59**, 381–384 (1987), doi:10.1103/PhysRevLett.59.381
6. Benshalom, G., White, E.L.: Quantification of thalamocortical synapses with spiny stellate neurons in layer IV of mouse somatosensory cortex. *J. Comp. Neurol.* **253**, 303–314 (1986), doi:10.1002/cne.902530303
7. Bojak, I., Liley, D.T.J.: Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* **71**, 041902 (2005), doi:10.1103/PhysRevE.71.041902
8. Bojak, I., Liley, D.T.J.: Self-organized 40-Hz synchronization in a physiological theory of EEG. *Neurocomp.* **70**, 2085–2090 (2007), doi:10.1016/j.neucom.2006.10.087

9. Bojak, I., Liley, D.T.J., Cadusch, P.J., Cheng, K.: Electrorhythmogenesis and anaesthesia in a physiological mean field theory. *Neurocomp.* **58-60**, 1197–1202 (2004), doi:10.1016/j.neucom.2004.01.185
10. Braitenberg, V., Schüz, A.: *Cortex: Statistics and geometry of neuronal connectivity*. Springer, New York, 2nd edn. (1998)
11. Bressler, S.L., Kelso, J.A.S.: Cortical coordination dynamics and cognition. *Trends Cogn. Sci.* **5**, 26–36 (2001), doi:10.1016/S1364-6613(00)01564-3
12. Bruno, R.M., Sakmann, B.: Cortex is driven by weak but synchronously active thalamocortical synapses. *Science* **312**, 1622–1627 (2006), doi:10.1126/science.1124593
13. Chatila, M., Milleret, C., Buser, P., Rougeul, A.: A 10 Hz “alpha-like” rhythm in the visual cortex of the waking cat. *Electroencephalogr. Clin. Neurophysiol.* **83**, 217–222 (1992), doi:10.1016/0013-4694(92)90147-A
14. Chatila, M., Milleret, C., Rougeul, A., Buser, P.: Alpha rhythm in the cat thalamus. *C. R. Acad. Sci. III, Sci. Vie* **316**, 51–58 (1993)
15. Chatrian, G.E., Bergamini, L., Dondey, M., Klass, D.W., Lennox-Buchthal, M.A., Petersén, I.: A glossary of terms most commonly used by clinical electroencephalographers. In: *International Federation of Societies for Electroencephalography and Clinical Neurophysiology (ed.), Recommendations for the practice of clinical neurophysiology*, Elsevier, Amsterdam (1983)
16. Ciulla, C., Takeda, T., Endo, H.: MEG characterization of spontaneous alpha rhythm in the human brain. *Brain Topogr.* **11**, 211–222 (1999), doi:10.1023/A:1022233828999
17. Coombes, S., Venkov, N.A., Shiau, L.J., Bojak, I., Liley, D.T.J., Laing, C.R.: Modeling electrocortical activity through improved local approximations of integral neural field equations. *Phys. Rev. E* **76**, 051901 (2007), doi:10.1103/PhysRevE.76.051901
18. Dafilis, M.P., Liley, D.T.J., Cadusch, P.J.: Robust chaos in a model of the electroencephalogram: Implications for brain dynamics. *Chaos* **11**, 474–478 (2001), doi:10.1063/1.1394193
19. Fisahn, A., Pike, F.G., Buhl, E.H., Paulsen, O.: Cholinergic induction of network oscillations at 40 Hz in the hippocampus in vitro. *Nature* **394**, 186–189 (1998), doi:10.1038/28179
20. Foster, B.L., Bojak, I., Liley, D.T.J.: Population based models of cortical drug response – insights from anaesthesia. *Cognitive Neurodyn.* **2** (2008), doi:10.1007/s11571-008-9063-z
21. Freeman, W.J., Holmes, M.D.: Metastability, instability, and state transition in neocortex. *Neural Netw.* **18**, 497–504 (2005), doi:10.1016/j.neunet.2005.06.014
22. Freeman, W.J., Rogers, L.J., Holmes, M.D., Silbergeld, D.L.: Spatial spectral analysis of human electrocorticograms including the alpha and gamma bands. *J. Neurosci. Methods* **95**, 111–121 (2000), doi:10.1016/S0165-0270(99)00160-0
23. Friston, K.J.: Transients, metastability, and neuronal dynamics. *NeuroImage* **5**, 164–171 (1997)
24. Friston, K.J.: The labile brain. I. Neuronal transients and nonlinear coupling. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **355**, 215–236 (2000), doi:10.1006/nimg.1997.0259
25. Gastaut, H.: Étude électrocorticographique de la réactivité des rythmes rolandiques. *Rev. Neurol. (Paris)* **87**, 176–182 (1952)
26. Gautama, T., Mandic, D.P., Van Hulle, M.M.: Indications of nonlinear structures in brain electrical activity. *Phys. Rev. E* **67**, 046204 (2003), doi:10.1103/PhysRevE.67.046204
27. Gloor, P.: Hans Berger on the electroencephalogram of man. *Electroencephalogr. Clin. Neurophysiol.* **S28**, 350 (1969)
28. Grillon, C., Buchsbaum, M.S.: Computed EEG topography of response to visual and auditory stimuli. *Electroencephalogr. Clin. Neurophysiol.* **63**, 42–53 (1986), doi:10.1016/0013-4694(86)90061-1
29. Hebb, D.O.: *The organization of behavior*. Wiley, New York (1949)
30. Hughes, S.W., Crunelli, V.: Thalamic mechanisms of EEG alpha rhythms and their pathological implications. *Neuroscientist* **11**, 357–372 (2005), doi:10.1177/1073858405277450
31. Hughes, S.W., Crunelli, V.: Just a phase they’re going through: the complex interaction of intrinsic high-threshold bursting and gap junctions in the generation of thalamic alpha and theta rhythms. *Int. J. Psychophysiol.* **64**, 3–17 (2007), doi:10.1016/j.ijpsycho.2006.08.004

32. International Federation of Societies for Electroencephalography and Clinical Neurophysiology: A glossary of terms commonly used by clinical electroencephalographers. *Electroencephalogr. Clin. Neurophysiol.* **37**, 538–548 (1974), doi:10.1016/0013-4694(74)90099-6
33. Jirsa, V.K., Haken, H.: Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996), doi:10.1103/PhysRevLett.77.960
34. Kantz, H., Schreiber, T.: *Nonlinear time series analysis*. Cambridge University Press, New York, 2nd edn. (2003)
35. Kelso, J.A.S.: *Dynamic patterns: The self-organization of brain and behavior*. The MIT Press (1995)
36. Kennedy, J.L., Gottsdanker, R.M., Armington, J.C., Gray, F.E.: A new electroencephalogram associated with thinking. *Science* **108**, 527–529 (1948), doi:10.1126/science.108.2811.527
37. Kristiansen, K., Courtois, G.: Rhythmic electrical activity from isolated cerebral cortex. *Electroencephalogr. Clin. Neurophysiol.* **1**, 265–272 (1949)
38. Liley, D.T.J., Alexander, D.M., Wright, J.J., Aldous, M.D.: Alpha rhythm emerges from large-scale networks of realistically coupled multicompartmental model cortical neurons. *Network: Comput. Neural Syst.* **10**, 79–92 (1999), doi:10.1088/0954-898X/10/1/005
39. Liley, D.T.J., Bojak, I.: Understanding the transition to seizure by modeling the epileptiform activity of general anesthetic agents. *J. Clin. Neurophysiol.* **22**, 300–313 (2005)
40. Liley, D.T.J., Cadusch, P.J., Dafilis, M.P.: A spatially continuous mean field theory of electrocortical activity. *Network: Comput. Neural Syst.* **13**, 67–113 (2002), doi:10.1088/0954-898X/13/1/303, see also [41]
41. Liley, D.T.J., Cadusch, P.J., Dafilis, M.P.: Corrigendum: A spatially continuous mean field theory of electrocortical activity. *Network: Comput. Neural Syst.* **14**, 369 (2003), doi:10.1088/0954-898X/14/2/601
42. Liley, D.T.J., Cadusch, P.J., Gray, M., Nathan, P.J.: Drug-induced modification of the system properties associated with spontaneous human electroencephalographic activity. *Phys. Rev. E* **68**, 051906 (2003), doi:10.1103/PhysRevE.68.051906
43. Liley, D.T.J., Cadusch, P.J., Wright, J.J.: A continuum theory of electro-cortical activity. *Neurocomp.* **26-27**, 795–800 (1999), doi:10.1016/S0925-2312(98)00149-0
44. Llinás, R.R.: The intrinsic electrophysiological properties of mammalian neurons: Insights into central nervous system function. *Science* **242**, 1654–1664 (1988), doi:10.1126/science.3059497
45. Lopes da Silva, F.H.: Dynamics of EEGs as signals of neuronal populations: Models and theoretical considerations. In: [51], pp. 85–106 (2005)
46. Lopes da Silva, F.H., Blanes, W., Kalitzin, S.N., Parra, J., Suffczyński, P., Velis, D.N.: Dynamical diseases of brain systems: Different routes to epileptic seizures. *IEEE Trans. Biomed. Eng.* **50**, 540–548 (2003), doi:10.1109/TBME.2003.810703
47. Lopes da Silva, F.H., van Lierop, T.H.M.T., Schrijer, C.F., van Leeuwen, W.S.: Essential differences between alpha rhythms and barbiturate spindles: Spectra and thalamo-cortical coherences. *Electroencephalogr. Clin. Neurophysiol.* **35**, 641–645 (1973), doi:10.1016/0013-4694(73)90217-4
48. Lopes da Silva, F.H., van Lierop, T.H.M.T., Schrijer, C.F., van Leeuwen, W.S.: Organization of thalamic and cortical alpha rhythms: Spectra and coherences. *Electroencephalogr. Clin. Neurophysiol.* **35**, 627–639 (1973), doi:10.1016/0013-4694(73)90216-2
49. Narici, L., Forss, N., Jousmäki, V., Peresson, M., Hari, R.: Evidence for a 7- to 9-Hz “sigma” rhythm in the human SII cortex. *NeuroImage* **13**, 662–668 (2001)
50. Niedermeyer, E.: The normal EEG of the waking adult. In: [51], pp. 167–192 (2005)
51. Niedermeyer, E., Lopes da Silva, F.H. (eds.): *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, Philadelphia, 5th edn. (2005)
52. Nikulin, V.V., Brismar, T.: Phase synchronization between alpha and beta oscillations in the human electroencephalogram. *Neuroscience* **137**, 647–657 (2006), doi:10.1016/j.neuroscience.2005.10.031

53. Novikov, E., Novikov, A., Shannahoff-Khalsa, D.S., Schwartz, B., Wright, J.: Scale-similar activity in the brain. *Phys. Rev. E* **56**, R2387–R2389 (1997), doi:10.1103/PhysRevE.56.R2387
54. Nunez, P.L.: *Electric fields of the brain: The neurophysics of EEG*. Oxford University Press, New York, 1st edn. (1981)
55. Nunez, P.L., Wingeier, B.M., Silberstein, R.B.: Spatial-temporal structures of human alpha rhythms: Theory, microcurrent sources, multiscale measurements, and global binding of local networks. *Hum. Brain. Mapp.* **13**, 125–164 (2001), doi:10.1002/hbm.1030
56. Oldeman, B.E., Krauskopf, B., Champneys, A.R.: Death of period doublings: Locating the homoclinic doubling cascade. *Physica D* **146**, 100–120 (2000), doi:10.1016/S0167-2789(00)00133-0
57. Peters, A., Payne, B.R.: Numerical relationships between geniculocortical afferents and pyramidal cell modules in cat primary visual cortex. *Cereb. Cortex* **3**, 69–78 (1993), doi:10.1093/cercor/3.1.69
58. Peters, A., Payne, B.R., Budd, J.: A numerical analysis of the geniculocortical input to striate cortex in the monkey. *Cereb. Cortex* **4**, 215–229 (1994), doi:10.1093/cercor/4.3.215
59. Pfurtscheller, G.: Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest. *Electroencephalogr. Clin. Neurophysiol.* **83**, 62–69 (1992), doi:10.1016/0013-4694(92)90133-3
60. Pfurtscheller, G., Lopes da Silva, F.H.: EEG event-related desynchronization (ERD) and event-related synchronization (ERS). In: [51], pp. 1003–1016 (2005)
61. Pfurtscheller, G., Neuper, C., Krausz, G.: Functional dissociation of lower and upper frequency mu rhythms in relation to voluntary limb movement. *Clin. Neurophysiol.* **111**, 1873–1879 (2000), doi:10.1016/S1388-2457(00)00428-4
62. Pritchard, W.S.: The brain in fractal time: 1/f-like power spectrum scaling of the human electroencephalogram. *Int. J. Neurosci.* **66**, 119–129 (1992)
63. Rabinovich, M.I., Huerta, R., Laurent, G.: Neuroscience. Transient dynamics for neural processing. *Science* **321**, 48–50 (2008), doi:10.1126/science.1155564
64. Rabinovich, M.I., Huerta, R., Varona, P., Afraimovich, V.S.: Transient cognitive dynamics, metastability, and decision making. *PLoS Comput. Biol.* **4**, e1000072 (2008), doi:10.1371/journal.pcbi.1000072
65. Robinson, P.A.: Patchy propagators, brain dynamics, and the generation of spatially structured gamma oscillations. *Phys. Rev. E* **73**, 041904 (2006), doi:10.1103/PhysRevE.73.041904
66. Robinson, P.A., Rennie, C.J., Wright, J.J.: Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* **56**, 826–840 (1997), doi:10.1103/PhysRevE.56.826
67. Schanze, T., Eckhorn, R.: Phase correlation among rhythms present at different frequencies: Spectral methods, application to microelectrode recordings from visual cortex and functional implications. *Int. J. Psychophysiol.* **26**, 171–189 (1997), doi:10.1016/S0167-8760(97)00763-0
68. Schreiber, T., Schmitz, A.: Discrimination power of measures for nonlinearity in a time series. *Phys. Rev. E* **55**, 5543–5547 (1997), doi:10.1103/PhysRevE.55.5443
69. Schreiber, T., Schmitz, A.: Surrogate time series. *Physica D* **142**, 346–382 (2000), doi:10.1016/S0167-2789(00)00043-9
70. Shaw, J.C.: *The brain's alpha rhythms and the mind*. Elsevier Sciences B. V., Amsterdam (2003)
71. Silva, L.R., Amitai, Y., Connors, B.W.: Intrinsic oscillations of neocortex generated by layer 5 pyramidal neurons. *Science* **251**, 432–435 (1991), doi:10.1126/science.1824881
72. Stam, C.J.: Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clin. Neurophysiol.* **116**, 2266–2301 (2005), doi:10.1016/j.clinph.2005.06.011
73. Stam, C.J., Pijn, J.P.M., Suffczynski, P., Lopes da Silva, F.H.: Dynamics of the human alpha rhythm: Evidence for nonlinearity? *Clin. Neurophysiol.* **110**, 1801–1813 (1999), doi:10.1016/S1388-2457(99)00099-1
74. Steriade, M.: Cellular substrates of brain rhythms. In: [51], pp. 31–83 (2005)
75. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W., Liley, D.T.J.: Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: Evidence for a general anesthetic-induced phase transition. *Phys. Rev. E* **60**, 7299–7311 (1999)

76. Tiesinga, P.H.E., Fellous, J.M., José, J.V., Sejnowski, T.J.: Computational model of carbachol-induced delta, theta, and gamma oscillations in the hippocampus. *Hippocampus* **11**, 251–274 (2001), doi:10.1002/hipo.1041
77. Tiihonen, J., Hari, R., Kajola, M., Karhu, J., Ahlfors, S., Tissari, S.: Magnetoencephalographic 10-Hz rhythm from the human auditory cortex. *Neurosci. Lett.* **129**, 303–305 (1991), doi:10.1016/0304-3940(91)90486-D
78. Tsuda, I.: Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behav. Brain Sci.* **24**, 793–810 (2001)
79. Uhl, C., Friedrich, R.: Spatiotemporal modeling based on dynamical systems theory. In: C. Uhl (ed.), *Analysis of Neurophysiological Brain Functioning*, pp. 274–306, Springer-Verlag, Berlin (1999)
80. van Rotterdam, A., Lopes da Silva, F.H., van den Ende, J., Viergever, M.A., Hermans, A.J.: A model of the spatialtemporal characteristics of the alpha rhythm. *Bull. Math. Biol.* **44**, 283–305 (1982)
81. van Veen, L., Liley, D.T.J.: Chaos via Shilnikov's saddle-node bifurcation in a theory of the electroencephalogram. *Phys. Rev. Lett.* **97**, 208101 (2006), doi:10.1103/PhysRevLett.97.208101
82. Walter, W.G., Walter, V.J.: The electrical activity of the brain. *Ann. Rev. Physiol.* **11**, 199–230 (1949), doi:10.1146/annurev.ph.11.030149.001215
83. Williamson, S.J., Kaufman, L.: Advances in neuromagnetic instrumentation and studies of spontaneous brain activity. *Brain Topogr.* **2**, 129–139 (1989), doi:10.1007/BF01128850

Chapter 7

Inducing transitions in mesoscopic brain dynamics

Hans Liljenström

7.1 Introduction

Brain structures are characterized by their complexity in terms of organization and dynamics. This complexity appears at many different spatial and temporal scales which, in relative terms, can be considered micro, meso, and macro scales. The corresponding dynamics may range from ion-channel kinetics, to spike trains of single neurons, to the neurodynamics of cortical networks and areas [6, 10]. The high complexity of neural systems is partly a result of the web of nonlinear interrelations between levels and parts with positive and negative feedback loops. This in turn introduces thresholds, lags and discontinuities in the dynamics, often leading to unpredictable and nonintuitive system behaviors [68].

Typical for complex systems in general, and for the nervous system in particular, is that different phenomena appear at different levels of spatial (and temporal) aggregation. New and unpredictable qualities emerge at every level, qualities that cannot be reduced to the properties of the components at the underlying level. In some cases, there is a hierarchical structure of a simple kind, where higher *macro* levels “control” lower ones (c.f., the so-called enslaving principle of Haken [43]). However, there could also be a more “bottom-up” interpretation of systems, where indeed the *micro* phenomena, through various mechanisms, set the frame for phenomena at higher structural levels. This interplay between micro and macro levels is part of what frames the dynamics of systems. Of special interest is the *meso* level, i.e., the level in between the micro and the macro, as this is where bottom-up meets top-down [30, 31, 68].

The activity of neural systems often seems to depend on nonlinear threshold phenomena: e.g., *microscopic* fluctuations may cause rapid and large *macroscopic* effects. There is a dynamical region between order and pure randomness that involves

Hans Liljenström

Research Group of Biometry and Systems Analysis, Department of Energy and Technology, SLU, SE-75007 Uppsala, Sweden; Agora for Biosystems, SE-19322 Sigtuna, Sweden.

e-mail: hans.liljenstrom@et.slu.se

a high degree of complexity, and which seems characteristic for neural processes. This dynamics is very unstable and shifts from one state to another within a few hundred milliseconds or less, typical of chaotic systems. (It may actually be more appropriate to refer to this behavior as “pseudo-chaotic”, since “true chaos”, as defined mathematically, requires “infinite” time for its development).

Despite at least a century of study, the functional significance of the neural dynamics at different levels is still not clear, nor is much known about the relation between activities at the different levels. However, it is reasonable to assume that different dynamical states correlate with different functional or mental states. This principle guides our research, and will be discussed further in the final section.

By studying transitions in brain dynamics, we may reveal fundamental properties of the brain and its constituents that relate to mental processes and transitions. Such transitions could, for example, involve various cognitive levels and conscious states that would be of interest not only to neuroscience, but also to psychology, psychiatry, and medicine.

In this chapter I present a range of computational models in which we investigate relations between structure, dynamics, and function of neural systems. My focus is on phase transitions in mesoscopic brain dynamics, since this type of dynamics constitutes a well-studied bridge between neural and mental processes [31]. These transitions can be induced by internal causes (noise and neuromodulation), but also by external causes (electric shocks and anesthetics). The functional significance of the model results are discussed in the concluding section.

7.1.1 Mesoscopic brain dynamics

In our description, *mesoscopic brain dynamics* refers to the neural activity or dynamics at intermediate scales of the nervous system, at levels between neurons and the entire brain. It relates to the dynamics of cortical neural networks, typically on the spatial order of a few millimetres to centimetres, and temporally on the order of milliseconds to seconds. This type of dynamics can be measured by methods such as ECoG (electrocorticography), EEG (electroencephalography), or MEG (magnetoencephalography).

We consider processes and structures studied with a microscope or microelectrodes as defining a *microscopic* scale of the nervous system; thus the micro scale could, for example, refer to ion channels or single neurons. The *macroscopic* scale, in this picture, corresponds to the largest measurable extent of brain activity. Typically, this could concern the dynamics of maps and areas, usually measured with PET or fMRI, or other brain-imaging techniques.

Mesoscopic brain dynamics, with its transitions, is partly a result of thresholds and the summed activity of a large number of elements interconnected with positive and negative feedback. It is also a result of the dynamic balance between opposing processes, influx and efflux of ions, inhibition and excitation, etc. Such interplay

between opposing processes often results in (transient or continuous) oscillatory and chaotic-like behaviour [6, 32, 44, 78].

The mesoscopic neurodynamics is naturally influenced and shaped by the activity at other scales. For example, it is often mixed with noise that is generated at a microscopic level by spontaneous activity of neurons and ion channels. It is also affected by macroscopic activity, such as slow rhythms generated by cortico-thalamic circuits or neuromodulatory influx from different brain regions. Transitions at these other levels could also be of relevance to the mesoscopic level. For example, at the microscopic level of ion channels, the kinetics assumes stochastic transitions between a limited number of static states. In spite of this, the kinetics can be given a deterministic, dynamic interpretation at a population level. Similarly, at the cellular level, there is regular or irregular spiking, or bursts of spikes, which form the basis for most mesoscopic and macroscopic descriptions of nerve activity. While the causal relations may be difficult to establish, transitions between different states of arousal, attention, or mood, could be seen as a top-down interaction from macroscopic activity to mesoscopic neurodynamics.

7.1.2 Computational methods

Computational approaches complement experimental methods in understanding the complexity of neural systems and processes. Computational methods have long been used in neuroscience, perhaps most successfully for the description of action potentials [49]. When investigating interactions between different neural levels, computational models are essential, and in some cases, may be the only method we have. (For an overview, see Refs. [3, 4, 30, 73]). In recent years, there has also been a growing interest in applying computational methods to problems in clinical neuroscience, with implications for psychology and psychiatry [29, 30, 36, 39, 42, 52, 53, 64, 74, 79, 80, 87, 88].

In our research, we use a computational approach to address questions regarding relations between structure, dynamics, and function of neural systems. Here, the focus is on understanding how transitions between different dynamical states can be implemented and interpreted. For this purpose, we present different kinds of computational models, at different scales and levels of detail, depending on the particular issues addressed.

In almost all cases, the emphasis is on network connectivity and hence there is, in general, a greater level of realism and detail for the network structures than for node characteristics. However, when microscopic details are important, or when model simulations are to be compared with data at molecular and cellular scales, such details need to be incorporated in the model, sometimes at the expense of details at the network level. Our aim is to use a level of description appropriate for the problem we address.

The first examples consider phase transitions in network dynamics arising from noise and neuromodulation. In this case, we use a three-layered paleocortical model

with simple network nodes of Hopfield type [50, 51]. Simulation results with this model are compared with LFP (local field potential) and EEG data from the olfactory cortex. For transitions due to attention, we want to compare our results with experimental data on spike trains, so we use a neocortical model with spiking neurons of Hodgkin–Huxley type [49].

In the case of electrical stimulation, we first use our paleocortical model, since we again compare with EEG data from experiments on animal olfactory cortex. The measured response is the summed activity of a very large number of neurons which will drown out single spikes, so there is no need for spiking neurons here.

When modeling and analyzing EEG related to electroconvulsive therapy, we use a neocortical network model with spiking neurons of FitzHugh–Nagumo type [25] (a simplification of the Hodgkin–Huxley description) to enable comparison against previous simulations with such model neurons [35].

In our final example, we investigate the mechanisms of anesthetics that block certain ion channels. We employ a network of Frankenhaeuser–Huxley neurons [54] because of their accurate description of ion-channel currents in cortical neurons. This microscopically detailed model allows us to compare our network results with those from single-neuron simulations for varying ion-channel composition [8].

7.2 Internally-induced phase transitions

The complex neurodynamics of the brain can be regulated by various neuromodulators, and presumably also by intrinsic noise levels, governed by thresholds for spontaneous activity. In addition, the state of arousal or attention may also change the cortical neurodynamics considerably, and even induce phase transitions that could affect the functional efficiency of cognitive processes. Such transitions may also be related to noncognitive mental processes and disorders, but that is beyond the scope of this discussion. In the following three sections, we will look at different possibilities for how intrinsic noise, neuromodulation, and attention may induce phase transitions in cortical structures.

7.2.1 *Noise-induced transitions*

Spontaneous activity, or neuronal noise, is normally seen as a naturally occurring side phenomenon without any functional role. However, it becomes increasingly clear that stochastic processes play a fundamental role in the nervous system, at least for keeping a baseline activity, but presumably also for increasing the efficiency in system performance (see e.g., Ref. [5], and Sect. 7.4 Discussion).

Noise appears primarily at the microscopic (subcellular and cellular) levels, but it is uncertain to what degree this noise normally is affecting meso- and macroscopic levels (networks and systems). Under certain circumstances, microscopic

noise can induce effects on mesoscopic and macroscopic levels, but the role of these effects is still unclear. Evidence suggests that even single channel openings can cause intrinsic spontaneous impulse generation in a subset of small hippocampal neurons [54].

In addition to the microscopic noise, irregular chaotic-like behavior, which may be indistinguishable from noise, could be generated by the interplay of neural excitatory and inhibitory activity at the network level. However, in contrast to a chaotic dynamics, where the dynamics can be controlled and easily shifted into an oscillatory or other state, stochastic noise is not equally controllable, and cannot shift into a completely different dynamics (even though its amplitude and frequency might vary as a result of neuromodulatory control).

7.2.1.1 A paleocortical network model

When studying how the neurodynamics of a cortical structure depends on various internal factors, including neuromodulation and intrinsic noise from spontaneously firing neurons, we use our previously constructed model of the olfactory cortex [60]. (With a few modifications, this model can also be used for the hippocampus, which has a similar structure). Paleocortex, primarily consisting of the olfactory cortex and hippocampus, is more primitive and simpler than neocortical structures, such as the visual cortex. It has a three-layered structure and a distributed connectivity pattern with extensive short- and long-range connections within a layer. Due to its simpler structure and well-studied neurodynamics, the olfactory cortex can be regarded as a suitable model system for the study of mesoscopic brain dynamics.

Our paleocortical model has network nodes with a continuous input–output relation, the output corresponding to the average firing frequency of neural populations [50, 51]. Three different types of nodes (neural populations) are organized in three layers, as seen in Fig. 7.1. The top layer consists of inhibitory feedforward interneurons, which receive inputs from the olfactory bulb, via the lateral olfactory tract (LOT), and from the excitatory pyramidal cells in the middle layer. The bottom layer consists of inhibitory feedback interneurons, receiving inputs only from the pyramidal cells and projecting back to those. The two sets of inhibitory cells are characterized by their different time-constants. In addition to the feedback from inhibitory cells, the pyramidal cells receive extensive inputs from each other and from the olfactory bulb, via the LOT. All connections are modeled with distance-dependent time-delays for signal propagation, corresponding to the geometry and fiber characteristics of the real cortex.

The time-evolution for a network of N network nodes (neural populations) is given by a set of coupled nonlinear first-order differential-delay equations for all the N internal states, u_i (corresponding to mean membrane potential of a population i). With external input, $I(t)$, characteristic time constant, τ_i , and connection weight w_{ij} between nodes i and j , separated with a time-delay δ_{ij} , we have for each node activity, u_i ,

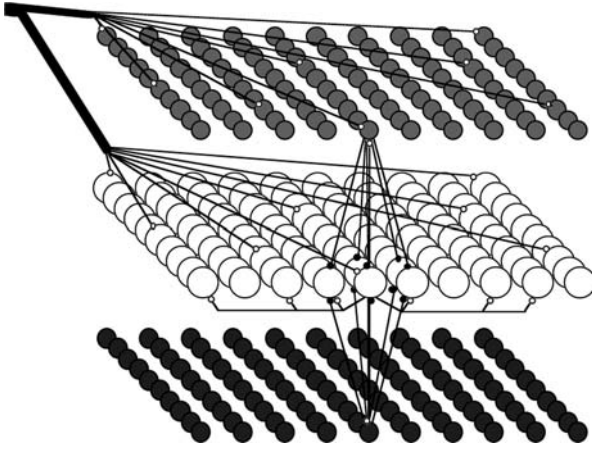


Fig. 7.1 Schematic of our model neural network that mimics the structure of the olfactory cortex. One layer of excitatory nodes, corresponding to populations of pyramidal cells (large circles in middle layer), is sandwiched between two layers of inhibitory nodes, corresponding to two different types of interneurons (smaller circles, top and bottom layers). External input (from “the olfactory bulb”) projects onto the two top layers in a fan-like fashion.

$$\frac{du_i}{dt} = -\frac{u_i}{\tau_i} + \sum_{j \neq i}^N w_{ij} g_j [u_j(t - \delta_{ij})] + I_i(t) + \xi(t). \quad (7.1)$$

The input–output function, $g_i(u_i)$, is a continuous sigmoid function, experimentally determined by Freeman [28]:

$$g_i = C Q_i \left\{ 1 - \exp \left[-\frac{\exp(u_i)}{Q_i} \right] \right\}. \quad (7.2)$$

The gain parameter Q_i determines the slope, threshold and amplitude of the input–output curve for node i . This gain parameter is associated with the level of arousal, which in turn may be linked to the level of a neuromodulator, such as acetylcholine (ACh). C is a normalisation constant.

The connection weights w_{ij} are initially set and constrained by the general connectivity principles for the olfactory cortex, but to allow for learning, the weights can be incrementally changed according to a learning rule of Hebbian type [61]. However, learning is not explicitly considered here, although it may well relate to the functional significance of phase transitions in cortical neurodynamics. (Our olfactory/hippocampal model has previously been used for studying the effects of neuromodulation and noise on the efficiency of information processing [61, 63, 69]).

Neuromodulatory effects are simulated by changing the Q -values for primarily the excitatory nodes. When neuromodulatory effects on synaptic transmission are included, we change separately a weight-constant that multiplies all connection strengths, w_{ij} . (Another way to implement neuromodulatory effects is by

multiplying the input–output function, g , with an exponential-decay function, representing neuronal adaptation, as has been described elsewhere [67]).

Noise, or spontaneous neural activity, is added in the last term of Eqn. (7.1) via a Gaussian noise function, $\xi(t)$, such that $\langle \xi(t) \rangle = 0$, and $\langle \xi(t)\xi(s) \rangle = 2A\delta(t-s)$. We have studied noise effects by increasing the level A . In some of the simulations, the noise level is changed equally for all network nodes, whereas in other simulations, the change takes place in only some of the network nodes.

7.2.1.2 Simulating noise-induced phase transitions

Simulations with our three-layered paleocortical model display a range of dynamic properties found in olfactory cortex and hippocampus. For example, the model accurately reproduces response patterns associated with a continuous random input signal, and with shock pulses applied to the cortex; see Figs. 7.2 and 7.7 [60].

For a constant, low-amplitude random input (noise), the network is able to oscillate with two separate frequencies simultaneously, around 5 Hz (theta rhythm) and 40 Hz (gamma rhythm). Under certain conditions, such as for high Q -values, the system can also display chaotic-like behaviour, similar to that seen in EEG traces (see Fig. 7.2). In associative memory tasks, the network may initially display a chaotic-like dynamics, which then converges to a near limit-cycle attractor when storing or retrieving a memory (activity pattern) [61, 69].

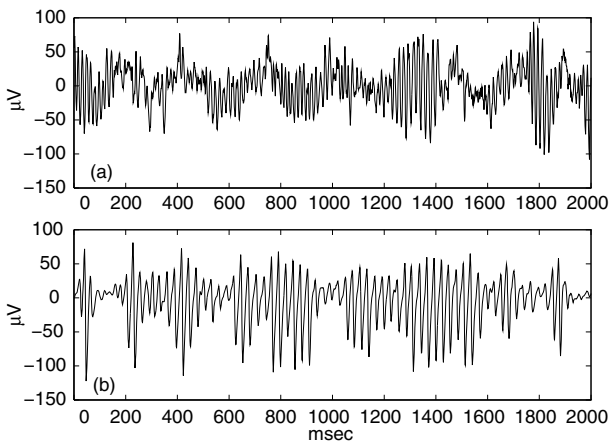


Fig. 7.2 (a) Real and (b) simulated EEG, showing the complex dynamics of cortical structures. Upper trace is from rat olfactory cortex (data courtesy of Leslie Kay); lower trace is from a simulation with the current model.

Simulations with various noise levels show that spontaneously active neurons can induce global, synchronized oscillations with a frequency in the gamma range (30–70 Hz) [62]. Even if only a few network nodes are noisy (i.e., have an increased

intrinsic random activity), and the rest are quiescent, coherent oscillatory activity can be induced in the entire network if connection weights are large enough [7, 62, 65]. The onset of global oscillatory activity depends on, for example, connectivity, noise level, number of noisy nodes, and duration of the noise activity [15]. The location and spatial distribution of these nodes in the network is also important for the onset and character of the global activity. For example, as the number or activity of noisy nodes is increased, or if the distance between them increases, the oscillations tend to change into irregular patterns. In Fig. 7.3, we show that global network activity can be induced if only five out of 1024 network nodes are noisy, and the rest are silent. After a short transient period of collective irregular activity, the entire network begins to oscillate, and collective activity waves move across the network. Even if there is only a short burst of noisy activity, this may be enough to induce global oscillations [15].

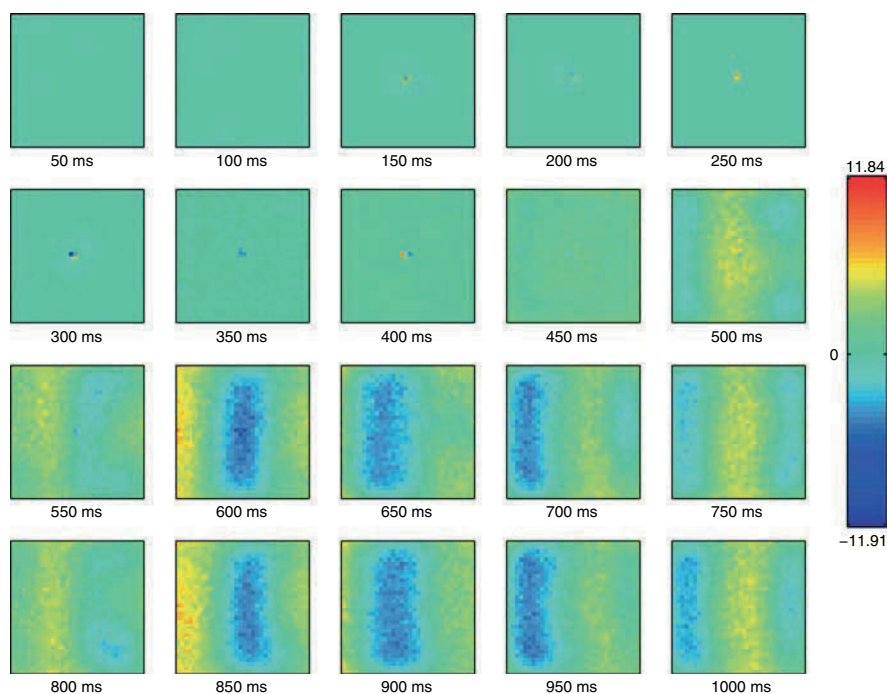


Fig. 7.3 [Color plate] Spatiotemporal activity of the excitatory layer of a three-layered paleocortical model, presented as snapshots of network activity (as mean membrane potential of neural populations) at 50-ms intervals. Five centrally-located noisy network nodes can induce collective waves of activity across the entire network. Simulations were made with a 32×32 grid of network nodes in each network layer, corresponding to a 10×10 -mm square of rat olfactory cortex. Activity is color-coded on a scale ranging from negative = blue to positive = red.

We have also studied the effects of spontaneously active feedforward inhibitory interneurons in the top layer, motivated by the experimental finding that single inhibitory neurons can synchronize the activity of up to 1000 pyramidal cells [21].

Our simulations demonstrated that even a single noisy network node in the feedforward layer could induce periods of synchronous oscillations in the excitatory layer with a frequency in the gamma range, interrupted by periods of irregular activity [15].

From the simulations, it is apparent that internal noise can cause various phase transitions in the network dynamics. An increased noise level in just a few network nodes can result in a transition from a stationary to an oscillatory state, or from an oscillatory to a chaotic state, or alternatively, a shift between two different oscillatory states [56, 69]. (A more thorough investigation—in which we studied the effects of varying the density of noisy nodes, the noise duration, and the noise level—is reported in [15].)

All of these phenomena depend critically on network structure, in particular on the feedforward and feedback inhibitory loops, and the long-range excitatory connections, modeled with distance-dependent time delays. In this model, details concerning neuron structure or spiking activity are not necessary for the neurodynamics under study. Instead, a balance between inhibition and excitation, in terms of connection strength and timing of events, is essential for coherent frequency and phase of the oscillating neural nodes.

7.2.2 Neuromodulatory-induced phase transitions

Brain activity is constantly changing due to sensory input, internal fluctuations, and neuromodulation. Neuromodulators, such as acetylcholine (ACh) and serotonin (5-HT), can change the excitability of a large number of neurons simultaneously, or the synaptic transmission between them [18], thus dramatically influencing brain dynamics. ACh can increase excitability by suppressing neuronal adaptation, an effect similar to that of increasing the gain in general. The concentration of these neuromodulators seems to be directly related to the level of arousal or motivation of the individual, and can have profound effects on the neural dynamics (e.g., an increased oscillatory activity), and on cognitive functions, such as associative memory [30].

We use the paleocortical model described in Sect. 7.2.1.1 to investigate how network dynamics can be regulated by neuromodulators, implemented in the model as a varied excitability of the network nodes, and modified connection strengths [67]. The frequencies of the network oscillations depend primarily on intrinsic time-constants and delays, whereas the amplitudes depend predominantly on connection weights and gains, which are under neuromodulatory control. Implementation of these neuromodulatory effects in the model cause dynamical changes analogous to those seen in physiological experiments.

In particular, a “cholinergic” increase in excitability together with suppression of synaptic transmission could induce theta (and/or gamma) rhythm oscillations within the model, even when starting from an initially quiescent state with no oscillatory activity. Fig. 7.4 shows how different oscillatory modes can be induced by

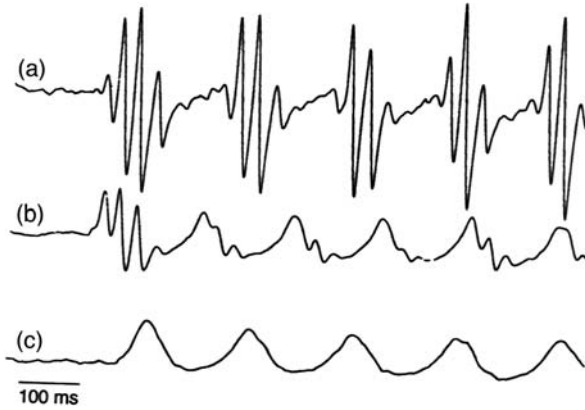


Fig. 7.4 Different oscillatory modes can be induced by cholinergic neuromodulatory effects that increase gain and decrease connection strengths. The activity evolution of one particular (arbitrarily chosen) excitatory network node is shown for three different levels of “cholinergic” action: (a) low; (b) intermediate; and (c) high.

neuromodulatory effects: increasing gain and decreasing connection weights. The activity evolution of one arbitrarily chosen excitatory network node is shown for three different levels of “ACh”. For example, if $Q = 10.0$ and $w_{exc} = w_{inh} = 1.0$ (i.e., no suppression of synaptic transmission; w_{exc} and w_{inh} are excitatory and inhibitory connection-weight factors respectively), we can get an oscillatory mode with two different frequencies (~ 5 Hz and 40 Hz) present simultaneously. This is shown in trace (a) of Fig. 7.4. If Q is kept constant ($= 10$) while w_{exc} and w_{inh} are reduced successively, the high-frequency component weakens and eventually can be totally eliminated. In trace (b), the connection strengths were decreased by 40% for all excitatory nodes (i.e., $w_{exc} = 0.6$), and by 60% for all inhibitory nodes ($w_{inh} = 0.4$). Trace (c) shows the result for $w_{exc} = 0.4$ and $w_{inh} = 0.2$. In the latter case, only the low-frequency component remains. If the excitatory connection strengths are decreased further, i.e. if $w_{exc} \leq 0.3$, oscillations disappear.

7.2.3 Attention-induced transitions

Related to the level of arousal, and apparently also under neuromodulatory control, is the phenomenon of attention, which plays a key role in perception, action selection, object recognition, and memory [46]. The main effect of visual attentional selection appears to be a modulation of the underlying competitive interaction between stimuli in the visual field. Studies of cortical areas V2 and V4 indicate that attention modulates the suppressive interaction between two or more stimuli presented simultaneously within the receptive field [22]. Visual attention has several effects on modulating cortical oscillations in terms of changes in firing rate [72], and gamma and beta coherence [34].

In selective-attention tasks, after the cue onset and before the stimulus onset, there is a delay-period during which a monkey's attention was directed to the place where the stimulus would appear [34]. During the delay, the dynamics was dominated by frequencies around 17 Hz, but with attention, this low-frequency synchronization decreased. During the stimulus period, there were two distinct bands in the power spectrum, one below 10 Hz and another at 35–60 Hz (gamma). With attention, there was a reduction in low-frequency synchronization and an increase in gamma-frequency synchronization.

At a meso-scale, each area of the visual cortex is conventionally divided into six layers, some of which can be further divided into several sub-layers, based on their detailed functional roles in visual information processing (such as orientation and retinotopic position).

The inter-scale network interactions of various excitatory and inhibitory neurons in the visual cortex generate oscillatory signals with complex patterns of frequencies associated with particular states of the brain. Synchronous activity at an intermediate and lower-frequency range (theta, delta, and alpha) between distant areas has been observed during perception of stimuli with varying behavioral significance [76, 84]. Rhythms in the beta (12–30 Hz) and the gamma (30–80 Hz) ranges are also found in the visual cortex, and are often associated with attention, perception, cognition and conscious awareness [23, 24, 34, 37, 38]. Data suggest that gamma rhythms are associated with relatively local computations, whereas beta rhythms are associated with higher-level interactions. Generally, it is believed that lower-frequency bands are generated by global circuits, while higher-frequency bands are derived from local connections.

7.2.3.1 A neocortical network model

In order to investigate how attentional neuromodulation can affect cortical neurodynamics, and cause the observed phase shifts discussed above, we use a neural network model of the visual cortex, based on known anatomy and physiology [41].

Although neocortex consists of six layers—in contrast to paleocortex with its three layers—for simplicity, we lump some of the neocortical layers together. Thus, our model has three functional layers, including layer 2/3, layer 4 and layer 5/6 of the visual cortex. Each layer contains 20×20 excitatory model neurons (pyramidal neurons in layer 2/3 and layer 5/6, and spiny stellate neurons in layer 4) in a quadratic lattice with lattice distance 0.2 mm. For each excitatory layer, there are also 10×10 inhibitory neurons in a quadratic lattice with lattice distance 0.4 mm. Thus, there are 20% inhibitory neurons, which roughly corresponds to the observed cortical distribution.

Figure 7.5 shows the schematic diagram of the network topology. The inhibitory neurons in each layer have interactions within their own layer only, while excitatory neurons have interactions within their own layer as well as between layers and areas. The within-layer connections between excitatory and inhibitory neurons is of “Mexican hat” shape, with an on-center and an off-surround lateral synaptic input

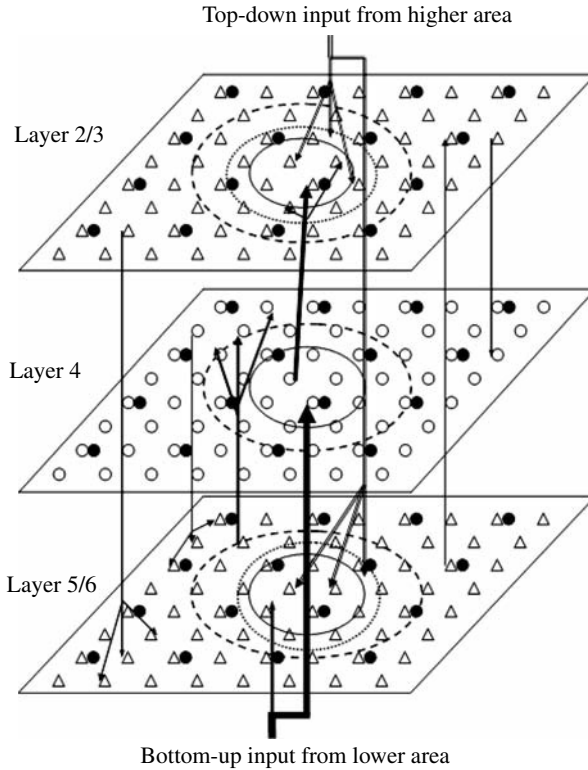


Fig. 7.5 Schematic diagram of the model architecture. Small triangles in layers 2/3 and 5/6 represent pyramidal neurons; small open circles in layer 4 are spiny stellate neurons; small filled circles in each layer are inhibitory neurons. Arrows show connection patterns between different layers and signal flows from other areas. Large solid open circles represent lateral excitatory connection radius; large dashed open circles represent inhibitory connection radius; dotted open circles in layers 2/3 and 5/6 denote the top-down attention modulation radius R_{modu} .

for each neuron, i.e., excitatory at short distance, and inhibitory at a long distance (see Ref. [41] for details).

Since we wish to compare model results against observed data from visual cortex—in particular, spike-triggered averages of local field potentials—we need to use spiking model neurons; this is in contrast to the paleocortical model, which uses network nodes corresponding to *populations* of neurons, resulting in a continuous non-spiking output. For the present case, all excitatory model neurons satisfy Hodgkin–Huxley equations of the form,

$$\begin{aligned}
 C \frac{dV}{dt} = & -g_L(V + 67) - g_{\text{Na}} m^3 h (V - 50) - g_K n^4 (V + 100) \\
 & - g_{\text{AHP}} w (V + 100) - I^{\text{syn}} + I^{\text{appl}}, \quad (7.3)
 \end{aligned}$$

where V is the membrane potential in mV; $C = 1 \mu\text{F}$ is the membrane capacitance; g_L is the leak conductance; $g_{\text{Na}} = 20 \text{ mS}$ and $g_K = 10 \text{ mS}$ are the maximal sodium and potassium conductances, respectively; g_{AHP} is the maximal slow potassium conductance of the after-hyperpolarization (AHP) current—this varies from 0 to 1.0 mS, depending on the attentional state: in an idle state, $g_{\text{AHP}} = 1.0 \text{ mS}$; with attention, $g_{\text{AHP}} \leq 1.0 \text{ mS}$. The variables m , h , n and w are calculated in a conventional way, and described more thoroughly in Ref. [41].

The inhibitory neurons have identical equations as above, except that there is no AHP current. The synaptic input current, I^{syn} of the pyramidal, stellate, and inhibitory neurons is described below.

In each layer j (where $j = 2/3, 4$, and $5/6$) of the local-area network, there are four types of interactions: (i) lateral excitatory–excitatory, (ii) excitatory–inhibitory, (iii) inhibitory–excitatory, and (iv) inhibitory–inhibitory, with corresponding connection strengths, $C_{j,kl}^{ee}$, $C_{j,kl}^{ie}$, $C_{j,kl}^{ei}$, and $C_{j,kl}^{ii}$, which vary with distance between neurons k and l .

The synaptic input current, $I_{4s,k}^{\text{syn}}(t)$, of the k^{th} stellate neuron in layer 4 at time t is composed of the ascending input from the pyramidal neurons in layer 5/6, descending input from the pyramidal neurons in layer 2/3, and lateral excitatory inputs from the on-centre neighboring stellate neurons in layer 4. It also includes lateral inhibitory inputs from the off-surround neighboring inhibitory neurons in the same layer, resulting in,

$$I_{4s,k}^{\text{syn}}(t) = (V_{4s,k}(t) - V_E) \left(\sum_l C_{4(5/6),kl}^{ee} s_{5/6,l}^e(t) + \sum_l C_{4(2/3),kl}^{ee} s_{2/3,l}^e(t) + \sum_l C_{4,kl}^{ee} s_{4,l}^e(t) \right) + (V_{4s,k}(t) - V_I) \sum_l C_{4,kl}^{ei} s_{4,l}^i(t). \quad (7.4)$$

The synaptic input current, $I_{4i,k}^{\text{syn}}(t)$, of the k^{th} inhibitory neuron in layer 4 is composed of the lateral excitatory inputs from neighboring stellate neurons and lateral inhibitory inputs from neighboring inhibitory neurons,

$$I_{4i,k}^{\text{syn}}(t) = (V_{4i,k}(t) - V_E) \sum_l C_{4,kl}^{ie} s_{4,l}^e(t) + (V_{4i,k}(t) - V_I) \sum_l C_{4,kl}^{ii} s_{4,l}^i(t). \quad (7.5)$$

The synaptic input currents for the other layers 2/3 and 5/6 are calculated in a similar way (see Ref. [41] for details). In addition, each neuron of the network receives an internal background noise current.

The excitatory and inhibitory presynaptic outputs in Eqs. (7.4) and (7.5) satisfy first-order differential equations (7.6) and (7.7), respectively:

$$\frac{d}{dt} s_{j,l}^e = 5(1 + \tanh(V_{j,l}/4))(1 - s_{j,l}^e) - s_{j,l}^e/2, \quad (7.6)$$

$$\frac{d}{dt} s_{j,l}^i = 2(1 + \tanh(V_{j,l}/4))(1 - s_{j,l}^i) - s_{j,l}^i/15, \quad (7.7)$$

where j refers to the layer, and l to the presynaptic neuron. $V_{j,l}$ corresponds to the membrane potential of presynaptic neuron l in layer j .

7.2.3.2 Simulating neurodynamical effects of visual attention

Our simulations are based on the visual attention experiment by Fries *et al.* [34]. Thus, in each of the three layers, we have groups of “attended-in” neurons, A_{in} (where attention is directed to a stimulus location inside the receptive field (RF) of these neurons), and groups of “attended-out” neurons, A_{out} (where attention is directed to a stimulus location outside the RF these neurons). During a stimulus period, two identical stimuli are presented: one appears at a location inside the RF of the A_{in} neurons and the other appears at a location inside the RF of the A_{out} neurons. The top-down modulation radius R_{modu} is taken as 0.6 mm, which is larger than the lateral excitatory connection radius of 0.5 mm, in each layer. In addition, each neuron of the network receives an internal background-noise input current.

When analyzing the simulated spike trains, we calculate power spectra of spike triggered averages (STAs) of the local field potential (LFP), representing the oscillatory synchronization between spikes and LFP. We investigate the dynamics and the effects of attention (cholinergic modulation) in an *idle* state, during *stimulation*, and during a *delay* period, as described in more detail below.

When attention is directed to a certain place, the prefrontal lobe sends cholinergic input signals via top-down pathways to layers 2/3 and 5/6 of the visual cortex, as shown in Fig. 7.5. To test various hypotheses about the mechanisms of attention modulation, we assume that the top-down signals may have *three different effects* on the pyramidal neurons, and on the local and global network connections in our simulations: (i) facilitation of extracortical top-down excitatory synaptic inputs to the pyramidal neurons (global connections); (ii) inhibition of certain intracortical excitatory and inhibitory synaptic conductances (local connections) [58, 59]; and (iii) modulation of the slow AHP current by decreasing the K-conductance, g_{AHP} , thus increasing excitability [19].

We simulated the attentional modulation effect of inhibition of intracortical excitatory and inhibitory synaptic inputs by decreasing the lateral excitatory and inhibitory conductances to zero (i.e., $g_j^{ee} = g_j^{ei} = 0$ mS) for the pyramidal neurons in the A_{in} neurons within R_{modu} in layers 2/3 and 5/6.

To simulate the dynamics during a stimulus period, we applied a pair of bottom-up sensory stimulation currents: a stronger current of $25 \mu\text{A}$, and one weaker current of $5 \mu\text{A}$. The stronger current was directly applied to layer-4 stellate neurons in both the A_{in} and the A_{out} groups. The weaker current was applied to layer-5/6 pyramidal neurons in both groups. In addition, top-down attention modulation was applied to the system.

Figure 7.6 shows the effects of attentional modulation on neuronal spikes, LFP, STA, and STA-power in a delay period (Fig. 7.6(a)), and in a stimulation period (Fig. 7.6(b)). The top traces show the LFP of A_{in} and A_{out} neurons, respectively. Below the LFP traces are the spike trains of a pyramidal cell in each of the A_{in} and

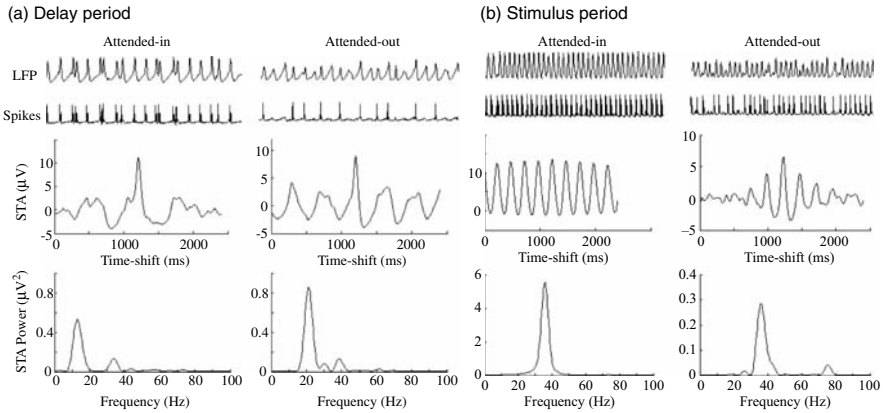


Fig. 7.6 Attentional modulation effects during (a) a *delay* period, and (b) a *stimulus* period. LFP (local field potential), spikes, STA (spike-triggered averages) and STA power of attended-in and attended-out groups, calculated for the superficial layer, when the excitatory connections and inhibitory connections to each pyramidal neuron in the attended-in group within R_{modu} in layers 2/3 and 5/6 are reduced to zero.

A_{out} groups. The computed STA and STA-power of the corresponding neurons in layer 2/3 are shown in the middle and bottom of the figure.

The simulation results show reduced beta synchronization with attention during a delay period (under certain modulation situations, see Fig. 7.6(a)), and enhanced gamma synchronization due to attention during a stimulation period (Fig. 7.6(b)). In comparison with an idle state for which the dominant frequencies are around 17 Hz, the bottom panel of Fig. 7.6(a) shows that the dominant frequency of the oscillatory synchronization and its STA power in the A_{in} group is decreased by inhibition of the intracortical synaptic inputs. This result agrees qualitatively with experimental findings that low-frequency synchronization is reduced during attention. In comparison with Fig. 7.6(a), the dominant frequency of the STA power spectrum of both A_{in} and A_{out} groups in Fig. 7.6(b) is shifted towards the gamma band due to the stimulation inputs. The STA power of the dominant frequency of the A_{in} group is higher than that of the A_{out} group.

It is apparent that many factors play important roles in the network neurodynamics. These include (i) the interplay of ion channel dynamics and neuromodulation at a micro-scale; (ii) the lateral connection patterns within each layer; (iii) the feed-forward and feedback connections between different layers at a meso-scale; and (iv) the top-down and bottom-up circuitries at a macro-scale. The interaction between the top-down attention modulation, and the lateral short-distance excitatory and long-range inhibitory interactions, all contribute to the beta synchronization decrease during the delay period, and to the gamma synchronization enhancement during the stimulation period in the A_{in} group. The top-down cholinergic modulation tends to enhance the excitability of the A_{in} group neurons. The Mexican-hat lateral interactions mediate the competition between A_{in} and A_{out} groups.

Other simulation results (not shown) demonstrate that the top-down attentional or cholinergic effects on individual neurons, and on local and global network connections, are quite different. The effect of facilitating global extracortical connections results in a slight shift of the dominant frequency in the STA power spectrum to higher beta in both the A_{in} and the A_{out} groups. In particular, the higher beta synchronization of the A_{in} group is much stronger than that of the A_{out} group.

7.3 Externally-induced phase transitions

In addition to various internal (natural) causes of phase transitions, there is a number of ways to induce neural phase transitions externally (artificially). Here, we will exemplify this by electrical stimulation and by application of anesthetics. Applying such external inputs may give a further clue to the dynamical features of the neural system under study, in much the same way as the response of any system to an external signal may reveal important system properties.

7.3.1 *Electrical stimulation*

By the 18th century, when the Italian physicists Galvani and Volta examined electrical properties of living tissues of frogs, it had become clear that nerves and muscles could respond to electrical stimulation. Since then, electricity has been used both to stimulate and to measure nerve activity in the body, and also in the brain itself. The possibility of measuring the electrical component of brain activity with external electrodes was discovered by Berger in the early 20th century [16], and it was not difficult to see that direct electrical stimulation also could affect brain activity. A variety of electric stimulations have been used not only for investigating brain response, but also to treat mental disorders such as depression [17, 42], schizophrenia [83], and neurological disorders such as Parkinson's disease [82]. In the following, we will give an example of how electrical stimulation can be used to study the relation between structure, dynamics, and function in a mammalian brain. A second example will illustrate how electrical stimulation is used in psychiatry.

7.3.1.1 **Electrical pulses to olfactory cortex**

When studying the dynamical properties of the olfactory cortex, Freeman and co-workers stimulated the lateral olfactory tract (LOT) of cats and rodents with electric shock pulses of varying amplitude and duration, then recorded the neural response via EEG [26, 27]. A strong pulse gives a biphasic response with a single fast wave moving across the surface, whereas a weak pulse results in an oscillatory response, showing up as a series of waves with diminishing amplitude. When a short pulse is applied to the LOT input corner of the network model, waves of activity move across the model cortex, consistent with corresponding global dynamic behavior

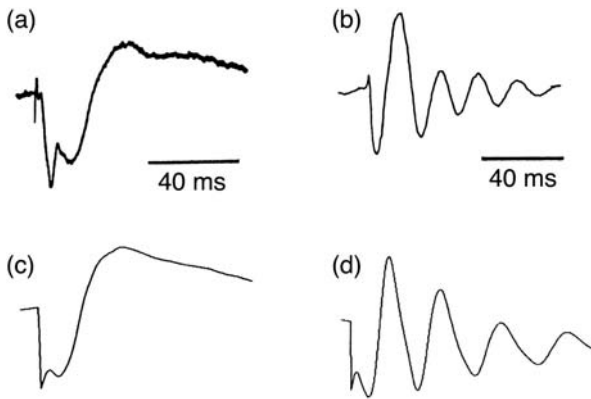


Fig. 7.7 Comparison of experimental data (a, b) from rodent olfactory cortex (courtesy of W.J. Freeman) with simulated data (c, d) from our paleocortical network model. Left traces show response to a strong shock pulse; right traces are response to a weak pulse.

of the functioning cortex. In Fig. 7.7, the experimentally measured responses are shown in the upper traces, and the model simulations are shown in the lower traces.

7.3.1.2 Electroconvulsive therapy

A more dramatic example of electrical stimulation comes from psychiatry, where electroconvulsive therapy (ECT) is one of the most successful treatments for depression and other mental disorders [17]. Despite its widespread use and successful results, it is still not known how ECT affects the brain neurologically. It has been suggested that it causes changes in the connectivity of cortical networks, either negatively, by destroying cells and/or synapses, or positively, by stimulating nerve-cell growth and sprouting [1, 86].

Clinical data show that the EEG of patients treated with ECT changes qualitatively over the treatment session, and displays some characteristic behaviors [42]. Due to the complexity of these time-series, analytical work has been difficult and scarce, and the anatomical and physiological basis for the dynamical patterns of post-ECT EEG remain to be elucidated.

In general, the EEG after ECT stimulation exhibits a specific pattern of seizures (see Fig. 7.8), but there are individual differences depending on seizure threshold, stimulus doses, and sub-diagnosis [39, 40, 42, 85]. Apparently, ECT stimulation can induce synchronous oscillations of neuronal populations over large parts of the brain where the oscillatory patterns depend on intrinsic properties, the external input and the treatment procedure. The dynamics of a recorded post-ECT, ictal, EEG time-series shifts between several phases [85]. Generally, in the clinical data one can find a sequence of phases such as preictal, polyspike (tonic), polyspike and slow-wave (clonic), termination, and postictal, respectively [17].

We apply computational methods to address the problem of how ECT might affect cortical structures and their dynamics. We have developed models of

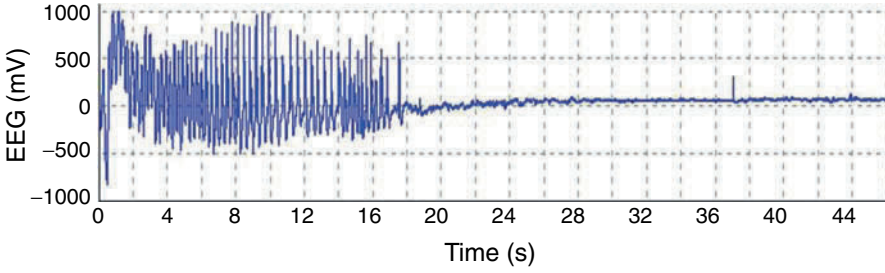


Fig. 7.8 EEG trace immediately after ECT stimulus in a patient with recurrent major depression.

neocortical structures to investigate and suggest possible mechanisms underlying the EEG signal, and in particular, how ECT-like input might influence the dynamics of the system. We are able to simulate qualitatively certain ECT EEG patterns [39, 40]. Considering the characteristics of the dynamical shifts between several phases of ECT EEG, we assume that the phase shifts are related to intrinsic local and global network properties, physiological parameters of the cortex, and external ECT stimulus.

We use various versions of a neocortical model similar to that of Sect. 7.2.3.1, but with differently modeled neurons, since we want to compare our results with previous simulations of ECT EEG by Giannakopoulos *et al* [35]. Network connectivity is varied in terms of cell types, number of neurons, and short- and long-distance connections. In particular, we investigate how a variation in the balance between excitation and inhibition affects the network dynamics. The guiding idea is that ECT primarily acts on network connectivity in stimulating nerve cell growth and sprouting [1, 86].

The model uses, as far as possible, physiological parameter values, and the same equations for describing the dynamics in all of the model variants. The network dynamics is described by Eq. (7.8), and the neurons are modeled as continuous output units of Fitzhugh–Nagumo type, as described by Eqs (7.9) and (7.10). The equations and parameter values are essentially the same as in Ref. [35], but in Eq. (7.8), we also include inputs from inhibitory neurons to other inhibitory and excitatory neurons.

$$\begin{aligned} \tau^{\text{ex(in)}} \frac{d}{dt} u_i^{\text{ex(in)}}(t) = & -u_i^{\text{ex(in)}}(t) + p^+ \sum_{k=1}^n c_{ik}^{\text{ex(in)/ex}} g(v_k^{\text{ex}}(t - T_{ik}^{\text{ex(in)/ex}})) \\ & + p^- \sum_{k=1}^n c_{ik}^{\text{ex(in)/in}} g(v_k^{\text{in}}(t - T_{ik}^{\text{ex(in)/in}})) + e_i^{\text{ex(in)}}(t - T^\sigma), \end{aligned} \quad (7.8)$$

$$\frac{d}{dt} v_i(t) = c(w_i(t) + v_i(t) - \frac{1}{3}v_i(t)^3) + \gamma_i u_i(t), \quad (7.9)$$

$$\frac{d}{dt} w_i(t) = (a - v_i(t) - bw_i(t))/c, \quad (7.10)$$

$$g(v) = \frac{M_g - m_g}{1 + \exp(-\alpha v)}. \quad (7.11)$$

Here, u_i is the postsynaptic potential of neuron i ; v_i is the membrane potential at the axon initial segment; w_i is an auxiliary variable; a , b , and c are appropriate positive constants which guarantee the existence of the oscillation interval; and $e_i^{\text{ex(in)}}$ is the external signal. The nonlinear function $g(v)$ describes the relation between the pre- and postsynaptic potential of the neurons, and is monotonically increasing ($\alpha > 0$) and nonnegative ($0 \leq m_g < M_g$). The elements of the connection matrix, c_{ik} , describe the topology of the network, and p^+ and p^- are the excitatory and inhibitory connection strengths, respectively. The neurons have time-constants τ^{ex} and τ^{in} . The total time-delay, T_{ik} , consists of a synaptic delay, T^σ , and the dendritic and axonal propagation time from neuron k to neuron i . The synaptic membrane conductance of neuron i is denoted by γ_i . The EEG signal is calculated as the mean membrane potential over all (excitatory) neurons.

The network connectivity mimics that of the six-layered neocortex, with columns connected via long-range lateral connections, and with a circuitry inspired by Szentagothai and others [75, 77, 81]. In our simulations, we use 100 neurons, of which 80 are excitatory of two types (pyramidal and spiny stellate neurons), and 20 are inhibitory of two types (large basket neurons and short-distance inhibitory interneurons). Each layer consists of 4×4 excitatory neurons in a quadratic lattice with lattice spacing 0.2 mm, and four randomly distributed inhibitory neurons. The distance between layers is 0.4 mm. The “regional” network connects four columns by long-distance excitatory connections in layers 2 and 3, with a distance between columns of 4 mm. (A more thorough description of the model is given in Ref. [39].)

In Fig. 7.9, the mean activity of simulated excitatory neurons in layers 2 to 6 is shown, going from top to bottom (layer 1 is considered to consist of fibers only). The duration of the ECT-like input is 200 ms. As seen from the figure, the neurons in each layer begin to oscillate synchronously during the ECT stimulation, but the collective oscillatory patterns vary from layer to layer, depending on the difference in connectivity.

In the left-hand traces of Fig. 7.9, the simulation shows the neural dynamics resulting from long-range inhibitory connections between basket cells in layer 3 of the four columns. In these traces, the mean membrane potential shows rather strong phase shifts in layers 2 and 3 due to the long-distance inhibitory connections in layer 3. In layer 4, mean membrane potential decreased abruptly, long before the ECT stimulus had ended. After the ECT input had ended, oscillations died out immediately in this layer, due to the lack of lateral excitatory connections here. The synchronous oscillations are comparatively strong in layers 5 and 6 due to a reduced neuronal density in these layers.

In the right-hand side of Fig. 7.9, we have replaced the long-range inhibitory connections by long-range excitatory lateral connections between the four pyramidal neurons in the centers of each column within layers 2 and 3. After the ECT stimulation, the synchronous oscillations in layers 2 and 3 show fewer phase shifts between high and low amplitude due to the long-distance excitatory connections. The activity in layer 4 is almost the same as for the case of long-range inhibition. In layers 5 and 6, the mean membrane potential shows more prominent phase shifts between synchronous and desynchronous oscillation after the ECT stimulus has ended.

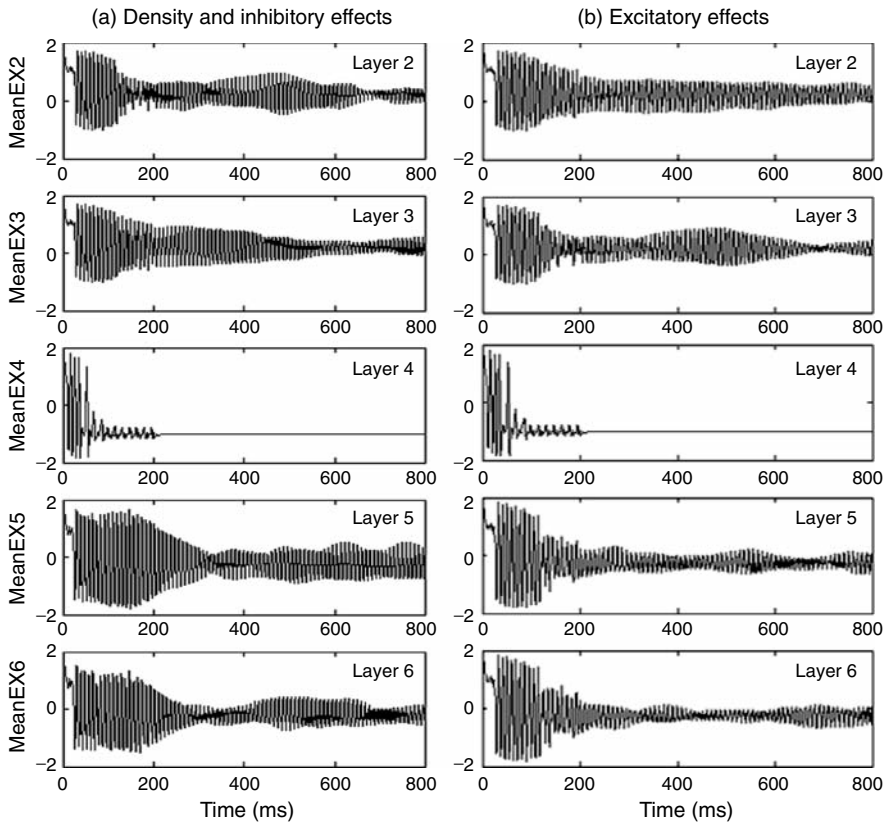


Fig. 7.9 Network response to simulated 200-ms ECT stimulus. From top to bottom, panels show mean membrane potential of excitatory neurons in layers 2–6 respectively. (a) Effect of density and long-distance inhibitory connections; (b) effect of long-distance excitatory lateral connections. Scale for y-axis is arbitrary.

When we decrease the excitatory connection strength of the network, the synchronous oscillations decrease in each layer. (The network dynamics can also change dramatically if, for example, the time-constants of excitatory and/or inhibitory neurons are changed slightly).

These studies demonstrate that the collective network dynamics varies with connection topology, neuron density in different layers, the balance between excitatory and inhibitory strength, neuronal intrinsic oscillatory properties, and external input.

Clinical EEG data from a series of six consecutive treatments [17] shows a transition from large amplitude oscillatory activity with apparent phase shifts, to low amplitude oscillations with fewer phase shifts. Comparing the model results of Fig. 7.9 with these findings, we may assume that the ECT stimuli could form new long-distance excitatory connections, as these lead to fewer phase shifts, while long-distance inhibitory connections induce more phase shifts. These results support the

notion that ECT stimuli can induce regeneration of neurons and the formation of new connections.

7.3.2 Anesthetic-induced phase transitions

Another way of artificially inducing phase transitions in cortical neurodynamics is by using neuroactive drugs, such as certain kinds of anesthetics and anti-epileptics, which clearly can induce transitions between mental states. An important principle in the action of these drugs is the selective blocking or activation of ion channels, which will have differing effects on the neurodynamics depending on the relative selectivity and intrinsic network activity [9, 48, 80]. Likewise, up-regulation of Na and K channels will induce different activity patterns, depending on their relative densities in the cell membrane.

The permeability constants, P_{Na}^* and P_{K}^* (defined as the permeability values for fully open ion channels), depend on the density of ion channels in the cell membrane, so they will be referred to as *channel densities* here. It has been shown that different combinations of these densities cause different oscillatory behaviors in single-cell dynamics at constant stimulation [8]. There are also combinations of Na and K channel density ($P_{\text{Na}}^*/P_{\text{K}}^*$) for which there are no oscillations at all.

If the stimulus applied to a given neuron is too strong, the potential cannot drop to the resting potential, and the neuron is not able to maintain an oscillatory activity; whereas if the stimulus is too weak, the neuron cannot be driven above the oscillation threshold. Both the upper and lower limit of the stimulus interval for which a neuron oscillates depend on the $P_{\text{Na}}^*/P_{\text{K}}^*$ ratio.

By constructing computational network models of neurons with different $P_{\text{Na}}^*/P_{\text{K}}^*$ values, we investigate how the network dynamics depend on the density of ion channels at the single-neuron level, thus relating microscopic properties of single neurons to mesoscopic brain dynamics. This is based on the notion that general anesthetics function by blocking specific K channels, thus shifting the affected neurons towards a larger Na:K permeability ratio [9, 33, 47].

7.3.2.1 Neural network model with spiking neurons

In this study, we use a neural network model with spiking neurons described by Frankenhaeuser-Huxley (FH) equations [8]; these deviate slightly from the classical Hodgkin-Huxley formalism, but are more accurate for cortical neurons and better for our purpose here. In our simulations, the only free parameters for the neuronal model are the permeability values (channel densities) P_{Na}^* and P_{K}^* . Using this model, we may study the effects of changes in ion-channel composition on the network dynamics as an assumed effect of certain anesthetics.

As a global activity measure (comparable to EEG), we use the arithmetic-mean field potential. Our network model here consists of 6×6 neurons, arranged in a

square lattice and connected in an all-to-all manner. We use a distance-dependent connectivity, with the connection strength decreasing with distance as $w \sim 1/r$.

Six (out of 36) homogenously distributed network neurons are inhibitory (with periodic boundary conditions), motivated by the fact that about 20% of the neo-cortical neurons in the mammalian brain are inhibitory (as in the previous models described above).

The synaptic input enters the single neuron model [8] as an additional input current, $I_i(t)$:

$$I_i(t) = \sum_j w_{ij} \sum_{(f)} (1/\tau_s) \exp[(t - t_{\text{syn}} - t_j^{(f)})/\tau_s], \quad (t - t_{\text{syn}} - t_j^{(f)}) > 0, \quad (7.12)$$

where w_{ij} is the synaptic weight between the neurons i and j , t_{syn} (1 ms) is the synaptic delay, and τ_s is the synaptic (membrane) time-constant (30 ms). The time $t^{(f)}$ refers to the arrival of an action potential.

Thus, in a network, the state equation for a neuron, with membrane potential v and capacitance C_M , becomes a sum of various currents:

$$C_M \frac{dv_i}{dt} = I_S(t) + I_G(t) + I_i(t) - I_{\text{Na}}(v_i, m_i, h_i) - I_{\text{K}}(v_i, n_i) - I_{\text{L}}(v_i). \quad (7.13)$$

I_S is the stimulation current, I_G is Gaussian noise, I_{Na} is the initial transient current through Na channels, I_{K} is the delayed sustained current through K channels, and I_{L} is the leak current. P_{Na}^* and P_{K}^* enter in the expressions for I_{Na} and I_{K} , respectively. (For more details of the model, see Ref. [45].)

7.3.2.2 Variation of network dynamics with channel-density composition

The network dynamics depends on the subcellular densities of Na and K channels (P_{Na}^* and P_{K}^*), and on the synaptic weight factor (w) at the network level; these are the only free parameters in our analysis. All neurons have the same initial conditions, but the spatial homogeneity is broken by the random component in the input. (The stimulus I_S is for every run given a value close to the oscillation threshold in each particular case).

The network consists of inhibitory and excitatory neurons with different $P_{\text{Na}}^*/P_{\text{K}}^*$ ratios. Keeping the excitatory neurons fixed at the channel density values, $P_{\text{Na}}^*/P_{\text{K}}^* = 15/7.5$, we vary the K-channel density in the inhibitory neurons. We model the effect of anesthetic by assuming that it blocks specific K channels, primarily in the inhibitory neurons. The arithmetic mean of the transmembrane potential, taken over all neurons, is used as a measure of the collective network dynamics (the ‘‘EEG’’).

The strength of the stimulus required to make a single neuron oscillate varies depending on the P_{Na}^* and P_{K}^* values for that neuron. There is a general trend that oscillation frequency increases with stimulus, and that neurons with low P_{K}^* values have a low oscillation threshold, but are also more sensitive to over-stimulation than neurons with high P_{K}^* values. (Here, we want to study the effect that these

findings have at a network level, where the stimulus varies over time due to synaptic interactions).

Since K channels are important regulators of firing patterns, and since K channels have been suggested to be the main targets for general anesthetics and anti-epileptics [33, 47], we explored the neurodynamical effects of reducing the K-channel density, in particular for the inhibitory neurons. In order to limit the number of simulations, we keep P_{Na}^* constant (at 15), varying only P_K^* . Fig. 7.10 shows a time series for a network, where the excitatory neurons have a low K-channel density ($P_K^* = 3.0$). The inhibitory neurons initially have a high density of K channels ($P_K^* = 12.5$), but the K channels were blocked in steps every 1000 ms, by decreasing P_K^* and shifting the inhibitory neurons from $P_K^* = 12.5$, to $P_K^* = 7.5$, and finally to $P_K^* = 3.0$. When the inhibitory neurons (middle trace) reach $P_K^* = 3.0$, both inhibitory and excitatory neurons alternate between periods of high amplitude activity, and periods with over-stimulation and potential drop. The mean network dynamics (bottom trace) is shifted towards a qualitatively different dynamical pattern. In this case, it is clear that the blocking of K channels in inhibitory neurons transforms unsynchronized, high-frequency oscillatory activity to an enveloped and steady slow-wave oscillation, qualitatively mimicking the transformation of EEG-patterns when applying general anesthetics [55].

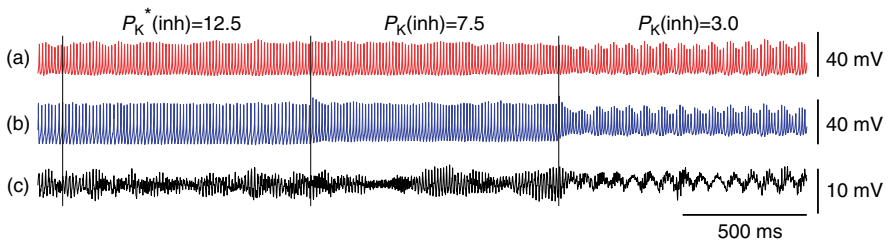


Fig. 7.10 [Color plate] Model response to stepped reductions in K-channel density in inhibitory neurons. For excitatory neurons, the densities of Na and K channels is kept fixed at the constant ratio $P_{Na}^*/P_K^* = 15/3$, while for inhibitory neurons the ratio is stepped consecutively from $P_{Na}^*/P_K^* = 15/12.5$, to $15/7.5$, and finally to $15/3$, by decreasing P_K^* every 1000 ms. The two upper time-series show the activity of (a) an excitatory neuron (red trace), and (b) an inhibitory neuron (blue trace); (c) lower trace (black) shows the network mean.

These simulations show that the mesoscopic network dynamics can be shifted into, or out of, different oscillatory states by small changes in the ion-channel densities, even for single neurons. Similar effects can also be obtained by changing connection strengths in the network model, which we have shown elsewhere [45]. Both of these phenomena are of pharmacological interest, since some drugs can affect the permeability of ion channels also in the synapses [48]. Our simulations demonstrate that the blocking of specific K channels, as a possible effect of some anesthetics, can change the global activity from high-frequency (awake) states to low-frequency (anesthetized) states, as apparent in recorded and simulated EEG.

7.4 Discussion

In this chapter, I have given a few examples of how computational models can be used to study phase transitions in mesoscopic brain dynamics. As examples of internally/naturally induced phase transitions, I have presented some models with intrinsic noise, neuromodulation, and attention, which in fact, may all be related. In particular, neuromodulation seems to be closely linked to the level of arousal and attention. It may also affect the internal noise level, e.g., by varying the threshold for firing. As examples of externally/artificially induced phase transitions, I have discussed electrical stimulation—both as electric shocks applied directly onto the olfactory bulb and cortex in an experimental setting with animals, and as electroconvulsive therapy applied in a clinical situation in treatment of psychiatric disorders. The final example was a network model testing how certain anesthetics may act on the brain dynamics through selective blocking of ion channels.

In all cases, the mesoscopic scale of cortical networks has been in focus, with an emphasis on network connectivity. The objective has been to investigate how structure is related to dynamics, and how the dynamics at one scale is related to that at another. Other than in passing, we have not discussed how structure and dynamics are related to function, since this is beyond the scope of this chapter, but the general notion is that mesoscopic brain dynamics reflects mental states and processes.

Our model systems have been paleocortical structures, the olfactory cortex and hippocampus, as well as neocortical structures, exemplified by the visual cortex. These structures display a complex dynamics with prominent oscillations in certain frequency bands, often interrupted by irregular, chaotic-like activity. In many cases, it seems that the collective cortical dynamics after external stimulation results from some kind of “resonance” between network connectivity (with negative and positive feedback loops), neuronal oscillators, and external input.

While our models are often aimed at mimicking specific cortical structures and network circuitry at a mesoscopic level, in some cases there is less realism in the connectivity than in the microscopic level of single neurons. The reason for this is that the objective in those cases has been to link the neuronal spiking activity with the collective activity of inter-connected neurons, irrespective of the detailed network structure. Model simulations then need to be compared with spike trains of single neurons, as captured with microelectrodes or patch-clamp techniques. In cases where the network connectivity is in focus, the network nodes may represent large populations of neurons, and their spiking activity is represented by a collective continuous output, more related to LFP or EEG activity.

Models should always be adapted to the problem they are supposed to address, with an appropriate level of detail at the spatial and temporal scales considered. In general, it could be wise to apply Occam’s razor in the modeling process, aiming at a model as simple as possible, and with few (unspecified) parameters. For the brain, due to its great complexity and our still rather fragmented knowledge, it is particularly hard to find an appropriate level of description and to decide which details to include. For example, different models may address the problem of neural computation at different levels, from the single-neuron level [57] to cortical networks

and areas [30, 74, 79, 89]. Even though the emphasis may be put at different levels, the different models can often be regarded as complementary descriptions, rather than mutually exclusive. At this stage, it is in general not possible to say which models give the best description, for example when trying to link neural and mental processes, in particular with regard to the significance of phase transitions.

Even though attempts have been made, it is a struggle to include several levels of descriptions in a single model, relating the activity at the different levels to each other [4, 10, 30, 31, 74, 88, 89]. In fact, relating different spatial and temporal scales in the nervous system, and linking them to mental processes, can be seen as the greatest challenges to modern neuroscience.

In the present work, I have focused on how to model phase transitions in mesoscopic brain dynamics, relating the presentation to anatomical and physiological properties, and I have not so much discussed the functional significance of such transitions, which has been done more thoroughly elsewhere [41, 45, 60, 63, 64]. Below, I will just briefly discuss some of these ideas.

The main question concerns the functional significance of the complex cortical neurodynamics described and simulated above, and in particular, the significance of the phase transitions between various oscillatory states and chaotic or noisy states. The electrical activity of the brain, as captured with EEG, is considered by many to be an epiphenomenon, without any information content or functional significance, but this view is challenged by the bulk of research presented, referenced, and discussed here.

Our computer simulations support the view that the complex dynamics makes the neural information processing more efficient, providing a fast and accurate response to external situations. For example, with an initial chaotic-like state, sensitive to small variations in the input signal, the system can rapidly converge to a limit-cycle attractor memory state [61, 62, 90]. Perhaps the most direct effect of cortical oscillations could be to enhance weak signals and speed up information processing, but it may also reflect collective, synchronous activity associated with various cognitive functions, including segmentation of sensory input, learning, perception, and attention.

In addition, a “recruitment” of neurons in oscillatory activity can eliminate the negative effects of noise in the input, by cancelling out the fluctuations of individual neurons. However, noise can also have a positive effect on system performance, as will be discussed briefly below. Finally, from an energy point of view, oscillations in the neuronal activity should be more efficient than if a static neuronal output (from large populations of neurons) was required.

The intrinsic noise found in all neural systems seems inevitable, but it may also have a functional role, being advantageous to the system. What, then, could be the functional role of the microscopic noise on the meso- and macroscopic dynamics? What, if any, could be the role of spontaneous activity in the brain? A traditional answer is that it generates baseline activity necessary for neural survival, and that it perhaps also brings the system closer to threshold for transitions between different neurodynamical states. It has also been suggested that spontaneous activity shapes synaptic plasticity during ontogeny (see references in Ref. [54]), and it has even

been argued that spontaneous activity plays a role for conscious processes [7, 11, 12, 70].

Internal system-generated fluctuations can apparently create state transitions, break down one kind of order to make place for and replacing it with a new kind of order. Externally-generated fluctuations can cause increased sensitivity in certain (receptor) cells through the phenomenon of *stochastic resonance* (SR) [2, 10, 20, 66, 69, 71]. The typical example of this is when a signal with the addition of noise overcomes a threshold, which results in an increased signal-to-noise relation.

The computer simulations we have described above demonstrate that “microscopic” noise can indeed induce global synchronous oscillations in cortical networks and shift the system dynamics from one dynamical state to another. This in turn can change the efficiency in the information processing of the system. Thus, in addition to the (pseudo-)chaotic network dynamics, the noise produced by a few (or many) neurons, could make the system more flexible, increasing the responsiveness of the system and avoiding getting stuck in any undesired oscillatory mode. In particular, we have shown that spontaneous activity can facilitate learning and associative memory. Indeed, simulations with our paleocortical model demonstrated that an increased neuronal noise level can reduce recall time in associative memory tasks, i.e., the time it takes for the system to recognize a distorted input pattern as any of the stored patterns. Consonant with SR theory [2, 20, 71], we found optimal noise values for which the recall-time reached a minimum [61, 62, 69].

In addition, our simulations also show that neuromodulatory control can be used in regulating the accuracy or rate of the recognition process, depending on current demands. Apparently, the complex dynamics of the brain can be regulated by neuromodulators, and perhaps also by noise. By such control, the neural system could be put into an appropriate state for the right response-action dependent on the environmental demand. Operating with a complex neurodynamics, shifting between various oscillatory and (pseudo-)chaotic states, the brain seems to balance between stability and flexibility, increasing performance efficiency and survival probability for the individual.

The kind of phase transitions discussed in this work may reflect transitions between different cognitive and mental levels or states, for example corresponding to various stages of sleep, anesthesia, or wake states with different levels of arousal, which in turn may affect the efficiency and rate of information processing. In some of our previous work, we have also added gap junctions to the ordinary synaptic connections in our paleocortical model, causing rapid synchronization of the network dynamics, and thus further improving neural information processing in associative memory tasks [13, 14].

Even though we are still at an early stage, I believe a combination of computational analysis and modeling methods of the kind discussed here can serve as an essential complement to experimental and clinical methods in furthering our understanding of neural and mental processes. In particular, when concerned with the inter-relation between structure, dynamics and function of the brain and its cognitive functions, this method may be the best way to make progress. The study of phase

transitions in the brain dynamics seems to be one of the most fruitful approaches in this respect.

Acknowledgments I would like to thank my co-workers, Peter Århem, Per Aronsson, Soumalee Basu, Yuqiao Gu, Geir Halmes, Björn Wahlund, and Xiangbao Wu. I also appreciate valuable discussions with Hans Braun, Walter Freeman, Hermann Haken, and Frank Moss. Grants from Vinova and the Swedish Research Council are gratefully acknowledged.

References

- Altar, C.A., Laeng, P., Jurata, L.W., Brockman, J.A., Lemire, A., Bullard, J., Bukhman, Y.V., Young, T.A., Charles, V., Palfreyman, M.G.: Electroconvulsive seizures regulate gene expression of distinct neurotrophic signaling pathways. *J. Neurosci.* **24**, 2667–2677 (2004), doi:10.1523/jneurosci.5377-03.2004
- Anishchenko, V.S., Neiman, A.B., Safanova, M.A.: Stochastic resonance in chaotic systems. *J. Stat. Phys.* **70**, 183–196 (1993), doi:10.1007/bf01053962
- Arbib, M.A. (ed.): *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, Mass. (1995)
- Arbib, M.A., Érdi, P., Szentágothai, J.: *Neural Organization Structure, Function and Dynamics*. MIT Press, Cambridge, Mass. (1998)
- Århem, P., Blomberg, C., Liljenström, H. (eds.): *Disorder Versus Order in Brain Function*. World Scientific, London (2000)
- Århem, P., Braun, H., Huber, M., Liljenström, H.: Nonlinear state transitions in neural systems: From ion channels to networks. In: H. Liljenström, U. Svedin (eds.), *Micro - Meso - Macro: Addressing Complex Systems Couplings*, pp. 37–72, World Scientific, London (2005)
- Århem, P., Johansson, S.: Spontaneous signalling in small central neurons: Mechanisms and roles of spike-amplitude and spike-interval fluctuations. *Int. J. Neural Syst.* **7**, 369–376 (1996), doi:10.1142/s0129065796000336
- Århem, P., Klement, G., Blomberg, C.: Channel density regulation of firing patterns in a cortical neuron model. *Biophys. J.* **90**, 4392–4404 (2006)
- Århem, P., Klement, G., Nilsson, J.: Mechanisms of anesthesia: Towards integrating network, cellular and molecular modeling. *Neuropsychopharmacology* **28**, S40–S47 (2003), doi:10.1038/sj.npp.1300142
- Århem, P., Liljenström, H.: Fluctuations in neural systems: From subcellular to network levels. In: F. Moss, S. Gielen (eds.), *Handbook of Biological Physics*, vol. 4, pp. 83–129, Elsevier, Amsterdam (2001)
- Århem, P., Liljenström, H.: Beyond cognition - on consciousness transitions. In: H. Liljenström, P. Århem (eds.), *Consciousness Transitions - Phylogenetic, Ontogenetic and Physiological Aspects*, pp. 1–25, Elsevier, Amsterdam (2007)
- Århem, P., Lindahl, B.I.B.: Neuroscience and the problem of consciousness: Theoretical and empirical approaches - an introduction. *Theor. Med.* **14**, 77–88 (1993), doi:10.1007/bf00997268
- Aronsson, P., Liljenström, H.: Non-synaptic modulation of cortical network dynamics. *Neurocomputing* **32-33**, 285–290 (2000), doi:10.1016/s0925-2312(00)00176-4
- Aronsson, P., Liljenström, H.: Effects of non-synaptic neuronal interaction in cortex on synchronization and learning. *Biosystems* **63**, 43–56 (2001), doi:10.1016/s0303-2647(01)00146-0
- Basu, S., Liljenström, H.: Spontaneously active cells induce state transitions in a model of olfactory cortex. *Biosystems* **63**, 57–69 (2001)

16. Berger, H.: Über das elektroencephalogramm des menschen. *Arch. Psychiatr. Nervenkrankh.* **87**, 527–570 (1929)
17. Beyer, J.L., Weiner, R.D., Glenn, M.D.: *Electroconvulsive Therapy*. American Psychiatric Press, London (1998)
18. Biedenbach, M.A.: Effects of anesthetics and cholinergic drugs on prepyriform electrical activity in cats. *Exp. Neurol.* **16**, 464–479 (1966), doi:10.1016/0014-4886(66)90110-5
19. Börgers, C., Epstein, S., Kopell, N.J.: Background gamma rhythmicity and attention in cortical local circuits: A computational study. *Proc. Natl. Acad. Sci. USA* **102**(19), 7002–7007 (2005), doi:10.1073/pnas.0502366102
20. Bulsara, A., Jacobs, E.W., Zhou, T., Moss, F., Kiss, L.: Stochastic resonance in a single neuron model: Theory and analog simulation. *J. Theor. Biol.* **152**, 531–555 (1991), doi:10.1016/s0022-5193(05)80396-0
21. Cobb, S.R., Buhl, E.H., Halasy, K., Paulsen, O., Somogyi, P.: Synchronization of neuronal activity in hippocampus by individual gabaergic interneurons. *Nature* **378**, 75–78 (1995), doi:10.1038/378075a0
22. Corchs, S., Deco, G.: Large-scale neural model for visual attention: Integration of experimental single-cell and fmri data. *Cerebr. Cortex* **12**, 339–348 (2002)
23. Crick, F., Koch, C.: Towards a neurobiological theory of consciousness. *Semin. Neurosci.* **2**, 263–275 (1990)
24. Eckhorn, R., Bauer, R., Jordon, W., Brosch, M., Kruse, W., Monk, M., Reitboeck, H.J.: Coherent oscillations: A mechanism of feature linking in the in the visual cortex? *Biol. Cybern.* **60**, 121–130 (1988), doi:10.1007/bf00202899
25. FitzHugh, R.: Mathematical models of threshold phenomena in the nerve membrane. *Bull. Math. Biophys.* **17**, 257–278 (1955), doi:10.1007/bf02477753
26. Freeman, W.J.: Distribution in time and space of prepyriform electrical activity. *J. Neurophysiol.* **22**, 644–665 (1959)
27. Freeman, W.J.: Linear models of impulse inputs and linear basis functions for measuring impulse responses. *Exp. Neurol.* **10**, 475–492 (1964), doi:10.1016/0014-4886(64)90046-9
28. Freeman, W.J.: Nonlinear gain mediating cortical stimulus-response relations. *Biol. Cybern.* **33**, 237–247 (1979), doi:10.1007/bf00337412
29. Freeman, W.J.: *Societies of Brains - A Study in the Neuroscience of Love and Hate*. Lawrence Erlbaum, Hillsdale, NJ (1995)
30. Freeman, W.J.: *Neurodynamics: An Exploration in Mesoscopic Brain Dynamics*. Springer, Berlin (2000)
31. Freeman, W.J.: The necessity for mesoscopic organization to connect neural function to brain function. In: H. Liljenström, U. Svedin (eds.), *Micro - Meso - Macro: Addressing Complex Systems Couplings*, pp. 25–36, World Scientific, London (2005)
32. Freeman, W.: *Mass Action in the Nervous System*. Academic Press, New York (1975)
33. Friedrich, P., Urban, B.W.: Interaction of intravenous anesthetics with human neuronal potassium currents in relation to clinical concentrations. *Anesthesiology* **91**, 1853–1860 (1999)
34. Fries, P., Reynolds, J.H., Rorie, A.E., Desimone, R.: Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* **291**, 1560–1563 (2001), doi:10.1126/science.1055465
35. Giannakopoulos, F., Bihler, U., Hauptmann, C., Luhmann, H.: Epileptiform activity in a neo-cortical network: a mathematical model. *Biol. Cybern.* **85**, 257–268 (2001), doi:10.1007/s004220100257
36. Gordon, E. (ed.): *Integrative Neuroscience: Bringing Together Biological Psychological and Clinical Models of the Human Brain*. Harwood Academic Press, New York (2000)
37. Gray, C.M., Konig, P., Engel, A.K., Singer, W.: Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* **338**, 334–337 (1989), doi:10.1038/338334a0
38. Gray, C.M., Singer, W.: Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA* **86**, 1698–1702 (1989)

39. Gu, Y., Haldnes, G., Liljenström, H., von Rosen, D., Wahlund, B., Liang, H.: Modelling ect effects by connectivity changes in cortical neural networks. *Neurocomputing* **69**, 1341–1347 (2006), doi:10.1016/j.neucom.2005.12.104
40. Gu, Y., Haldnes, G., Liljenström, H., Wahlund, B.: A cortical network model for clinical EEG data analysis. *Neurocomputing* **58–60**, 1187–1196 (2004), doi:10.1016/j.neucom.2004.01.184
41. Gu, Y., Liljenström, H.: A neural network model of attention-modulated neurodynamics. *Cognitive Neurodynamics* **1**, 275–285 (2007), doi:10.1007/s11571-007-9028-7
42. Gu, Y., Wahlund, B., Liljenström, H., von Rosen, D., Liang, H.: Analysis of phase shifts in clinical EEG evoked by ect. *Neurocomputing* **65**, 475–483 (2005), doi:10.1016/j.neucom.2004.11.004
43. Haken, H.: *Synergetics: An Introduction*. Springer-Verlag, Berlin (1983)
44. Haken, H.: *Principles of Brain Functioning*. Springer, Berlin (1996)
45. Haldnes, G., Liljenström, H., Århem, P.: Density dependent neurodynamics. *Biosystems* **89**, 126–134 (2007), doi:10.1016/j.biosystems.2006.06.010
46. Hamker, F.H.: A dynamic model of how feature cues guide spatial attention. *Vision Res.* **44**, 501–521 (2004), doi:10.1016/j.visres.2003.09.033
47. Harris, T., Shahidullah, M., Ellingson, J., Covarrubias, M.: General anesthetic action at an internal protein site involving the S4-S5 cytoplasmic loop of a neuronal K⁺ channel. *J. Biol. Chem.* **275**, 4928–4936 (2000)
48. Hille, B.: *Ion Channels of Excitable Membranes*. Sinauer, Sunderland, Mass., 3rd edn. (2001)
49. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**, 500–544 (1952)
50. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558 (1982)
51. Hopfield, J.J.: Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. USA* **81**, 3088–3092 (1984)
52. Huber, M.T., Braun, H.A., Krieg, J.C.: Consequences of deterministic and random dynamics for the course of affective disorders. *Biol. Psychiatr.* **46**, 256–262 (1999), doi:10.1016/s0006-3223(98)00311-4
53. Huber, M.T., Braun, H.A., Krieg, J.C.: Effects of noise on different disease states of recurrent affective disorders. *Biol. Psychiatr.* **47**, 634–642 (2000), doi:10.1016/s0006-3223(99)00174-2
54. Johansson, S., Århem, P.: Single-channel currents trigger action potentials in small cultured hippocampal neurons. *Proc. Natl. Acad. Sci. USA* **91**, 1761–1765 (1994)
55. John, E.R., Prichep, L.S.: The anesthetic cascade: A theory of how anesthesia suppresses consciousness. *Anesthesiology* **102**, 447–471 (2005)
56. Kelso, S.: Fluctuations in the coordination dynamics of brain and behaviour. In: P. Århem, C. Blomberg, H. Liljenström (eds.), *Disorder versus Order in Brain Function*, pp. 185–204, World Scientific, London (2000)
57. Koch, C.: *Biophysics of Computation: Information Processing in Single Neurons*. Oxford University Press, New York (1999)
58. Korchounov, A., Ilic, T., Schwinge, T., Ziemann, U.: Modification of motor cortical excitability by an acetylcholinesterase inhibitor. *Exp. Brain Res.* **164**, 399–405 (2005), doi:10.1007/s00221-005-2326-6
59. Kuczewski, N., Aztiria, E., Gautam, D., Wess, J., Domenici, L.: Acetylcholine modulates cortical synaptic transmission via different muscarinic receptors, as studied with receptor knock-out mice. *J. Physiol.* **566**(3), 907–919 (2005), doi:10.1113/jphysiol.2005.089987
60. Liljenström, H.: Modeling the dynamics of olfactory cortex using simplified network units and realistic architecture. *Int. J. Neural Syst.* **2**, 1–15 (1991), doi:10.1142/S0129065791000029
61. Liljenström, H.: Autonomous learning with complex dynamics. *Int. J. Intell. Syst.* **10**, 119–153 (1995), doi:10.1002/int.4550100109
62. Liljenström, H.: Global effects of fluctuations in neural information processing. *Int. J. Neural Syst.* **7**, 497–505 (1996), doi:10.1142/S0129065796000488

63. Liljenström, H.: Cognition and the efficiency of neural processes. In: P. Århem, H. Liljenström, U. Svedin (eds.), *Matter Matters? On the Material Basis of the Cognitive Aspects of Mind*, pp. 177–213, Springer, Heidelberg (1997)
64. Liljenström, H.: Neural stability and flexibility - a computational approach. *Neuropsychopharmacology* **28**, S64–S73 (2003), doi:10.1038/sj.npp.1300137
65. Liljenström, H., Århem, P.: Investigating amplifying and controlling mechanisms for random events in neural systems. In: J.M. Bower (ed.), *Computational Neuroscience*, pp. 711–716, Plenum Press, New York (1997)
66. Liljenström, H., Halmes, G.: Noise in neural networks— in terms of relations. *Fluct. Noise Lett.* **4**(1), L97–L106 (2004), doi:10.1142/S0219477504001707
67. Liljenström, H., Hasselmo, M.E.: Cholinergic modulation of cortical oscillatory dynamics. *J. Neurophysiol.* **74**, 288–297 (1995)
68. Liljenström, H., Svedin, U. (eds.): *Micro-Meso-Macro: Addressing Complex Systems Couplings*. World Scientific, London (2005)
69. Liljenström, H., Wu, X.: Noise-enhanced performance in a cortical associative memory model. *Int. J. Neural Systems* **6**, 19–29 (1995), doi:10.1142/S0129065795000032
70. Lindahl, B.I.B., Århem, P.: Mind as a force field: Comments on a new interactionistic hypothesis. *J. Theor. Biol.* **171**, 111–122 (1994), doi:10.1006/jtbi.1994.1217
71. Mandell, A., Selz, K.: Brain stem neuronal noise and neocortical resonance. *J. Stat. Phys.* **70**, 355–373 (1993), doi:10.1007/bf01053973
72. McAdams, C., Maunsell, J.: Effects of attention on orientation-tuning functions of single neurons in macaque cortical area v4. *J. Neurosci.* **19**, 431–441 (1999)
73. Moss, F., Gielen, S. (eds.): *Neuro-Informatics and Neural Modelling*, vol. 4 of *Handbook of Biological Physics*. Elsevier, Amsterdam (2001)
74. Robinson, P.A., Rennie, C.J., Rowe, D.L., O'Connor, S.C., Wright, J.J., Gordon, E.: Neurophysical modeling of brain dynamics. *Neuropsychopharmacology* **28**, S74–S79 (2003), doi:10.1038/sj.npp.1300143
75. Shepherd, G.M.: *The Synaptic Organization of the Brain*. Oxford University Press, Oxford (1998)
76. Siegel, M., Kording, K., König, P.: Integrating top-down and bottom-up sensory processing by somato-dendritic interactions. *J. Comput. Neurosci.* **8**, 161–173 (2000), doi:10.1023/a:1008973215925
77. Sirosh, J., Miikkulainen, R.: Self-organizing feature maps with lateral connections: Modeling ocular dominance. In: M.C. Mozer, P. Smolensky, D.S. Touretzky, J.L. Elman, A.S. Weigend (eds.), *Proceedings of the 1993 Connectionist Models Summer School*, pp. 31–38, CMSS-93, Boulder, Colorado (1994)
78. Skarda, C.A., Freeman, W.J.: How brains make chaos in order to make sense of the world. *Behav. Brain Sci.* **10**, 161–195 (1987)
79. Steyn-Ross, D.A., Steyn-Ross, M.L., Sleigh, J.W., Wilson, M.T., Gillies, I.P., Wright, J.J.: The sleep cycle modelled as a cortical phase transition. *J. Biol. Phys.* **31**, 547–569 (2005), doi:10.1007/s10867-005-1285-2
80. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W.: Modelling general anaesthesia as a first-order phase transition in the cortex. *Progr. Biophys. Mol. Biol.* **85**, 369–385 (2004), doi:10.1016/j.pbiomolbio.2004.02.001
81. Szentagothai, J.: Local neuron circuits of the neocortex. In: F. Schmitt, F. Worden (eds.), *The Neurosciences 4th Study Program*, pp. 399–415, MIT Press, Cambridge, Mass. (1979)
82. Tass, P.A.: Desynchronizing double-pulse phase resetting and application to deep brain stimulation. *Biol. Cybern.* **85**(5), 343–354 (2001), doi:10.1007/s004220100268
83. Tharyan, P., Adams, C.E.: Electroconvulsive therapy for schizophrenia. *Cochrane Db. Syst. Rev.* **2**, CD000076 (2005), doi:10.1002/14651858.cd000076
84. von Stein, A., Chiang, C., König, P.: Top-down processing mediated by interareal synchronization. *Proc. Natl. Acad. Sci. USA* **97**(26), 14748–14753 (2000)

85. Wahlund, B., Piazza, P., von Rosen, D., Liberg, B., Liljenström, H.: Seizure (ictal)-EEG characteristics subgroup depressive disorder in patients receiving ECT– A preliminary study and multivariate approach. *Comput. Intell. Neurosci.* (2009), [In press]
86. Wahlund, B., von Rosen, D.: ECT of major depressed patients in relation to biological and clinical variables: A brief overview. *Neuropsychopharmacology* **28**, S21–S26 (2003), doi:10.1038/sj.npp.1300135
87. Wright, J.J., Bourke, P.D., Chapman, C.L.: Synchronous oscillation in the cerebral cortex and object coherence: Simulation of basic electrophysiological findings. *Biol. Cybern.* **83**, 341–353 (2000), doi:10.1007/s004220000155
88. Wright, J.J., Liley, D.T.J.: Dynamics of the brain at global and microscopic scales. *Neural networks and the EEG. Behav. Brain Sci.* **19**, 285–320 (1996)
89. Wright, J.J., Rennie, C.J., Lees, G.J., Robinson, P.A., Bourke, P.D., Chapman, C.L., Gordon, E., Rowe, D.L.: Simulated electrocortical activity at microscopic, mesoscopic, and global scales. *Neuropsychopharmacology* **28**, S80–S93 (2003), doi:10.1038/sj.npp.1300138
90. Wu, X., Liljenström, H.: Regulating the nonlinear dynamics of olfactory cortex. *Netw. Comput. Neural Syst.* **5**, 47–60 (1994), doi:10.1088/0954-898x/5/1/003

Chapter 8

Phase transitions in physiologically-based multiscale mean-field brain models

P.A. Robinson, C.J. Rennie, A.J.K. Phillips, J.W. Kim, and J.A. Roberts

8.1 Introduction

Brain dynamics involves interactions across many scales—spatially from microscopic to whole-brain, and temporally from the sub-millisecond range to seconds, or even years. Except under artificial conditions that isolate a single scale, these multiscale aspects of the underlying physiology and anatomy must be included to model the behavior adequately at any scale. In particular, microscale behavior must be included to understand large-scale phase transitions, because the theory of critical phenomena implies that their properties are strongly constrained by the symmetries and conservation properties of the system's microscopic constituents [2].

In condensed matter physics, where they are most familiar, phase transitions arise at the macroscale in systems of atoms, molecules, nuclear spins, or other microscopic constituents. Phase transitions are intrinsically collective properties that are typically analyzed in the thermodynamic limit of infinitely many constituents. They become apparent through discontinuous changes in large-scale order parameters or their derivatives. Examples include the sudden change in density at vaporization or

Peter A. Robinson

School of Physics, University of Sydney, NSW 2006, Australia; Brain Dynamics Centre, Westmead Millennium Institute and Westmead Hospital, Westmead, NSW 2145, Australia; Faculty of Medicine, University of Sydney, NSW 2006, Australia.

e-mail: robinson@physics.usyd.edu.au

Christopher J. Rennie

School of Physics, University of Sydney, NSW 2006, Australia; Department of Medical Physics, Westmead Hospital, Westmead, NSW 2145, Australia; Brain Dynamics Centre, Westmead Millennium Institute and Westmead Hospital, Westmead, NSW 2145, Australia.

e-mail: chris_rennie@wmi.usyd.edu.au

Andrew J. K. Phillips · Jong W. Kim · James A. Roberts

School of Physics, University of Sydney, NSW 2006, Australia; Brain Dynamics Centre, Westmead Millennium Institute and Westmead Hospital, Westmead, NSW 2145, Australia.

e-mail: ajp@physics.usyd.edu.au jwkim@physics.usyd.edu.au jamesr@physics.usyd.edu.au

freezing, with an associated nonzero specific heat, or the divergent magnetic susceptibility (derivative of magnetization) at the transition of iron from its nonmagnetic state to a ferromagnetic one with falling temperature, where there is no associated specific heat. Such transitions are termed first-order and second-order, respectively, and occur at *critical points* at which thermodynamic variables such as temperature, pressure, etc., take on highly specific values [2].

Other systems with phase transitions have been identified. Notable are self-organized critical (SOC) systems which self-organize to the critical point, rather than requiring external tuning parameters to be set independently as in thermodynamic transitions. One example is an idealized sandpile growing by the continuous addition of grain, whose slope (the tuning parameter) adjusts itself automatically to very near the critical value at which avalanches commence. The critical slope is then maintained by balance between addition of grains and their loss via avalanches. Criticality in plasma systems has also been shown to be closely associated with microinstabilities of the system that lead to macroscopic changes in the system state [10, 11]. Phase transitions are accompanied by divergent correlation lengths of fluctuations, $1/f$ power-law spectra, and power-law probability distributions of fluctuation amplitudes, the latter two effects implying that critical states inherently involve a wide range of scales in their dynamics.

Mean-field theories provide a natural basis for modeling and analyzing phase transitions in neural systems. Moreover, links to measurements become easy to include—an essential point because most measurement processes aggregate over many neurons and all modify signals in some way. Mean-field theories that incorporate the measurement function are a natural bridge between theoretical and experimental results.

In the class of models described here, averages are taken over microscopic neural structure to obtain mean-field descriptions on scales from tenths of a millimeter up to the whole brain, incorporating representations of the anatomy and physiology of separate excitatory and inhibitory neural populations, nonlinear neural responses, multiscale interconnections, synaptic, dendritic, cell-body, and axonal dynamics, and corticothalamic feedback [4, 7, 12, 14, 16, 17, 24–27, 30–38, 42, 48, 50]. These models readily include measurement effects such as the volume conduction that acts to spatially smooth EEG signals, and the hemodynamic response that temporally filters the BOLD signal that underlies functional MRI.

Essential features of any realistic neurodynamic model are that it: (i) be based on physiology and anatomy, including the salient features at many spatial and temporal scales; (ii) be quantitative with predictions that can be calculated analytically or numerically, including measurement effects; (iii) have parameters that directly relate to physiology and anatomy, and that can be measured, or at least constrained in value, via independent experiments (this does not exclude the theory itself enabling improved estimates of parameters); (iv) be applicable to multiple phenomena and data types, rather than being a theory of a single phenomenon or experimental modality; and (v) be invertible, if possible, allowing parameters to be deduced by fitting model predictions to data (the parameters obtained must be consistent with independent measurements). These criteria rule out (among others) highly idealized

models of abstract neurons, as are sometimes used in computer science, theories of single phenomena, or models with parameters highly tailored to single phenomena, models with completely free parameters, and models that take no account of measurement effects.

We have developed a physiologically based mean-field model of brain dynamics that satisfies the above criteria. When applied to the corticothalamic system, it reproduces and unifies many features of EEGs, including background spectra and the spectral peaks seen in waking and sleeping states [32, 34, 37], evoked response potentials [25], measures of coherence and spatiotemporal structure [19, 20, 26, 27], and generalized epilepsies and low-dimensional seizure dynamics [4, 31]. Our approach averages over microstructure to yield mean-field equations in a way that complements cellular-level and neural-network analyses.

In Sect. 8.2 we outline our model, including its physiological and anatomical foundations, basic predictions, and its connection to measurements. In Sects 8.3 and 8.4 we then discuss a range of predictions that relate to neural phase transitions in several regimes, and compare them with experimental data on normal arousal states, epilepsies, and sleep dynamics. Section 8.5 summarizes and discusses the material. We also take the opportunity (in Sects 8.2.1 and 8.3.1) to address a number of fallacies surrounding mean-field theory and its applications, and to highlight open questions (in Sect. 8.5).

8.2 Mean-field theory

In this section we briefly review our model and its connections with measurable quantities. More detailed discussion and further generalizations can be found elsewhere [24, 25, 28, 29, 37].

8.2.1 Mean-field modeling

The brain contains multiple populations of neurons, which we distinguish by a subscript a that designates both the structure in which a given population lies (e.g., a particular nucleus) and the type of neuron (e.g., interneuron, pyramidal cell). We average their properties over scales of ~ 0.1 mm and seek equations for the resulting mean-field quantities.

The perturbation $V_a(\mathbf{r}, t)$ to the mean soma potential is approximated as the sum of contributions $V_{ab}(\mathbf{r}, t)$ arriving as a result of activity at each type of (mainly) dendritic synapse b , where b denotes both the population and neurotransmitter type, \mathbf{r} denotes the spatial location, and t the time. This gives

$$V_a(\mathbf{r}, t) = \sum_b V_{ab}(\mathbf{r}, t). \quad (8.1)$$

The potential V_{ab} is generated when synaptic inputs from afferent neurons are temporally low-pass filtered and smeared out in time as a result of receptor dynamics and passage through the dendritic tree (i.e., by dynamics of ion channels, membranes, etc.). It approximately obeys a differential equation [28, 32, 34, 37]

$$D_{ab}V_{ab}(\mathbf{r}, t) = N_{ab}s_{ab}\phi_b(\mathbf{r}, t - \tau_{ab}), \quad (8.2)$$

$$D_{ab} = \frac{1}{\alpha_{ab}\beta_{ab}} \frac{d^2}{dt^2} + \left(\frac{1}{\alpha_{ab}} + \frac{1}{\beta_{ab}} \right) \frac{d}{dt} + 1, \quad (8.3)$$

where $1/\beta_{ab}$ and $1/\alpha_{ab}$ are the rise and decay times of the cell-body potential produced by impulse at a dendritic synapse. The right of Eq. (8.2) describes the influence of the firing rates ϕ_b from neuronal populations b , in general delayed by a time τ_{ab} due to discrete anatomical separations between different structures. The quantity N_{ab} is the mean number of synapses from neurons of type b to type a , and s_{ab} is the time-integrated strength of the response in neurons of type a to a unit signal from neurons of type b , implicitly weighted by the neurotransmitter release probability. Note that we ignore the dynamics of s_{ab} , which can be driven by neuromodulators, firing rate, and other effects; however, such dynamics can be incorporated straightforwardly [5]. An alternative representation of the dynamics in Eq. (8.2) is as a convolution in which

$$V_{ab}(\mathbf{r}, t) = \int_{-\infty}^t L_{ab}(t - t') N_{ab}s_{ab} \phi_b(\mathbf{r}, t - t' - \tau_{ab}) dt', \quad (8.4)$$

$$L_{ab}(u) = \frac{\alpha_{ab}\beta_{ab}}{\beta_{ab} - \alpha_{ab}} (e^{-\alpha_{ab}u} - e^{-\beta_{ab}u}). \quad (8.5)$$

Equation (8.4) is a good approximation to the soma response to a spike input at the dendrites.

In cells with voltage-gated ion channels, action potentials are produced at the axonal hillock when the soma potential exceeds a threshold. In effect, V_a acts as a control variable for the fast spike dynamics, taking the place of the applied current (apart from a capacitive proportionality) characteristic of single-neuron experiments. Spikes in most cortical cells arise via a saddle-node bifurcation in a set of Hodgkin-Huxley-like equations for ionic currents [49]. As such, spikes are produced only for V_a above an individual threshold $\tilde{\theta}_a$, at a mean rate

$$Q_a \propto (V_a - \tilde{\theta}_a)^{1/2}, \quad (8.6)$$

for low Q_a [47], leveling off due to saturation effects at higher V_a [49]. Individual cells differ slightly from the mean in the number and strength of ion channels and, hence, in $\tilde{\theta}_a$. Moreover, fluctuations in V_a affect the difference in (8.6). Hence, the dependence (8.6) must be both modified to include saturation and convolved with an approximately normal distribution of individual deviations to obtain the population-average response function

$$Q_a(\mathbf{r}, t) = S[V_a(\mathbf{r}, t)], \quad (8.7)$$

where S is a sigmoidal function that increases from 0 to Q_{\max} as V_a increases from $-\infty$ to $+\infty$ [7, 28, 37]. We use the form

$$S[V_a(\mathbf{r}, t)] = \frac{Q_{\max}}{1 + \exp\{-[V_a(\mathbf{r}, t) - \theta_a]/\sigma'\}}, \quad (8.8)$$

where we assume a common mean neural firing threshold θ relative to resting, with $\sigma'\pi/\sqrt{3}$ being its standard deviation (these quantities and Q_{\max} are assumed to be the same for all populations for simplicity). When in the linear regime, we make the approximation

$$Q_a(\mathbf{r}, t) = \rho_a V_a(\mathbf{r}, t), \quad (8.9)$$

where ρ_a is the derivative of the sigmoid at an assumed steady state of the system in the absence of perturbations (we discuss the existence and stability of such states in later sections).

Each neuronal population a within the corticothalamic system produces a field ϕ_a of pulses, that travels to other neuronal populations at a velocity v_a through axons with a characteristic range r_a . These pulses spread out and dissipate if not regenerated. To a good approximation, this type of propagation obeys a damped wave equation [12, 17, 37]:

$$\mathcal{D}_a \phi_a(\mathbf{r}, t) = S[V_a(\mathbf{r}, t)], \quad (8.10)$$

$$\mathcal{D}_a = \left(\frac{1}{\gamma_a^2} \frac{\partial^2}{\partial t^2} + \frac{2}{\gamma_a} \frac{\partial}{\partial t} + 1 - r_a^2 \nabla^2 \right) \phi_a(\mathbf{r}, t), \quad (8.11)$$

where the damping coefficient is $\gamma_a = v_a/r_a$. Equations (8.10) and (8.11) yield propagation ranges in good agreement with anatomical results [3], and with other phenomena. It is sometimes erroneously claimed that this propagation is only an approximation to propagation with delta-function delays of the form $\delta(t - |\mathbf{r}|/v_a)$, and Eq. (8.11) has even been “derived” from the latter under certain assumptions; however, in reality, both are approximations to the true physical situation in the brain.

Equations (8.1)–(8.3), (8.7), (8.8), (8.10), and (8.11) form a closed nonlinear set, which can be solved numerically, or examined analytically in various limits (see Sect. 8.3).

Once a set of specific neural populations has been chosen, and physiologically realistic values have been assigned to their parameters, these equations can be used to make predictions of neural activity. It should be noted that these equations govern spatiotemporal dynamics of firing rates, not of the individual spike dynamics. The two are tightly correlated, but the nonlinearities of our equations are weaker than those that produce the spikes themselves, at least in the sense that they only produce effects on much longer timescales than those of spikes. We stress that the oscillations predicted from our equations are collective oscillations of the rate of spiking, whose frequencies do not directly relate to the frequency of spiking itself—a common misunderstanding of mean-field models by those more familiar with spiking neurons.

8.2.2 Measurements

Once neural activity has been predicted from stimuli, one must relate it to measurements to interpret experimental results. The limited spatiotemporal resolution of such measurements often provides an additional justification for the use of mean-field modeling, since finer-scale structure is not resolvable.

In the case of EEG measurements, the effects of volume conduction on the propagation of neural potential changes to the scalp have been incorporated into our model, via attenuation and spatial filtering parameters [20, 32, 34, 38]. These are included in the bulk of the results reviewed here; space limitations preclude a detailed discussion, but their effects on spectral shape, for example, are slight at frequencies below about 20 Hz, since these correspond to the longest wavelengths. We have also shown how to include the effects of reference electrode and multielectrode derivations [8, 27]. It should also be noted that scalp potentials are primarily generated by excitatory (mainly pyramidal) neurons owing to their greater size and degree of alignment compared to other types [17–19, 25]. For any given geometry, in the linear regime at least, the scalp potential is proportional to the cortical potential, which is itself proportional to the mean cellular membrane currents, which are in turn proportional to ϕ_e . Hence, apart from a (dimensional) constant of proportionality, and the spatial low-pass filtering effects of volume conduction, scalp EEG signals correspond to ϕ_e to a good approximation in the linear domain [36].

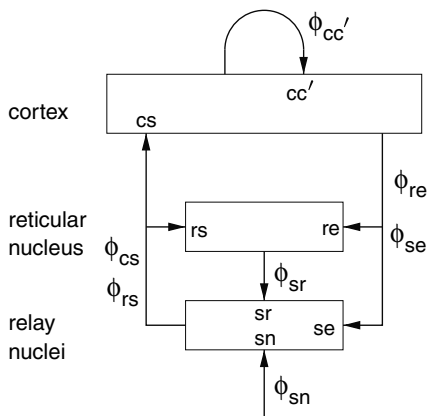
8.3 Corticothalamic mean-field modeling and phase transitions

Much work has been done on applications of mean field theory to cortical and corticothalamic systems. Here we consider the latter system since, as discussed below, inclusion of the thalamus is essential if phenomena at typical EEG frequencies are to be successfully modeled.

8.3.1 Corticothalamic connectivities

Figure 8.1 shows the large-scale structures and connectivities incorporated in the model, including the thalamic reticular nucleus r , which inhibits relay (or *specific*) nuclei s , and is lumped here with the perigeniculate nucleus, which has an analogous role [40, 43]. Relay nuclei convey external stimuli ϕ_n to the cortex, as well as passing on corticothalamic feedback. In this section we consider long-range excitatory cortical neurons ($a = e$), short-range mainly inhibitory cortical neurons ($a = i$), neurons in the reticular nucleus of the thalamus ($a = r$), neurons of thalamic relay nuclei ($a = s$), and external inputs ($a = n$) from non-corticothalamic neurons. These populations are discussed further below. Application of these methods to brainstem and hypothalamic structures is discussed in Sect. 8.4.

Fig. 8.1 Schematic of corticothalamic interactions, showing the locations at which the v_{ab} of Eq. (8.12) and linear gains G_{ab} act, where $c, c' = e, i$ denote cortical quantities.



A point that is sometimes overlooked or mistaken in the literature is that mean-field models do not need to divide the cortex into discrete pieces. In particular, there is no need to divide the cortex into hypercolumns, and this is actually likely to be a poor approximation. Indeed, this procedure as it is often implemented is highly misleading, since it imposes sharp hypercolumn boundaries where no such boundaries exist in nature [9]. This is because an anatomical hypercolumn qualitatively corresponds to the region around any given cortical neuron to which that neuron is most strongly connected. A neuron near the boundary of this hypercolumn (which is not sharp in any case) will be strongly connected to neurons on both sides of the boundary (i.e., each neuron lies at the center of its own hypercolumn). So hypercolumn boundaries are not like the walls of a honeycomb, with a fixed physical location, and theoretical approaches that discretize by laying down fixed boundaries must be viewed with some suspicion.

A related misunderstanding in the literature is the idea that short-range and long-range interactions must be treated by different means. This is often encapsulated in a division into short-range connections within hypercolumns and long-range cortico-cortical connections between hypercolumns, often treated by different mathematical methods. In fact, all connections can be handled using the same formalism, with different ranges simply incorporated via separate neural populations with different axonal range parameters (which does not preclude approximations being made when these ranges are very small) [28].

8.3.2 Corticothalamic parameters

If intracortical connectivities are proportional to the numbers of neurons involved—the *random connectivity approximation*—and $s_{ib} = s_{eb}$, $L_{ib} = L_{eb}$ for each b , then $V_i = V_e$ and $Q_i = Q_e$ [37, 50], which lets us concentrate on excitatory quantities, with inhibitory ones derivable from them. The short range of i neurons and the small size

of the thalamic nuclei enables us to assume $r_a \approx 0$ and, hence, $\gamma_a \approx \infty$ for $a = i, r, s$ for many purposes. The only nonzero discrete delays are $\tau_{es} = \tau_{se} = \tau_{re} = t_0/2$, where t_0 is the time for signals to pass from cortex to thalamus and back again. We also assume that all the synaptodendritic time constants are equal, for simplicity, and set $\alpha_{ab} = \alpha$ and $\beta_{ab} = \beta$ for all a and b in what follows; this allows us to drop the subscripts ab in Eqs (8.2), (8.3), and (8.5) and write D_α in place of D_{ab} .

Including only the connections shown in Fig. 8.1 and making the approximations mentioned above, we find that our nonlinear model has 16 parameters (and not all of these appear separately in the linear limit). By defining

$$v_{ab} = N_{ab}s_{ab}, \quad (8.12)$$

these are Q_{\max} , θ , σ' , α , β , γ_e , r_e , t_0 , v_{ee} , v_{ei} , v_{es} , v_{se} , v_{sr} , v_{sn} , v_{re} , and v_{rs} . These are sufficient in number to allow adequate representation of the most important anatomy and physiology, but few enough to yield useful interpretations and to enable reliable determination of values by fitting theoretical predictions to data. The parameters are approximately known from experiment [28, 29, 32, 34, 38] leading to the indicative values in Table 8.1. We use only values compatible with physiology. Sensitivities of the model to parameter variations have been explored in general [34] and in connection with variations between sleep, wake, and other states [31]. In the present work we concentrate on results for which the model parameters are assumed to be spatially uniform, but where the activity is free to be nonuniform; generalization to include spatial nonuniformities is straightforward [36].

Table 8.1 Indicative parameters for the alert, eyes-open state in normal adults [32]. Parameters used in some figures in this chapter are similar, but not identical.

Quantity	Nominal	Unit
Q_{\max}	340	s^{-1}
v_e	10	$m s^{-1}$
r_e	86	mm
θ	13	mV
σ'	3.8	mV
γ_e	116	s^{-1}
α	80	s^{-1}
β	500	s^{-1}
t_0	85	ms
v_{ee}	1.6	mV s
$-v_{ei}$	1.9	mV s
v_{es}	0.4	mV s
v_{se}	0.6	mV s
$-v_{sr}$	0.45	mV s
v_{sn}	0.2	mV s
v_{re}	0.15	mV s
v_{rs}	0.03	mV s
$\phi_n^{(0)}$	16	s^{-1}

An important implication of the parameters above is that the corticothalamic loop delay t_0 places any oscillations that involve this loop at frequencies of order 10 Hz. This means that inclusion of the thalamus and the dynamics of these loops is essential to understand phenomena at frequencies below ~ 20 Hz. At very low frequencies ($\ll 10$ Hz) it is sufficient to include a static corticothalamic feedback strength to the cortex, and at very high frequencies ($\gg 10$ Hz) the corticothalamic feedback is too slow to affect the dynamics strongly. As we will see in the next section, thalamic effects dominate much of the dynamics at intermediate frequencies.

8.3.3 Specific equations

The above connectivities and parameters imply, using Eqs (8.1)–(8.3),

$$D_\alpha V_e(t) = v_{ee}\phi_e(t) + v_{ei}\phi_i(t) + v_{es}\phi_s(t - t_0/2), \quad (8.13)$$

$$D_\alpha V_i(t) = v_{ie}\phi_e(t) + v_{ii}\phi_i(t) + v_{is}\phi_s(t - t_0/2), \quad (8.14)$$

$$D_\alpha V_r(t) = v_{re}\phi_e(t - t_0/2) + v_{rs}\phi_s(t), \quad (8.15)$$

$$D_\alpha V_s(t) = v_{se}\phi_e(t - t_0/2) + v_{sr}\phi_r(t) + v_{sn}\phi_n(t), \quad (8.16)$$

whence $V_i = V_e$ and $Q_i = Q_e$, as asserted above. The right-hand sides of Eqs (8.13)–(8.16) describe, for each population, the spatial summation of all afferent activity (including via self-connections), and D_α on the left describes temporal dynamics. The short ranges of the axons i , r , and s imply that the corresponding damping rates are large and that $D_\alpha \approx 1$ for these populations, further implying

$$\phi_a = Q_a = S(V_a), \quad (8.17)$$

for $a = i, r, s$. For the remaining e population, Eqs (8.10) and (8.11) yield

$$\left(\frac{1}{\gamma_e^2} \frac{\partial^2}{\partial t^2} + \frac{2}{\gamma_e} \frac{\partial}{\partial t} + 1 - r_e^2 \nabla^2 \right) \phi_e(\mathbf{r}, t) = S[V_e(\mathbf{r}, t)], \quad (8.18)$$

with $\gamma_e = v_e/r_e$. Collectively, Eqs (8.13)–(8.18) describe our corticothalamic model.

8.3.4 Steady states

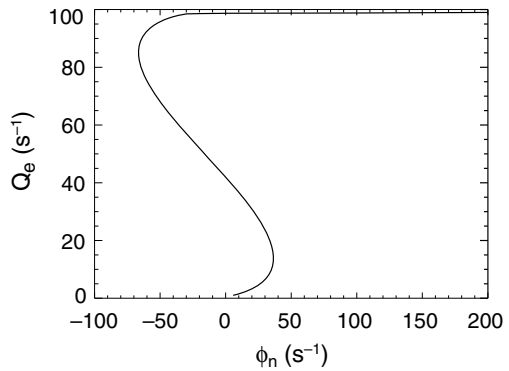
We can find spatially uniform steady states of our system by setting all the spatial and temporal derivatives to zero in Eqs (8.13)–(8.18). The resulting equations can be rearranged to yield a single equation for the steady state value of ϕ_e [32]:

$$0 = S^{-1} \left(\phi_e^{(0)} \right) - (v_{ee} + v_{ei})\phi_e^{(0)} - v_{es}S \left(v_{se}\phi_e^{(0)} + v_{sn}\phi_n^{(0)} \right) + v_{sr}S \left[v_{re}\phi_e^{(0)} + \frac{v_{rs}}{v_{es}} \left\{ S^{-1} \left(\phi_e^{(0)} \right) - (v_{ee} + v_{ei})\phi_e^{(0)} \right\} \right], \quad (8.19)$$

where S^{-1} denotes the inverse of the sigmoid function S . The function on the right of Eq. (8.19) is continuous and asymptotes to $-\infty$ as $\phi_e^{(0)} \rightarrow 0$ and to $+\infty$ as $\phi_e^{(0)} \rightarrow Q_{\max}$. Hence, it has an odd number of zeros, and thus at least one zero [32, 37]. Typically, there is either a single zero or there are three zeros, two stable separated by one unstable in the latter case. For very restricted parameter sets, five zeros (three stable and two unstable at $\omega = 0$, in alternation) are possible, and the addition of neuromodulatory feedbacks on synaptic strengths s_{ab} in Eq. (8.12) can also increase the number of zeros and broaden this parameter range [5]. We mention these generalizations further later, but restrict attention to the main case of three zeros for now.

When there are three zeros, one stable zero occurs at low $\phi_e^{(0)}$, and we identify this as the baseline activity level of normal brain function. The other stable zero is at high $\phi_e^{(0)}$ with all neurons firing near to their physiological maximum. This would thus represent some kind of seizure state, but would require further physiology (e.g., of hemodynamics and hypoxia at these high activity levels) to be treated adequately. The states are shown in Fig. 8.2, where they are linked by the unstable fixed point to form a “fold”. It should be noted that other authors have identified the pair of stable states as representing anesthesia/sleep, sleep/wake, or non-REM sleep/REM sleep, often using parameters that lower $\phi_e^{(0)}$ in the upper state to acceptable levels [44–46]. However, they do not seem to have made an overall identification of cases with branches to unify all these possibilities. As we show in Sect. 8.4, brainstem states must be taken into account in this context, so any final identification is probably premature and the above possibilities are not necessarily mutually exclusive.

Fig. 8.2 Q_e vs ϕ_n , showing the stable states with low firing rates ($< 15 \text{ Hz}^{-1}$) and with firing rates near saturation ($> 85 \text{ Hz}^{-1}$). These two branches are linked by an unstable branch to form a “fold”. Note that the negative steady state values of ϕ_n in the figure are physical, provided this variable is considered to embody inhibitory neuromodulation, as well as tonic sensory activity.



8.3.5 Transfer functions and linear waves

Small perturbations relative to steady states can be treated using linear analysis. A stimulus $\phi_n(\mathbf{k}, \omega)$ of angular frequency ω ($= 2\pi f$, where f is the usual frequency in Hz) and wave vector \mathbf{k} ($= 2\pi/\lambda$ in magnitude, where λ is the wavelength) has the transfer function to $\phi_e(\mathbf{k}, \omega)$

$$\frac{\phi_e(\mathbf{k}, \omega)}{\phi_n(\mathbf{k}, \omega)} = \frac{G_{es}L}{1 - G_{ei}L} \frac{G_{sn}Le^{i\omega t_0/2}}{1 - G_{srs}L^2} \frac{1}{q^2(\omega)r_e^2 + k^2r_e^2}, \quad (8.20)$$

$$q^2(\omega)r_e^2 = (1 - i\omega/\gamma_e)^2 - \frac{L}{1 - G_{ei}L} \left[G_{ee} + \frac{(G_{ese} + G_{esre}L)L}{1 - G_{srs}L^2} e^{i\omega t_0} \right], \quad (8.21)$$

$$G_{ab} = \frac{\phi_a^{(0)}}{\sigma'} \left(1 - \frac{\phi_a^{(0)}}{Q_{\max}} \right) v_{ab}, \quad (8.22)$$

where $L = (1 - i\omega/\alpha)^{-1}(1 - i\omega/\beta)^{-1}$ embodies the lowpass filter characteristics of synaptodendritic dynamics and $\phi_a^{(0)}$ is the steady-state value of ϕ_a . The ratio (8.20) is the cortical excitatory response per unit external stimulus, and encapsulates the relative phase via its complex value [25, 28, 34]; it is the key to linear properties of the system. The gain G_{ab} is the differential output produced by neurons a per unit change in input from neurons b , and the static gains for loops in Fig. 8.1 are $G_{ese} = G_{es}G_{se}$ for feedback via relay nuclei only, $G_{esre} = G_{es}G_{sr}G_{re}$ for the loop through reticular and relay nuclei, and $G_{srs} = G_{sr}G_{rs}$ for the intrathalamic loop.

Waves obey the dispersion relation [37]

$$q^2(\omega) + k^2 = 0, \quad (8.23)$$

which corresponds to singularity of the transfer function (8.20). Solutions of this equation satisfy $\omega = kv_e - i\gamma_e$ at high frequencies [37]. At lower frequencies, their dispersion has been investigated in detail previously [19, 24, 37].

8.3.6 Spectra

The EEG frequency spectrum is obtained by squaring the modulus of $\phi_e(\mathbf{k}, \omega)$ and integrating over \mathbf{k} . It can be written in terms of the transfer function (8.20) as

$$P_e(\omega) = \int \left| \frac{\phi_e(\mathbf{k}, \omega)}{\phi_n(\mathbf{k}, \omega)} \right|^2 |\phi_n(\mathbf{k}, \omega)|^2 d^2\mathbf{k}. \quad (8.24)$$

If we make the assumption that under conditions of spontaneous EEG the field of external stimuli $\phi_n(\mathbf{k}, \omega)$ is so complex that it can be approximated by spatiotemporal white noise, this gives $|\phi_n(\mathbf{k}, \omega)|^2 = \text{const}$. In the white noise case

$$P_e(\omega) = \frac{\langle \phi_n^2 \rangle}{4\pi r_e^4} \left| \frac{G_{esn}L^2}{(1 - G_{ei}L)(1 - G_{srs}L^2)} \right|^2 \frac{\text{Arg } q^2}{\text{Im } q^2}, \quad (8.25)$$

where $\langle \phi_n^2 \rangle$ is the mean-square noise level. Figure 8.3 shows excellent agreement of Eq. (8.25) with an observed spectrum over several decades. The features reproduced include the alpha and beta peaks 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, including the strong increase in low-frequency activity in sleep, where our model predicts a steepening of the spectrum from $1/f$ to $1/f^3$ [34]. Notably, each of the features can be related to underlying anatomy and physiology. The low-frequency $1/f$ behavior is a signature of marginally stable, near-critical dynamics, which allow complex behavior [31, 34, 37], while the steep high-frequency fall-off results from low-pass filtering by synaptodendritic dynamics. Corticothalamic loop resonances account for the alpha and beta peaks, their relative frequencies, the correlated changes in spectral peaks between sleep and waking, and splitting of the alpha peak, for example [31, 34, 36]. Suggested alternative mechanisms, including pacemakers and purely cortical resonances, can account for some features of the data, but the trend in mode frequency predicted for purely cortical eigenmodes tends to be in the opposite direction to that observed, although this is not unequivocal. Likewise, the pacemaker hypothesis is ad hoc, with a new pacemaker proposed for every spectral peak [17, 30, 36]. Overall, the evidence is now strong that the thalamus must be included to account for most salient EEG features at frequencies below about 20 Hz. The advantage of its inclusion is underlined by the ability of the resulting theory to simultaneously account for the wide range of phenomena mentioned in Sect. 8.1.

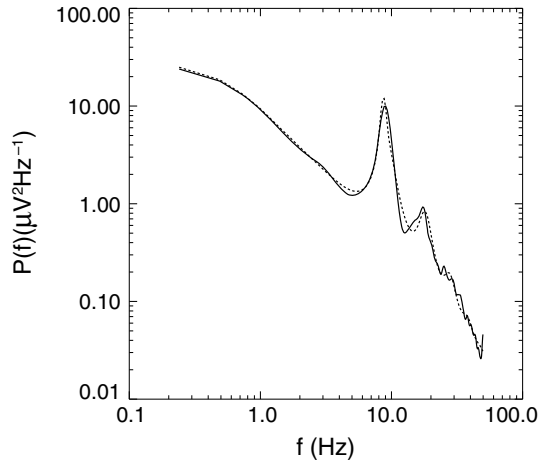


Fig. 8.3 Example spectrum (solid) and model fit (dashed) from a typical adult subject in the eyes-closed state.

One key aspect of phase transitions is the divergent correlation length near the critical point, mentioned above. Correlations and coherence can be computed using our theory. Specifically, the Wiener–Khinchine theorem implies that the correlation

function is the Fourier transform of the power spectrum, which yields long-range correlations at sharp spectral peaks, with the correlation length increasing in proportion to the quality factor of the peak [27]. This accords with these waves being weakly damped (and thus close to instability) and so able to propagate large distances at high amplitudes.

The cross spectrum $P_e(\mathbf{r}, \mathbf{r}', \omega)$ is the phase average of $\phi_e(\mathbf{r}, \omega)\phi_e^*(\mathbf{r}', \omega)$, which can be computed via the spatial Fourier transform of $\phi_e(\mathbf{k}, \omega)$. The coherence function is then

$$\gamma^2(\mathbf{r}, \mathbf{r}', \omega) = \frac{[P_e(\mathbf{r}, \mathbf{r}', \omega)]^2}{P_e(\mathbf{r}, \mathbf{r}, \omega)P_e(\mathbf{r}', \mathbf{r}', \omega)}. \quad (8.26)$$

This result has been shown to give good agreement with observations of γ^2 as a function of frequency at fixed separation for model parameters close to those used in obtaining the other plots in this work [27, 41]. Particular features are that coherence peaks correspond to spectral peaks, reflecting the fact that weakly damped waves can reach high amplitudes (hence a spectral peak) and propagate far before dissipating (hence high coherence).

8.3.7 Stability zone, instabilities, seizures, and phase transitions

Linear waves obey the dispersion relation (8.23), with instability boundaries occurring where this equation is satisfied for real ω [31, 34, 37]. In most circumstances, waves with $k = 0$ (i.e., spatially uniform) are the most unstable [37], and it is found that only the first few (i.e., lowest frequency) spectral resonances can become unstable. Analysis of stability of perturbations relative to the steady state that represents normal activity for realistic parameter ranges finds just four $k = 0$ instabilities, leading to global nonlinear dynamics [4, 31, 33]: (a) Slow-wave instability ($f \approx 0$) via a saddle–node bifurcation that leads to a low frequency spike-wave limit cycle; (b) theta instability, via a supercritical Hopf bifurcation that saturates in a nonlinear limit cycle near 3 Hz, with a spike-wave form unless its parameters are close to the instability boundary; (c) alpha instability, via a subcritical Hopf bifurcation, giving a limit cycle near 10 Hz; and (d) spindle instability at $\omega \approx (\alpha\beta)^{1/2}$, leading to a limit cycle at 10–15 Hz (the nature of this bifurcation has not yet been investigated). The boundaries defined by these instabilities are interpreted as corresponding to onsets of generalized seizures, as discussed in more detail below [4, 31, 33].

The occurrence of only a few instabilities, at low frequencies, enables the state and physical stability of the brain to be represented in a 3-D space with axes

$$x = G_{ee}/(1 - G_{ei}), \quad (8.27)$$

$$y = (G_{ese} + G_{esre})/[(1 - G_{srs})(1 - G_{ei})], \quad (8.28)$$

$$z = -G_{srs}\alpha\beta/(\alpha + \beta)^2, \quad (8.29)$$

which parameterize cortical, corticothalamic, and thalamic stability, respectively [4, 31]. In terms of these quantities, parameters corresponding to linearly stable brain states lie in a stability zone illustrated in Fig. 8.4. The back is at $x = 0$ and the base is at $z = 0$. A pure spindle instability occurs at $z = 1$, which couples to the alpha instability, with spindle dominating at top and left, and alpha at right. At small z , the left surface is defined by a theta instability [4, 31]. The front right surface corresponds to slow-wave instability at $x + y = 1$.

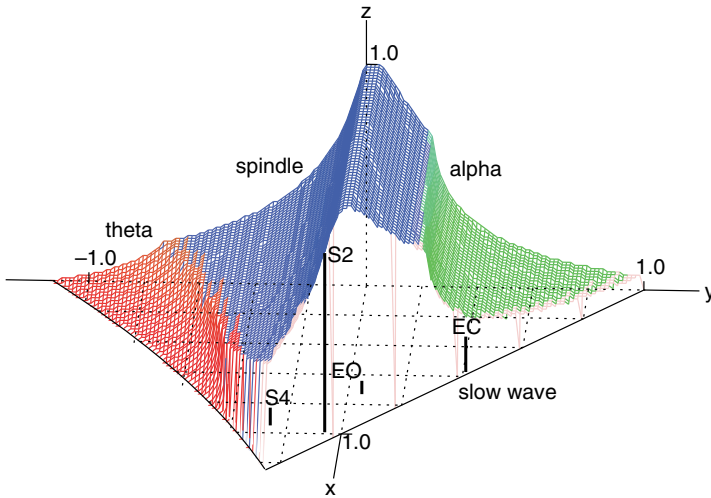


Fig. 8.4 [Color plate] Brain stability zone. The surface is shaded according to instability, as labeled (blue = spindle, green = alpha, red = theta), with the front right-hand face left transparent as it corresponds to a slow-wave instability. Approximate locations are shown of alert eyes-open (EO), relaxed, eyes-closed (EC), sleep stage 2 (S2), and sleep stage 4 (S4) states, with each state located at the top of its bar, whose (x, y) coordinates can be read from the grid.

Non-seizure states lie within the stability zone in Fig. 8.4. Detailed arguments regarding the sign of feedback via the thalamus, proximity between neighboring behavioral states, and the results of explicit fitting to data (which is enabled by using the present model), place the arousal sequence, from alert eyes-open (EO) to deep sleep, including relaxed eyes-closed (EC) and sleep stages 1–4 (S1–S4), as shown in Fig. 8.4 [31]. In future, it is expected that known differences between EEG spectra for subjects with differing disorders will also enable classification of these conditions into different parts of the stability zone.

Two of the most common generalized epilepsies are absence and tonic-clonic seizures. In absence epilepsy, seizures last 5–20 s, cause loss of consciousness, show a spike-wave cycle which starts and stops abruptly across the whole scalp, and the subject reaches a post-seizure state similar to the pre-seizure one. Tonic-clonic seizures display a tonic phase of roughly 10 Hz oscillations lasting about 10 s, followed by a clonic phase of similar duration dominated by polyspike-wave

complexes, with an unresponsive post-seizure state very different from the pre-seizure one [4, 15, 41]. Figures 8.5(a) and (b) show results from our model under conditions for theta and alpha instability, respectively. In Fig. 8.5(a) the onset of an approximately 3-Hz spike-wave cycle is seen as the system is forced across the instability boundary by ramping one of its parameters, in this case v_{se} . This closely resembles observed absence time series [4, 6, 31, 33]. If the destabilizing parameter is ramped back, the system returns smoothly to very nearly its initial state, consistent with clinical observations. Figure 8.5(b) shows good agreement with generalized tonic-clonic seizure dynamics near 10 Hz. However, in this case, the limit cycle sets in with nonzero amplitude. Moreover, when the control parameter is ramped back, hysteresis is observed, with the limit cycle terminating to yield a different final state, with a quiescent time series, consistent with clinical observations [4, 15].

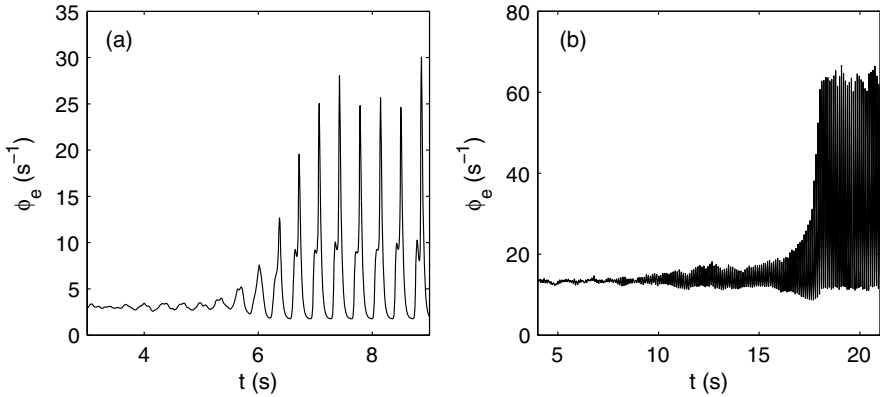


Fig. 8.5 Sample time series from the model in regimes corresponding to onset of (a) an absence seizure, and (b) a tonic-clonic seizure.

Each of the above instabilities can be seen as a phase transition. The saddle–node bifurcation is marked by a spectral divergence at $f = 0$, a $1/f$ spectrum at low f , and long-range correlations and coherence. There is also a divergence of the variance of ϕ_e , which can be approximated by integrating $P_e(\omega)$ over ω to yield the scaling

$$\left\langle \left[\phi_e - \phi_e^{(0)} \right]^2 \right\rangle \propto \left[V_{SN} - V_n^{(0)} \right]^{-1/2}, \quad (8.30)$$

where the angle brackets denote an average, the mean external input $V_n^{(0)}$ is the control parameter for the transition, and V_{SN} is its value at the bifurcation. This result accords with numerical results for such a transition [44] and related analysis of single neurons [45] (see also Sect. 8.4).

One recently explored feature of the nonzero- f limit cycles is that these can be initiated in localized regions of the system, and then spread to other areas, qualitatively consistent with clinical observations of secondary seizure generalization from

a focus [13]. In this case, the boundary between seizing and normal zones propagates in a manner akin to a domain boundary between solid and liquid in a spatially nonuniform melting/freezing transition. An example is shown in Fig. 8.6 [13].

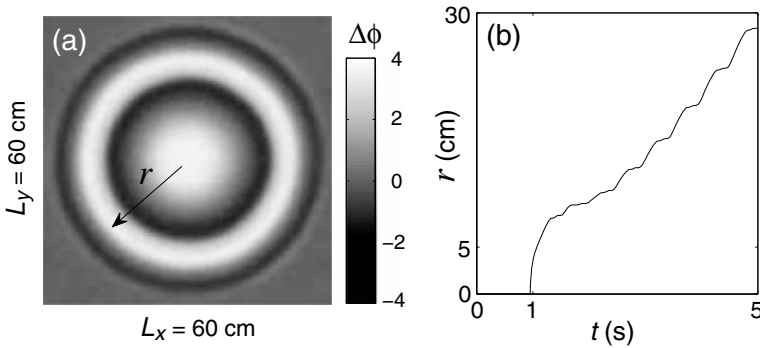


Fig. 8.6 Spreading of seizure activity from an initial focus. The figure shows (a) a snapshot of $\phi(\mathbf{r}, t)$, and (b) the linear spread of the wave following a stimulus at $t = 1$ s.

8.4 Mean-field modeling of the brainstem and hypothalamus, and sleep transitions

Wake-sleep transitions are primarily governed by the nuclei of the ascending arousal system of the brainstem and hypothalamus, that project diffusely to the corticothalamic system. As we will see shortly, these nuclei are also capable of undergoing instabilities and phase transitions in their dynamics. Hence, a full description of sleep-wake transitions and their EEG correlates requires an integrated model of both the ascending arousal system and the corticothalamic system (at least), including their mutual interactions. This section briefly describes how the nuclei of the Ascending Arousal System (AAS) are modeled using the same methods as above, and outlines the direction of integration of the two models, currently under way. In this section, observables consist of arousal states (sleep vs. wake), so other measurement effects need not be taken into account.

8.4.1 Ascending Arousal System model

The most important nuclei to model in the AAS are well established from detailed physiological investigations, and are shown in Fig. 8.7. These include the monoaminergic (MA) group and the ventrolateral preoptic nucleus (VLPO), which mutually inhibit one another, resulting in flip-flop dynamics if the interaction is sufficiently strong—only one can be active at a time, and it suppresses the other

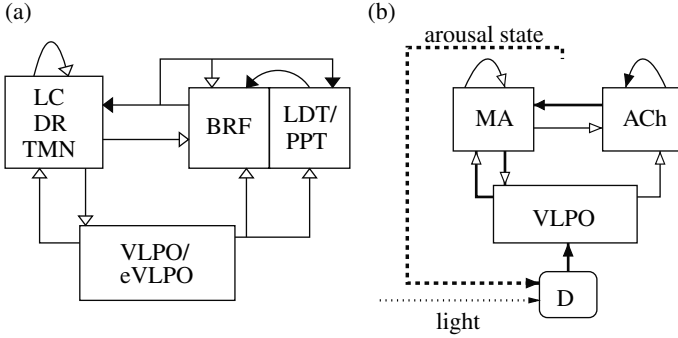


Fig. 8.7 Parts (a) and (b) show schematics of the actual AAS populations, and our sleep model, respectively. Excitatory inputs are represented by solid arrow heads, and inhibitory by open ones. In each case, the top left box is the MA group, and the top right box is the ACh group. In (a), the MA group consists of the LC, DR (dorsal raphe) and TMN (tuberomamillary nucleus); and the ACh group consists of cholinergic LDT/PPT, and glutamergic BRF. The VLPO/eVLPO GABAergic inhibits other AAS nuclei. In (b) the drive D is shown, which consists of circadian (C) and homeostatic (H) components. In our model the thick-lined interactions in (b) are used.

[39]. During wake, the MA group is dominant, while the VLPO is dominant in sleep. Transitions between states are driven by inputs to the VLPO, which include the circadian drive C (mainly from light exposure), and the homeostatic sleep drive H arising from net buildup of metabolic byproducts (mostly adenosine) during wake, and their net clearance during sleep. There is also an input to the MA group from cholinergic and orexinergic nuclei, as shown [21, 39].

Until recently, models of AAS dynamics have been either nonmathematical (e.g., based on sleep diaries or qualitative considerations) or abstract (mathematical, but not derived directly from physiology). The widely known *two-process model* is of the latter form, and includes circadian and homeostatic influences [1]. In this section, which summarizes our recent model of the AAS [22], we use the same methods as in Sects 8.2–8.3 to model the dynamics of the AAS nuclei, viewing them as the assemblies of neurons they are. Several simplifications and approximations are appropriate: the nuclei are small, so $r_a \approx 0$ and $\gamma_a \rightarrow \infty$ in Eq. (8.11), implying that Eq. (8.17) applies for these nuclei. Also, since the transitions take place on timescales of many seconds to minutes, first-order in time versions of Eq. (8.3) can be used. We also assume that, since the system spends little time in transitions, the generation rate of H has just two values—one for wake and one for sleep—and that its clearance rate is proportional to H , while the variation of C is approximated as sinusoidal. These approximations yield

$$\tau \frac{dV_v}{dt} + V_v = v_{vm} Q_m + D, \quad (8.31)$$

$$\tau \frac{dV_m}{dt} + V_m = v_{mv} Q_v + A, \quad (8.32)$$

$$\chi \frac{dH}{dt} + H = \mu Q_m, \quad (8.33)$$

$$Q_a = S(V_a), \quad (8.34)$$

$$C = c_0 + \cos(\Omega t), \quad (8.35)$$

$$D = v_{vc}C + v_{vh}H, \quad (8.36)$$

where the time constants τ of the nuclear responses have been assumed equal [these replace $1/\alpha$ in (8.3), with $\beta \rightarrow \infty$ formally], χ is the adenosine clearance time, v denotes VLPO, m denotes monoaminergic nuclei, the v_{ab} , V_a , and Q_a have the same meanings as in previous sections, μ gives the proportionality between monoaminergic activity and adenosine generation rate, C_A is the amplitude of the C cycle, and $\Omega = 2\pi/(1 \text{ day})$.

In the above form the model has 12 physiological parameters: τ , χ , v_{vm} , v_{mv} , A , μ , c_0 , v_{vc} , v_{vh} , Q_{\max} , θ , and σ' , whose nominal values are given in Table 8.2. These values were determined by a combination of physiological constraints from the literature, and comparison of the dynamics with behavior in a restricted set of sleep experiments on normal sleep and sleep deprivation [22, 23]. The theory then predicts other phenomena in regimes outside those of the calibration experiments.

In the context of the present chapter, the key result is that the steady states of Eqs (8.31)–(8.36) display a “fold” as a function of the total drive D . The upper and lower branches represent wake and sleep, respectively, with an unstable branch in between. Cyclic variations in D cause the system to move around the hysteresis loop shown in Fig. 8.8, with saddle–node bifurcations from wake to sleep and back again. In the presence of noise added to D on the right of Eq. (8.31), Fig. 8.9 shows that these are preceded by divergences in V_m fluctuations that satisfy the same power-law scaling as Eq. (8.30) for subthreshold noise, and lead to what appear to be microsleeps and microwakes in the vicinity of the transition for larger amplitude noise. Narcolepsy, with its lack of stability of wake and sleep is then interpreted as resulting from a reduction or disappearance of the hysteresis loop [22].

Table 8.2 Nominal parameter values for the ascending arousal system model.

Quantity	Nominal	Unit
$-v_{vc}$	2.9	mV
v_{vh}	1.0	mV nM ⁻¹
χ	45	h
μ	4.4	nM s
c_0	4.5	–
Q_{\max}	100	s ⁻¹
θ	10	mV
σ'	3	mV
A	1.3	mV
$-v_{vm}$	2.1	mV s
$-v_{mv}$	1.8	mV s
τ	10	s

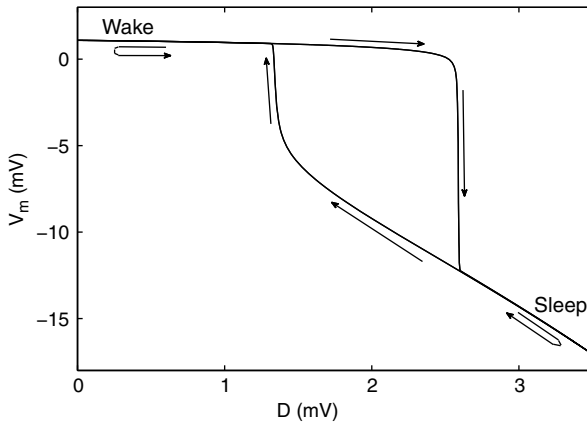


Fig. 8.8 Plot of V_m versus the sleep drive D across a 24 h period. As D oscillates across the day, V_m cycles around a hysteresis loop between wake and sleep states.

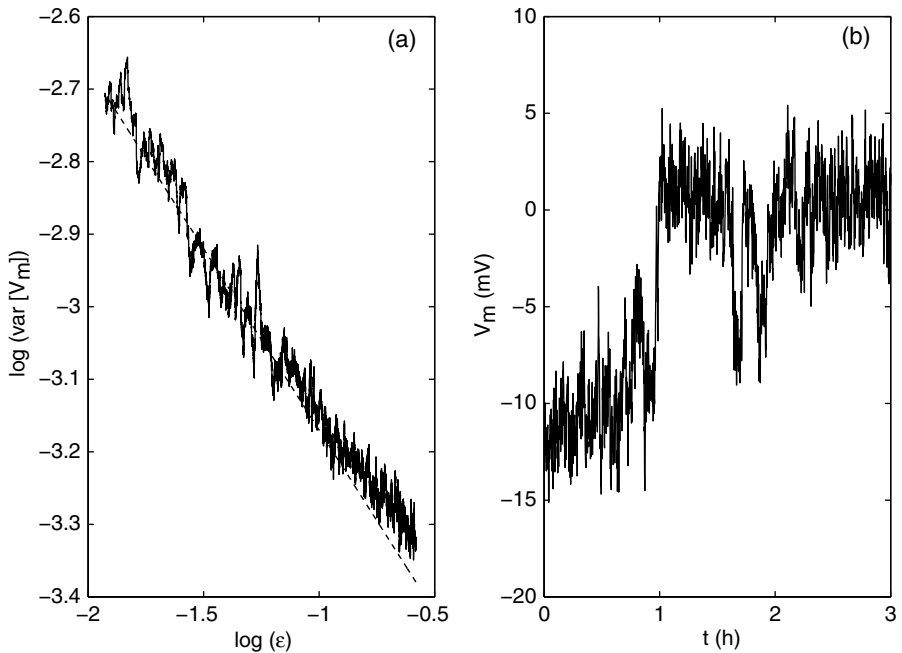


Fig. 8.9 (a) Log-log plot of the variance of V_m in the presence of low amplitude noise (solid), versus $\epsilon = D - D_0$, where D_0 is the value of D for which the wake state loses stability. D is increased linearly at a rate of $7 \times 10^{-5} \text{ h}^{-1}$, and variance is calculated in moving windows of length 17 h. The asymptotic gradient of -0.5 is shown as a dashed line. (b) Transitions between high V_m (wake) and low V_m (sleep) in the presence of high amplitude noise.

This set of outcomes implies that inclusion of the dynamics of the AAS is essential to understand sleep–wake cycles, although work is still under way to incorporate the ascending projections to the corticothalamic system quantitatively, feedback in the reverse direction, and quantitative models of the circadian pathway, involving the suprachiasmatic nucleus (SCN).

8.5 Summary and discussion

Physiologically based mean-field theories of the brain are able to incorporate essential physiology and anatomy across the many scales necessary to treat phase transitions and other phenomena involving neural activity. They can achieve this for physiologically realistic parameters, and yield numerous predictions that accord with observations using a variety of experimental methods in both the linear and nonlinear regimes (see Sect. 8.1). Moreover, they do this in a way that unifies what have hitherto been disparate subfields and measurement modalities within a single framework, and which permits parameter determination via fits of model predictions to experimental data. In addition to these specific results, major qualitative conclusions that are reached using such models include the necessity of incorporating the thalamus to understand EEG phenomena at frequencies below about 20 Hz, and the need to include the ascending arousal system to understand sleep–wake dynamics.

In the area of phase transitions, mean-field modeling successfully predicts the connections between transitions, instabilities, long-range correlations and coherence, spectral peaks, and divergences of variance in a number of regimes. However, much remains to be done in directions such as the fuller integration of multiple brain subsystems into unified models, exploration of the dynamics of neuromodulators and behavioral feedbacks, and application to other putative phase transitions in areas such as visual rivalry and perception, parkinsonian tremor onset, and possibly bipolar disorder. One could also investigate whether some Hopf bifurcations (e.g., supercritical ones) correspond to second-order phase transitions, as opposed to the first-order ones investigated here, and whether the variance divergences seen near criticality have a role in control or prevention of phase transitions.

Acknowledgments The Australian Research Council supported this work.

References

1. Achermann, P., Borbély, A.A.: Mathematical models of sleep regulation. *Front. Biosci.* **8**, s683–s693 (2003)
2. Binney, J.J., Dowrick, N.J., Fisher, A.J., Newman, M.E.J.: *The Theory of Critical Phenomena*. Clarendon Press, Oxford (1992)

3. Braitenberg, V., Shüz, A.: *Anatomy of the Cortex: Statistics and geometry*. Springer, Berlin (1991)
4. Breakspear, M., Roberts, J.A., Terry, J.R., Rodrigues, S., Mahant, N., Robinson, P.A.: A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cerebral Cortex* **16**, 1296–1313 (2006), doi:10.1093/cercor/bhj072
5. Clearwater, J.M., Rennie, C.J., Robinson, P.A.: Mean field model of acetylcholine mediated dynamics in the cerebral cortex. *Biol. Cybernetics* **97**, 449–460 (2007), doi:10.1007/s00422-007-0186-9
6. Feucht, M., Möller, U., Witte, H., Schmidt, K., Arnold, M., Benninger, F., Steinberger, K., Friedrich, M.H.: Nonlinear dynamics of 3 Hz spike-and-wave discharges recorded during typical absence seizures in children. *Cerebral Cortex* **8**(6), 524–533 (1998)
7. Freeman, W.J.: *Mass Action in the Nervous System*. Academic Press, New York (1975)
8. Henderson, J.A., Phillips, A.J.K., Robinson, P.A.: Multielectrode electroencephalogram power spectra: Theory and application to approximate correction of volume conduction effects. *Phys. Rev. E* **73**, 051918 (2006), doi:10.1103/PhysRevE.73.051918
9. Horton, J.C., Adams, D.L.: The cortical column: A structure without a function. *Philos. Trans. Roy. Soc. Lond. Ser. B* **360**, 837–862 (2005), doi:10.1098/rstb.2005.1623
10. Ivanov, A.V., Cairns, I.H., Robinson, P.A.: Wave damping as a critical phenomenon. *Phys. Plasmas* **10**, 4649–4661 (2004), doi:10.1063/1.1785789
11. Ivanov, A.V., Vladimirov, S.V., Robinson, P.A.: Criticality in a Vlasov-Poisson system: A fermionlike universality class. *Phys. Rev. E* **71**, 056406 (2005), doi:10.1103/PhysRevE.71.056406
12. Jirsa, V.K., Haken, H.: Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996)
13. Kim, J.W., Roberts, J.A., Robinson, P.A.: Dynamics of epileptic seizures: Evolution, spreading, and suppression. *J. Theor. Biol.* **257**(4), 527–532 (2009), doi:10.1016/j.jtbi.2008.12.009
14. Lopes da Silva, F.H., Hoeks, A., Smits, H., Zetterberg, L.H.: Model of brain rhythmic activity. T alpha-rhythm of the thalamus. *Kybernetik* **15**, 27–37 (1974)
15. Niedermeyer, E.: The normal EEG of the waking adult. In: E. Niedermeyer, F. Lopes da Silva (eds.), *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, pp. 149–173, Williams & Wilkins, Baltimore, fourth edn. (1999)
16. Nunez, P.L.: The brain wave equation: A model for the EEG. *Math. Biosci.* **21**, 279–297 (1974)
17. Nunez, P.L.: *Neocortical Dynamics and Human EEG Rhythms*. Oxford University Press, New York (1995)
18. Nunez, P.L., Srinivasan, R.: *Electric Fields of the Brain : The Neurophysics of EEG*. Oxford University Press, New York, 2nd edn. (2006)
19. O'Connor, S.C., Robinson, P.A.: Wave-number spectrum of electrocorticographic signals. *Phys. Rev. E* **67**, 051912 (2003), doi:10.1103/PhysRevE.67.051912
20. O'Connor, S.C., Robinson, P.A., Chiang, A.K.I.: Wave-number spectrum of electroencephalographic signals. *Phys. Rev. E* **66**, 061905 (2002), doi:10.1103/PhysRevE.66.061905
21. Pace-Schott, E.F., Hobson, J.A.: The neurobiology of sleep: Genetics, cellular physiology and subcortical networks. *Nature Rev. Neurosci.* **3**, 591–605 (2002), doi:10.1038/nrn895
22. Phillips, A.J.K., Robinson, P.A.: A quantitative model of sleep-wake dynamics based on the physiology of the brainstem ascending arousal system. *J. Biol. Rhythms* **22**(2), 167–179 (2007), doi:10.1177/0748730406297512
23. Phillips, A.J.K., Robinson, P.A.: Sleep deprivation in a quantitative physiologically-based model of the ascending arousal system. *J. Theor. Biol.* **255**(4), 413–423 (2008), doi:10.1016/j.jtbi.2008.08.022
24. Rennie, C.J., Robinson, P.A., Wright, J.J.: Effects of local feedback on dispersion of electrical waves in the cerebral cortex. *Phys. Rev. E* **59**(3), 3320–3329 (1999)
25. Rennie, C.J., Robinson, P.A., Wright, J.J.: Unified neurophysical model of EEG spectra and evoked potentials. *Biol. Cybernetics* **86**, 457–471 (2002), doi:10.1007/s00422-002-0310-9

26. Robinson, P.A.: Interpretation of scaling properties of electroencephalographic fluctuations via spectral analysis and underlying physiology. *Phys. Rev. E* **67**, 032902 (2003), doi:10.1103/PhysRevE.67.032902
27. Robinson, P.A.: Neurophysical theory of coherence and correlations of electroencephalographic and electrocorticographic signals. *J. Theor. Biol.* **222**, 163–175 (2003), doi:10.1016/j.jtbi.2004.07.003
28. Robinson, P.A.: Propagator theory of brain dynamics. *Phys. Rev. E* **72**, 011904 (2005), doi:10.1103/PhysRevE.72.011904
29. Robinson, P.A.: Patchy propagators, brain dynamics, and the generation of spatially structured gamma oscillations. *Phys. Rev. E* **73**, 041904 (2006), doi:10.1103/PhysRevE.73.041904
30. Robinson, P.A., Loxley, P.N., O'Connor, S.C., Rennie, C.J.: Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. *Phys. Rev. E* **63**(4), 041909 (2001), doi:10.1103/PhysRevE.63.041909
31. Robinson, P.A., Rennie, C.J., Rowe, D.L.: Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Phys. Rev. E* **65**(4), 041924 (2002), doi:10.1103/PhysRevE.65.041924
32. Robinson, P.A., Rennie, C.J., Rowe, D.L., O'Connor, S.C.: Estimation of multiscale neurophysiologic parameters by electroencephalographic means. *Hum. Brain Mapp.* **23**, 53–72 (2004), doi:10.1002/hbm.20032
33. Robinson, P.A., Rennie, C.J., Rowe, D.L., O'Connor, S.C., Wright, J.J., Gordon, E., Whitehouse, R.W.: Neurophysical modeling of brain dynamics. *Neuropsychopharmacology* **28**, s74–s79 (2003), doi:10.1038/sj.npp.1300143
34. Robinson, P.A., Rennie, C.J., Wright, J.J., Bahramali, H., Gordon, E., Rowe, D.L.: Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E* **63**(2), 021903 (2001), doi:10.1103/PhysRevE.63.021903
35. Robinson, P.A., Rennie, C.J., Wright, J.J., Bourke, P.: Steady states and global dynamics of electrical activity in the cerebral cortex. *Phys. Rev. E* **58**(3), 3557–3571 (1998)
36. Robinson, P.A., Whitehouse, R.W., Rennie, C.J.: Nonuniform corticothalamic continuum model of electroencephalographic spectra with application to split-alpha peaks. *Phys. Rev. E* **68**, 021922 (2003), doi:10.1103/PhysRevE.68.021922
37. Robinson, P.A., Rennie, C.J., Wright, J.J.: Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* **56**(1), 826–840 (1997)
38. Rowe, D.L., Robinson, P.A., Rennie, C.J.: Estimation of neurophysiological parameters from the waking EEG using a biophysical model of brain dynamics. *J. Theor. Biol.* **231**, 413–433 (2004), doi:10.1016/j.jtbi.2004.07.004
39. Saper, C.B., Chou, T.C., Scammell, T.E.: The sleep switch: hypothalamic control of sleep and wakefulness. *Trends Neurosci.* **24**, 726–731 (2001), doi:10.1016/S0166-2236(00)02002-6
40. Sherman, S.M., Guillery, R.W.: *Exploring the Thalamus*. Academic Press (2001)
41. Srinivasan, R., Nunez, P.L., Silberstein, R.B.: Spatial filtering and neocortical dynamics: Estimates of EEG coherence. *IEEE Trans. Biomed. Eng.* **45**, 814–826 (1998)
42. Steriade, M., Gloor, P., Llinás, R.R., Lopes da Silva, F.H., Mesulam, M.M.: Basic mechanisms of cerebral rhythmic activities. *Electroenceph. Clin. Neurophysiol.* **76**, 481–508 (1990)
43. Steriade, M., Jones, E.G., McCormick, D.A. (eds.): *Thalamus* (2 vols). Elsevier, Amsterdam (1997)
44. Steyn-Ross, D.A., Steyn-Ross, M.L., Seigh, J.W., Wilson, M.T., Gillies, I.P., Wright, J.J.: The sleep cycle modelled as a cortical phase transition. *J. Biol. Phys.* **31**, 547–569 (2005), doi:10.1007/s10867-005-1285-2
45. Steyn-Ross, D.A., Steyn-Ross, M.L., Wilson, M.T., Sleight, J.W.: White-noise susceptibility and critical slowing in neurons near spiking threshold. *Phys. Rev. E* **74**, 051920 (2006), doi:10.1103/PhysRevE.74.051920
46. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleight, J.W., Liley, D.T.J.: Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: Evidence for a general anesthetic-induced phase transition. *Phys. Rev. E* **60**(6), 7299–7311 (1999)

47. Strogatz, S.H.: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Addison-Wesley, Reading, Mass. (1994)
48. Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973)
49. Wilson, H.R.: *Spikes, Decisions, and Actions: The Dynamical Foundations of Neurosciences*. Oxford University Press, Oxford, New York (1999)
50. Wright, J.J., Liley, D.T.J.: Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. *Behav. Brain Sci.* **19**, 285–309 (1996)

Chapter 9

A continuum model for the dynamics of the phase transition from slow-wave sleep to REM sleep

J.W. Sleigh, M.T. Wilson, L.J. Voss, D.A. Steyn-Ross, M.L. Steyn-Ross, and X. Li

9.1 Introduction

The cortical transition from the slow-wave pattern of sleep (SWS) to the rapid-eye-movement (REM) pattern is a dramatic feature of the somnogram. Indeed, the change in the electrocorticogram (ECoG) is so abrupt that the moment of transition usually can be identified with a time-resolution of about one second [8, 37]. Although the neuromodulatory environment and electroencephalographic patterns recorded during the steady states of SWS and REM have been well described [16, 30], the dynamics of the transition itself has been described only in a qualitative observational fashion [12], and has not been the focus of detailed quantitative modeling.

In SWS, the rat cortex shows predominant activity in the delta ($\sim 1\text{--}4$ Hz) band. This pattern shifts to an intermediate sleep state (IS)—sometimes termed “pre-REM”—where the cortical activity shows features of both SWS and REM, lasting 10–30 seconds [7, 12, 25, 28]. This is followed by an abrupt transition to the REM state, characterized by strong theta ($\sim 5\text{--}8$ Hz) oscillation [2], and loss of delta power. The main effector of the cortical transition from SWS to REM is believed to be a linear progressive increase in cholinergic input into the cortex from the brainstem (mainly from the pedunculo-pontine tegmentum area), acting via the thalamus or basal forebrain [18, 34].

Jamie W. Sleigh · Logan J. Voss

Department of Anaesthesia, Waikato Clinical School, University of Auckland, Waikato Hospital, Hamilton 3204, New Zealand.

e-mail: sleighj@waikatodhb.govt.nz VossL@waikatodhb.govt.nz

Marcus T. Wilson · D. Alistair Steyn-Ross · Moira L. Steyn-Ross

Department of Engineering, University of Waikato, P.B. 3105, Hamilton 3240, New Zealand.

e-mail: m.wilson@waikato.ac.nz asr@waikato.ac.nz msr@waikato.ac.nz

Xiaoli Li

School of Computer Science, The University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK. e-mail: X.Li@cs.bham.ac.uk

Several studies [8, 32, 37] have measured the activity of the cortically-projecting pontine cholinergic REM-on neurons during the SWS–REM transition. These studies have shown a progressive, linear increase in firing rate, starting 10–60 s before onset of REM sleep, and plateauing about 60 s after transition. The predominant effect of increasing cholinergic input is to raise cerebral cortical arousal by acting on muscarinic (mainly M1) receptors to close potassium channels, causing a depolarizing shift in the cortical resting membrane potential. The increase in cholinergic tone also results in a small decrease in the amplitude of the excitatory postsynaptic potential (EPSP) [13, 23, 27].

A recent paper by Lu and others has highlighted the fact that the pontine cholinergic nuclei are themselves under the influence of other orexinergic and gamma-aminobutyric acid (GABA)-ergic switching circuits [24]. They suggest that the “flip-flop” arrangement of these brainstem and midbrain circuits could explain the abrupt changes in state observed in the cortex.

In contrast, we suggest that the abrupt changes seen in ECoG during SWS-to-REM transition could be explained in terms of an abrupt cortical response to a gradual change in the underlying subcortical neuromodulator activity. As described below, we enhance a previously-published continuum model of interactions between excitatory and inhibitory populations of cortical neurons [39, 40, 47–49], and compare its output with experimentally-derived data recorded from rats during the SWS-to-REM transition.

9.2 Methods

9.2.1 Continuum model of cortical activity

We use a continuum model of the interactions between populations of inhibitory and excitatory cortical neurons to describe the features of the SWS-to-REM transition. The continuum (or mean-field) approach assumes that the important dynamics of neuronal activity can be captured by the averages of small populations of ($\sim 10\,000$ – $100\,000$) neurons contained within a “macrocolumn” (defined by the spatial extent of the typical pyramidal neuron’s dendritic arborization). Continuum modeling of the cortex originated with Wilson and Cowan [46], and has been progressively refined since then by the inclusion of more neurobiologically realistic terms and parameters [22, 39, 50].

Our version of the model has been described in detail by Wilson *et al.* [49]. It is cast in the form of a set of stochastic differential equations; these equations incorporate (i) spike-rate input from: neighboring cortical neurons (dependent on local membrane potential), distant cortical neurons (dependent on distant membrane potential), and subcortical structures (independent of cortical membrane potential); (ii) dendritic time-evolution and magnitudes of fast inhibitory and excitatory synaptic potentials (including the effects of reversal potentials); (iii) a sigmoid

relationship between soma potential and neuronal firing rate; and (iv) cortical connectivity that drops off (approximately exponentially) with increasing spatial separation. The mathematical details of the theoretical model are outlined in the Appendix.

In general, parameter values are drawn from experimentally-derived measurements reported in the literature, so are physiologically plausible. The actual values used here are similar to those presented in an earlier paper modeling the seizurogenic effects of enflurane in humans [47], with some alterations to better represent the smaller rat cortex (see Table 9.1).

Table 9.1 Parameters for rat cortex model

Symbol	Description	Value
$\tau_{e,i}$	Membrane time constant	20, 20 ms
$Q_{e,i}$	Maximum firing rates	60, 120 s ⁻¹
$\theta_{e,i}$	Sigmoid thresholds	-58.5, -58.5 mV
$\sigma_{e,i}$	Standard deviation of thresholds	4, 6 mV
$\rho_{e,i}$	Gain per synapse at resting voltage	0.001, -0.000863 mV s
$V_{e,i}^{\text{rev}}$	Cell reversal potential	0, -70 mV
$V_{e,i}^{\text{rest}}$	Cell resting potential	-64, -64 mV
N_{eb}^{α}	Long-range $e \rightarrow e, i$ connectivity	3710
N_{eb}^{β}	Short-range $e \rightarrow e, i$ connectivity	410
N_{ib}^{β}	Short-range $i \rightarrow e, i$ connectivity	800
$\langle \phi_{eb}^{\text{sc}} \rangle$	Mean $e \rightarrow e, i$ subcortical flux	50 s ⁻¹
$\langle \phi_{ib}^{\text{sc}} \rangle$	Mean $i \rightarrow e, i$ subcortical flux	50 s ⁻¹
γ_{eb}	Excitatory synaptic rate constant	949 s ⁻¹
γ_{ib}	Inhibitory synaptic rate constant	100 s ⁻¹
$L_{x,y}$	Spatial length of cortex	2 cm
Λ_{eb}	Inverse-length connection scale	0.4 mm ⁻¹
v	Mean axonal conduction speed	1400 mm s ⁻¹
ΔV_e^{rest}	Effect of altering extrasynaptic ion channels	-5 → +5 mV
λ	Scaling for EPSP amplitude	see Eqs (9.2, 9.3)

The primary output from the model is the spatially-averaged soma potential. The fluctuation of this voltage with time is assumed to be the source of the experimentally observable ECoG signal [50]. In its present form, the model considers the cortex to be a two-dimensional sheet; the model ignores within-cortex microanatomical layering, and does not include synaptic plasticity. (The dynamic implications of these finer biological details may be the subject of future investigation.)

The changes in interneuronal population activity during sleep can be represented on a two-parameter domain (see Fig. 9.1), chosen to make explicit the effects of selected neuromodulators on the cortex. It is assumed that these neuromodulations are slow processes (\sim seconds to minutes) compared to the much faster time-scale of synaptic neurotransmission and conduction of action potentials (\sim ones to tens of milliseconds).

This separation of time-scales allows neuromodulator action to be incorporated into the model as a pair of slowly-varying control parameters that represent excursions along the mutually-orthogonal horizontal directions in Fig. 9.1 labeled, respectively, λ and ΔV_e^{rest} . Here, λ represents EPSP synaptic strength, and ΔV_e^{rest} represents relative neuronal excitability (displacement of resting membrane potential of the pyramidal neurons relative to a default background value).

This arrangement allows the effects of neuromodulators on synaptic function (the λ -axis) to be separated from their effects on extrasynaptic leak currents that will alter resting voltage (ΔV_e^{rest} -axis). We assume that SWS is associated with decreased levels of aminergic and cholinergic arousal from the brainstem, together with elevated concentrations of somnogens such as adenosine [31]; both effects will tend to hyperpolarize the membrane voltage by increasing outward potassium leak current.

The gradual rise in neuromodulator-induced membrane polarization is incorporated into the model as a slowly-increasing ΔV_e^{rest} parameter (see Eq. (9.1) below) that can be visualized as a trajectory (thick gray arrow) superimposed on the Fig.-9.1 manifold of stationary states. The mean excitatory soma potential V_e is depicted

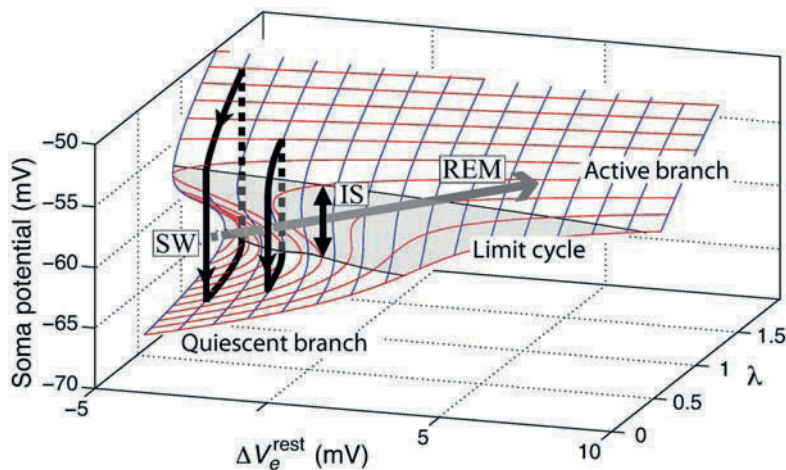


Fig. 9.1 Manifold of equilibrium states for homogeneous model cortex for different stages of sleep. Vertical axis is excitatory soma potential; horizontal axes are λ (dimensionless EPSP-amplitude scalefactor), and ΔV_e^{rest} (deviation of resting membrane potential above default value of -64 mV). Shaded-gray area shows region of instability giving rise to theta-frequency limit-cycle oscillations. The gray arrow shows a trajectory for transition from slow-wave (SW) to intermediate (IS) to REM sleep caused by a gradual increase in resting soma potential (see Eq. (9.1)). Black arrows denote the details of the trajectory in response to the dynamic modulation in EPSP (Eq. (9.3)).

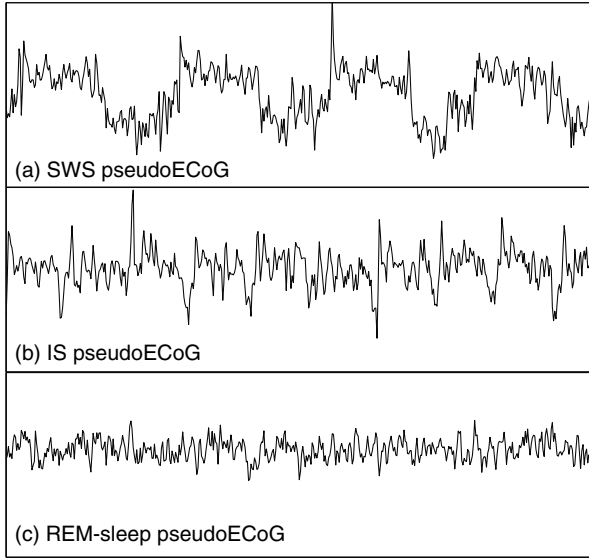


Fig. 9.1 (*cont.*) Model-generated ECoG time-series at the three labeled points on the sleep manifold: (a) slow-wave sleep (SWS); (b) intermediate sleep (IS); (c) rapid-eye-movement (REM) sleep. Duration of each time-series = 6 s; vertical-axis extent for each graph = 0.4 mV.

on the vertical axis of Fig. 9.1. Where there is a “fold” in the manifold, there can exist up to *three* steady-state values for V_e at a given $(\lambda, \Delta V_e^{\text{rest}})$ coordinate: for these cases we label the lower state (lying under the fold) as “quiescent”, and the upper state (located on top of the fold) as “active”. We note that the upper and lower stationary states are *not necessarily stable*. In fact, it is the transition to instability that gives rise to oscillatory behavior in the model.

9.2.2 Modeling the transition to REM sleep

The transition to REM sleep is characterized by a progressive increase in cholinergic activity from the brainstem that occurs over a time-course of a few minutes. This increase in cholinergic tone simultaneously depolarizes the cortex, and reduces (slightly) the excitatory synaptic gain (via reduction in the area of the excitatory postsynaptic potential (EPSP)). We model the gradual rise in cortical depolarization by imposing a linear increase in the excitatory resting-potential *offset* ΔV_e^{rest} , from -5 mV (i.e., resting voltage set 5 mV *below* nominal) to $+5$ mV (resting voltage set 5 mV *above* nominal), over a period of 4 min (240 s),

$$\Delta V_e^{\text{rest}}(t) = -5 \text{ mV} + 10 \text{ mV}(t/240 \text{ s}), \quad (9.1)$$

where t is the elapsed time in seconds. The nominal resting voltage is -64 mV (see Table 9.1). At the same time, the EPSP gain-factor λ_1 (dimensionless) decreases

linearly over 4 min to its nominal value ($\lambda_1 = 1.00$) from a starting value set 5% above nominal,

$$\lambda_1(t) = 1.05 - 0.05(t/240 \text{ s}). \quad (9.2)$$

Here, λ_1 is modeling the synaptic effect of the steady increase of acetylcholine concentration as the cortex transits from SWS to REM sleep. This slow acetylcholine change (λ_1) will be combined with the synaptic-gain adaptations (λ_2), described below, brought about by the SWS cortical oscillations between “down” and “up” states.

9.2.3 Modeling the slow oscillation of SWS

Because slow-wave sleep is characterized by low levels of cholinergic tone, the excitatory synaptic gain in the cortex is influenced by the presynaptic firing-rate, resulting in a form of slow spike-frequency adaptation: when the presynaptic neuron is in a high-firing state, consecutive EPSP events decrease in magnitude exponentially over the time-course of a few-hundred milliseconds as described by Tsodyks [43]; conversely, once the presynaptic neuron becomes quiescent, it becomes relatively more sensitive to input. Under conditions of low-cholinergic effect, this fluctuation in excitatory synaptic gain induces the cortex to undergo a slow oscillation between the distinct “down” (quiescent) and “up” (activated) states that are observed in SWS [26, 38]. This approach to understanding the cortical slow oscillation is broadly equivalent to previous SWS models [1, 6, 14, 17, 33, 36, 41] which rely on time-varying changes in the Na^+ and K^+ ion-channel conductances to cycle the cortex between “up” and “down” states.

This up–down cycling is in addition to the slower modulation due to brainstem changes (Eq. (9.2)), so we write the total synaptic-gain factor as $\lambda = \lambda_1 + \lambda_2$, where λ_1 corresponds to very slow brainstem effects, and λ_2 to the slow-oscillation effects. We drive parameter λ_2 through a cycle between down- and up-states by raising λ_2 (increasing EPSP) slightly in response to a low firing rate, and reducing it (lowering EPSP) in response to a high firing rate,

$$\frac{d\lambda_2}{dt} = -k(\lambda_2 - \lambda_{\text{aim}}), \quad (9.3)$$

where the shunting rate-constant k is 2 s^{-1} , and the steady-state target value λ_{aim} is determined by the firing rate,

$$\lambda_{\text{aim}} = \begin{cases} +0.2, & Q_e < 10 \text{ s}^{-1}, \\ -0.2, & Q_e \geq 10 \text{ s}^{-1}. \end{cases} \quad (9.4)$$

The full set of cortical equations are listed in the Appendix on p.215. The total synaptic gain $\lambda = \lambda_1 + \lambda_2$ from Eqs (9.2–9.3), and the ΔV_e^{rest} depolarization from Eq. (9.1), are applied as modulatory effects to the differential equation for excitatory soma potential (Appendix, Eq. (9.5)). The model behavior resulting from the cycling and modulation in λ , and the modulation of ΔV^{rest} , is summarized by the arrowed paths in Fig. 9.1.

9.2.4 Experimental Methods

9.2.4.1 Animals

Four male Sprague-Dawley rats, weighing 300–400 g at the time of surgery, served as subjects. The rats were maintained on a 12:12-hr light–dark cycle, were individually housed following surgery, and had *ad libitum* access to food and water. Ethical approval for this study was granted by the Ruakura and University of Auckland Animal Ethics Committees.

9.2.4.2 Surgery

Animals were anesthetised with ketamine/xylazine (75/10 mg/kg, i.p.), and mounted in a stereotaxic instrument with the skull held level. Four holes were drilled in the exposed skull: three for stainless-steel skull screws (positioned over the cerebellum and bilaterally over the parietal cortex), and one for implantation of a tungsten stereotrode pair (Micro Probe Inc, Potomoc, USA) into the parietal cortex for two-channel electrocorticogram (ECoG) recording. The stereotrode consisted of two insulated microelectrodes (3- μm diameter) separated by 200 μm . The stereotrode was lowered into the cortex to a depth of 0.5 mm and cemented to one of the anchor screws with rapid-setting glue. The skull screws also served as reference and ground electrodes for the cortical local-field-potential recordings. Insulated wires from the screws, along with the stereotrode electrodes, were terminated in a plastic nine-pin socket, the base of which was embedded in dental acrylic (GC Corporation, Tokyo, Japan). The animals were allowed to recover for at least seven days prior to testing.

9.2.4.3 Data recording

There were two ECoG recording channels. The two parietal skull screws served as the common reference for the two cortical electrodes, and the cerebellar screw was used as the common ground. The leads were connected to two differential amplifiers (A-M systems Inc, Carisborg, USA) via a tether and electrical swivel (Stoelting Co, Illinois, USA), allowing free movement of the animal within the recording enclosure. The two cortical field-potential channels were digitized at 10000 samples/s (CED Power 1401, Cambridge, England), high- and lowpass filtered at 1 and 2500 Hz, respectively, and 50-Hz notch-filtered. The data were

displayed and recorded continuously on computer using Spike2 software (CED, Cambridge, England). The animals were video-recorded during all sessions to aid offline sleep-staging (described below). The video was synchronized with the electrophysiological recordings. Data were collected for up to six hours while the animals slept naturally.

9.2.4.4 Sleep staging

Sleep-staging was performed offline using accepted electrophysiological and behavioral criteria [44]. Slow-wave sleep (SWS) was characterized by a large-voltage, low-frequency ECoG waveform and regular respiratory pattern (observed on video). During SWS, the rats typically lay on their abdomen in a reclined posture.

Rapid eye movement (REM) sleep was characterized by a low-voltage, high-frequency ECoG waveform, and a respiratory pattern that was irregular, with frequent short apneas. REM sleep was confirmed by the observation of phasic phenomena such as eye movements and whisker twitches (observed on video) [44]. The rats often assumed a curled posture before entering REM sleep. Transitions from SWS to REM sleep were identified offline; two minutes of ECoG spanning each transition point were extracted for later analysis.

9.3 Results

The numerical model was implemented in MATLAB (Mathworks, Natick, MA, USA), simulating a 2×2 -cm square of cortex on a 16×16 grid with toroidal boundaries. We used a time-step of $50 \mu\text{s}$, chosen sufficiently small to ensure numerical stability. All grid points were driven by small-amplitude spatiotemporal white noise to simulate nonspecific (unstructured) flux activity from the subcortex. The primary output was the time-course of the mean-soma potential at selected grid points on the cortical sheet.

The soma-voltage predictions for the effect of neuromodulator-induced changes in excitability (ΔV_e^{rest}), and synaptic efficiency (λ), are illustrated in the manifold of equilibrium states in Fig. 9.1. Superimposed on the figure is a gray-arrowed hypothetical trajectory that tracks the influence of increasing cholinergic tone occurring during SWS-to-REM transition. The voltage-vs-time graphs show typical samples of model-generated “pseudoECoG” time-series at three selected ($\lambda, \Delta V_e^{\text{rest}}$) coordinates representing (a) slow-wave sleep (SWS); (b) intermediate sleep (IS); and (c) REM sleep.

With no cholinergic input, the SWS pattern (a) is associated with a slow cycling (~ 0.5 to 2 Hz) between the “up” (activated, upper stable region of the manifold) and “down” (quiescent, lower stable region) states (Fig. 9.1, area “SW”).

The effect of the increasing acetylcholine is modeled as a depolarizing drift of the resting membrane potential V_e^{rest} , a slight decrease in EPSP gain, and a loss of

frequency adaptation [13]. With increasing cholinergic tone, the trajectory moves to the right to a point where the up-states are close to the area in the phase space where an ~ 8 -Hz oscillatory state exists (shaded region in Fig. 9.1). At the upper-branch stability boundary, a subcritical Hopf bifurcation occurs [48]. Just beyond this point, the up-state is no longer stable (the real part of the dominant eigenvalue is positive), so cortical excursions to the upper-state lead to oscillations in the theta-frequency band. PseudoECoG time-series generated in this unstable region show spectral features characteristic of intermediate sleep: simultaneous delta and theta oscillations (Fig. 9.1, area “IS”; time-series (b)). The delta oscillation arises from the continuing presence of the up- and down-states, and the theta oscillation from the instability of the upper branch.

As the effects of the cholinergic modulation increase further, the pseudocortex eventually becomes so depolarized that only the up-state is available. The cortex now acts as an ~ 8 -Hz narrow-band filter of the nonspecific white noise input; this is our model’s representation of the REM theta-oscillation state (Fig. 9.1, area “REM”; time-series (c)).

Because we were unable to locate descriptions in the research literature for ECoG spectra at the SWS-to-REM transition, we elected to compare our model-generated spectra with those obtained from our ECoG recordings of sleep-transitioning rats. We computed time–frequency spectrograms for the model time-series (Fig. 9.2(a)) and for the rat ECoG recordings (Fig. 9.2(b)), and also calculated time–frequency coscalograms showing the spectral coherence between a pair of adjacent grid positions on the pseudocortex (Fig. 9.3(a)), comparing this with the spectral coherence between the pair of electrodes comprising the stereotrode that sensed rat ECoG activity (Fig. 9.3(b)).¹

The spectrograms (Fig. 9.2(a, b)) and coscalograms (Fig. 9.3(a, b)) both show predominant delta activity in SWS, appearance of co-existing theta activity in IS, and abrupt loss of delta activity marking the start of REM sleep.

The two-channel coherence scalograms of Fig. 9.3 exhibit distinct frequency banding. The peak frequency of the theta oscillation in the model (Fig. 9.3(a)) starts at ~ 4.5 Hz in intermediate sleep, increasing to ~ 6.5 Hz at the transition into REM, while the rat data show a broader low-frequency spectrum with transient tongues of coherent activity that extend almost into the theta range.

Table 9.2 compares the rat and model-generated ECoG data in terms of the mean and standard deviation for two-point wavelet coherence in the delta and in the theta wavebands. Coherences for both rat and simulated data exhibit similar trends across the transition from slow-wave sleep to REM sleep: both show a more than fourfold *decrease* in delta-band coherence, simultaneous with a fourfold *increase* in theta-band coherence. Across the three sleep stages, the absolute difference in coherences between rat and model data is better than 0.2 for delta-band, and better than 0.1 for theta-band. These coherence trends are illustrated in Fig. 9.4.

¹ See Appendix (Sects 9.5.2 and 9.5.3) for details of data processing, and calculation of coherence estimates from the Morlet wavelet transform.

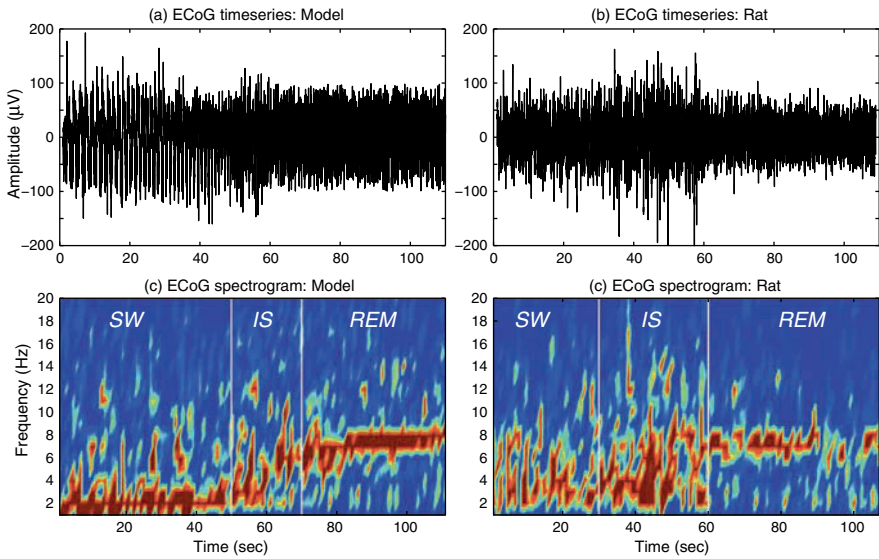


Fig. 9.2 [Color plate] Time-series and spectrograms of the model pseudoECoG signal (left), and of a typical example of rat ECoG (right) across the transition from slow-wave to REM sleep (*SW* = slow wave sleep; *IS* = intermediate sleep; *REM* = REM sleep). In both spectrograms, theta-band (5–8 Hz) activity first appears during early IS, while delta-band activity (1–4 Hz) is lost by the end of IS.

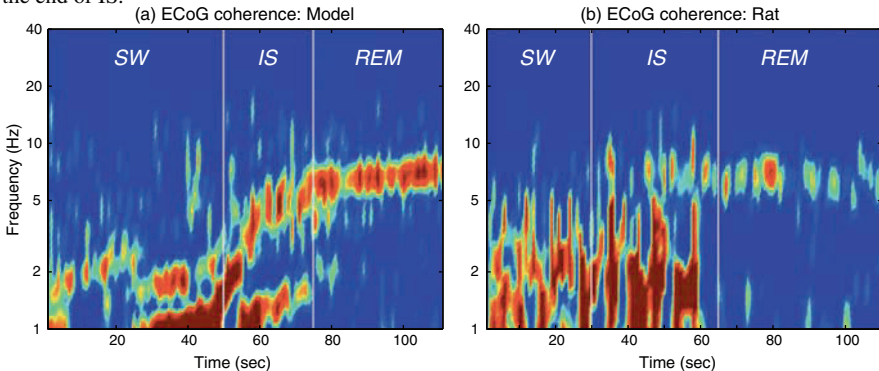


Fig. 9.3 [Color plate] Two-point temporal-coherence for two channels of pseudoECoG generated by the mean-field cortical model (left), compared with two-point coherence for rat stereotrode ECoG recording (right). Coherence is calculated using the Morlet continuous-wavelet transform (see Eq. (9.17)). In both model and experiment, there are coherent oscillations in theta- and delta-bands during the IS transition into REM sleep. (The nonlinear frequency axis is derived from the inverse of the wavelet-scale axis, and is thus distorted by the reciprocal transformation.)

9.4 Discussion

Most current neurobiological modeling of changes in cortical state involves simulation of various ion currents in assemblies of discrete Hodgkin–Huxley or

Table 9.2 Comparison of changes in two-point wavelet coherence (mean (SD)) for measured rat ECoG versus simulated pseudoECoG generated by the numerical model

Sleep stage	Delta band (1–4 Hz)		Theta band (5–8 Hz)	
	Rat	Model	Rat	Model
Slow-wave sleep	0.60 (0.02)	0.49 (0.06)	0.07 (0.06)	0.09 (0.04)
Intermediate sleep	0.48 (0.03)	0.63 (0.10)	0.28 (0.05)	0.22 (0.07)
REM sleep	0.14 (0.01)	0.09 (0.04)	0.29 (0.05)	0.36 (0.04)

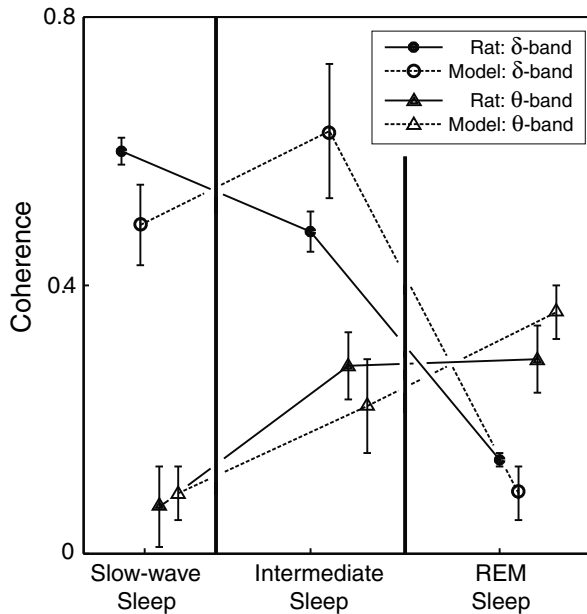


Fig. 9.4 Changes in two-point wavelet coherence across the SW-to-REM sleep transition, comparing recorded rat ECoG with model-generated ECoG. See Table 9.2 for coherence values.

integrate-and-fire neurons—the “neuron-by-neuron” approach [1, 6]. In contrast, the continuum philosophy assumes that, on average, neighboring neurons have very similar activity, so neural behavior can be approximated by population means that have been averaged over a small area of cortex. This assumption is in agreement with measured anatomic and functional spatial correlations [35]. The continuum method makes tractable the problem of quantifying global cortical phenomena, such as states of sleep, general anesthesia, and generalized seizures. Since the averaged electrical activity of populations of neurons is a commonly-measured experimental signal (ECoG), the accuracy of continuum models can be verified directly from clinical and laboratory observations, and many of the global phenomena of the cortex can be explained simply using the continuum approach. For example, if the cortex is envisaged as a network of single neurons, it is hard to explain the widespread

zero-lag spatial synchrony detected in SWS oscillations by Volgushev and colleagues [45]. In that paper, Volgushev postulated gap-junction coupling as the origin of the tight synchrony. In contrast, when the continuum formulation is used to model SWS, the high spatial synchrony arises naturally from the present set of cortical equations that do not contain diffusive couplings.

We have focused on a single question: “What is the basis for the abrupt changes in cortical activity that occur during the SWS-to-REM sleep transition?” There are two opposed explanations. One possibility is that there might be a massive, sudden increase in subcortical excitatory input into the cortex. In this scenario, the underlying cause of the change occurs *subcortically*, and the cortex is responding *linearly* to that input—this is the picture that is implicit in most earlier qualitative descriptions of the SWS-to-REM transition.

Our modeling supports the contrary view. We suggest that there is only a modest change in subcortical input occurring over a time-course of a few minutes, but that this modest change in stimulus triggers a secondary *nonlinear change in cortical self-interaction*, causing an abrupt jump to a new mode of electrophysiological (and cognitive) behavior. This cortical change of state is analogous to a phase-transition in physics, and can be described as a bifurcation in dynamical systems theory.

The earliest semi-quantitative model for SWS–REM cycling was developed by Hobson [15] who suggested that the changes in sleep state are driven by cycling in the brainstem, alternating between monoaminergic and cholinergic states. This theory did not address how the brainstem cycling determines the ECoG cortical response. More recently, Lu and others [24] have postulated that a bistable (“flip-flop”) system exists in the brainstem (the mesopontine tegmentum) to control SWS–REM transitions. The flip-flop is driven by mutual negative-feedback between “REM-on” and “REM-off” GABAergic areas. However, the primary effector pathways between the REM-on/REM-off flip-flop and the cortex were not well described. Part of the REM-on area contains glutamatergic neurons that form a *localized* ascending projection to the medial septum area of the basal forebrain, and hence to the hippocampus; thus the widespread *global* cortical activation seen in REM sleep is not explained. There is a weight of other evidence suggesting the pre-eminent role of cholinergic activation as the final common pathway of REM-sleep cortical arousal [5, 7].

The origin of the cortical theta-oscillation in the REM-state is the source of some debate. It had been assumed that the neocortical theta arises from volume conduction of the strong theta-oscillation that is observed in the hippocampus during REM sleep. However, there is evidence that the neocortical theta-rhythm may arise from the neocortex, *independent* of the hippocampus [3, 5, 20]. Cantero *et al.* [4] showed that cortical and hippocampal theta-rhythms often have different phases, and concluded that there are many different generators of theta-rhythm in the brain; the hippocampal theta is in turn dependent on the degree of activation of various brainstem structures [29].

Our theoretical model demonstrates that it is possible for an 8-Hz rhythm in the neocortex to emerge naturally from the internal dynamics of the equations. This oscillation derives from the lags introduced into the system by the inhibitory dendritic

input into the pyramidal neuronal population. Whether this is the source of all theta-oscillations observed during REM sleep in rats remains an open question. One problem with our model is that the amplitude of the theta-oscillation is sometimes as large as the slow oscillation, making the SWS-to-REM transition less clear on the pseudoECoG timeseries—although the transition remains obvious in the spectral views provided by the spectrogram and coscalogram.

As we have confirmed in the present study, one of the characteristics of the SWS–REM transition is high temporal coherence—in particular frequency bands—for spatially-separated electrodes. In rats, coherent oscillations have been observed across widespread areas of the brain that include not only cortex, but also thalamus, hippocampus and striatum [10]. Interestingly, the pattern of spindle activity observed during IS resembles that of forebrain preparations completely isolated from the brainstem [12]. In addition, examination of the amplitudes of sensory evoked-potentials shows that IS is associated with the lowest level of thalamocortical transfer of any sleep state. Thus IS is the most “introspective” of all sleep states, consistent with a role for IS in functional “in-house” coupling to support organization and integration of information between separated brain regions.

The continuum model provides a plausible theoretical basis for the phenomenon of intermediate sleep in rats. However, the model requires further experimental validation to test its predictions. We envisage a detailed and systematic experimental exploration of the model parameter space, using *in vitro* cortical-slice methods to manipulate the position of the cortical dynamics on the equilibrium manifold. Here is an example of the approach: It is known that the addition of triazolam (a GABAergic drug) markedly increases the duration of the IS stage, at the expense REM sleep [11]. We tested this finding against our continuum model by running simulations with enhanced GABAergic effect. We found that the duration of IS increases linearly with IPSP (inhibitory postsynaptic potential) *decay-time*, and increases markedly with even modest increases in the IPSP *magnitude*. Conversely, a drug which shortens IPSP decay-time will reduce the extent of the unstable “tongue” on the edge of the upper branch of the manifold, and, with sufficient reduction, should eliminate the IS stage entirely. This is an unambiguous and testable prediction.

9.5 Appendix

9.5.1 Mean-field cortical equations

The model describes interactions between cortical populations of excitatory pyramidal neurons (subscript e) and inhibitory interneurons (subscript i).² We use a left-to-right double-subscripting convention so that subscript ab (where a and b are labels standing for either e or i) implies $a \rightarrow b$, that is, the direction of transmission

² Note that we have not included a representation of the thalamus in the model, because the slow oscillation of sleep can be generated in the cortex alone—without thalamic input [33, 41, 45].

in the synaptic connections is from the presynaptic nerve a to postsynaptic nerve b . Superscript “sc” indicates subcortical input that is independent of the cortical membrane potential.

The time-evolution of the population-mean membrane potential (V_a) in each neuronal population, in response to synaptic input $\rho_a \Psi_{ab} \Phi_a$, is given by,

$$\tau_e \frac{\partial V_e}{\partial t} = V_e^{\text{rest}} + \Delta V_e^{\text{rest}} - V_e + \lambda \rho_e \Psi_{ee} \Phi_{ee} + \rho_i \Psi_{ie} \Phi_{ie}, \quad (9.5)$$

$$\tau_i \frac{\partial V_i}{\partial t} = V_i^{\text{rest}} - V_i + \lambda \rho_e \Psi_{ei} \Phi_{ei} + \rho_i \Psi_{ii} \Phi_{ii}, \quad (9.6)$$

where τ_a are the time constants of the neurons, and ρ_a are the strengths of the postsynaptic potentials (proportional to the total charge transferred per PSP event). Here we modulate the resting-voltage offset ΔV_e^{rest} and the synaptic-gain scalefactor λ as set out in Eqs (9.1–9.3). The Ψ_{ab} are weighting functions that allow for the effects of AMPA and GABA reversal potentials V_a^{rev} ,

$$\Psi_{ab} = \frac{V_a^{\text{rev}} - V_b}{V_a^{\text{rev}} - V_a^{\text{rest}}}. \quad (9.7)$$

The Φ_{ab} are synaptic-input spike-rate fluxes described by Eqs (9.8–9.9),

$$\left(\frac{\partial^2}{\partial t^2} + 2\gamma_{eb} \frac{\partial}{\partial t} + \gamma_{eb}^2 \right) \Phi_{eb} = \gamma_{eb}^2 \left(N_{eb}^\alpha \phi_{eb} + N_{eb}^\beta Q_e + \phi_{eb}^{\text{sc}} \right), \quad (9.8)$$

$$\left(\frac{\partial^2}{\partial t^2} + 2\gamma_{ib} \frac{\partial}{\partial t} + \gamma_{ib}^2 \right) \Phi_{ib} = \gamma_{ib}^2 \left(N_{ib}^\beta Q_i + \phi_{ib}^{\text{sc}} \right), \quad (9.9)$$

where γ_{ab} are the synaptic rate-constants, N^α are the number of long-range connections, and N^β the number of local, within-macrocolumn connections. The spatial interactions between macrocolumns are described by damped wave equations,

$$\left(\frac{\partial^2}{\partial t^2} + 2v\Lambda_{eb} \frac{\partial}{\partial t} + v^2\Lambda_{eb}^2 - v^2\nabla^2 \right) \phi_{eb} = v^2\Lambda_{eb}^2 Q_e, \quad (9.10)$$

where v is the mean axonal velocity, and $1/\Lambda_{eb}$ is the characteristic length-scale for axonal connections.

The population firing-rate of the neurons is related to the population-mean soma potential by a sigmoidal mapping,

$$Q_a(V_a) = \frac{Q_a^{\text{max}}}{1 + \exp[-\pi(V_a - \theta_a)/\sqrt{3}\sigma_a]} \quad (9.11)$$

where θ_a is the population-average threshold voltage, and σ_a is its standard deviation. The parameters and ranges used in our simulations are listed in Table 9.1.

9.5.2 Comparison of model mean-soma potential and experimentally-measured local-field potential

In order to translate from a grid simulation of soma potential to a pseudoECoG, we introduce three virtual electrodes that sample the fields from different sets of neurons. The first virtual electrode serves as a “common reference” that samples the local field $V_{j,k}$ from all grid-points (j,k) equally, giving a reference voltage $V^{\text{ref}}(t) = \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N V_{j,k}(t)$, where $N = 16$ in our simulations. This construction is intended to be broadly equivalent to the common reference electrodes utilized in our rat experiments—these were located bilaterally over the parietal cortex, responding to the local field potential across a wide spatial extent of neurons.

The other two virtual electrodes, numbered (1) and (2), we assume to be much more localized, sampling the field from a pair of adjacent grid-points (j,k) , $(j+1,k)$ in the simulation, but each containing a small voltage contribution, say 1%, coming from the spatial average over all 256 grid-points, leading to a pair of electrode potentials, relative to ground, of

$$\begin{aligned} V^{(1)}(t) &= V_{j,k}(t) + 0.01 V^{\text{ref}}(t), \\ V^{(2)}(t) &= V_{j+1,k}(t) + 0.01 V^{\text{ref}}(t). \end{aligned}$$

The two pseudoECoG voltages are then formed as the respective differences between $V^{(1),(2)}$ and V^{ref} ,

$$\text{pECoG}^{(1)}(t) = V_{j,k}(t) - 0.99 V^{\text{ref}}(t), \quad (9.12)$$

$$\text{pECoG}^{(2)}(t) = V_{j+1,k}(t) - 0.99 V^{\text{ref}}(t). \quad (9.13)$$

We have found that this subtractive fraction of 99% of the spatial-average soma potential produces physically reasonable pseudoECoG traces which have a strong contribution from the local voltage fluctuations at the specified grid point, but retain a weak contribution from the global activity that is common to the entire grid.

To reduce memory requirements, grid-simulation soma potentials were recorded every 250 timesteps with $\Delta t = 50 \mu\text{s}$, giving an effective sampling rate of $1/(250\Delta t) = 80 \text{ s}^{-1}$. A Butterworth highpass filter was applied to remove fluctuation energy below 0.5 Hz. ECoG voltages recorded from rat cortex were bandpass filtered by the acquisition hardware to eliminate spectral content outside the range 1–2500 Hz. The effective sampling rate was reduced from $10\,000 \text{ s}^{-1}$ to 80 s^{-1} using `decimate` to lowpass filter the time-series to 32 Hz, then subsample by a factor of 125.

9.5.3 Spectrogram and coscalogram analysis

To track the spectral changes in ECoG voltage activity over the course of the SWS-to-REM transition, we computed Hanning-windowed spectrograms with a 1-Hz

resolution using the MATLAB `pwelch` function. This spectral analysis was applied to the rat ECoG time-series (Fig. 9.2(b)), and also to the pseudoECoG time-series generated by our mean-field numerical simulations (Fig. 9.2(a)).

The synchronous interactions *between* the two ECoG channels were quantified using wavelet coherence [19, 21]. Because optimal time–frequency localization was required, we devised a new coherence measure—based on the Morlet continuous wavelet transform—to investigate the spatiotemporal relationships of the two ECoG series during the rat transition into sleep.

Given a time-function $x(t)$, its continuous wavelet transform (CWT) is defined,

$$W(s, \tau) = \frac{1}{\sqrt{s}} \int x(t) \Psi^* \left(\frac{t - \tau}{s} \right) dt, \quad (9.14)$$

where s and τ denote the temporal scale and translation respectively; $W(s, \tau)$ are the wavelet coefficients; $\Psi(t)$ is the wavelet function; and superscript (*) denotes complex conjugation. In this study, a Morlet wavelet function,

$$\Psi_0(u) = \pi^{-1/4} e^{i\omega_0 u} e^{-\frac{1}{2}u^2}, \quad (9.15)$$

is applied. Here, ω_0 is a nondimensional central angular frequency; a value of $\omega_0 = 8$ is considered optimal for good time–frequency resolution [9]. Because the Morlet wavelet retains amplitude and phase information, the degree of synchronization between neural activity simultaneously recorded at two sites can be measured.

Given a pair of ECoG time-series, X and Y , their Morlet wavelet transforms are denoted by $W_X(s, n)$ and $W_Y(s, n)$, respectively, where s is the scale and n the time-index. Their coscogram is defined

$$|W_{XY}(s, n)| \equiv |W_X(s, n) W_Y^*(s, n)|. \quad (9.16)$$

The coscogram illustrates graphically the coincident events between two time-series, at each scale s and at the each time index n . To quantify the degree of synchronization between the two time-series, we compute the wavelet coherence,

$$[\text{coh}(s, n)]^2 = \frac{|\langle s^{-1} W_{XY}(s, n) \rangle|^2}{\langle s^{-1} |W_{XX}(s, n)|^2 \rangle \langle s^{-1} |W_{YY}(s, n)|^2 \rangle}. \quad (9.17)$$

The coherence ranges from 0 to 1, and provides an accurate representation for the covariance between two EEG time-series. The angle-brackets $\langle \cdot \rangle$ indicate smoothing in time and scale; the internal factor s^{-1} is required to convert to an energy density. The smoothing in time is achieved using a Gaussian function $\exp(-\frac{1}{2}(t/s)^2)$; the smoothing in scale is done using a boxcar filter of width 0.6 (see Ref. [42]). Because of the smoothing, wavelet coherence effectively provides an ensemble averaging localized in time, thereby reducing the variance from noise [19].

References

1. Bazhenov, M., Timofeev, I., Steriade, M., Sejnowski, T.J.: Model of thalamocortical slow-wave sleep oscillations and transitions to activated states. *J. Neurosci.* **22**(19), 8691–704 (2002)
2. Benington, J.H., Kodali, S.K., Heller, H.C.: Scoring transitions to REM sleep in rats based on the EEG phenomena of pre-REM sleep: an improved analysis of sleep structure. *Sleep* **17**(1), 28–36 (1994)
3. Borst, J., Leung, L.W., MacFabe, D.: Electrical activity of the cingulate cortex. ii. cholinergic modulation. *Brain Res* **407**, 81–93. (1987), doi:10.1016/0006-8993(87)91221-2
4. Cantero, J., Atienza, M., Stickgold, R., Kahana, M., Madsen, J., Kocsis, B.: Sleep-dependent theta oscillations in the human hippocampus and neocortex. *J. Neurosci.* **23**(34), 10897–10903 (2003)
5. Cape, E., Jones, B.: Effects of glutamate agonist versus procaine microinjections into the basal forebrain cholinergic cell area upon gamma and theta EEG activity and sleep-wake state. *Eur. J. Neurosci.* **12**, 2166–2184 (2000), doi:10.1046/j.1460-9568.2000.00099.x
6. Compte, A., Sanchez-Vives, M.V., McCormick, D.A., Wang, X.J.: Cellular and network mechanisms of slow oscillatory activity (<1 Hz) and wave propagations in a cortical network model. *J. Neurophys.* **89**(5), 2707–2725 (2003), doi:10.1152/jn.00845.2002
7. Datta, S., Siwek, D.F.: Excitation of the brainstem pedunculopontine tegmentum cholinergic cells induces wakefulness and REM sleep. *J. Neurophys.* **77**(6), 2975–2988 (1997)
8. El Mansari, M., Sakai, K., Jouvet, M.: Unitary characteristics of presumptive cholinergic tegmental neurons during the sleep-waking cycle in freely moving cats. *Exp. Brain Res.* **76**, 519–529 (1989), doi:10.1007/BF00248908
9. Farge, M.: Wavelet transforms and their applications to turbulence. *Ann. Rev. Fluid. Mech.* **24**, 395–457 (1992), doi:10.1146/annurev.fl.24.010192.002143
10. Gervasoni, D., Lin, S.C., Ribeiro, S., Soares, E.S., Pantoja, J., Nicolelis, M.A.: Global forebrain dynamics predict rat behavioral states and their transitions. *J. Neurosci.* **24**(49), 11137–11147 (2004), doi:10.1523/jneurosci.3524-04.2004
11. Gottesmann, C.: The transition from slow-wave sleep to paradoxical sleep: Evolving facts and concepts of the neurophysiological processes underlying the intermediate stage of sleep. *Neurosci. Biobehav. Rev.* **20**(3), 367–387 (1996), doi:10.1016/0149-7634(95)00055-0
12. Gottesmann, C., Gandolfo, G., Arnaud, C., Gauthier, P.: The intermediate stage and paradoxical sleep in the rat: Influence of three generations of hypnotics. *Eur. J. Neurosci.* **10**(2), 409–14 (1998), doi:10.1046/j.1460-9568.1998.00069.x
13. Hasselmo, M., McGaughy, J.: High acetylcholine levels set circuit dynamics for attention and encoding and low acetylcholine levels set dynamics for consolidation. *Prog. Brain Res.* **145**, 207–231 (2004), doi:10.1016/S0079-6123(03)45015-2
14. Hill, S., Tononi, G.: Modeling sleep and wakefulness in the thalamocortical system. *J. Neurophysiol.* **93**, 1671–1698 (2005), doi:10.1152/jn.00915.2004
15. Hobson, J.A., McCarley, R.W., Wyzinski, P.W.: Sleep cycle oscillation: Reciprocal discharge by two brainstem neuronal groups. *Science* **189**(4196), 55–58 (1975), doi:10.1126/science.1094539
16. Hobson, J.A., Pace-Schott, E.F.: The cognitive neuroscience of sleep: Neuronal systems, consciousness and learning. *Nat. Rev. Neurosci.* **3**(9), 679–693 (2002), doi:10.1038/nrn915
17. Holcman, C., Tsodyks, M.: The emergence of up and down states in cortical networks. *PLoS. Comput. Biol.* **2**(3), 174–181 (2006), doi:10.1371/journal.pcbi.0020023
18. Jones, B.: Activity, modulation and role of basal forebrain cholinergic neurons innervating the cerebral cortex. *Prog. Brain Res.* **145**, 157–169 (2004), doi:10.1016/S0079-6123(03)45011-5
19. Lachaux, J.P., Rudrauf Lutz, A., Cosmelli, D., Le Van Quyen, M., Martinerie, J., Varela, F.: Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence. *Clin. Neurophysiol.* **32**(3), 157–174 (2002), doi:10.1016/S0987-7053(02)00301-5

20. Lee, M., Hassani, O., Alonso, A., Jones, B.: Cholinergic basal forebrain neurons burst with theta during waking and paradoxical sleep. *J. Neurosci.* **25**(17), 4365–4369. (2005), doi:10.1523/jneurosci.0178-05.2005
21. Leinekugel, X., Khazipov, R., Cannon, R., Hirase, H., Ben-Ari, Y., Buzsáki, G.: Correlated bursts of activity in the neonatal hippocampus in vivo. *Science* **296**, 2049–2052 (2002), doi:10.1126/science.1071111
22. Liley, D., Cadusch, P., Wright, J.: A continuum theory of electro-cortical activity. *Neurocomp.* **26-27**, 795–800 (1999), doi:10.1016/S0925-2312(98)00149-0
23. Linster, C., Hasselmo, M.E.: Neuromodulation and the functional dynamics of piriform cortex. *Chem. Senses* **26**(5), 585–594 (2001), doi:10.1093/chemse/26.5.585
24. Lu, J., Sherman, D., Devor, M., Saper, C.: A putative flip-flop switch for control of REM sleep. *Nature* **441**, 583–591. (2006), doi:10.1038/nature04767
25. Mandile, P., Vescia, S., Montagnese, P., Romano, F., Giuditta, A.: Characterization of transition sleep episodes in baseline EEG recordings of adult rats. *Physiol. Behav.* **60**(6), 1435–9 (1996), doi:10.1016/S0031-9384(96)00301-0
26. Massimini, M., Huber, R., Ferrarelli, F., Hill, S., Tononi, G.: The sleep slow oscillation as a traveling wave. *J. Neurosci.* **24**(31), 6862–6870 (2004), doi:10.1523/jneurosci.1318-04.2004
27. Metherate, R., Cox, C.L., Ashe, J.H.: Cellular bases of neocortical activation: modulation of neural oscillations by the nucleus basalis and endogenous acetylcholine. *J. Neurosci.* **12**(12), 4701–4711 (1992)
28. Morrissey, M.J., Anch, A.M., Duntley, S.P.: An evaluation of the use of seizure prone rats when investigating intermediate stage sleep. *J. Neurosci. Meth.* **143**(2), 159–162 (2005), doi:10.1016/j.jneumeth.2004.09.026
29. Nunez, A., Cervera-Ferri, A., Olucha-Bordonau, F., Ruiz-Torna, A., Teruel, V.: Nucleus incertus contribution to hippocampal theta rhythm generation. *Eur. J. Neurosci.* **23**, 2731–2738 (2006), doi:10.1111/j.1460-9568.2006.04797.x
30. Pace-Schott, E.F., Hobson, J.A.: The neurobiology of sleep: genetics, cellular physiology and subcortical networks. *Nat. Rev. Neurosci.* **3**(8), 591–605 (2002)
31. Porkka-Heiskanen, T., Alanko, L., Kalinchuk, A., Stenberg, D.: Adenosine and sleep. *Sleep Med. Rev.* **6**(4), 321–332 (2002), doi:10.1053/smr.2001.0201
32. Saito, H., Sakai, K., Jouvet, M.: Discharge patterns of the nucleus parabrachialis lateralis neurons of the cat during sleep and waking. *Brain Res.* **134**(1), 59–72 (1977), doi:10.1016/0006-8993(77)90925-8
33. Sanchez-Vives, M.V., McCormick, D.A.: Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nat. Neurosci.* **3**(10), 1027–1034 (2000), doi:10.1038/79848
34. Saper, C., Scammell, T., Lu, J.: Hypothalamic regulation of sleep and circadian rhythms. *Nature* **437**, 1257–1263 (2005), doi:10.1038/nature04284
35. Shepherd, G., Stepanyants, A., Bureau, I., Chklovskii, D., Svoboda, K.: Geometric and functional organization of cortical circuits. *Nat. Neurosci.* **8**(6), 782–791 (2005), doi:10.1038/nn1447
36. Shu, Y., Hasenstaub, A., McCormick, D.A.: Turning on and off recurrent balanced cortical activity. *Nature* **423**(6937), 288–293 (2003), doi:10.1038/nature01616
37. Steriade, M., Datta, S., Pare, D., Oakson, G., Curro Dossi, R.C.: Neuronal activities in brainstem cholinergic nuclei related to tonic activation processes in thalamocortical systems. *J. Neurosci.* **10**(8), 2541–2559 (1990)
38. Steriade, M., Timofeev, I., Grenier, F.: Natural waking and sleep states: A view from inside neocortical neurons. *J. Neurophys.* **85**(5), 1969–1985 (2001)
39. Steyn-Ross, D.A., Steyn-Ross, M.L., Sleigh, J.W., Wilson, M.T., Gillies, I.P., Wright, J.J.: The sleep cycle modelled as a cortical phase transition. *J. Biol. Phys.* **31**, 547–569 (2005), doi:10.1007/s10867-005-1285-2
40. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W., Wilson, M.T., Wilcocks, L.C.: Proposed mechanism for learning and memory erasure in a white-noise-driven sleeping cortex. *Phys. Rev. E* **72**, 061910 (2005), doi:10.1103/PhysRevE.72.061910

41. Timofeev, I., Grenier, F., Steriade, M.: Impact of intrinsic properties and synaptic factors on the activity of neocortical networks in vivo. *J. Physiology-Paris* **94**(5-6), 343–355 (2000), doi:10.1016/S0928-4257(00)01097-4
42. Torrence, C., Webster, P.J.: Interdecadal changes in the ENSO-monsoon system. *J. Climate* **12**(8), 2679–2690 (1999), doi:10.1175/1520-0442(1999)012<2679:icitem>2.0.CO;2
43. Tsodyks, M.V., Markram, H.: The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *P. Natl. Acad. Sci. USA* **94**(2), 719–723 (1997), doi:10.1073/pnas.94.2.719
44. van Beteray, J.N., Vossen, J.M., Coenen, A.M.: Behavioural characteristics of sleep in rats under different light/dark conditions. *Physiol. Behav.* **50**(1), 79–82 (1991), doi:10.1016/0031-9384(91)90501-E
45. Volgushev, M., Chauvette, S., Mukovski, M., Timofeev, I.: Precise long-range synchronization of activity and silence in neocortical neurons during slow-wave sleep. *J. Neurosci.* **26**(21), 5665–5672 (2006), doi:10.1523/jneurosci.0279-06.2006
46. Wilson, H.R., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**(1), 1–24 (1972)
47. Wilson, M.T., Sleigh, J.W., Steyn-Ross, D.A., Steyn-Ross, M.L.: General anaesthetic-induced seizures can be explained by a mean-field model of cortical dynamics. *Anesthesiology* **104**(3), 588–593 (2006), doi:10.1097/00000542-200603000-00026
48. Wilson, M.T., Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W.: Predictions and simulations of cortical dynamics during natural sleep using a continuum approach. *Phys. Rev. E* **72**(5), 051910 (2005), doi:10.1103/PhysRevE.72.051910
49. Wilson, M., Steyn-Ross, D., Sleigh, J., Steyn-Ross, M., Wilcocks, L., Gillies, I.: The K-complex and slow oscillation in terms of a mean-field cortical model. *J. Comput. Neurosci.* **21**, 243–257 (2006), doi:10.1007/s10827-006-7948-6
50. Wright, J.J., Liley, D.T.: Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. *Behav. Brain Sci.* **19**, 285–320 (1996)

Chapter 10

What can a mean-field model tell us about the dynamics of the cortex?

M.T. Wilson, M.L. Steyn-Ross, D.A. Steyn-Ross, J.W. Sleigh, I.P. Gillies, and D.J. Hailstone

10.1 Introduction

In this chapter we consider the dynamics of the cortex. We will show that mean-field models can be applied to predict and explain large-scale features of the electroencephalogram (EEG), such as seizures and K-complexes. The changes between these states can be viewed as phase transitions. A cortical mean-field model is perhaps well suited for investigating EEG behavior, given that the EEG is a result of sampling large numbers of neurons. The model we consider follows the continuum approach introduced by Nunez [12] and Freeman [2]. This style of model has been developed by, amongst others, Wright and Liley [23], Robinson *et al.* [16], Liley *et al.* [9], and Rennie *et al.* [14]. Here we implement a two-dimensional model in the manner of Liley *et al.* [9], that specifically incorporates the dynamics of neuron somas, synapses and axonal propagation.

Mean-field models have a history of being used to describe phase transitions—for example, the Weiss model of ferromagnetism. Although it is known that mean-field models are lacking in their description of correlations (and these can influence the critical exponents of second-order phase transitions), such a basis is a natural place to begin for describing phase transitions in the cortex. In the limit of spatially symmetric perturbations, the model exhibits stable and unstable nodes, and undergoes Hopf bifurcations into limit cycles. We identify these limit cycles as seizures. If the symmetry is relaxed, traveling-wave solutions can be excited, reminiscent of the slow-waves and K-complexes found in sleep. These solutions are related to the

Marcus T. Wilson · Moira L. Steyn-Ross · D. Alistair Steyn-Ross · Iain-Patrick Gillies · David J. Hailstone

Dept of Engineering, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand.
e-mail: m.wilson@waikato.ac.nz msr@waikato.ac.nz asr@waikato.ac.nz

Jamie W. Sleigh

Waikato Clinical School, University of Auckland, Waikato Hospital, Hamilton 3204, New Zealand.
e-mail: sleighj@waikatodhb.govt.nz

spatially symmetric limit cycles. Also, spiral-wave solutions have been found from numerical simulation.

10.2 A mean-field model of the cortex

The cortex consists of a large number of macrocolumns, each containing around 10^5 neurons in a volume of $\sim 1 \text{ mm}^2$ area by $\sim 1 \text{ mm}$ thickness. We model the cortex as a two-dimensional continuous area of macrocolumns. We follow the mean-field continuum approach of Liley *et al.* [9] in using a set of coupled differential equations in time and space to describe the excitatory and inhibitory soma potentials V_e and V_i within the cortex, and the time-evolution of postsynaptic potentials (PSPs). This approach, in which the postsynaptic fluxes Φ_{jk} (where j and k can correspond to excitatory e or inhibitory i neuron populations) are described by differential equations, is equivalent to approaches that describe the build-up of potentials in terms of time-integrated inputs, for example that of Jirsa and Haken [6]. We use a standard wave-equation, in the manner of Robinson *et al.* [16], to describe the propagation of presynaptic fluxes ϕ_{jk} from one part of the cortex to another. We model the subcortical input with white noise.

It is important to note that individual firing events are not modeled explicitly, instead the effects of a firing of a *population* of neurons are considered.

The complete set of equations describing the macrocolumn averages of soma potential and synaptic fluxes, as a function of space and time, are described in Refs [18], [19], and [22], and are given below. Note that later, Eqs (10.1) and (10.2) will be modified slightly by the introduction of neuromodulators into the model.

$$\tau_e \frac{dV_e}{dt} = V_e^{\text{rest}} - V_e + \rho_e \psi_{ee} \Phi_{ee} + \rho_i \psi_{ie} \Phi_{ie}, \quad (10.1)$$

$$\tau_i \frac{dV_i}{dt} = V_i^{\text{rest}} - V_i + \rho_e \psi_{ei} \Phi_{ei} + \rho_i \psi_{ii} \Phi_{ii}, \quad (10.2)$$

$$\left(\frac{d^2}{dt^2} + 2\gamma_{ee} \frac{d}{dt} + \gamma_{ee}^2 \right) \Phi_{ee} = \gamma_{ee}^2 \left(N_{ee}^\alpha \phi_{ee} + N_{ee}^\beta Q_e + \phi_{ee}^{\text{sc}} \right), \quad (10.3)$$

$$\left(\frac{d^2}{dt^2} + 2\gamma_{ei} \frac{d}{dt} + \gamma_{ei}^2 \right) \Phi_{ei} = \gamma_{ei}^2 \left(N_{ei}^\alpha \phi_{ei} + N_{ei}^\beta Q_e + \phi_{ei}^{\text{sc}} \right), \quad (10.4)$$

$$\left(\frac{d^2}{dt^2} + 2\gamma_{ie} \frac{d}{dt} + \gamma_{ie}^2 \right) \Phi_{ie} = \gamma_{ie}^2 \left(N_{ie}^\beta Q_i + \phi_{ie}^{\text{sc}} \right), \quad (10.5)$$

$$\left(\frac{d^2}{dt^2} + 2\gamma_{ii} \frac{d}{dt} + \gamma_{ii}^2 \right) \Phi_{ii} = \gamma_{ii}^2 \left(N_{ii}^\beta Q_i + \phi_{ii}^{\text{sc}} \right), \quad (10.6)$$

$$\left(\frac{\partial^2}{\partial t^2} + 2v\Lambda_{ee} \frac{\partial}{\partial t} + v^2\Lambda_{ee}^2 - v^2\nabla^2 \right) \phi_{ee} = v^2\Lambda_{ee}^2 Q_e, \quad (10.7)$$

$$\left(\frac{\partial^2}{\partial t^2} + 2v\Lambda_{ei} \frac{\partial}{\partial t} + v^2\Lambda_{ei}^2 - v^2\nabla^2 \right) \phi_{ei} = v^2\Lambda_{ei}^2 Q_e. \quad (10.8)$$

In these equations, V_e^{rest} and V_i^{rest} are the excitatory and inhibitory neuron resting potentials, and ρ_e and ρ_i are the strengths of the EPSP and IPSP response functions (i.e., the *area* of the plot of PSP response function against time). Note that the inhibitory effect is modeled with a *negative* ρ_i . The variables ψ_{ee} , ψ_{ei} , ψ_{ie} and ψ_{ii} are weighting functions dependent upon the soma potentials. They are given by:

$$\psi_{ab} = \frac{V_a^{\text{rev}} - V_b}{V_a^{\text{rev}} - V_b^{\text{rest}}}. \quad (10.9)$$

Here, V_a^{rev} is the reversal potential of the type a synapse, due to the concentrations of the neurotransmitters AMPA and GABA. The suffices a and b can take on the labels e and i .

The terms τ_e and τ_i describe the time-constants for the e and i neurons. The γ_{ab} terms are synaptic rate-constants; their reciprocals give the time-scales over which the EPSPs and IPSPs occur. The mean axonal velocity for long-range interactions is given by v , and the characteristic length for long-range interactions is given by $1/\Lambda_{ea}$. Short-range interactions are not modeled with axonal propagation but are assumed to be instantaneous; the N_{jk}^β terms couple directly with the population firing rates Q_k in equations (10.3)–(10.6).

The sigmoidal functions Q_e and Q_i , describing the population firing-rate of neurons, are given by:

$$Q_e(V_e) = \frac{Q_e^{\text{max}}}{1 + \exp[-\pi(V_e - \theta_e)/\sqrt{3}\sigma_e]}, \quad (10.10)$$

$$Q_i(V_i) = \frac{Q_i^{\text{max}}}{1 + \exp[-\pi(V_i - \theta_i)/\sqrt{3}\sigma_i]}. \quad (10.11)$$

Here we have introduced further parameters Q_e^{max} and Q_i^{max} , the maximum firing rates for the excitatory and inhibitory neurons respectively; θ_e and θ_i , the inflexion-point voltage; and σ_e and σ_i , the standard deviation of the threshold potential.

Finally, the N_{ab}^β represent numbers of *local* intra-macrocolumn connections from type a neurons to type b (again a and b can take on the labels e and i), and the N_{ea}^α represent the number of *long-range* connections from type e neurons to type a (note that inhibitory neurons have no long-range projections). Noise enters the model through the ϕ_{ab}^{sc} terms.

The list of standard parameters used for a human cortex is given in Table 10.1.

We now modify the equations by introducing two parameters to describe the effects of neuromodulators. Following Steyn-Ross *et al.* [18], we introduce a term ΔV_e^{rest} to the excitatory resting potential in Eq. (10.1) to model the archetypal

Table 10.1 The standard parameters for a human cortex used throughout this chapter, except where stated otherwise. In this table the suffix a can take on the labels e and i . The values are taken mostly from the paper of Rennie *et al.* [14]. Although there is considerable uncertainty in these parameters, they form a plausible set that is sufficient for the purposes of elucidating much of the physics of the cortical model. It is quite possible that further physical effects can be produced by varying these parameters sufficiently.

Parameter	Description	Standard value
$\tau_{e,i}$	membrane time constants	0.04, 0.04 s ⁻¹
$Q_{e,i}$	maximum firing rates	30, 60 s ⁻¹
$\theta_{e,i}$	sigmoid thresholds	-58.5, -58.5 mV
$\sigma_{e,i}$	standard deviation for threshold	4.0, 6.0 mV
$\rho_{e,i}$	gain per synapse at resting voltage	0.001, -0.00105 mV · s
$V_{e,i}^{\text{rev}}$	reversal potentials at synapse	0, -70 mV
$V_{e,i}^{\text{rest}}$	cell resting potential	-64, -64 mV
N_{ea}^{α}	long-range e to e or i connectivity	3710
N_{ea}^{β}	short-range e to e or i connectivity	410
N_{ia}^{β}	short-range i to e or i connectivity	800
$\langle \phi_{ea}^{\text{sc}} \rangle$	mean e to e or i subcortical flux	750 s ⁻¹
$\langle \phi_{ia}^{\text{sc}} \rangle$	mean i to e or i subcortical flux	1500 s ⁻¹
γ_{ea}	excitatory synaptic rate constant	300 s ⁻¹
γ_{ia}	inhibitory synaptic rate constant	65 s ⁻¹
$L_{x,y}$	spatial length of cortex in model	500 mm
a_{mc}	area of macrocolumn	1 mm ²
Λ_{ea}	characteristic inverse length-scale for connections	0.2 mm ⁻¹
v	mean axonal conduction speed	1400 mm s ⁻¹

somnogen adenosine, and a multiplier λ to the excitatory synaptic strength to account for the effect of acetylcholine (abundant in rapid-eye-movement (REM) sleep, but absent in slow-wave sleep). Equations (10.1) and (10.2) now become:

$$\tau_e \frac{dV_e}{dt} = V_e^{\text{rest}} + \Delta V_e^{\text{rest}} - V_e + \lambda \rho_e \psi_{ee} \Phi_{ee} + \rho_i \psi_{ie} \Phi_{ie}, \quad (10.12)$$

$$\tau_i \frac{dV_i}{dt} = V_i^{\text{rest}} - V_i + \lambda \rho_e \psi_{ei} \Phi_{ei} + \rho_i \psi_{ii} \Phi_{ii}. \quad (10.13)$$

10.3 Stationary states

The full two-dimensional dynamics of the model will be explored by simulation; however, it is instructive to calculate the stationary (equilibrium) states of the model (i.e., solving for the state variables V_j , Φ_{jk} , ϕ_{ej} , where $j, k = e$ or i) such that their time- and spatial-derivatives are zero. The solutions are computed numerically as a function of the two parameters ΔV_e^{rest} and λ . Figure 10.1 is a plot of the solution

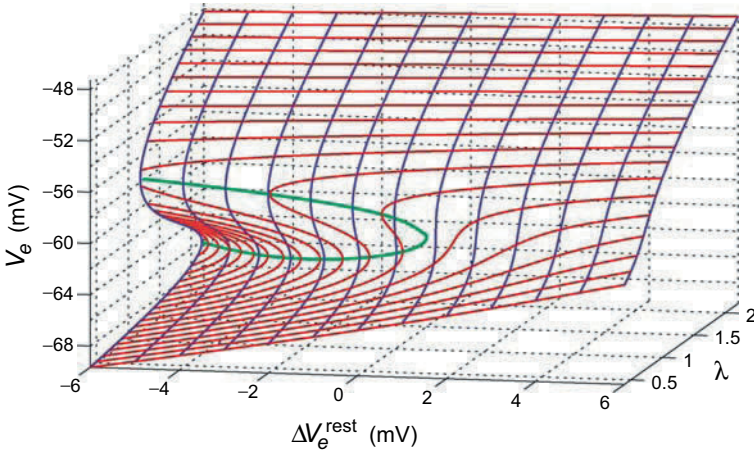


Fig. 10.1 Plot of V_e for stationary states in the sleep domain, using Table 10.1 parameter values. Multiple stationary states exist in a region of the sleep domain at negative ΔV_e^{rest} . The thick-green curve marks turning points (where the gradient is infinite). (Modified from Ref. [18]).

for V_e against these two parameters (we could also construct associated plots for V_i , Φ_{ee} , etc). It is immediately clear that there exists a region where there are multiple stationary states. This is suggestive of the presence of phase transitions—e.g., in Ref. [18], Steyn-Ross *et al.* identified the lower (more negative V_e) branch as slow-wave sleep (SWS) and the upper branch as REM sleep.

10.4 Hopf bifurcations

In this section we investigate the stability of the stationary solutions. Specifically, we will show the presence of Hopf bifurcations that lead to instabilities in the stationary states. Where there is no stable *stationary* state, a stable spatially-symmetric limit-cycle *oscillatory* state exists. The dynamic solutions to the equations are presented in the next section.

10.4.1 Stability analysis

To carry out a stability analysis, we decompose the six second-order equations (Eqs. (10.3)–(10.8)) into pairs of first-order equations, which, along with Eqs. (10.12) and (10.13), give a total of fourteen coupled first-order differential equations in time. Then we perform a first-order series expansion in time about the stationary state. We define uniquely any state of the system by its 14-dimensional state-vector

$\mathbf{y}(\mathbf{r})$, which contains the variables $V_e, V_i, \Phi_{ee}, d\Phi_{ee}/dt, \Phi_{ei}, d\Phi_{ei}/dt, \Phi_{ie}, d\Phi_{ie}/dt, \Phi_{ii}, d\Phi_{ii}/dt, \phi_{ee}, d\phi_{ee}/dt, \phi_{ei}$, and $d\phi_{ei}/dt$. These are all functions of space (\mathbf{r}), and Eqs. (10.3)–(10.8), (10.12) and (10.13) describe explicitly how these evolve with time t . We then assume a small plane-wave perturbation of the system in two-dimensional space of wave vector \mathbf{q} , about the stationary point \mathbf{y}_{eqm} . That is,

$$\mathbf{y}(\mathbf{r}) = \mathbf{y}_{\text{eqm}} + \Delta\mathbf{y}^{\mathbf{q}} \exp(i\mathbf{q}\cdot\mathbf{r}), \quad (10.14)$$

where the dynamics can be written as a simple matrix equation

$$\frac{d}{dt}(\Delta\mathbf{y}^{\mathbf{q}}) = A(q)\Delta\mathbf{y}^{\mathbf{q}}. \quad (10.15)$$

Here, A is a sparse 14×14 matrix which contains all the dynamics of the system; it is dependent upon the modulus of the wave vector $q = |\mathbf{q}|$. The vector $\Delta\mathbf{y}^{\mathbf{q}}$ denotes the amplitude of the plane-wave perturbation.

In order for the system to be stable to small perturbations about a given stationary point, we require the real parts of all the eigenvalues of A to be negative, for all q . We do not present the matrix A here, but have described the method in more detail in Ref. [22].

The choice of wave vector q is important, since some eigenvalues are particularly sensitive to q . In all cases, we find that increasing q *decreases* the real part of the eigenvalues (i.e., makes them less positive or more negative). The perturbations that are most likely to lead to instabilities are therefore those with $q = 0$, corresponding to spatially symmetric “breathing” modes. This observation agrees with that made by Robinson *et al.* using a different model [15], and is confirmed by simulation. In what follows, we assess the stability of the system by constructing the matrix $A(q = 0)$, then finding its eigenvalues. If one or more eigenvalues has a positive real part, then Eq. (10.15) predicts an instability.

10.4.2 Stability of the stationary states

We begin by presenting a typical result for the stability of the sleep domain.

Figure 10.1 showed graphically the stationary values for V_e : across most of the domain, there is just one stationary state; however, for a small region there are three stationary states. Figure 10.2 shows the stability of these states. There is a small lake-like region where the linearized system has a pair of eigenvalues with positive real part, corresponding to an instability. The nature of this region is discussed in Ref. [22]; notably the edge of the lake-like region corresponds to a supercritical Hopf bifurcation. The real parts of these eigenvalues are most positive when $q = 0$, and reduce monotonically as q increases.

Let us now comment on how the situation changes with a different parameter set. We find that the size of the unstable “lake” of Fig. 10.2 is particularly dependent on the choice of γ_i (we assume $\gamma_{ie} = \gamma_{ii}$, so we can use a single index without ambiguity). This is to be expected since it is this term that governs the rate at which

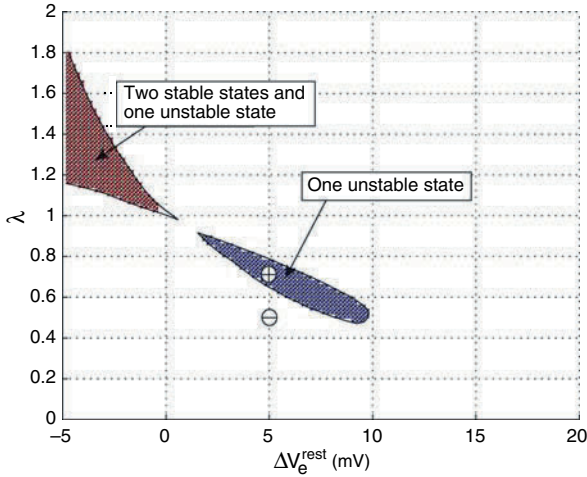


Fig. 10.2 Plot of the sleep domain showing a region of multiple stationary states (shaded area to upper-left), and a region of single, unstable stationary states (shaded “lake”), for the parameters of Table 10.1. The system is *stable* in the unshaded regions (e.g., at the point marked \ominus), and *unstable* at the point marked \oplus . (Reprinted from [22] with permission.)

negative (stabilizing) feedback is applied to the system. A small γ_i corresponds to an IPSP that is spread-out in time; this allows positive feedback from the excitatory–excitatory route to build up before being quenched. Figure 10.3 shows the stability of the domain when γ_i is reduced from 65 s^{-1} to 15 s^{-1} . The stationary solutions remain unchanged; however, the unstable lake has now grown into the area of three stationary states. This means that either or both of the upper and lower branches, in addition to the mid-branch, can become unstable. There is the possibility of the system having multiple stationary states, with none of them being stable.

The limit cycles associated with these instabilities are found by simulation, and are described in the next section.

10.5 Dynamic simulations

Numerical simulations of the equations on a two-dimensional grid allow the exploration of a range of different dynamic behaviors. We first outline how this is done, then consider spatially symmetric and asymmetric solutions. We tentatively suggest how these solutions might correspond with gross features of the EEG. The parameter values used here are broadly representative of human cortex; parameters and bifurcation structures more appropriate for rat cortex are discussed in the chapter by Sleigh *et al.*¹

¹ See Ch. 9 of the present volume.

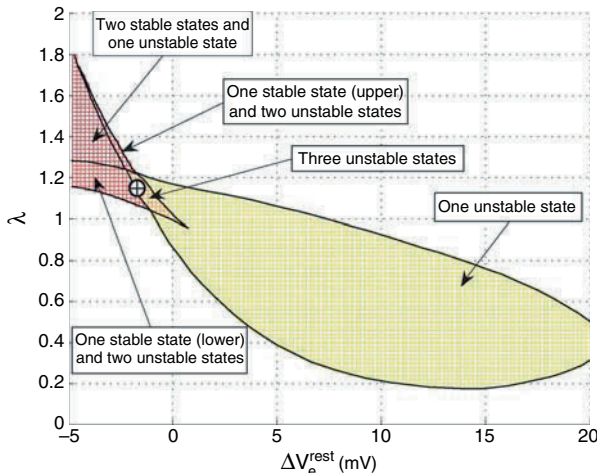


Fig. 10.3 Stability of the sleep domain for a decrease in γ_i to 15 s^{-1} . The unstable lake on the right has grown into the region of multiple stationary states. Five possibilities exist: one stable state; one unstable state; three unstable states (e.g., at point \oplus); one stable state and two unstable; two stable states and one unstable. (The mid-branch state is always unstable.) The unmarked region of the figure corresponds to a single, stable state. (Reprinted from [22] with permission.)

In one dimension, a study of spatial and temporal structures such as traveling waves in a similar model has been performed by Hutt *et al.* [5] using both a stability analysis similar to that in Sect. 10.4.1, and numerical simulations.

We have performed grid simulations on the full set of equations in 2-D space. We used a regular square grid to represent the cortex in space, and iterated the full equations in time using an order-1 Euler predictor–corrector method [7], as detailed in reference [22]. Noise is fed into the system through the subcortical terms, but does not affect gross dynamic behaviors that dominate when there is no stable stationary state available.

10.5.1 Breathing modes

We now look at the stability and dynamics of the system in more detail. We consider vertical slices across the domains of Figs. 10.2 and 10.3: we fix ΔV_e^{rest} , and show the mean soma potential as a function of λ . The slice-plot for Fig. 10.2 (with $\gamma_i = 65 \text{ s}^{-1}$) is shown in Fig. 10.4; and the slice-plot for Fig. 10.3 ($\gamma_i = 15 \text{ s}^{-1}$) is shown in Fig. 10.5. Also indicated on these plots is the stability of the solution (solid = stable, dashed = unstable), and the extent of “breathing mode” (spatially symmetric) limit cycles. The breathing-mode cycle, corresponding to $q = 0$, is always the most unstable perturbation, and is usually the cycle reached in the limit of large times. The dominance of the $q = 0$ instability arises because the model

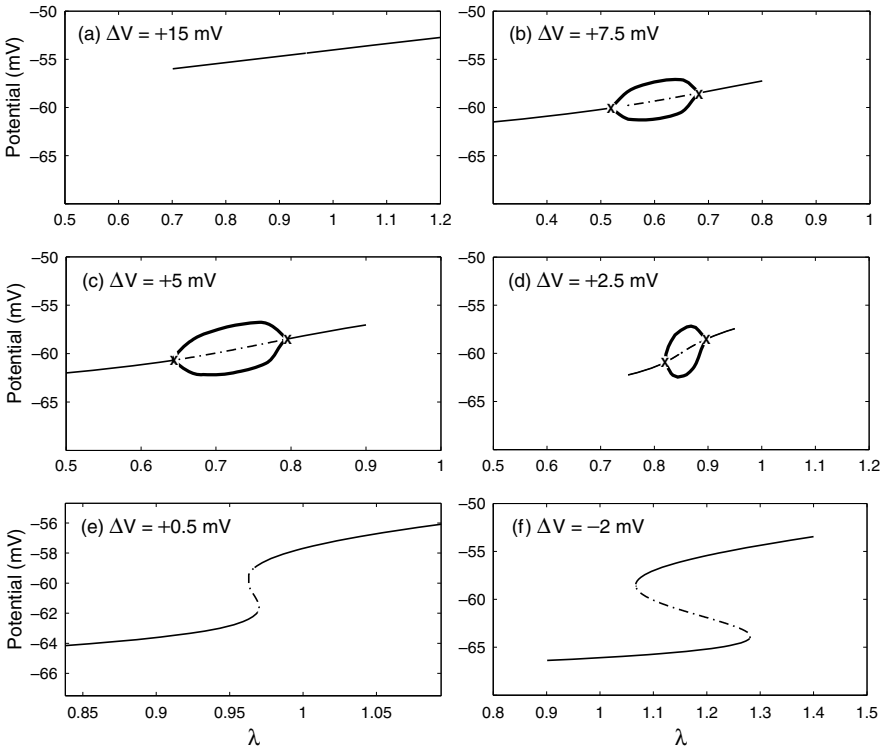


Fig. 10.4 Stationary states, bifurcations (found by calculation), and breathing modes (found by simulation) corresponding to six vertical slices through Fig. 10.2 with $\gamma_i = 65 \text{ s}^{-1}$. Graphs plot excitatory soma potential V_e versus excitatory synaptic strength λ for six values of ΔV_e^{rest} (displayed in the subtitles as ΔV). Thin line indicates stationary states of V_e : solid = stable; dotted = unstable. Supercritical Hopf bifurcations are marked with \times . Thick-solid line denotes maximum and minimum extent of the associated limit cycle (only stable part of the limit cycle is shown).

used in this chapter has no spatial diffusion; when diffusive coupling in space is introduced—for example, via gap-junctions—instabilities at nonzero q can dominate, as discussed in the chapter by Steyn-Ross *et al.*²

The limit cycles were found by simulation, and are indicated by plotting thick-solid lines in Figs 10.4 and 10.5 to show the maximum and minimum values of the mean excitatory soma potential during the cycle; unstable limit cycles associated with these Hopf bifurcations are not shown.

We start with Fig. 10.4 ($\gamma_i = 65 \text{ s}^{-1}$), and discuss the solutions with reference to stepped decreases in the ΔV_e^{rest} parameter. For ΔV_e^{rest} larger than about +10 mV, there is a single stationary state, and this is stable (e.g., see Fig. 10.4(a)). When ΔV_e^{rest} is reduced, a small unstable region appears, associated with a pair of Hopf bifurcations (marked with “ \times ” on Fig. 10.4(b, c, d)). If a simulation is performed

² See Ch. 12 of the present volume.

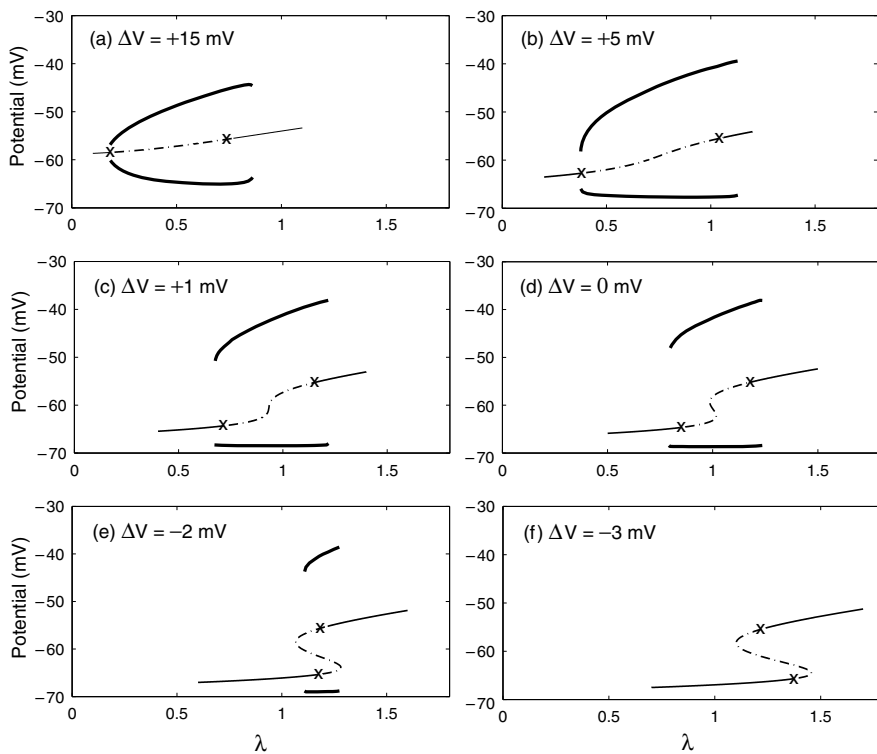


Fig. 10.5 Stationary states, bifurcations, and breathing modes corresponding to six vertical slices through Fig. 10.3: γ_i has been reduced to 15 s^{-1} . Graphs plot V_e versus λ for six values of ΔV_e^{rest} (displayed in the subtitles as ΔV). Thin line indicates stationary states of V_e : solid = stable; dotted = unstable. Subcritical Hopf bifurcations are marked with \times . Thick-solid line denotes maximum and minimum extent of the associated stable limit cycle.

in the unstable region, a spatially-symmetric (breathing mode) limit cycle results, with an amplitude extent as marked on the plot with thick-solid lines. Note the supercritical nature of the Hopf bifurcations for this value of γ_i .

This region of instability disappears when ΔV_e^{rest} is reduced to about 1 mV. At the same time the slope of membrane potential against λ increases; and for ΔV_e^{rest} less than about 1 mV (see Fig. 10.4(e, f)), multiple stationary states can exist (i.e., there exists a saddle–node bifurcation for variations in λ .) The middle branch of the three solutions is unstable. As ΔV_e^{rest} decreases further, the extent of the multiple solutions increases.

We now look at the case of a time-lengthened IPSP, where $\gamma_i = 15 \text{ s}^{-1}$. Figure 10.3 shows that this delayed feedback leads to a large increase in instability; the corresponding slice-plots are shown in Fig. 10.5. There are three key differences from Fig. 10.4. First, we note that the extent of the unstable area (the region between Hopf bifurcation pairs) has increased. Second, the bifurcations have become subcritical—i.e., the limit cycle extends beyond the region bracketed by the Hopf

bifurcations. This means that the system can have both a stable node and a stable limit cycle available to it simultaneously. Finally, we see that the unstable region co-exists with the region of multiple stationary state solutions. This can result in a rich variety of solutions; for example, at $\Delta V_e^{\text{rest}} = -2.5$ mV and $\lambda = 1.25$ there are two stable stationary states and a stable limit cycle available (see Fig. 10.6). For negative ΔV_e^{rest} , an unstable “lip” remains on the top branch of the stationary states, and a very small unstable lip remains on the bottom branch. (See also Fig. 3).

The Fig. 10.5 limit cycles arise as a result of small inhibitory synaptic rate-constant (i.e., large synaptic time-constant), leading to an effective delay in application of negative feedback. The formation of this instability has been linked with seizures (e.g., Kramer *et al.* [8]), and in Ref. [20] we have suggested that the tendency for the anesthetic enflurane to promote seizures is due to its lengthening of the inhibitory postsynaptic potential, causing Hopf bifurcation to a limit cycle. It is unlikely that a healthy brain would operate in such a region.

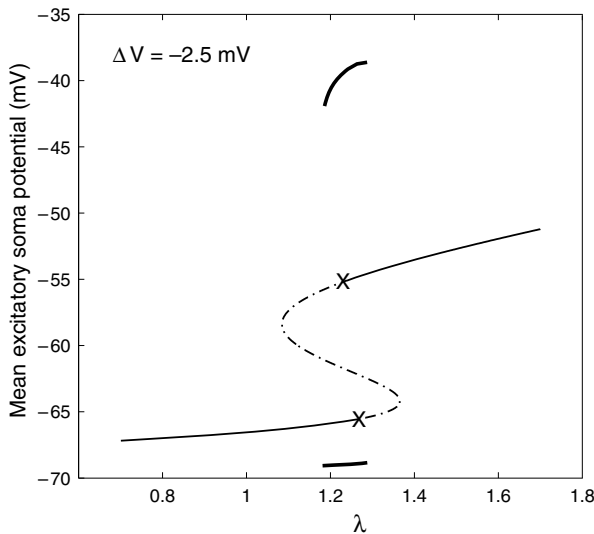


Fig. 10.6 Stationary states for the excitatory membrane potential V_e for the case $\Delta V_e^{\text{rest}} = -2.5$ mV, $\gamma_i = 15$ s $^{-1}$. Stable stationary states are marked with a thin-solid line, unstable states with a dot-dashed line. Subcritical Hopf bifurcations are marked with (\times). The limit cycle is indicated by thick-solid lines showing the upper and lower extents of the excitatory membrane potential. Note that for $\lambda \approx 1.25$, the system has access to two stable stationary states and a stable limit cycle.

10.5.2 Response to localized perturbations

To investigate the dynamic, spatially asymmetric response of this model, we look specifically at the response to small “kicks” to the cortex, which physically could be a result of thalamic pathways to the cortex, or spontaneous neurotransmitter

release. As will be shown below, these kicks can produce traveling-wave structures which we will identify with slow waves or K-complexes [21].

We have carried out simulations on a 64×64 spatial grid, for different values of the control parameter λ for the case $\gamma_i = 15 \text{ s}^{-1}$ and $\Delta V_e^{\text{rest}} = -2.5 \text{ mV}$. This choice of parameters leads to the stationary-state structure of Fig. 10.6, and gives rich dynamic behavior. In these simulations we place the system initially in a stationary state (where multiple stationary states exist we start on the lower-branch). The noise input results in the soma potential V_e undergoing fluctuations about this stationary state, as described by Wilson *et al.* [22]. Here, we examine the response of the system to a sharp kick applied at a single point in space. This has been done by making a momentary (0.1 s) increase in ΔV_e^{rest} (by 10 mV) at a *single* point on the simulation grid (e.g., ΔV_e^{rest} becomes +7.5 mV for 0.1 s, before returning to -2.5 mV). This corresponds to a large increase in cortical excitation at this point in space only.

With reference to the stationary states shown in Fig. 10.6, for $\gamma_i = 15 \text{ s}^{-1}$, we will now present results for simulations at various points along the λ axis. We start with λ small, where there is just one stationary state, and then raise λ to move into the region of multiple stationary states, then into the region where the limit cycle exists. Results are presented as grayscale plots (white = high, black = low) of the excitatory soma potential V_e for a slice through the cortex (incorporating the position of the kick) against time. We look first at $\lambda = 0.8$ where there is only a single, stable stationary state available that is well away from the region of multiple stationary states. Figure 10.7 shows that a kick of 10 mV for 0.1 s does not lead to

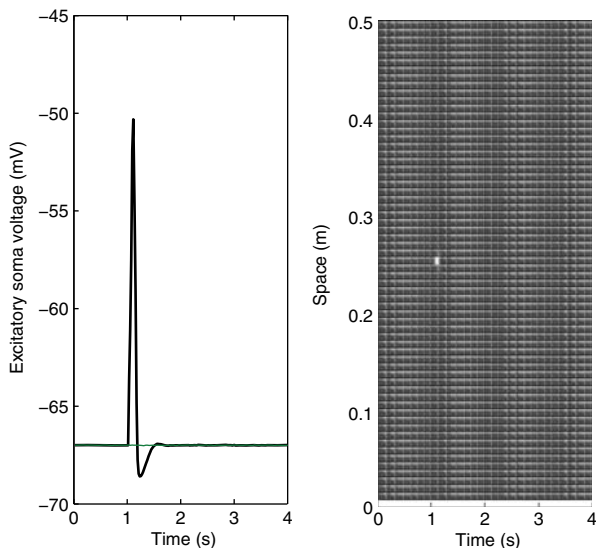


Fig. 10.7 Excitatory soma potential as a function of time in response to a short, localized impulse when only one stationary state is accessible to the system. Parameter values: $\Delta V_e^{\text{rest}} = -2.5 \text{ mV}$, $\gamma_i = 15 \text{ s}^{-1}$, with $\lambda = 0.8$. The effect of the kick does not propagate. (Reprinted from [21] with permission.)

any interesting dynamics—the soma potential rises at the location of the kick only, and then decays quickly back to the stationary state.

However, the dynamics changes markedly as we approach the region of multiple stationary states. We look at the response of the system to the same kick, but with the control parameter increased to $\lambda = 1.1$; there are now three stationary states, but only the lower-branch state is stable. The results are shown in Fig. 10.8. We see that the response propagates as a wave of constant velocity over the whole cortex. The system is pushed from its stable lower branch, beginning at the point where the impulse is applied, and heads towards the unstable upper branch. However, since the upper branch is unstable, it cannot remain there, and instead returns to the stable lower branch, where it finally comes to rest (the kick having been removed). This process generates a single circular traveling wave that propagates outwards from the point of the initial kick—every point in space leaves the lower branch, heads towards the upper branch, overshoots, then falls down towards the lower branch, overshoots this, and eventually settles back on the lower branch. The shape of the plot of Fig. 10.8 is similar to that of the K-complex seen in EEGs. Indeed, we will follow Massimini *et al.* [10] and identify this traveling wave with a K-complex. Golomb and Amitai have shown that a disinhibited network model can also give a similar result [3]—the significant difference between their work and

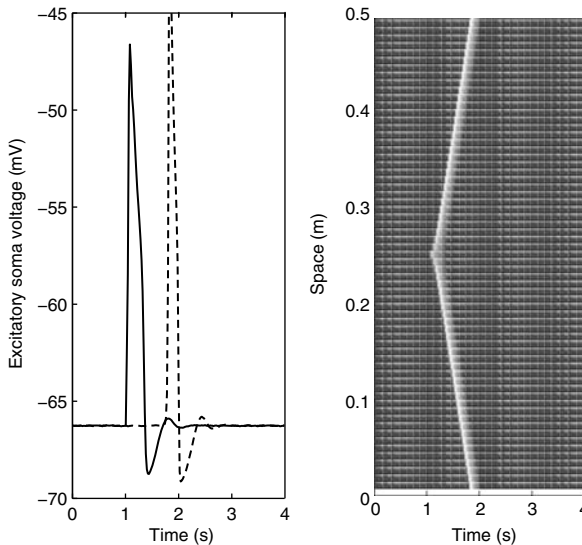


Fig. 10.8 Excitatory soma potential as a function of time for a short kick at one point in space when lower branch is stable, upper branch unstable. Parameter values: $\Delta V_e^{\text{rest}} = -2.5$ mV, $\gamma_i = 15$ s $^{-1}$, with $\lambda = 1.1$. Left-hand graph shows V_e for two points on the cortex (the point of initial kick, and 25 cm from this point) versus time. Right-hand image shows V_e as grayscale (white = high, black = low) against space and time for a line through the position of the kick. The resulting traveling wave is similar to a K-complex. (Reprinted from [21] with permission.)

ours is that inhibitory neurons are included here, and multiple stationary states are involved.

The cause of the K-complex in the EEG is uncertain—it can occur either as a response to a specific stimulus (*e.g.* a sound), or sometimes spontaneously with no obvious cause. In his summary paper, Colrain [1] describes the K-complex, with reference to the earlier work of Numminen *et al.* [11] as,

—a nonspecific reaction of the cortex to various stimuli during sleep, via activation of projection pathways from the nonspecific thalamic nuclei.

We now raise λ further, to $\lambda = 1.25$. From Fig. 10.6 we note that, in addition to the stable stationary state of the lower branch, there is now a stable limit cycle available to the spatially-homogeneous system. When we apply a kick to the system that is initially on the lower branch, the system is provoked into this limit cycle, as illustrated in Fig. 10.9. The firing rate moves from near zero to near maximum, in a similar way to spike–wave seizures [17]. Figure 10.9 also shows the limit cycle becoming progressively more synchronized in space as time increases, corresponding to the $q = 0$ breathing mode.

To explore the possibilities further, we look at the crucial significance of the inhibitory synaptic rate-constant, γ_i , in determining the stability of the cortical model. An increase in γ_i to 65 s^{-1} leads to a stabilization of the stationary states—*i.e.*, the Hopf bifurcations and limit cycles are removed from Fig. 10.6. In this case a kick produces no traveling waves—*i.e.*, the existence of the traveling wave (K-complex) state depends upon having a sufficiently long inhibitory time-constant.

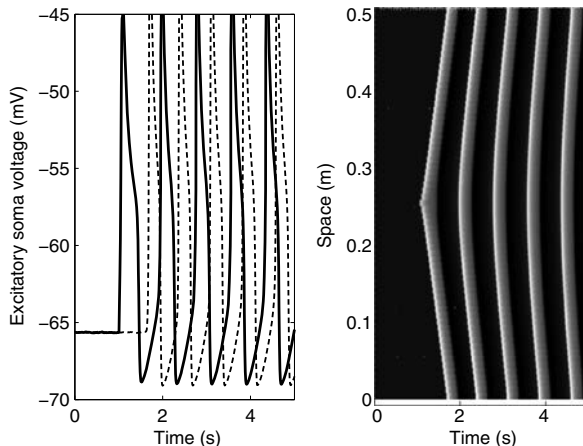


Fig. 10.9 Excitatory soma potential as a function of time when there is a limit cycle available to the system. Parameter settings: $\lambda = 1.25$, $\Delta V_e^{\text{rest}} = -2.5 \text{ mV}$, $\gamma_i = 15 \text{ s}^{-1}$. Note how the limit cycle begins to synchronize in space as time increases. (Reprinted from [21] with permission.)

10.5.3 K-complex revisited

We return to the situation that produced the single traveling-wave solution which we identify as a K-complex [21], and explore the response to impulsive kicks of varying size. In Fig. 10.10 we have summarized the four possible dynamic outcomes of a voltage kick of given size at a given value of the parameter λ . Where there are multiple stationary states, we consider the system to be initially on the lower branch. We set $\Delta V_e^{\text{rest}} = -2.5$ mV and $\gamma_i = 15$ s⁻¹, i.e., the stability diagram of Fig. 10.6 applies.

At large λ (corresponding to $\lambda > 1.27$), a small kick results in the system moving to the upper state where the stability is greater. Indeed, for $\lambda > 1.35$ (region (d) of Fig. 10.10), there is no stable lower state, and the system moves spontaneously to the upper state without any applied kick.

For $1.0 < \lambda < 1.27$ (region (c)) a kick of sufficient size will result in a traveling wave of the form of Fig. 10.8. Note that the size of the kick is important; if it is not large enough, the disturbance will rapidly die away rather than propagating over

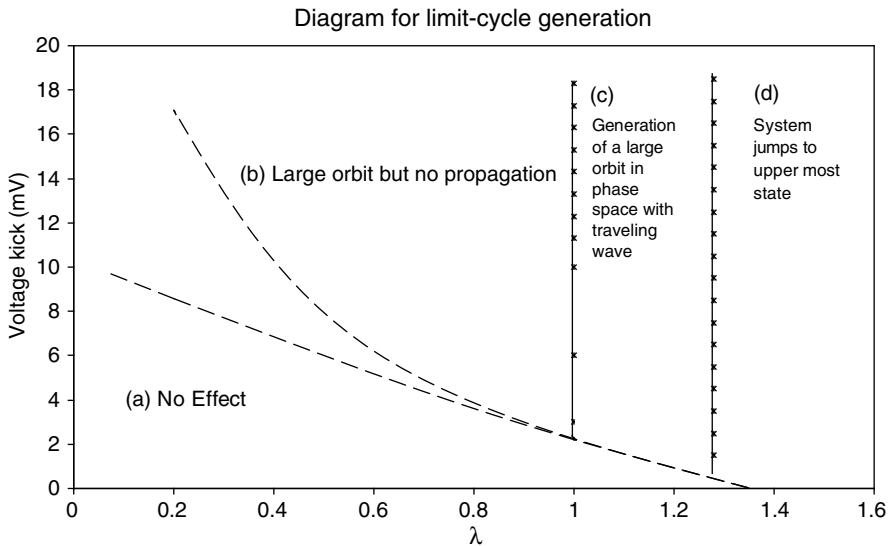


Fig. 10.10 Summary of the effect of disturbing the cortical system at one point in space. Parameter settings: $\Delta V_e^{\text{rest}} = -2.5$ mV, $\gamma_i = 15$ s⁻¹. (a) For small kicks, except when the lower branch is unstable, the generated disturbance quickly dies away. (b) At low λ , a large kick can result in a large, localized response (an orbit in phase space similar to Fig. 10.11) but with no propagation away from the site of disturbance. (c) In the vicinity of the region where there are multiple stationary states, with the top state being unstable, a sufficiently-sized kick can generate a large disturbance that propagates as a wave. Note that this can occur even when there is only one stationary state (e.g., $\lambda=1.0$); however the system has to be close to the saddle–node bifurcation. The boundary between a (c) propagating and (b) nonpropagating disturbance is very distinct. (d) At high values of λ , a small kick will displace the system onto the upper state, where it will remain. Again, a distinct boundary separates regions (c) and (d).

space, as in Fig. 10.7. The required size of the kick reduces as λ increases; the lower stationary state has become less stable. Note that the boundary between regions (c) and (d) is very distinct.

For $\lambda < 1.0$, no kick can generate a traveling wave. However, a large enough kick (region (b) of Fig. 10.10) will result in initial local growth of the disturbance, but one that fails to propagate entirely to other regions. At very low λ , the boundary between the region of initial growth (b) and of initial decay (a) is quite indistinct; it is marked on Fig. 10.10 by dashed lines. However, the boundary between regions (b) and (c) is very sharp.

This abrupt triggering of a K-complex in region (c) can be explained with reference to the orbits of the spatially-homogeneous system in phase-space. In Fig. 10.11(a), for $\Delta V_e^{\text{rest}} = -2.5$ mV, $\lambda = 1.15$, we plot a projection of the trajectories in phase-space of the spatially-homogeneous system. The top graph is for a small inhibitory synaptic rate-constant, $\gamma_i = 15$ s⁻¹; the bottom graph is for a larger rate of 65 s⁻¹. All state variables have been started at their equilibrium values except for Φ_{ee} and Φ_{ii} . The equations (10.1–10.8), if written as a set of first-order differential equations, have fourteen dynamic variables; the two variables that illustrate the situation most clearly are Φ_{ee} , the $e \rightarrow e$ synaptic flux, and Φ_{ii} , the $i \rightarrow i$ synaptic flux. For this reason, these two variables have been plotted here. For the top case, since there is only one stable solution, all of the trajectories eventually end on the lower-branch solution (Os in lower left-hand corner). However, looking at the trajectories starting in this vicinity, it is clear that two initial points very close together can generate manifestly different trajectories in order to return to the single stable solution. In one case, the trajectory returns quickly to the stable solution of the lower branch; in the other, first the $e \rightarrow e$ synaptic flux grows, then the $i \rightarrow i$ flux, the $e \rightarrow e$ flux diminishes, and finally the $i \rightarrow i$ flux diminishes, and the trajectory returns to the stable lower branch (Os). This path takes the system around the unstable solution of the upper branch (Δu in top-right corner). This divergence of trajectories explains why a tiny increase in the size of the kick (e.g., from region (a) to region (c) of Fig. 10.10) can result in a very different solution to the dynamical equations.

Why does the numerical simulation produce a traveling wave? Different points in space are coupled through the ∇^2 -term of Eqs (10.7) and (10.8). If a single point in space is kicked from its stable equilibrium onto the trajectory that travels around the upper branch, as happens in Fig. 10.8, the ∇^2 spatial-coupling will pull the neighboring points from their position on the lower stable branch onto this trajectory too. These in turn influence their neighbors, sending out a traveling disturbance. After the wave has passed, all points return to the original, lower-branch stable steady state (Os).

In Fig. 10.11(b) we see the effect of restoring the inhibitory rate-constant to $\gamma_i = 65$ s⁻¹, thereby removing the instability from the upper branch of Fig. 10.6. A *large* kick, large enough to take the system past the unstable mid-state (\times), will cause the system to move to the upper stationary state (Δs). This removes the possibility of producing a K-complex, and instead, a sufficiently large kick, applied to the bottom branch, results in the system moving directly to the top branch.

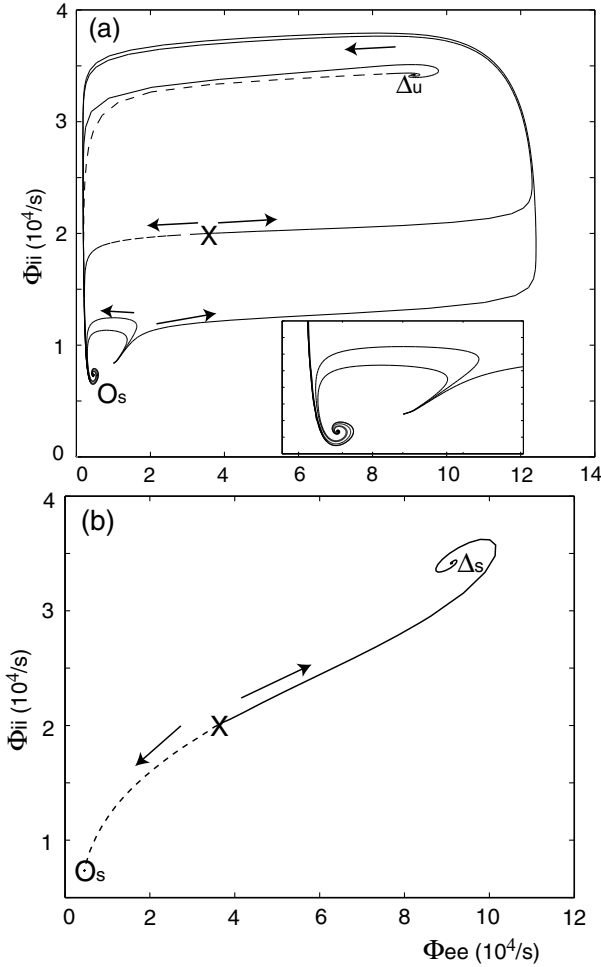


Fig. 10.11 Phase-space trajectories for parameter settings $\Delta V_e^{\text{rest}} = -2.5 \text{ mV}$, $\lambda = 1.15$, and (a) $\gamma_i = 15 \text{ s}^{-1}$; (b) $\gamma_i = 65 \text{ s}^{-1}$. Panel (a): Lower, stable stationary state is in the bottom left-hand corner of the figure (marked “ O_s ”); the unstable upper state in the top-right corner (Δ_u); the unstable mid-branch solution is in the center (\times). Two trajectories have been started on each of the upper- and mid-branches—these eventually reach the lower, stable branch (O_s). Other trajectories have been started close to the lower state. Two of these return quickly to this state, but the third, initially displaced by a very small distance from the first two, exhibits a markedly different trajectory that loops around the upper state (Δ_u). We identify this trajectory with a K-complex. The vicinity of the lower state has been expanded in the subpanel for clarity (but the different starting points are still indistinguishable). Panel (b): Trajectories have changed markedly with a reduced synaptic response-time. The stationary states remain in the same positions, but the upper state is no longer unstable (now marked “(Δ_s)”). There is no longer divergence of trajectories as in (a). The solid and dotted lines denote different trajectories leaving the same unstable initial point. (Reprinted from [21] with permission.)

10.5.4 Spiral waves

Finally, we remark on another limit-cycle that is available to an oscillatory system, namely that of spiral waves. These persistent features sometimes can be generated when a simulation is run through an unstable region of the sleep domain. These waves are spatially structured states consisting of pairs of counter-rotating spirals. Figure 10.12 shows a gray-scale plot of $V_e(\mathbf{r})$ at a given time. Specifically, this simulation involved starting on a stable, upper-branch solution and then, by reducing λ , moving the system through an unstable region into the region where the lower branch is stable. These limit-cycles are extremely persistent—to quench them, a large reduction in λ or ΔV_e^{rest} is required (i.e., a large reduction in excitatory component). Experimentally, spirals have been observed in disinhibited cortical slices [4], and demonstrated in neuron models with no inhibition [4, 13]. However, their presence in the cortex, and their relationship with states such as epileptic seizures, is unclear.

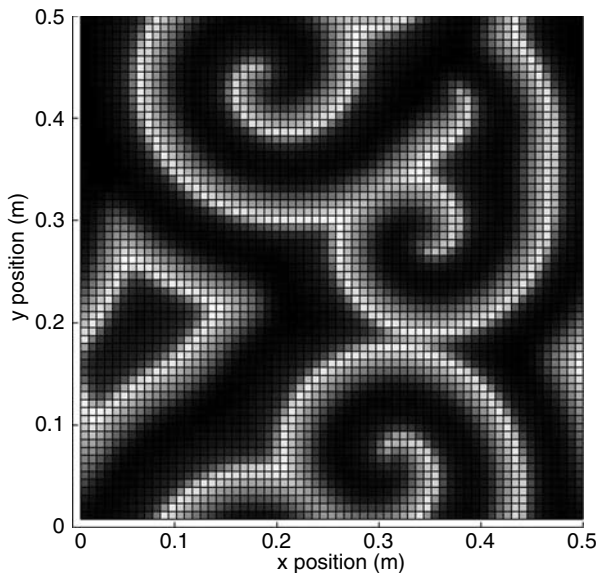


Fig. 10.12 Snapshot of a spiral wave generated by the cortical model (white = high V_e ; black = low V_e). To trigger spiral formation, the system was started on the stable region of the upper branch ($\Delta V_e^{\text{rest}} = 0.5$ mV, $\lambda = 1.75$, $\gamma_i = 33$ s $^{-1}$), then λ was rapidly lowered (over ~ 7 s) through the unstable region and into the region where only a single stable state exists. Resulting spiral wave is extremely persistent, and requires a considerable further reduction in λ to destroy it. (Reprinted from [22] with permission.)

10.6 Conclusions

In this chapter we have demonstrated some of the dynamic features associated with a nonlinear mean-field cortical model. In particular, spatially symmetric limit cycles are reminiscent of seizure-like states [8, 20], and arise as a result of a delay in negative feedback through a lengthened inhibitory postsynaptic potential. This delay may explain of the tendency of some anesthetic drugs, such as enflurane, to promote seizures.

For certain parameter sets, traveling waves of activity can be produced. These are associated with the combination of saddle–node bifurcations and Hopf bifurcations, and are reminiscent of the K-complexes and slow oscillations of slow-wave sleep. Such waves can be activated by a point-like disturbance of sufficient magnitude—a below-threshold disturbance will fail to propagate.

In some limited conditions, spiral waves are generated. These are extremely persistent once established. The biological significance of such waves is not entirely clear, although, in general, the ability of a 2-D nonlinear system to produce spiral waves is not surprising.

References

1. Colrain, I.M.: The K-complex: A 7-decade history. *Sleep* **28**, 255–273 (2006)
2. Freeman, W.J.: Predictions on neocortical dynamics derived from studies in paleocortex. In: E. Basar, T.H. Bullock (eds.), *Induced Rhythms of the Brain*, chap. 9, pp. 183–199, Birkhaeuser, Boston (1992)
3. Golomb, D., Amitai, Y.: Propagating neuronal discharges in neocortical slices: Computational and experimental study. *J. Neurophys.* **78**, 1199–1211 (1997)
4. Huang, X., Troy, W.C., Yang, Q., Ma, H., Laing, C.R., Schiff, S.J., Wu, J.Y.: Spiral waves in disinhibited mammalian neocortex. *J. Neurosci.* **24**, 9897–9902 (2004), doi:10.1523/jneurosci.2705-04.2004
5. Hutt, A., Bestehorn, M., Wennekers, T.: Pattern formation in intracortical neuronal fields. *Network* **14**, 351–368 (2003)
6. Jirsa, V.K., Haken, H.: A field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996), doi:10.1103/PhysRevLett.77.960
7. Kloeden, P.E., Platen, E.: *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin (1992)
8. Kramer, M.A., Kirsch, H.E., Szeri, A.J.: Pathological pattern formation and epileptic seizures. *J. Roy. Soc. Interface* **2**, 113 (2005), doi:10.1098/rsif.2004.0028
9. Liley, D.T.J., Cadusch, P.J., Wright, J.J.: A continuum theory of electro-cortical activity. *Neurocomp.* **26–27**, 795–800 (1999), doi:10.1016/S0925-2312(98)00149-0
10. Massimini, M., Huber, R., Ferrarelli, F., Hill, S., Tononi, G.: The sleep slow oscillation as a traveling wave. *J. Neurosci.* **24**, 6862–6870 (2004), doi:10.1523/jneurosci.1318-04.2004
11. Numminen, J., Makela, J.P., Hari, R.: Distributions and sources of magnetoencephalographic K-complexes. *Electroencephal. Clin. Neurophysiol.* **99**, 544–555 (1996), doi:10.1016/S0013-4694(96)95712-0
12. Nunez, P.L.: The brain wave function: A model for the EEG. *Math. Biosci.* **21**, 279–297 (1974), doi:10.1016/0025-5564(74)90020-0

13. Osan, R., Ermentrout, B.: Two dimensional synaptically generated travelling waves in a theta-neuron neural network. *Neurocomp.* **38–40**, 789–795 (2001)
14. Rennie, C.J., Wright, J.J., Robinson, P.A.: Mechanisms for cortical electrical activity and emergence of gamma rhythm. *J. Theor. Biol.* **205**, 17–35 (2000), doi:10.1006/jtbi.2000.2040
15. Robinson, P.A., Rennie, C.J., Rowe, D.L., O'Connor, S.C., Wright, J.J., Gordon, E., Whitehouse, R.W.: Neurophysical modeling of brain dynamics. *Neuropsychopharmacol.* **28**, S74–S79 (2003), doi:10.1038/sj.npp.1300143
16. Robinson, P.A., Rennie, C.J., Wright, J.J.: Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* **56**, 826–840 (1997), doi:10.1103/PhysRevE.56.826
17. Steriade, M., Timofeev, I., Grenier, F.: Natural waking and sleep states: A view from inside neocortical neurons. *J. Neurophysiol.* **85**, 1969–1985 (2001)
18. Steyn-Ross, D.A., Steyn-Ross, M.L., Sleigh, J.W., Wilson, M.T., Gillies, I.P., Wright, J.J.: The sleep cycle modelled as a cortical phase transition. *J. Biol. Phys.* **31**, 547–569 (2005), doi:10.1007/s10867-005-1285-2
19. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W.: Modelling general anaesthesia as a first-order phase transition in the cortex. *Progress in Biophysics and Molecular Biology* **85**, 369–385 (2004), doi:10.1016/j.pbiomolbio.2004.02.001
20. Wilson, M.T., Sleigh, J.W., Steyn-Ross, D.A., Steyn-Ross, M.L.: General anesthetic-induced seizures can be explained by a mean-field model of cortical dynamics. *Anesthesiology* **104**, 588–593 (2006), doi:10.1097/00000542-200603000-00026
21. Wilson, M.T., Steyn-Ross, D.A., Sleigh, J.W., Steyn-Ross, M.L., Wilcocks, L.C., Gillies, I.P.: The K-complex and slow oscillation in terms of a mean-field cortical model. *J. Comp. Neurosci.* **21**, 243–257 (2006), doi:10.1007/s10827-006-7948-6
22. Wilson, M.T., Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W.: Predictions and simulations of cortical dynamics during natural sleep using a continuum approach. *Phys. Rev. E* **72**, 051910 1–14 (2005), doi:10.1103/PhysRevE.72.051910
23. Wright, J.J., Liley, D.T.J.: Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. *Behav. Brain Sci.* **19**, 285–316 (1996)

Chapter 11

Phase transitions, cortical gamma, and the selection and read-out of information stored in synapses

J.J. Wright

11.1 Introduction

This chapter attempts a unification of phenomena observable in electrocortical waves, with mechanisms of synaptic modification, learning and recall. In this attempt I draw upon concepts advanced by Freeman and colleagues, and on studies of the mechanism of gamma oscillation in cortex, and of synaptic modification and learning. I use a recently published system of state equations, and numerical simulations, to illustrate the relevant properties.

In Walter Freeman's early work [12, 21], he derived pulse-to-wave and wave-to-pulse conversion equations, and described ensembles of neurons in "K-sets". In recent work [4, 13–20, 22], he has sought a widely embracing theory of perception and cognition by integrating a body of intermediate results, revealing the occurrence of sequential, transiently synchronised and spatially organized electrocortical fields, occurring in the beta and gamma bands, and associated with traveling waves organized into "phase (in the Fourier sense) cones". The origin of these phenomena he and colleagues associate with "phase (in the thermodynamic sense) transitions" in cortical neural activity, marked by signatures in Hilbert-transformed ECoG (electrocorticogram)—termed "phase (in the Fourier sense) slip" and "null spikes".

Freeman's work has features in common with other theoretical approaches [2, 3, 5, 63] but has most in common with those described as mean-field, continuum, or population approximation [26, 41, 66, 68, 75], and is thus in line with the simulation results to be described below.

A large body of other work [7, 8, 11, 23–25, 37, 40, 53, 54] also indicates that gamma oscillation and synchronous oscillation are of central importance. The persuasive link for physiologists, whose emphasis is on unit action potentials, is the

James J. Wright

Liggins Institute, and Department of Psychological Medicine, University of Auckland, Auckland, New Zealand; Brain Dynamics Centre, University of Sydney, Sydney, Australia.
e-mail: jj.w@xtra.co.nz

finding that field potentials and pulse activity are strongly correlated in the gamma frequency band [57]. Yet the mechanisms of origin and the control of gamma oscillation are not fully determined. During gamma oscillation an average lead-vs-lag relation exists between local excitatory and inhibitory cells, as predicted by Freeman [12] and subsequently experimentally observed [27]. Recent analyses of the cellular dynamics [38] conclude that recurrent inhibition from fast-spiking inhibitory cells is largely responsible for maintaining the rhythmic drive—but the role played by excitatory processes in modulating or driving the oscillations remains undetermined. 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. Without clarification of these mechanisms it remains difficult to define the link of gamma activity to information storage, retrieval and transmission, and between thermodynamic and cognitive or informational perspectives.

There are also theoretically important relations of synchrony to information storage and retrieval [32, 33, 44, 67] including recent considerations of synaptic self-organization [72, 73]. These models relate directly or indirectly to wave aspects of cortical signal transmission [8, 9, 29, 37, 45–51, 56, 69, 75]. The present chapter attempts to make these links more explicit, by drawing upon properties demonstrated in a continuum cortical model. Properties of the simulations have been previously reported [70, 71]. A single parameter set obtained *a priori* is used to reproduce appropriate scale-dependent ECoG effects. Receptor dynamics of three major neurotransmitter types and effects of action potential retrograde propagation into the dendritic tree are included, although these features are inessential to the minimalist goal of replication of the ECoG. However, their inclusion assists in linking field properties to information storage and exchange, which is the principal matter to be further explored in the present chapter.

11.2 Basis of simulations

Figures 11.1 and 11.2 show the context of the simulation in relation to qualitative physiological and anatomical features. Crucial anatomical aspects in Fig. 11.1(a) are the local interaction of excitatory (pyramidal) and inhibitory cells in the cortical mantle and extension of the pyramidal tree into the upper cortical layer, to receive nonspecific afferents, while Fig. 11.1(b) sketches the interaction of the cortex with subcortical systems, over recurrent polysynaptic pathways, to regulate the spatial pattern of cortical activation, and hence attentional state [1].

Figure 11.2 indicates microscopic features, and a block diagram of connections, captured in the state equations.

Figure 11.3 shows graphically the steady-state functions, and unit-impulse response functions associated with the simulation parameters. For ease of presentation, the state equations themselves are given in the Appendix (Sect. 11.6.1).

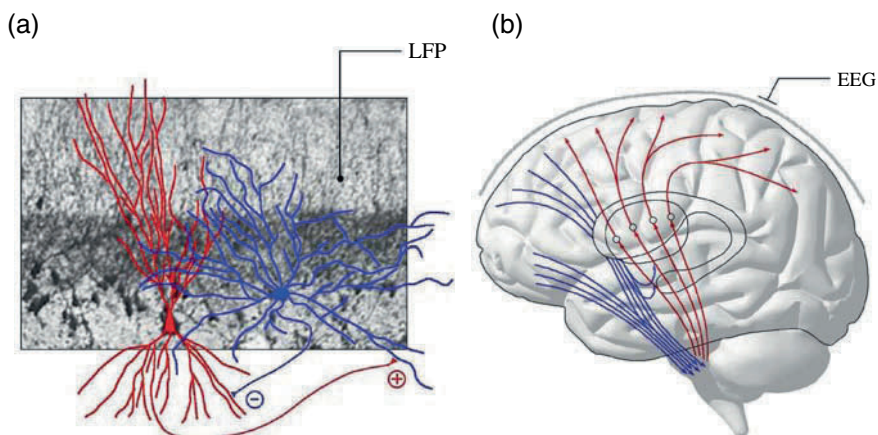


Fig. 11.1 [Color plate] Anatomical aspects of the model. (a) Excitatory (red) and inhibitory (blue) cells in the cerebral cortex. (b) Major cortical and subcortical interactions mediated by descending (blue) and ascending (red) connections. (Reproduced from [71] with permission.)

11.3 Results

11.3.1 Nonspecific flux, transcortical flux, and control of gamma activity

The simulation's state equations and parameter set have the advantage that they can be applied consistently in simulations at two scales—that of a unit macrocolumn about $300\ \mu\text{m}$ across [39]—or that of cortex at the scale of many centimeters, up to the size of the human cortex. In this model, gamma oscillation is defined by analogy to properties of physiological gamma, and comparison of properties at the two scales permits the local effects and the distant influences upon gamma oscillation to be discerned. Effects of parameter variation show that the onset and offset of gamma is subject to variation of all influences upon the excitatory and inhibitory balance, but chief among those influences is the finding that nonspecific flux (NSF) and transcortical flux (TCF) exert opposite effects upon the onset and offset of autonomous gamma oscillation.

At centimetric scale, where longer physiological conduction delays are applied, there is a persistence of stability to higher levels of NSF, in contrast to the macrocolumnar case, with its shorter axonal conduction delays. The two scales of simulation yield a consistent result: increasing levels of uniform NSF over wide extents of cortex suppress the onset of gamma activity. This action of NSF is mediated secondarily by the spread of TCF, which acts upon both excitatory and inhibitory compartments. In contrast, the action of focal NSF, as observed in the macrocolumnar-scale simulations, is to trigger gamma activity, as is next described.

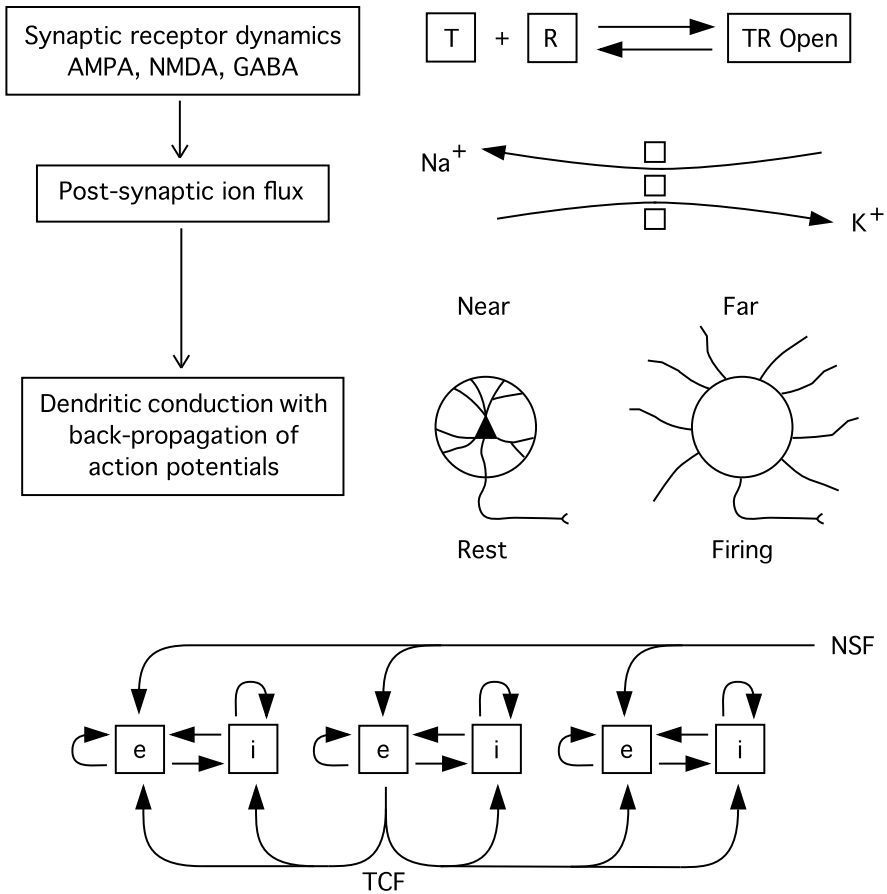


Fig. 11.2 State equations capture the features indicated in descending rows of the figure. (a) Mass-action interaction of neurotransmitters and receptors, thus regulating the opening and closing of ion channels; (b) consequent ionic flux in postsynaptic dendritic membrane; (c) dynamic variation in synaptic weights, consequent to changes in relative and absolute depolarisation of synapses in the proximal and distal dendritic trees, induced by back-propagation of action potentials; (d) cortical architecture, shown in one row of a square array of elements, as blocks of excitatory (*e*) and inhibitory (*i*) cell groups and axo-synaptic linkages. Site of action of the reticular activation system, on *e*-components only, is indicated as NSF (nonspecific flux); the lateral spread of excitatory signals to both *e*- and *i*-components is indicated as TCF (transcortical flux). (Reproduced from [70] with permission.)

11.3.2 Transition to autonomous gamma

Figure 11.4 quantifies the inverse actions of NSF and TCF on the transition to gamma in a macrocolumnar element. It can be seen that the mean value of cell firing increases with NSF, with little sensitivity to TCF. In contrast, over an operating

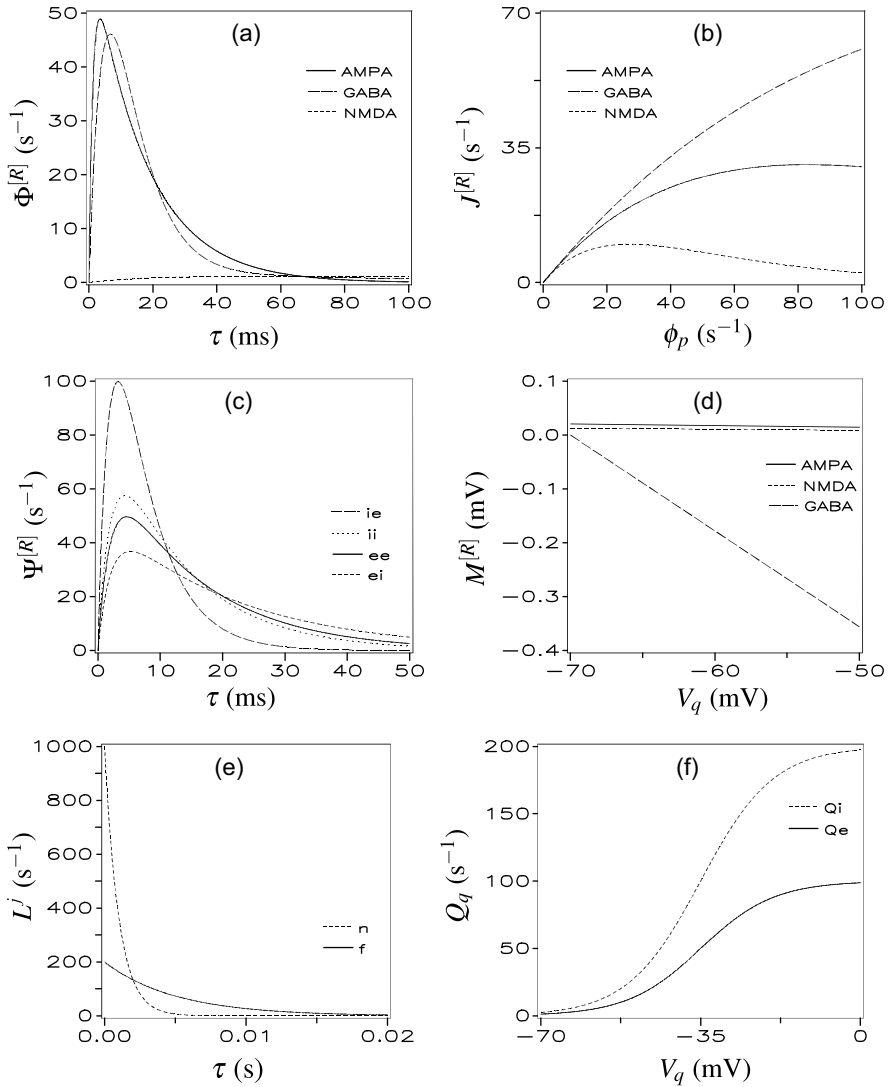


Fig. 11.3 Steady-state and unit-impulse response functions associated with the simulation parameters: (a) normalised receptor time responses; (b) receptor steady states; (c) normalised post-synaptic membrane time responses; (d) postsynaptic membrane steady states at $\phi_p = 20 s^{-1}$; (e) normalised dendritic tree delays; (f) action potential generation. (Reproduced from [70] with permission.)

range of NSF, increasing TCF suppresses autonomous gamma, offering a mechanism for negative feedback between concurrently active patches of cortical gamma oscillation. From lag correlations between the pulse densities of the excitatory and inhibitory compartments of a single element, a measure, $\Theta = 2\pi |l_+| / (|l_+| + |l_-|)$,

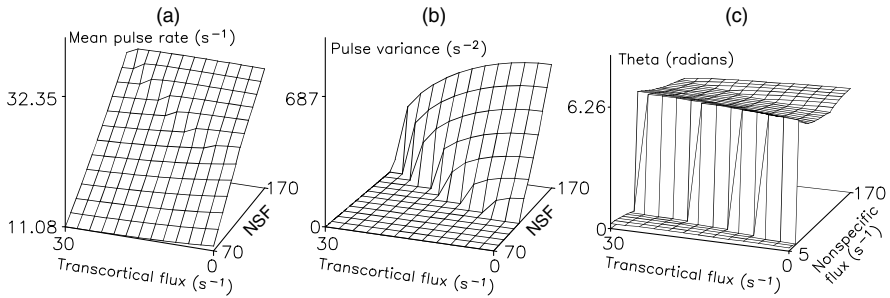


Fig. 11.4 Gamma oscillation as a function of transcortical flux (TCF) and nonspecific flux (NSF) for macrocolumnar-scale simulations. (a) Mean value of pulse rate in excitatory cells; (b) variance of pulse rate (oscillation power) in excitatory cells; (c) relative phase-lead, Θ , for excitatory vs inhibitory oscillations. (Modified from [70].)

where $|\ell_+|$ and $|\ell_-|$ are the respective lags from zero in the leading and lagging direction to the first peaks in the lag correlation function, was calculated. This measure shows that as the threshold of gamma oscillation is approached, excitatory and inhibitory compartments exhibit lagged correlation over a critical, threshold, range of NSF and TCF. As threshold is exceeded, excitatory and inhibitory compartments begin to fire in phase.

11.3.3 Power spectra

Macrocolumnar scale

Figures 11.5(a, c) show the impact of reciprocal variation of NSF and TCF on the spectrum of gamma oscillation. In all conditions, the spectrum is strongly peaked near the gamma range, with variation from the high-beta range to the high-gamma, with both gamma and high-gamma (~ 100 Hz) activity being seen above transition. There is no evidence of a $1/f^2$ spectrum at this scale.

Centimetric scale

Figures 11.5(b, d) show that at centimetric scale and low NSF, a $1/f^2$ spectrum is apparent, and this is greatly enhanced in amplitude when the system is driven by low frequency modulations. As NSF is increased, resonance in the gamma band increases.

11.3.4 Selective resonance near the threshold for gamma oscillation

Figures 11.5 (a, c) show that the spectral form of simulated gamma varies systematically above and below the threshold of transition. To study spectra nearer the

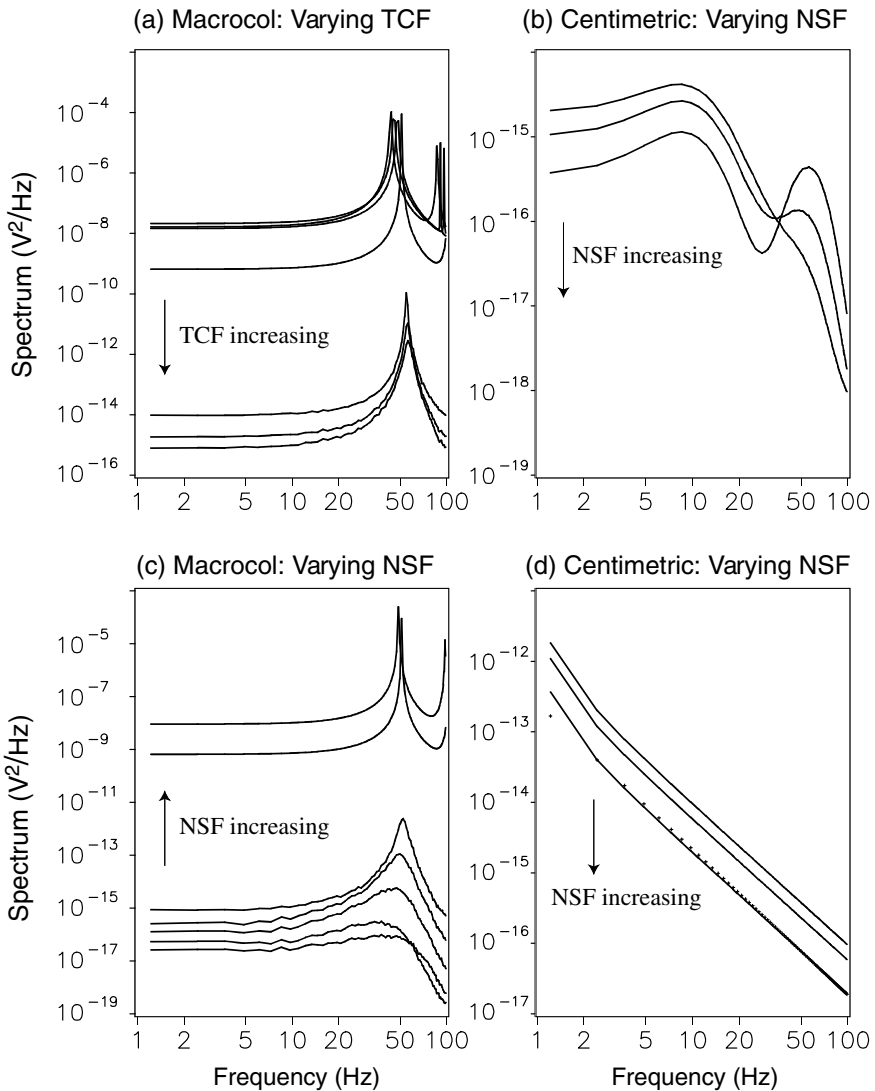


Fig. 11.5 Log-log plots of power spectra at macrocolumnar and centimetric scales. (a) and (c): *Macrocolumnar scale*; upper groups of lines indicate spectra associated with spontaneous oscillation; lower groups of lines indicate spectra for damped resonance. (a) NSF set at 141 s^{-1} ; TCF stepped through values 0, 5, 10, 15, 20, 25, 30 s^{-1} , for spectra in order of *decrementing* power at lowest frequencies. (c) TCF set at 15 s^{-1} ; NSF stepped through values 0, 28.2, 56.5, 85, 113, 141, 169 s^{-1} , for spectra in order of *incrementing* power at lowest frequencies. (b) and (d): *Centimetric scale*; NSF was set to 0, 8, 16 s^{-1} in order of decrementing power at lower frequencies. (b) Synchronous white-noise input to all elements in the driven row. (d) Driving noise band-limited from 0.002 to 0.954 Hz. Dotted line shows least-squares linear best-fit, with slope -2.07 . (Modified from [70].)

transition, power spectra were computed as transition was approached, by applying near-critical levels of NSF. Figure 11.6 shows a representative outcome at macro-columnar scale, with $\text{TCF} = 15 \text{ s}^{-1}$ applied. (Similar results were found at all levels of applied TCF.) It can be seen that as threshold is approached, a broad spectrum centred in the gamma range becomes supplemented by a sharper peak, also in the gamma range. Comparison against the upper curves of Figs. 11.5 (a) and (c) indicates how the sharp gamma peak of sustained oscillation is supplemented by a harmonic in the high gamma range.

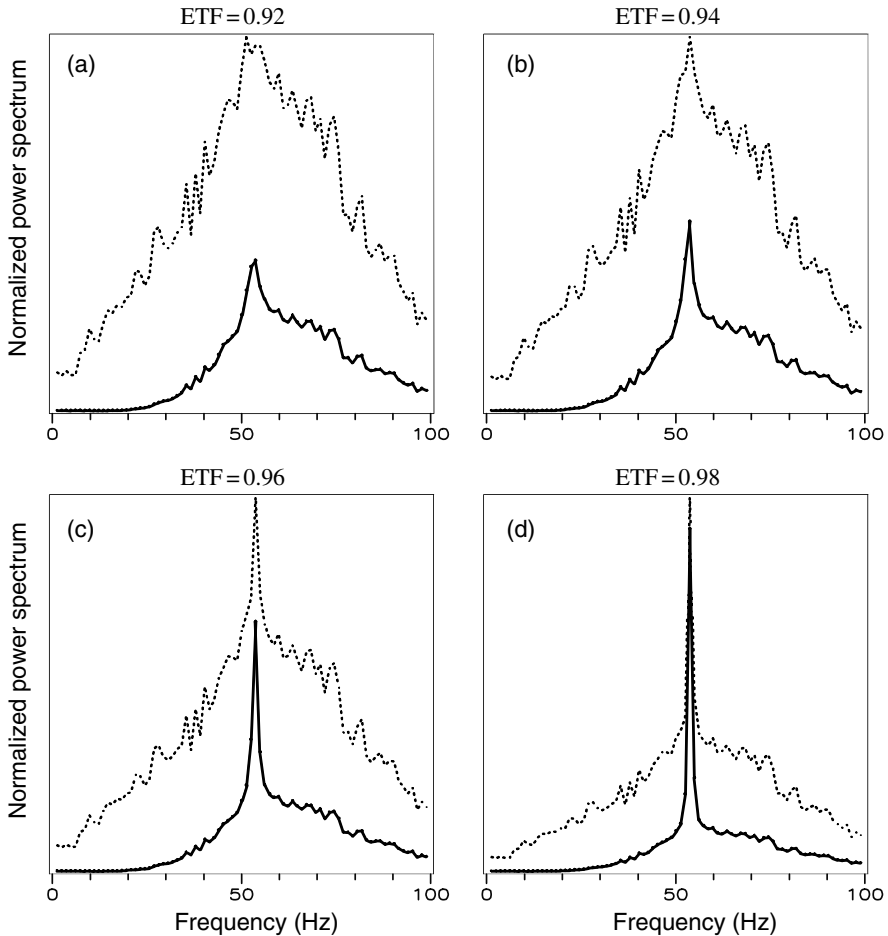


Fig. 11.6 Power spectral response as transition to autonomous gamma activity is approached (macro-columnar scale). Spatiotemporal white noise delivered to all excitatory cell components. ETF = excitatory transition fraction (the fraction of NSF input required to reach threshold for spontaneous oscillation). Solid lines: excitatory compartment; dotted lines: inhibitory compartment. (Reproduced from [71] with permission.)

Figure 11.7 shows firing-rate covariance and phase differences, for centimetric scale, between the excitatory and inhibitory cell continua, during the transition from damped gamma oscillation to autonomous gamma. It can again be seen from the top graph how system sensitivity to small noise inputs increases markedly, while a sharp change in excitatory/inhibitory phase relations characterises the transition, after a premonitory gradual increase in phase difference as transition is approached.

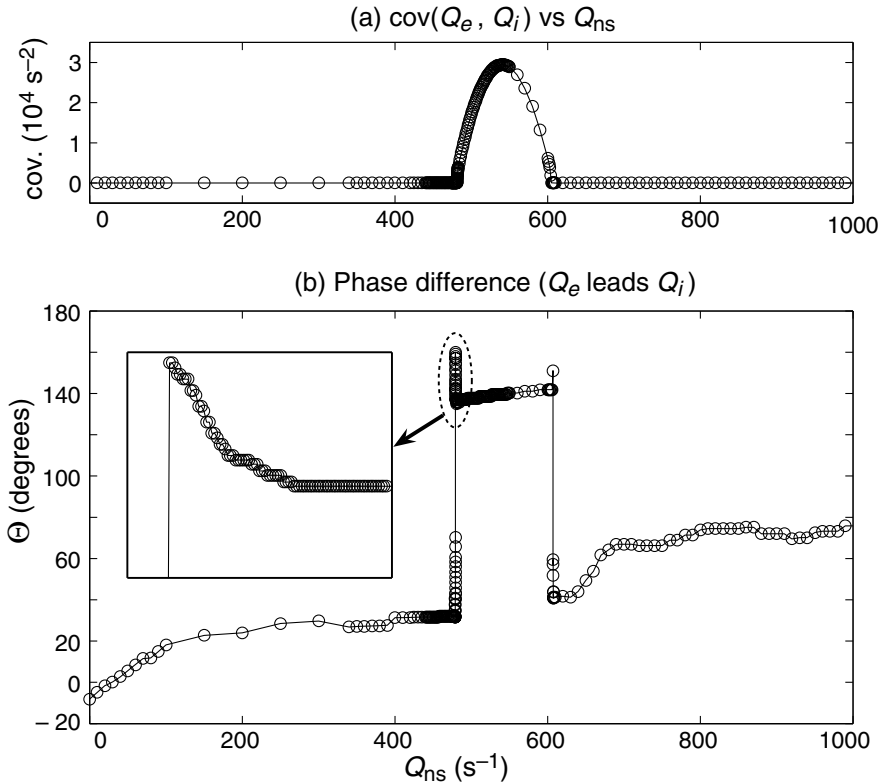


Fig. 11.7 Excitatory/inhibitory covariance and phase relations (centimetric scale) as transition to gamma oscillation is approached and exceeded. Q_{ns} is the level of nonspecific flux (NSF). (a) Zero-lag cross-power of excitatory and inhibitory firing rates at (row-10, col-10) on the cortical sheet; (b) Q_e vs Q_i phase-difference calculated as in Sect. 3.2.

11.3.5 Synchronous oscillation and traveling waves

Two points on a simulated macrocolumn were driven with independent white-noise inputs. Figure 11.8 shows lag covariance between excitatory compartments in two reference elements, adjacent to each of the two driven elements. Amplitudes of the

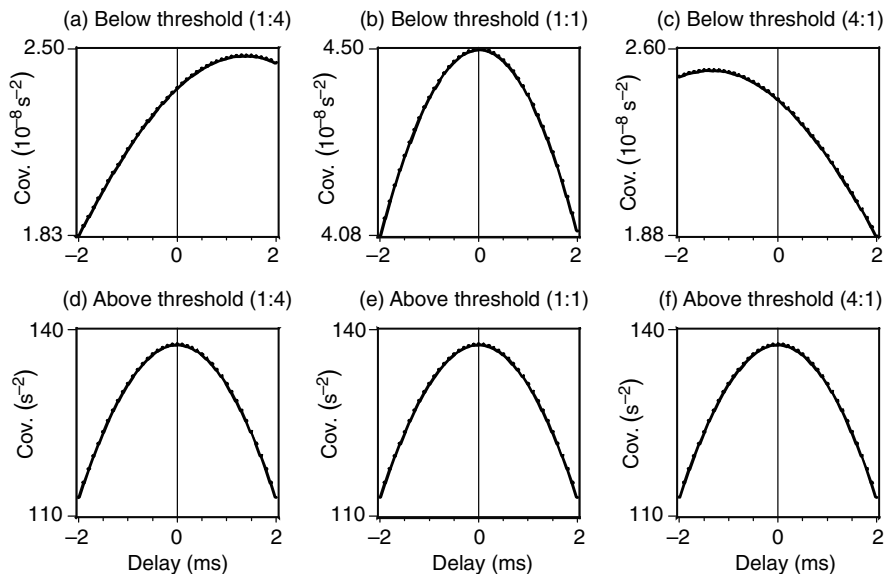


Fig. 11.8 Pulse density covariance versus conduction delay in a macrocolumnar simulation with $\text{TCF} = 0 \text{ s}^{-1}$. Asynchronous white-noise inputs are delivered to each of two elements, situated at (row, col.) = (10, 16) and (10, 10); lag covariances are computed between elements adjacent to the sites of input. RMS amplitudes of the input signals are in the ratios 1:4, 1:1, 4:1. Top row: Traveling waves apparent below threshold of oscillation ($\text{NSF} = 61 \text{ s}^{-1}$). Bottom row: Synchronous fields apparent above threshold of oscillation ($\text{NSF} = 124 \text{ s}^{-1}$). (Modified from [70].)

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 traveling 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 traveling waves despite disparity of input magnitudes. Similar effects 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.

11.4 Comparisons to experimental results, and an overview of cortical dynamics

The limited two-scale results, with their static restraints on NSF and TCF, taken in combination with results from related simulations [9, 47, 48, 74, 76], are consistent with a wide range of experimental findings. The simulation results do not include peaked resonances at the theta, alpha and beta rhythms, the representation

of which requires incorporation of a model of the thalamus [48, 50, 59]. A fuller model would also require consideration of spatiotemporal variation of NSF and TCF.

11.4.1 Comparability to classic experimental data

Simulation results show a good match to widely observed electro-cortical phenomena, despite experimental uncertainties in the values of many parameters, and the use of simplified cortical architecture. The $1/f^2$ background spectrum and the resonance peaks are in accord with EEG in a number of species [8, 18, 43]. The predominance of the $1/f^2$ background when very low frequency inputs are introduced mimics the slow fluctuations in cortical activation associated with variation of the field of attention [1]. There is good correspondence with gamma activity including increasing covariance of pulses and waves as autonomous oscillation is approached [57], and pulse rates accord with those found in awake cortex [55]. Simulation properties shared with related simulations [74], mimic the classic synchrony findings of moving visual-bar experiments [11, 24].

11.4.2 Intracortical regulation of gamma synchrony

It appears that both local and global factors contribute to the control of gamma activity. Locally, all influences on excitatory/inhibitory balance 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 this global control is apparent when it is recalled that NSF was delivered to excitatory cells only, while TCF was delivered to both excitatory and inhibitory components.

A general account of gamma activity and information transfer in the cortex can be advanced by placing the simulation findings at the two scales in their anatomical contexts, as shown in Fig. 11.1. Cortical/subcortical interactions produce a changing pattern of NSF inputs to the cortex, mirrored in the $1/f^2$ background, and facilitating specific spatially organized transitions into autonomous gamma. Patches of autonomous gamma activity inject information into the wider cortical field, leading to the generation of fresh synchronous interactions. Linkages between excitatory cells—including voltage-dependent transmitters—link together patches of gamma activity in synchronous fields, whereas long-range excitatory flux to both excitatory and inhibitory cells (TCF) acts to suppress gamma oscillation, as is shown by the results in Fig. 11.4, (a) and (b), 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 [1], acts to favor particular synchronous fields, thus permitting a large set of possible states of attention. The inclusion of other, slower, cortical resonances in the theta and alpha bands, imposed upon the continuous interaction of time-varying synchronous fields, may be expected to lead to a shutter-like intermittency—as proposed in Freeman’s

cinematographic mechanism of perception [17]. The results in Figs. 11.5 and 11.6 show that, near transition levels of NSF, spectral tuning of gamma resonance below threshold, and the spectrum of autonomous gamma above threshold, are well matched, favoring selective information transfer among patches of gamma activity, even if widely separated.

Results in Fig. 11.4(c) are broadly in accord with Freeman's hypothesis of the origin of gamma [12], and help to explain experimental data on inhibitory/excitatory phase relations in gamma [27, 38]. However, these results 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 a stable phase relationship with each other. Thus, in the alert state, cortex may be normally poised near transition to oscillation, and sharp changes of excitatory/inhibitory relations may occur when the cortex moves into a locally autonomous mode. Fig. 11.7 emphasises the sudden change of phase at the transition—a change which is literally a “phase slip” (see below). This is consistent with Freeman's concepts of gamma activity as akin to thermodynamic phase transitions—although it is problematic whether phase transition in the formal thermodynamic sense is wholly applicable to a process involving alternation between a linear stochastic state and a nonlinear oscillating state.

11.4.3 Synchrony, traveling waves, and phase cones

The results shown in Figure 11.8 highlight the relationship between traveling waves and synchronous fields 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 as autonomous oscillation supervenes. These effects arise because intersecting cortical traveling waves exhibit annihilation of their anti-phase (odd) components and superposition of their in-phase (even) components [9], resulting in synchrony between equally co-active and reciprocally linked points, generated in times nearly as short as the one-way axonal conduction delay. Waves not intersecting with other co-active sites continue to propagate as pure traveling waves. Figure 11.9 shows the basis of synchrony as a consequence of superposition and cancelation of waves.

The roughly equal numbers of inwardly directed, and outwardly directed radiating waves, identified as phase cones by Freeman and colleagues [13–16, 18, 19], may be the two-dimensional equivalents of the unidirectional traveling waves shown in Fig. 11.8. This interpretation arises from consideration of the self-similar temporal character of the $1/f^2$ background, along with the multicentric variations of NSF supposed to arise from cortical/subcortical interactions. Experimentally and theoretically, electrocortical waves are approximately nondispersive and self-similar

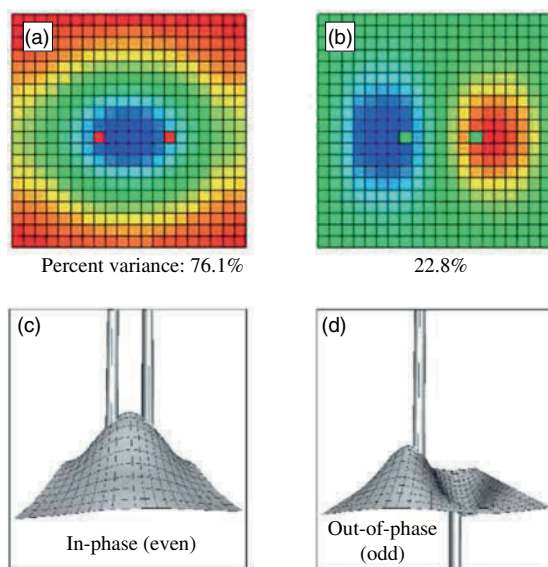


Fig. 11.9 [Color plate] Top row: First and second spatial eigenmodes of waves in a simulated cortical field [33] driven by two independent white-noise time-series of equal variance, applied at the points marked in (a) with red dots. Bottom row: A freeze-frame of motion associated with each eigenmode, showing that the dominant mode (c) arises from summation of inputs of even character at all dendritic summing junctions, while the minor eigenmode (d) arises from reciprocally canceling components.

[13, 43, 77], so there may be self-similarity, or at least considerable spatial complexity, in the fields of autonomous gamma triggered by continuous variation in perception and attention. Depending upon both their scale and position relative to recording electrodes, the fields of synchrony and traveling waves being continually generated and suppressed may be registered as phase cones with traveling wave components radiating either inwards or outwards, as is shown in Fig. 11.10. Association of traveling waves with transient synchronous fields, cone generation at frequencies corresponding to the cerebral rhythms, variation of cone size, origination at multiple foci, association with phase velocities less than, or of similar magnitude to, the conduction velocities of cortical axons, all follow as consequences, and accord with the experimental results. Synchrony, traveling waves, and phase cones.

11.4.4 Phase transitions and null spikes

In recent works Freeman and colleagues [17, 19, 20, 22] have drawn attention to the occurrence of behaviorally linked episodes in ECoG in which analytic power, measured by computing the Hilbert transform, drops to zero, accompanied by a

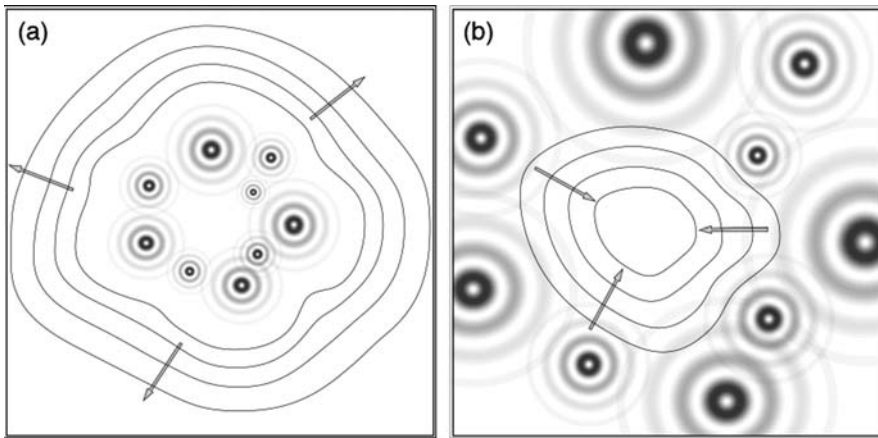


Fig. 11.10 Outwardly- and inwardly-radiating waves generating phase cones. Gray concentric circles represent cortical fields of synchronous oscillation, generated at two scales in a self-similar field of transitions to and from autonomous gamma. Contour lines of similar phase mark average radiation of traveling waves (a) outward, or (b) inward, in surrounding subthreshold field. (Reproduced from [71] with permission.)

sharp step in the accompanying analytic phase. These events they term “null spike”, and “phase slip”, respectively, and have considered these to be markers of phase transition, in the thermodynamic sense. They also draw attention to the occurrence of similar null spikes in brown noise as a purely random event [22], so it remains uncertain why these apparently random events may be behaviorally linked, and what the association with phase transition actually is.

Figure 11.11 indicates schematically how null spikes and phase slips may be systematically linked to transitions into autonomous gamma activity. The peak of the inverted “Mexican hat” represents a focus on the cortex, which is undergoing a transition into autonomous gamma activity. The spreading transcortical flux stabilizes the surrounding cortex, suppressing oscillation or excursion of the ECoG without significant effect on the average firing rate, as shown in Figs 11.4(a) and (b). This suppression will have the effects of lowering the amplitude, and of reducing the ECoG voltage towards its mean value over time—making zero-crossings of the time-base more likely. At the same time, any ECoG signal recorded from the burst of gamma oscillation occurring at the site of transition may exhibit brief epochs in which the form of the signal is symmetrical about some point in time. These signal characteristics, whether occurring concurrently or not, are each likely to produce an analytic null spike, which therefore has a somewhat ambiguous character since similar events may occur at random. A similar ambiguity may pertain for a phase slip, which might reflect merely the inaccurate calculation of phase near a zero-crossing, or may indicate a sharp change of state like the change from damped resonance to autonomous oscillation. A simple mathematical justification for these deductions is given in the Appendix (Sect. 11.6.2).

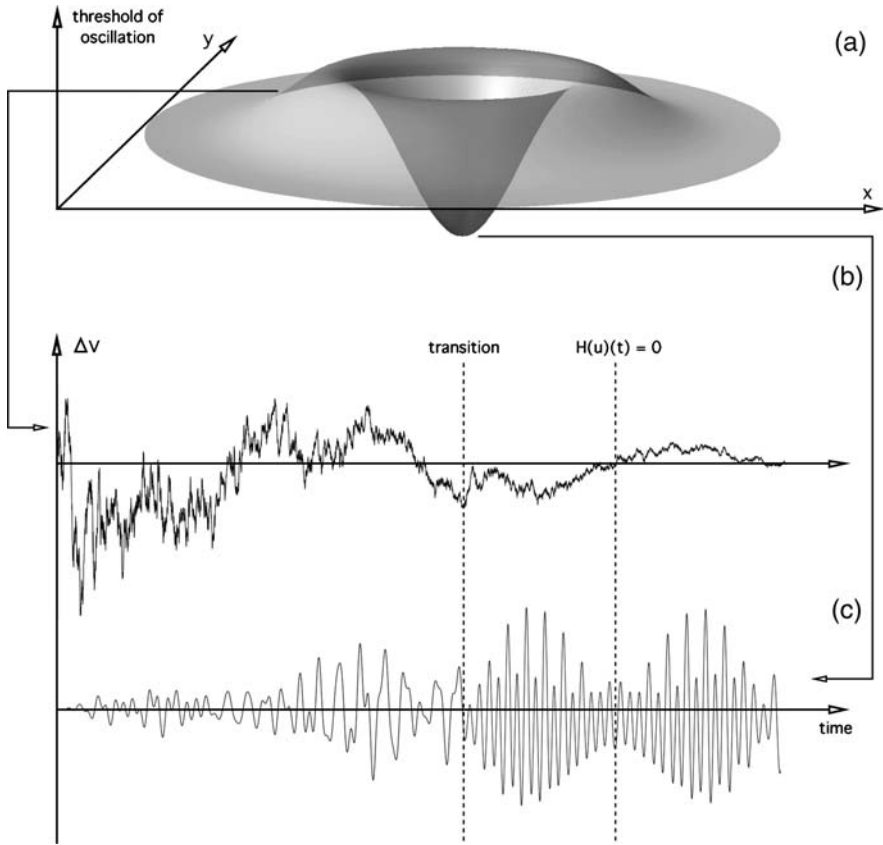


Fig. 11.11 Transition to autonomous gamma, and circumstances combining to favor the occurrence of null spikes and phase slips, measured from ECoG. (a) Schematic representation of a focus of autonomous gamma, suppressing oscillation in the surrounding field via TCF. (b) ECoG of the surrounding field. Zero-crossing in Gaussian noise favored by the spreading TCF. Zero-mean stochastic background ($1/f^2$ noise) favors $H(u)(t) = 0$ (subject to variation in the time-series sampled). (c) ECoG close to the focus of transition to autonomous gamma. Transition to oscillation with recurrent, time-symmetric character favors $H(u)(t) = 0$, with or without associated zero-crossing. (Reproduced from [71] with permission.)

11.5 Implications for cortical information processing

Preferential selection for input signals in the gamma range as patches of cortex approach transition, and the spectral similarity of the (relatively nonlinear) oscillations in the gamma range, offer a mechanism for tuned information exchange in cortex. The shift to coherent phase relations between excitatory and inhibitory cells, and facilitation of synapses in the far dendritic trees as autonomous gamma is generated, which would permit the ordered readout of information stored in the distal dendritic tree (see Fig. 11.12) while the specific set of inputs sensed in the

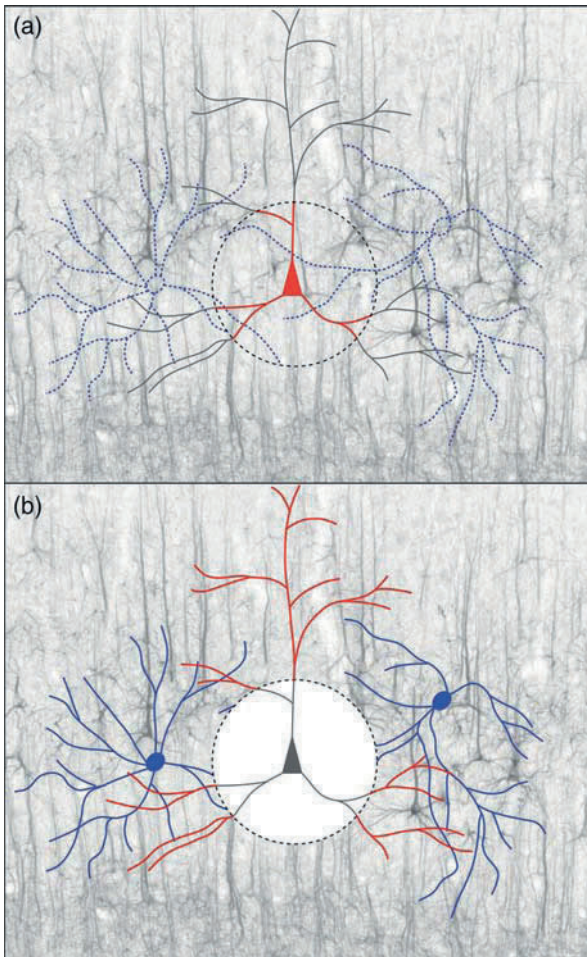


Fig. 11.12 [Color plate] Selection and recall of stored information. (a) Average neuronal state in stochastic background firing conditions. Pyramidal cells (red) are principally susceptible to inputs in the proximal dendritic tree. Inhibitory surround cells (blue) fire with variable timing relative to pyramidal cells. (b) Ordered neuronal state during gamma oscillation. Backpropagating action potentials block proximal synapses. Inhibitory and excitatory cells fire in phase. Synapses of distal dendritic trees mediate spatial and temporal patterns of synchronous oscillation.

proximal dendritic trees would select particular spatiotemporal patterns of synchronous oscillation. These patterns could be very complicated and various, as they draw on synaptic connections in the distal dendritic trees, which are effectively silent until they are brought into play by proximal synapses, which were themselves activated by distant fields of synchronously active cells. The transient synchronous fields, and the traveling wave patterns, can therefore be treated as the

signatures of changes of state in a finite state machine. This raises the important question of whether these dynamic properties can be linked to processes modifying individual synapses, to provide a general model of learning and memory, integrating cognitive activity at both individual cellular and populations scales in the brain.

Recent studies on synaptic modifications in hippocampal cells [31, 64, 65] have led to a distinction between synapses in the proximal dendritic tree, and those in the distal dendritic tree. Those of the proximal tree appear to follow a version of the Hebb rule in which synaptic gain is increased if presynaptic activity precedes postsynaptic, and is decreased if postsynaptic precedes presynaptic. This corresponds to a strengthening of input patterns mediated by one-way transmission—i.e., traveling waves. A second rule, termed the spatiotemporal learning rule (STLR) has been found to apply to the distal dendritic tree, where synapses appear little affected by the postsynaptic state, but mutually facilitate consolidation in neighboring, and co-active, synapses thus favoring the establishment of synapses linking cells engaged in synchronous activity.

The STLR and modified Hebb learning rules correspond in turn to distinctions made in the Coherent Infomax theory advanced by Kay and Phillips and colleagues [32, 33, 44]. This theoretical model also distinguishes two types of synaptic connection—CF connections which are assumed to mediate synchronous activity in adjacent cells, and RF connections, originally named to correspond to receptive field inputs in the visual cortex, but also applicable to any other feed-forward connections. Thus, CF and RF might be equated with synapses on the distal and proximal dendritic trees respectively. Kay and Phillips show that such a network can maximize the storage of information of both individual features transmitted by RF connections, and mutual information transmitted by CF connections, in multiple RF streams. They then show the consolidation of learning requires a learning rule similar to the Hebb rule (although not identical), but can proceed only under the influence of an activation function. The activation function requires that CF activity alone is not sufficient for learning to take place to learn contextual information: cells must be activated by RF connections as well as receiving CF input. Kay and Phillips identify this effect—similar to a gain control—to voltage-sensitive channels of the NMDA type. The idea of activation of the distal dendritic tree by back-propagating action potentials offers a more general mechanism for gain control. Applying the Coherent Infomax concept to more realistic neuronal models, the Relevant Infomax model advanced by Kording and Konig [34] also invokes the effect of back-propagation in the dendritic tree, and distinguishes synapses which largely determine firing from those that gate plasticity—a concept somewhat different from that advanced here.

At present it has not been demonstrated that the physiologically appropriate learning rules and the Coherent Infomax principle can be successfully transplanted to the dynamical model described here. Although the analogies seem clear enough, the information-theoretic ideas have as yet been applied only to small “toy” neuron sets. However, if the concepts do hold for more complex dynamics and connection systems, then a theoretical model of optimised contextual learning, the release of

complex spatiotemporally organized and stored memory sequences, and their relation to electrocortical fields, may be practicable in future.

Acknowledgments The author thanks Nick Hawthorn, Paul Bourke, Alistair Steyn-Ross, and Yanyang Xu for their help, and the Bioengineering Institute of the University of Auckland for computing resources.

11.6 Appendix

11.6.1 Model equations

Conventions

State variables are average membrane potentials, $V_{p,q}(\mathbf{r},t)$, pulse densities, $Q_{p,q}(\mathbf{r},t)$, and afferent synaptic flux $\phi_{p,q}(\mathbf{r},t)$. 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 $[R] = [\text{NDMA}], [\text{AMPA}], [\text{GABA}_a]$ indicate neurotransmitter receptor types. State equations are steady-state functions of state variables, and lag-response functions in $\tau = n\delta t$, where δt is the time-step, and $n = 1, 2, \dots$. Lag-response functions are normalised so that $\int_0^\infty f(\tau) d\tau = 1$.

Parameter values and references to their derivations are given in Wright (2009) [70, 71] and are based upon largely independent primary and secondary sources [5, 6, 10, 28, 30, 35, 36, 39, 41, 42, 45–47, 52, 55, 58, 60–62].

Afferent synaptic flux

The distribution of the neuron cell bodies giving rise to afferents at a cortical point, \mathbf{r} , is $f(\mathbf{r}, \mathbf{r}')$, where $\{\mathbf{r}'\}$ are all other points in the field. Connection densities are reciprocal for all $\{\mathbf{r}, \mathbf{r}'\}$. The *afferent flux density*, $\phi_p(\mathbf{r}, t)$, the population average input pulse rate per synapse, is given by

$$\phi_p(\mathbf{r}, t) = \int_0^\infty f(\mathbf{r}, \mathbf{r}') Q_p(\mathbf{r}, t - |\mathbf{r} - \mathbf{r}'|/v_p) d^2\mathbf{r}', \quad (11.1)$$

where $Q_p(\mathbf{r}, t)$ are mean pulse rates of neurons at \mathbf{r}' , at t , also termed *pulse density*, and v_p is the velocity of axonal conduction. Here, $f(\mathbf{r}, \mathbf{r}')$ describes intracortical and cortico-cortical connections, approximated as Gaussian [6]:

$$f(\mathbf{r}, \mathbf{r}') = \frac{1}{2\pi\gamma^2} \exp(-|\mathbf{r} - \mathbf{r}'|^2/2\gamma^2), \quad (11.2)$$

where γ is the standard deviation axonal range. Equation (11.2) can be applied separately to the short intracortical excitatory and inhibitory fibers, and to the long

range, wholly excitatory, cortico-cortical connections. Steps necessary to compute depolarization as a function of afferent flux density (see Eqs. (11.10) and (11.11)) and the subsequent regeneration of pulses (Eq. (11.12)) are next described.

Transcortical flux (TCF) and nonspecific flux (NSF)

A major distinction is made between the afferent synaptic flux transmitted by cortico-cortical fibres, here termed the *transcortical flux* (TCF; see Fig. 11.2), and excitatory synaptic flux delivered wholly to the excitatory cortical cells, from the reticular activation system, and termed *nonspecific afferent flux* (NSF; see Fig. 11.2). In the macrocolumnar-scale simulations, NSF is introduced by adding a given value of synaptic flux, in spikes per second, weighted by $N_{ee,ns}/N_{TOT}$, the fraction of synapses per excitatory cortical cell attributable to subcortical afferents, to $Q_e(\mathbf{r}, t)$, the pulse densities of the excitatory cells.

Synaptic receptor dynamics

The postsynaptic impact of $\phi_p(\mathbf{r}, t)$ is modified by changes in the conformation of ion channels. The open-channel steady state is

$$J^{[R]}(\phi_p) = \exp\left(-\lambda^{[R]}\phi_p\right)\phi_p, \quad (11.3)$$

and $\Phi^{[R]}(\tau)$ describes the rise and fall of receptor adaptation to a brief afferent stimulus

$$\Phi^{[R]}(\tau) = \left[\sum_n B_n^{[R]} / \beta_n^{[R]} - \sum_m A_m^{[R]} / \alpha_m^{[R]} \right]^{-1} \times \left[\sum_n B_n^{[R]} \exp\left(-\beta_n^{[R]}\tau\right) - \sum_m A_m^{[R]} \exp\left(-\alpha_m^{[R]}\tau\right) \right], \quad (11.4)$$

where $\{\lambda^{[R]}, B_n^{[R]}, A_m^{[R]}, \beta_n^{[R]}, \alpha_m^{[R]}\}$ with $m, n = 1, 2, 3, \dots$, are derived from transmitter/receptor models.

Postsynaptic membrane gain

Afferent synaptic flux, modified by synaptic adaptation, generates a change in average membrane potential, V_q , with a steady state solution

$$M^{[R]}(V_q, \phi_p) = g_p^{[R]} \left(\frac{V_p^{\text{rev}} - V_q}{V_p^{\text{rev}} - V_q^{[0]}} \right) J^{[R]}. \quad (11.5)$$

Here, $g_p^{[R]}$ is the synaptic gain at resting membrane potential, V_p^{rev} is the excitatory or inhibitory reversal potential, and $V_q^{[0]}$ is the resting membrane potential. Since V_q and ϕ_p are serially-dependent state variables, $V_q(t)$ must be substituted by $V_q(t - \delta t)$ in computation.

Dendritic time- and space-response

The rise and fall of postsynaptic membrane potential at the sites of synaptic input is given by,

$$\Psi^{[R]}(\tau) = \frac{a_{qp}b_{qp}}{b_{qp} - a_{qp}} (\exp[-a_{qp}\tau] - \exp[-b_{qp}\tau]), \quad (11.6)$$

where $\{a_{qp}, b_{qp}\}$ are constants. Postsynaptic depolarization, transferred by cable effects, reaches the action potential trigger points after delays which are greater from synapses in the distal dendritic trees than from proximal dendritic trees. The relative magnitudes of depolarization reaching the trigger points over a spread of arrival times is given by,

$$L^j(\tau) = \mathcal{A}^j \exp[-\mathcal{A}^j \tau]. \quad (11.7)$$

The \mathcal{A}^j are constants, and $j = n, f$ indicates relationship to the near or far dendritic trees respectively.

Effects of action potential back-propagation

At the release of an action potential, anterograde and retrograde propagation takes place, the latter depolarizing the membrane throughout the proximal dendritic tree [58]. It is assumed that when the neuron is fully repolarized, the greatest weight in the generation of a subsequent action potential can be ascribed to activity at the near synapses, because of their weighting by proximity to the axon hillock. On the release of an action potential, the near synapses become reduced in efficacy to zero during the absolute refractory period, and the distal synaptic trees become partially depolarized, so that determination of whether or not a subsequent action potential is generated at the conclusion of the relative refractory period is then relatively weighted toward activity at the far synapses.

The fractions of neurons A^n, A^f , having respective biases toward activation from the near or far dendritic trees, are

$$A^f(t) = Q_q(t)/Q_q^{\text{max}}, \quad (11.8)$$

$$A^n(t) = 1 - Q_q(t)/Q_q^{\text{max}}, \quad (11.9)$$

where Q_q^{max} is the maximum firing rate of neurons and reflects the refractory period.

Fractional distributions, $r^{n[R]} + r^{f[R]} = 1$, 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 $r^{f[\text{NMDA}]} = 1$.

Aggregate depolarization

The voltage at the trigger points for action potential generation, ψ_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_{qp} ,

$$\psi_q(t) = \sum_p \sum_j \sum_{[R]} N_{qp} A^j r^{j[R]} \left(((M^{[R]} \otimes \Phi^{[R]}) \otimes \Psi^{[R]}) \otimes L^j \right), \quad (11.10)$$

where \otimes indicates convolution in time. In the population average,

$$V_q(t) \approx V_q^{[0]} + \psi_q(t). \quad (11.11)$$

Equation (11.11) establishes $V_q(t - \delta t)$ for the next time-step in Eq. (11.5).

Action potential generation

From Eq. (11.11), the mean firing rate is calculated from

$$Q_q(t) = Q_q^{\max} / (1 + \exp[-\pi(V_q - \theta_q) / \sqrt{3}\sigma_q]), \quad (11.12)$$

yielding the pulse densities of neurons required in Eq. (11.1). Here, θ_q is the mean value of V_q at which 50% of neurons are above threshold for the emission of action potentials; and σ_q approximates one standard deviation of probability of emission of an action potential in a single cell, as a function of V_q . For a comparison with standard EEG and local field potential (LFP) data, we also assume $\text{LFP} \equiv V_e(t)$.

Application at mesoscopic (macrocolumnar) and macroscopic (centimetric) spatial scales

The above equations are 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 position \mathbf{r} , surrounded by other elements at positions $\{\mathbf{r}'\}$, coupled as in Eq. (11.1) with delays, $\delta_p = |\mathbf{r} - \{\mathbf{r}'\}| / v_p$, and $f(\mathbf{r}, \{\mathbf{r}'\})$ chosen as the sum of two-dimensional Gaussian distributions of connections, each Gaussian term appropriate to intracortical and cortico-cortical connections, of excitatory and inhibitory types. The grid can be used to represent the cortex at any chosen spatial scale, by applying

a physiologically appropriate value for v_p , and setting excitatory/inhibitory axonal connection ranges appropriate to scale. Two configurations are used:

- a centimetric scale, which treats inhibitory and short intracortical connections as local to each element, with cortico-cortical connections linking elements. Nonspecific afferent flux (NSF) is applied uniformly to all excitatory elements, and thus provides a single control parameter;
- a macrocolumnar scale, which connects elements together by both intracortical excitatory and inhibitory connections, forming a “Mexican hat” connection field of approximately 300- μm diameter. At this scale, no cortico-cortical connections are represented explicitly. Instead, the transcortical flux is introduced as a spatially uniform input to all elements of the simulation. Thus, at macrocolumnar scale, there are two control parameters—the nonspecific flux (NSF), and the transcortical flux (TCF).

Numerical considerations

Simulation time-step was 0.1 ms. Individual simulation runs were considered to have reached a statistically stationary state at $t = 200$ s after initialization with state variables set to zero. 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 independently obtained 0.8192-s final epochs were analyzed for all spectral and correlation analyses. Noise inputs were applied as zero-mean signals added to the applied constant values of NSF, and unless otherwise stated, were applied to the top row (row-0) of the simulated grid of units, while excitatory cell potentials were recorded from the element at row-10, column-10.

Simulated gamma oscillation was defined as oscillation with a peak frequency in the 30–60-Hz band, associated with a threshold reached with increasing NSF producing transition from a static steady-state (in the absence of noise) or dampened oscillation (in the presence of driving noise), to an autonomous oscillation, with the transition occurring below a mean excitatory firing rate of 20 s^{-1} , and not associated with excursions of membrane potential encountering reversal-potential bounds. Normal firing patterns of cortical neurones other than gamma were equated with stochastic background and identified with the simulation’s non-oscillating states.

11.6.2 Hilbert transform and null spikes

The Hilbert transform is given by

$$H(u)(t) = -\frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \int_{\varepsilon}^{\infty} \frac{u(t + \tau') - u(t - \tau')}{\tau'} d\tau', \quad (11.13)$$

where $u(t)$ is a stationary, continuous, and infinite-duration signal, and τ' is the temporal lag. In a discrete approximation for a time-limited epoch,

$$H(u)(t) = \frac{1}{\pi} \sum_{\tau=1}^m \frac{u(t+\tau) - u(t-\tau)}{\tau} \quad (11.14)$$

where t is now a dimensionless time-index, $\tau = \tau'/\varepsilon$, ε is the time-step, and m is the number of forward time-steps in the epoch. Let $\bar{u}(t)$ be the mean value of the set $S \equiv \{u(t+\tau) - u(t-\tau)\}$, and let

- $\{\tilde{u}(t, \tau)\} \subseteq S$ be the subset whose members are equal to $\bar{u}(t)$,
- $\{\hat{u}(t, \tau)\} \subseteq S$ be the subset whose members are each greater than $\bar{u}(t)$,
- $\{\check{u}(t, \tau)\} \subseteq S$ be the subset whose members are each less than $\bar{u}(t)$,

then let $\sum \tilde{u}(t, \tau)/\tau$, $\sum \hat{u}(t, \tau)/\tau$, $\sum \check{u}(t, \tau)/\tau$ be the sums of terms, each weighted by τ , in the respective subsets. Thus,

$$H(u)(t) = \frac{1}{\pi} \left(\sum \frac{\tilde{u}(t, \tau)}{\tau} + \sum \frac{\hat{u}(t, \tau)}{\tau} + \sum \frac{\check{u}(t, \tau)}{\tau} \right). \quad (11.15)$$

Special cases in which $H(u)(t) = 0$ are:

1. For all $\tau \geq \varepsilon$, $u(t+\tau) = u(t-\tau)$, hence

$$\sum \tilde{u}(t, \tau)/\tau = \sum \hat{u}(t, \tau)/\tau = \sum \check{u}(t, \tau)/\tau = 0; \quad (11.16)$$

2. When elements of S are randomly distributed about $\bar{u}(t)$, with $\bar{u}(t) = 0$, hence

$$\sum \hat{u}(t, \tau)/\tau = -\sum \check{u}(t, \tau)/\tau, \quad \text{and} \quad \sum \tilde{u}(t, \tau)/\tau = 0. \quad (11.17)$$

Therefore analytic power, $H^2(u)(t) + u^2(t)$, can approach zero where:

1. $u(t)$ is symmetric about some t , and $u(t) \rightarrow 0$; (this includes the special case where $u(t) = 0$ for all t);
2. $u(t)$ is close to or identical with a zero-crossing in a sample of a zero-mean Gaussian noise.

Analytic phase, $\tan^{-1}[H(u)(t)/u(t)]$, is ill-defined where $u(t) \rightarrow 0$, but a sharp change in analytic phase might also be detected in association with physiological equivalent of the step-like transition shown in Fig. 11.7.

References

1. Alexander, G.E., Crutcher, M.D., DeLong, M.R.: Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, prefrontal and limbic functions. In: H.B.M. Uylings (ed.), *The Prefrontal Cortex: Its Structure, Function, and Pathology*, Elsevier, Amsterdam (1990)
2. Amit, D.J.: *Modelling Brain Function*. Cambridge University Press, Cambridge (1989)

3. Arbib, M.A. (ed.): *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, Massachusetts (1995)
4. Barrie, J.M., Freeman, W.J., Lenhart, M.D.: Spatiotemporal analysis of prepyriform, visual, auditory, and somesthetic surface EEGs in trained rabbits. *J. Neurophysiol.* **76**, 520–539 (1996)
5. Bower, J., Beeman, D.: *The Book of GENESIS. Exploring Realistic Neural Models with the GENeral NEural SIMulation System*. Springer, New York, 2nd edn. (1998)
6. Braitenberg, V., Schuz, A.: *Anatomy of the Cortex: Statistics and Geometry*. Springer-Verlag, Berlin, New York (1991)
7. Bressler, S.L., Coppola, R., Nakamura, R.: Episodic multiregional cortical coherence at multiple frequencies during visual task performance. *Nature* **366**, 153–156 (1993), doi:10.1038/366153a0
8. Buzsaki, G., Draguhn, A.: Neuronal oscillations in cortical networks. *Science* **304**, 1926–1929 (2004), doi:10.1126/science.1099745
9. Chapman, C.L., Bourke, P.D., Wright, J.J.: Spatial eigenmodes and synchronous oscillation: Coincidence detection in simulated cerebral cortex. *J. Math. Biol.* **45**, 57–78 (2002), doi:10.1007/s002850200141
10. Dominguez-Perrot, C., Feltz, P., Poulter, M.O.: Recombinant GABA_A receptor desensitization: The role of the gamma2 subunit and its physiological significance. *J. Physiol.* **497**, 145–159 (1996)
11. Eckhorn, R., Bauer, R., Jordon, W., Brosch, M., Kruse, W., Monk, M., Reitboeck, H.J.: Coherent oscillations: A mechanism of feature linking in the in the visual cortex? *Biol. Cybern.* **60**, 121–130 (1988), doi:10.1007/BF00202899
12. Freeman, W.J.: Predictions on neocortical dynamics derived from studies of paleocortex. In: *Induced Rhythms of the Brain*, Birkhauser, Boston (1991)
13. Freeman, W.J.: Origin, structure and role of background EEG activity. Part 1: Analytic amplitude. *Clin. Neurophysiol.* **115**, 2077–2088 (2004), doi:10.1016/j.clinph.2004.02.029
14. Freeman, W.J.: Origin, structure and role of background EEG activity. Part 2: Analytic phase. *Clin. Neurophysiol.* **115**, 2089–2107 (2004), doi:10.1016/j.clinph.2004.02.028
15. Freeman, W.J.: Origin, structure and role of background EEG activity. Part 3: Neural frame classification. *Clin. Neurophysiol.* **116**, 1118–1129 (2005), doi:10.1016/j.clinph.2004.12.023
16. Freeman, W.J.: Origin, structure and role of background EEG activity. Part 4: Neural frame simulation. *Clin. Neurophysiol.* **117**, 572–589 (2006), doi:10.1016/j.clinph.2005.10.025
17. Freeman, W.J.: Proposed cortical ‘shutter’ mechanism in cinematographic perception. In: L.I. Perlovsky, R. Kozma (eds.), *Neurodynamics of Cognition and Consciousness*, pp. 11–38, Springer, Heidelberg (2007), doi:10.1007/978-3-540-73267-9_2
18. Freeman, W.J., Barrie, J.M.: Analysis of spatial patterns of phase in neocortical gamma EEGs in rabbit. *J. Neurophysiol.* **84**, 1266–1278 (2000)
19. Freeman, W.J., Holmes, M.D., West, G.A., Vanhatalo, S.: Fine spatiotemporal structure of phase in human intracranial EEG. *Clin. Neurophysiol.* **117**, 1228–1243 (2006), doi:10.1016/j.clinph.2006.03.012
20. Freeman, W.J., Vitiello, G.: Dissipation and spontaneous symmetry breaking in brain dynamics. *J. Phys. A: Gen. Phys.* **41**, 17p (2008), doi:10.1088/1751-8113/41/30/304042
21. Freeman, W.: *Mass Action in the Nervous System*. Academic Press, New York (1975)
22. Freeman, W., O’Nuillain, S., Rodriguez, J.: Simulating cortical background activity at rest with filtered noise. *J. Integr. Neurosci.* **7**(3), 337–344 (2008), doi:10.1142/S0219635208001885
23. Gray, C.M., Engel, A.K., Konig, P., Singer, W.: Synchronization of oscillatory neuronal responses in cat striate cortex: Temporal properties. *Vis. Neurosci.* **8**, 337–347 (1992)
24. Gray, C.M., Konig, P., Engel, A.K., Singer, W.: Oscillatory responses in cat visual cortex exhibit intercolumnar synchronisation which reflects global stimulus properties. *Nature* **388**, 334–337 (1989), doi:10.1038/338334a0

25. Gray, C.M., Singer, W.: Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1698–1702 (1989), doi:10.1073/pnas.86.5.1698
26. Haken, H.: *Principles of Brain Functioning*. Springer, Berlin (1996)
27. Hasenstaub, A., Shu, Y., Haider, B., Kraushaar, U., Duque, A., McCormick, D.A.: Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron* **47**, 423–435 (2005), doi:10.1016/j.neuron.2005.06.016
28. Hausser, W., Roth, A.: Dendritic and somatic glutamate receptor channels in rat cerebellar purkinje cells. *J. Physiol.* **501**, 77–95 (1997), doi:10.1111/j.1469-7793.1997.077bo.x
29. Jirsa, V.K., Haken, H.: Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996), doi:10.1103/PhysRevLett.77.960
30. Kandel, E.R., Schwartz, J.H., Jessell, T.M.: *Principles of Neural Science*. Prentice-Hall International, London, 3 edn. (1991)
31. Kaneki, K., Ariki, O., Tsukada, M.: Dual synaptic plasticity in the hippocampus: Hebbian and spatiotemporal learning dynamics. *Cogn. Neurodyn.* (**in Press**) (2008), doi:10.1007/s11571-008-9071-z
32. Kay, J., Floreano, D., Phillips, W.A.: Contextually guided unsupervised learning using local multivariate binary processors. *Neural Netw.* **11**(1), 117–140 (1998), doi:10.1016/S0893-6080(97)00110-X
33. Kay, J., Phillips, W.A.: Activation functions, computational goals and learning rules for local processors with contextual guidance. *Neural Comp.* **9**, 763–768 (1997), doi:10.1162/neco.1997.9.4.895
34. Kording, K.P., Konig, P.: Learning with two sites of synaptic integration. *Network: Comp. Neural Sys.* **11**, 25–39 (2000), doi:10.1088/0954-898X/11/1/302
35. Lester, R.A., Jahr, C.E.: NMDA channel behavior depends on agonist affinity. *J. Neurosci.* **12**, 635–643 (1992)
36. Liley, D.T.J., Wright, J.J.: Intracortical connectivity of pyramidal and stellate cells: Estimates of synaptic densities and coupling symmetry. *Network: Comp. Neural Sys.* **5**, 175–189 (1994), doi:10.1088/0954-898X/5/2/004
37. Miltner, W.H., Braun, C., Arnold, M., Witte, H., Taube, E.: Coherence of gamma-band EEG activity as a basis for associative learning. *Nature* **397**, 434–436 (1999), doi:10.1038/17126
38. Morita, K., Kalra, R., Aihara, K., Robinson, H.P.C.: Recurrent synaptic input and the timing of gamma-frequency-modulated firing of pyramidal cells during neocortical “up” states. *J. Neurosci.* **28**, 1871–1881 (2008), doi:10.1523/JNEUROSCI.3948-07.2008
39. Mountcastle, V.B.: An organizing principle for cerebral function: The unit module and the distributed system. In: *The Neurosciences 4th Study Program*, MIT Press, Cambridge, Mass (1979)
40. Neuenschwander, S., Singer, W.: Long-range synchronisation of oscillatory light responses in the cat retina and lateral geniculate nucleus. *Nature* **379**, 728–733 (1996), doi:10.1038/379728a0
41. Nunez, P.L.: *Electric Fields of the Brain*. Oxford University Press, New York (1981)
42. Nunez, P.L.: *Neocortical Dynamics and Human EEG Rhythms*. Oxford University Press, New York (1995)
43. O’Connor, S.C., Robinson, P.A.: Wave-number spectrum of electrocorticographic signals. *Phys. Rev. E* **67**, 1–13 (2003), doi:10.1103/PhysRevE.67.051912
44. Phillips, W.A., Singer, W.: In search of common foundations for cortical computation. *Behav. Brain Sci.* **20**, 657–722 (1997), doi:10.1017/S0140525X9700160X
45. Rennie, C.J., Robinson, P.A., Wright, J.J.: Effects of local feedback on dispersion of electrical waves in the cerebral cortex. *Phys. Rev. E* **59**, 3320–3329 (1999), doi:10.1103/PhysRevE.59.3320
46. Rennie, C.J., Robinson, P.A., Wright, J.J.: Unified neurophysical model of EEG spectra and evoked potentials. *Biol. Cybern.* **86**, 457–471 (2002), doi:10.1007/s00422-002-0310-9

47. Rennie, C.J., Wright, J.J., Robinson, P.A.: Mechanisms of cortical electrical activity and the emergence of gamma rhythm. *J. Theoretical. Biol.* **205**, 17–35 (2000), doi:10.1006/jtbi.2000.2040
48. Robinson, P.A., Rennie, C.J., Rowe, D.L., O'Connor, S.C., Wright, J.J., Gordon, E.: Neurophysical modelling of brain dynamics. *Neuropsychopharmacol.* **28**, S74–S79 (2003), doi:10.1038/sj.npp.1300143
49. Robinson, P.A., Rennie, C.J., Wright, J.J.: Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* **56**, 826–840 (1997), doi:10.1103/PhysRevE.56.826
50. Robinson, P.A., Rennie, C.J., Wright, J.J., Bahramali, H., Gordon, E., Rowe, D.L.: Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E* **63**, 1–18 (2001), doi:10.1103/PhysRevE.63.021903
51. Robinson, P.A., Wright, J.J., Rennie, C.J.: Synchronous oscillations in the cerebral cortex. *Phys. Rev. E* **57**, 4578–4588 (1998), doi:10.1103/PhysRevE.57.4578
52. Scholl, D.A.: *The Organization of the Cerebral Cortex*. Wiley, New York (1956)
53. Singer, W.: Putative functions of temporal correlations in neocortical processing. In: C. Koch, J.L. Davis (eds.), *Large Scale Neuronal Theories of the Brain*, MIT Press, Cambridge Mass., London (1994)
54. Singer, W., Gray, C.M.: Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* **18**, 555–586 (1995), doi:10.1146/annurev.ne.18.030195.003011
55. Steriade, M., Timofeev, I., Grenier, F.: Natural waking and sleep states: A view from inside cortical neurons. *J. Neurophysiol.* **85**, 1969–1985 (2001)
56. Steyn-Ross, D.A., Steyn-Ross, M.L., Sleight, J.W., Wilson, M.T., Gillies, I.P., Wright, J.J.: The sleep cycle modelled as a cortical phase transition. *J. Biol. Phys.* **31**, 547–569 (2005), doi:10.1007/s10867-005-1285-2
57. Stryker, M.P.: Is grandmother an oscillation? *Nature* **388**, 297–298 (1989), doi:10.1038/338297a0
58. Stuart, G.J., Sakmann, B.: Active propagation of somatic action potentials into neocortical cell pyramidal dendrites. *Nature* **367**, 69–72 (1994), doi:10.1038/367069a0
59. Suffczynski, P., Kalitzin, S., Pfurtscheller, G., Lopes da Silva, F.H.: Computational model of thalamocortical networks: Dynamical control of alpha rhythms in relation to focal attention. *Int. J. Psychophysiol.* **43**, 25–40 (2001), doi:10.1016/S0167-8760(01)00177-5
60. Szentagothai, J.: Local neuron circuits of the neocortex. In: F. Schmitt, F. Worden (eds.), *The Neurosciences 4th Study Program*, pp. 399–415, MIT Press, Cambridge, Mass (1979)
61. Thomson, A.M.: Activity dependent properties of synaptic transmission at two classes of connections made by rat neocortical pyramidal neurons in vitro. *J. Physiol.* **502**, 131–147 (1997), doi:10.1111/j.1469-7793.1997.131bl.x
62. Thomson, A.M., West, D.C., Hahn, J., Deuchars, J.: Single axon IPSPs elicited in pyramidal cells by three classes of interneurons in slices of rat neocortex. *J. Physiol.* **496**, 81–102 (1996)
63. Traub, R.D., Whittington, M.A., Stanford, I.M., Jeffereys, J.G.R.: A mechanism for generation of long-range synchronous fast oscillations in the cortex. *Nature* **383**, 621–624 (1996), doi:10.1038/383621a0
64. Tsukada, M., Aihara, T., Saito, H.: Hippocampal LTP depends on spatial and temporal correlation of inputs. *Neural Netw.* **9**, 1357–1365 (1996), doi:10.1016/S0893-6080(96)00047-0
65. Tsukada, M., Yamazaki, Y., Kojima, H.: Interaction between the spatio-temporal learning rule (STLR) and Hebb in single pyramidal cells in the hippocampal CA1 area. *Cogn. Neurodyn.* **1**, 305–316 (2007), doi:10.1007/s11571-006-9014-5
66. van Rotterdam, A., Lopes da Silva, F.H., van den Ende, J., Viergever, M.A., Hermans, A.J.: A model of the spatio-temporal characteristics of the alpha rhythm. *Bull. Math. Biol.* **44**, 283–305 (1982), doi:10.1016/S0092-8240(82)80070-0
67. von der Malsburg, C.: How are nervous structures organised? In: E. Basar (ed.), *Synergetics of the Brain*, Springer, Berlin, Heidelberg, New York (1983)
68. Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973), doi:10.1007/BF00288786

69. Wright, J.J.: Reticular activation and the dynamics of neuronal networks. *Biol. Cybern.* **62**, 289–298 (1990), doi:10.1007/BF00201443
70. Wright, J.J.: Cortical phase transitions: Properties demonstrated in continuum simulations at mesoscopic and macroscopic scales. *J. New Math. Nat. Comput.* (**in press**) (2009)
71. Wright, J.J.: Generation and control of cortical gamma: Findings from simulation at two scales. *Neural Netw.* (**in press**) (2009)
72. Wright, J.J., Alexander, D.M., Bourke, P.D.: Contribution of lateral interactions in V1 to organization of response properties. *Vision Res.* **46**, 2703–2720 (2006), doi:10.1016/j.visres.2006.02.017
73. Wright, J.J., Bourke, P.D.: An outline of functional self-organization in V1: Synchrony, str and hebb rules. *Cogn. Neurodyn.* **2**, 147–157 (2008), doi:10.1007/s11571-008-9048-y
74. Wright, J.J., Bourke, P.D., Chapman, C.L.: Synchronous oscillation in the cerebral cortex and object coherence: Simulation of basic electrophysiological findings. *Biol. Cybern.* **83**, 341–353 (2000), doi:10.1007/s004220000155
75. Wright, J.J., Liley, D.T.J.: Dynamics of the brain at global and microscopic scales. *Neural networks and the EEG.* *Behav. Brain Sci.* **19**, 285–320 (1996)
76. Wright, J.J., Rennie, C.J., Lees, G.J., Robinson, P.A., Bourke, P.D., Chapman, C.L., Gordon, E., Rowe, D.L.: Simulated electrocortical activity at microscopic, mesoscopic, and global scales. *Neuropsychopharmacol.* **28**, 80–93 (2003), doi:10.1038/sj.npp.1300138
77. Wright, J.J., Sergejew, A.A.: Radial coherence, wave velocity and damping of electrocortical waves. *Electroencephalogr. Clin. Neurophysiol.* **79**, 403–412 (1991), doi:10.1016/0013-4694(91)90205-1

Chapter 12

Cortical patterns and gamma genesis are modulated by reversal potentials and gap-junction diffusion

M.L. Steyn-Ross, D.A. Steyn-Ross, M.T. Wilson, and J.W. Sleigh

12.1 Introduction

Continuum models of the cortex aim to describe those interactions of neural populations that generate the electrical fluctuations and rhythms able to be detected directly, with scalp and cortical EEG (electroencephalogram) electrodes, or remotely, using their magnetic counterpart, via MEG (magnetoencephalogram) sensors. Because the numbers of neurons involved in these cooperative behaviors is so vast, the continuum, or *mean-field*, approach makes no attempt to model the detailed biophysics of individual neurons, nor does it attempt to track the birth and axonal propagation of individual spike events. Instead, neuronal properties are represented as spatial averages, averaged, say, over the population of neurons sampled by a small EEG electrode, with spiking activity being represented as an average *firing rate* for the population-average neuron.

When constructing a theoretical model for the cerebral cortex, there is an unavoidable tension between the competing requirements of biophysical accuracy (leading to increased complexity) versus mathematical tractability (arguing for simplicity). In the end, we must make a pragmatic assessment of model quality by asking: Is the model fit for purpose—i.e., Is the model able to make predictions that can be tested against biological reality? And: Does the model provide fresh insight?

In this chapter we will argue that incorporation of two biophysical features—namely, cell-reversal potentials, and direct diffusive coupling between inhibitory neurons—has important implications for emergent nonlinear behavior with respect to oscillatory rhythms and pattern formation in the cortex. Specifically, we show that

Moira L. Steyn-Ross · D. Alistair Steyn-Ross · Marcus T. Wilson
Department of Engineering, University of Waikato, P.B. 3105, Hamilton 3240, New Zealand.
e-mail: msr@waikato.ac.nz asr@waikato.ac.nz m.wilson@waikato.ac.nz
<http://phys.waikato.ac.nz/cortex/>

Jamie W. Sleigh
Waikato Clinical School, University of Auckland, Waikato Hospital, Hamilton 3204, New Zealand.
e-mail: sleighj@waikatodhb.govt.nz

the manner in which reversal potentials enter the model determines whether cortical oscillations appear in the delta (~ 2 -Hz) or in the gamma (~ 30 -Hz) frequency range. Further, we will demonstrate that inclusion of gap-junction diffusive connections modifies the strength and spatial extent of the Turing and standing-wave firing-rate patterns that can form in the cortical sheet.

12.1.1 Continuum modeling of the cortex

While continuum models of the cortex have evolved considerably since the foundation work of Wilson and Cowan [31], Nunez [18], and Freeman [9], present mean-field models continue to share three simplifying assumptions: (i) neural properties can be represented as spatial averages, (ii) neural inter-connectedness decays with distance, (iii) neural firing rates can vary between zero and some maximum value, with a sigmoidal mapping from membrane voltage to firing rate.

Continuum models are expressed either as coupled partial differential equations (PDEs), or as integro-differential equations (IDEs), or as purely integral equations, with the choice of representation being determined by the type and dimensionality (1-D or 2-D) of the connectivity kernel. The PDE forms have the advantage of speed and ease of analysis, but have access to a restricted range of connectivity kernels, e.g., exponential decay in 1-D [11, 31], modified Bessel (Macdonald)-function decay in 2-D [21]. The integral forms are slower to compute numerically, and require a large amount of storage, but have the advantage that the kernel can be chosen at will; Wright and Liley [35] use a Gaussian to represent the decreasing synaptic density with distance. For the integro-differential forms, “Mexican hat” connectivity kernels are frequently used [4–6, 14, 15].

In the PDE-based cortical model we present here, flux activity generated by excitatory and inhibitory neural populations is received at a dendritic synapse whose transmission efficiency is modulated by the difference between the membrane voltage and its reversal potential [16, 20]. Following Robinson *et al.* [21], axonal flux transmission is assumed to obey a 2-D wave equation with Macdonald-function connectivity. The net neuron voltage is determined not only by axono-dendritic activity at chemical synapses, but also by diffusive currents from adjacent neurons that are directly coupled to the target neuron via gap junctions. Our parameter values for the chemical-synaptic component of the model largely match those of Rennie *et al.* [20], but we have chosen to retain the symbols and labeling conventions used in our earlier sleep [23] and anesthesia modeling [24], which drew on work by Liley *et al.* [16]. For the gap-junction component of the model, we adopt the values and notation we introduced in Ref. [26].

12.1.2 Reversal potentials

The size and direction of the postsynaptic potential evoked at a chemical synapse by incoming spike activity depends on the voltage state of the receiving neuron,

and, in particular, on $(V^{\text{rev}} - V)$, the voltage of the receiving dendrite relative to its reversal potential. If this difference is large, spike events will be more effective at transferring charge across, and eliciting a voltage response in, the post-synaptic membrane; this efficiency diminishes to zero as V approaches V^{rev} , the reversal potential being ~ 0 mV for excitatory events (mediated by AMPA receptors) and ~ -70 mV for inhibitory events (mediated by GABA receptors).

Although a standard feature in all Hodgkin–Huxley [13] conductance-based neuron models, surprisingly few mean-field cortical models include excitatory and inhibitory reversal potentials [16, 20, 25, 36]. The neglect of these biophysical constraints might be justifiable if the voltage fluctuations about resting equilibrium ($V^{\text{rest}} \approx -60$ mV) remain sufficiently small that the reversal potentials are effectively infinite. But if the fluctuations grow sufficiently large—as can happen when the equilibrium state destabilizes in favor of a Hopf, Turing, or wave instability—then the existence of finite reversal potentials could have a significant impact on neural behavior. In fact, we will show that a subtle change in the way in which reversal potentials are incorporated into the model leads to qualitative change in its stability properties.

12.1.3 Gap-junction diffusion

The traditional picture of neural communication requires *active* propagation of action potentials from the axon of the transmitting neuron to the dendrite of the receiving neuron via release of neurotransmitters at the chemical-synaptic interface.

There is accumulating evidence, however, that subthreshold voltage fluctuations can be *passively* communicated from neuron to neuron via electrical synapses formed from gap-junction proteins that make direct resistive connections between neighboring cells at their points of dendritic contact. This is particularly so for inhibitory neurons in the cat visual cortex where the measured density of connexin-36 (Cx36) gap-junctions is so high that Fukuda *et al.* [10] described the result as establishing a dense and widespread network of interneurons able to be traced in a boundless chain. In addition, researchers have detected copious gap-junction couplings between interneurons and their supporting glial cells (via Cx32 connexin), and between pairs of glial cells (via Cx43) [1, 17], suggesting that diffusive neuronal coupling may be augmented by glial-cell “bridges”. To date, there are no reports of dense gap-junction connectivity between pairs of *excitatory* neurons, suggesting that, for reasons unknown, neural tissue has evolved to strongly favor inhibitory-to-inhibitory diffusion over excitatory-to-excitatory diffusion.

In Ref. [26], we used the Fukuda measurements to estimate an upper bound for D_2 , the inhibitory coupling strength, $D_2 \approx 0.6$ cm², then investigated the impact of incorporating inhibitory diffusion into a mean-field model of the cortex based on chemical synapses. We found that, provided that the D_2 inhibitory diffusion is sufficiently large, a homogeneous cortical sheet will spontaneously destabilize in favor of cm-scale stationary Turing patterns of intermixed regions of high- and low-firing activity.

In this chapter we extend this work by demonstrating that gap-junction diffusion $D_2 \nabla^2 V$ can interact with the $(V^{\text{rev}} - V)$ reversal-potential terms to generate two distinct types of spatiotemporal instability: either (a) stationary Turing structures when the feedback from soma to dendrite is delayed (“slow-soma” model); or (b) standing waves of gamma-band cortical activity when the soma-to-dendrite feedback is prompt (“fast-soma” model). We develop the background theory for the slow- and fast-soma models in Sect. 12.2; analyze their respective linearized stability characteristics in Sect. 12.3, then verify these predictions with a series of 2-D grid simulations of the full nonlinear equations. We follow this with a comparison against an earlier model due to Rennie and colleagues [20] in Sect. 12.4, then comment on the possible biological significance of the slow- and fast-soma forms.

12.2 Theory

We present the equations of motion for a continuum model of the cortex that consists of mutually interacting populations of excitatory and inhibitory neurons, each population receiving flux inputs from spiking events arriving at excitatory and inhibitory chemical synapses. The transmission efficiency of an excitatory (inhibitory) synapse is modulated by the voltage state of the post-synaptic dendrite relative to the AMPA (GABA) reversal potential. We consider two alternative schemes for incorporating the dendritic reversal potentials, leading to the slow-soma (Sect. 12.2.1.1) and fast-soma (Sect. 12.2.1.2) variants of the model. In both cases we assume that axonal flux propagation obeys damped 2-D wave equations (Sect. 12.2.1.3), with slower local (unmyelinated gray-matter) connections and faster long-range (myelinated white-matter) connections. The cortex is stimulated by nonspecific tonic activity generated by the subcortex (Sect. 12.2.1.4). Finally, in Sect. 12.2.2 we complete the model with the addition of diffusive voltage perturbations transmitted via electrical (gap-junction) synapses.

12.2.1 Input from chemical synapses

In deriving the equations of motion for V_e and V_i , the soma voltages for the excitatory and inhibitory neural populations, we assume that a pre-synaptic spike event will induce a *post-synaptic potential* (PSP), a momentary voltage change in the receiving dendrite, whose *shape* can be modeled either as a biexponential (first line of Eq. (12.1)) or as an alpha-function (second line),

$$H(t) = \begin{cases} \frac{\alpha\beta}{\beta-\alpha}(e^{-\alpha t} - e^{-\beta t}), & \alpha \neq \beta \\ \alpha^2 t e^{-\alpha t}, & \alpha = \beta \end{cases} \quad (12.1)$$

for $t > 0$, where α and β are positive constants.

If the incoming presynaptic spike rate is M [spikes/s], then the net voltage disturbance [in mV] at the dendrite will be given by ρU , where ρ is the *synaptic strength* [mV·s]; and U is the *post-synaptic response rate* [s⁻¹] given by the temporal convolution-integral of the input flux M with the dendrite filter response H , scaled by a dimensionless synaptic reversal-potential factor ψ ,

$$U(t) = \psi(t) [H(t) \otimes M(t)]. \quad (12.2)$$

We follow the earlier work of Wright *et al.* [36], Liley *et al.* [16], and Rennie *et al.* [20] in defining the ψ scaling factor to be unity when the neuron is at rest ($V = V^{\text{rest}}$), and zero when the membrane voltage matches the relevant synaptic reversal potential ($V_e^{\text{rev}} = 0$ mV for excitatory (AMPA) receptors; $V_i^{\text{rev}} = -70$ mV for inhibitory (GABA) receptors),

$$\psi_{ab}(t) = \frac{V_a^{\text{rev}} - V_b(t)}{V_a^{\text{rev}} - V_b^{\text{rest}}}, \quad a, b \in \{e, i\}. \quad (12.3)$$

Here we have introduced subscript labels a, b , each of which stands for either e (excitatory) or i (inhibitory), indicating that there are four reversal-potential functions: $\psi_{ee}, \psi_{ei}, \psi_{ie}, \psi_{ii}$, where, for example, ψ_{ei} is the scaling function for excitatory flux entering an inhibitory neuron. Corresponding double-subscripts are also to be attached to the H, M , and U appearing in Eq. (12.2).

The excitatory and inhibitory voltage disturbances at the dendrite are then integrated at the soma by convolving with the exponential soma impulse-response L ,

$$L_b(t) = \frac{1}{\tau_b} e^{-t/\tau_b}, \quad t > 0, \quad (12.4)$$

where τ_b is the soma time-constant for neurons of type b (e or i). This second integration results in a pair of integral equations of motion for V_e and V_i , the soma voltages for the excitatory and inhibitory neuron populations,

$$V_e(t) = V_e^{\text{rest}} + L_e(t) \otimes [\rho_e U_{ee}(t) + \rho_i U_{ie}(t)], \quad (12.5)$$

$$V_i(t) = V_i^{\text{rest}} + L_i(t) \otimes [\rho_e U_{ei}(t) + \rho_i U_{ii}(t)]. \quad (12.6)$$

The $\rho_{e,i}$ synaptic strengths are signed quantities, with $\rho_e > 0$ for excitatory post-synaptic potential (EPSP) events, and $\rho_i < 0$ for inhibitory postsynaptic potentials (IPSPs).

We wish to draw attention to the assumption, implicit in Eq. (12.2) regarding the construction of the post-synaptic rate U , by asking the question: Should the ψ -scaling by the reversal-potential weight be performed *after* the $H \otimes$ dendrite integration of input flux M —as written in Eq. (12.2)—

$$\begin{aligned} U_{ab}(t) &= \psi_{ab}(t) \cdot [H_{ab}(t) \otimes M_{ab}(t)] \\ &= \psi_{ab}(t) \int_0^t H_{ab}(t-t') M_{ab}(t') dt', \quad (\text{“slow soma”}), \end{aligned} \quad (12.7)$$

or, should the ψ -scaling be applied *directly* to the input flux, so that it is the weighted product $\psi \cdot M$ that is integrated at the dendrite? This alternative ordering leads to a revised post-synaptic rate U ,

$$\begin{aligned} U_{ab}(t) &= H_{ab}(t) \otimes [\psi_{ab}(t) \cdot M_{ab}(t)] \\ &= \int_0^t H_{ab}(t-t') \psi_{ab}(t') M_{ab}(t') dt', \quad (\text{“fast soma”}). \end{aligned} \quad (12.8)$$

We refer to the Eq. (12.7) form for the U post-synaptic rate as the “*slow soma*” case, since the somata voltage of the neuron is presumed to vary on a time-scale that is much slower than that of the synaptic input events. This limit should be valid when the synaptic inputs are sparse or weak, so that the ψ reversal-potential feedback from soma to dendrite is slow to arrive. On the other hand, if synaptic activity is strong, or if the soma voltage changes on a time-scale similar to that of dendritic integration, then the reversal-potential feedback onto the dendrite will be prompt. In this case, the M flux input at time t should be scaled by the ψ reversal-potential weight at time t , then integrated at the dendrite; this limit gives the Eq. (12.8) “*fast-soma*” form for the U post-synaptic rate. These slow- and fast-soma variants are block-diagrammed in the flow-charts of Fig. 12.1.

We find that swapping the order of the $\psi \cdot$ and $H \otimes$ operations has surprising implications for cortical stability that may have biological significance. If the ψ -weighting occurs *after* the $H \otimes$ dendrite integration (i.e., Eq. (12.7): *slow-soma*), then, as reported in [26], the homogeneous 2-D cortex can destabilize in the presence of inhibitory diffusion to form Turing patterns—stationary spatial patterns of activated and inactivated patches of cortical tissue.¹ But if the $\psi \cdot$ weighting is applied *prior* to the $H \otimes$ convolution (i.e., Eq. (12.8): *fast-soma*), we will see that the stationary Turing patterns are replaced by standing-wave patterns of similar spatial frequency but whose temporal frequency lies within the gamma band (~ 30 – 80 Hz) of EEG oscillations.

12.2.1.1 Slow-soma limit

In the limit of a slowly varying membrane potential, the soma-voltage equations (12.5) and (12.6) become,

$$\begin{aligned} V_b(t) &= V_b^{\text{rest}} + L_b(t) \otimes [\rho_e U_{eb}(t) + \rho_i U_{ib}(t)] \\ &= V_b^{\text{rest}} + L_b(t) \otimes [\rho_e \psi_{eb}(t) \cdot \Phi_{eb}(t) + \rho_i \psi_{ib}(t) \cdot \Phi_{ib}(t)], \end{aligned} \quad (12.9)$$

where the Φ_{eb} , Φ_{ib} ($b = e, i$) represent the four slow-soma flux convolutions of flux input M against dendrite filter H ,

¹ Our prior modeling of anesthetic induction [24, 25, 32] and state transitions in natural sleep [23, 33, 34] assumed a slow-soma limit; gap-junction effects were not included.

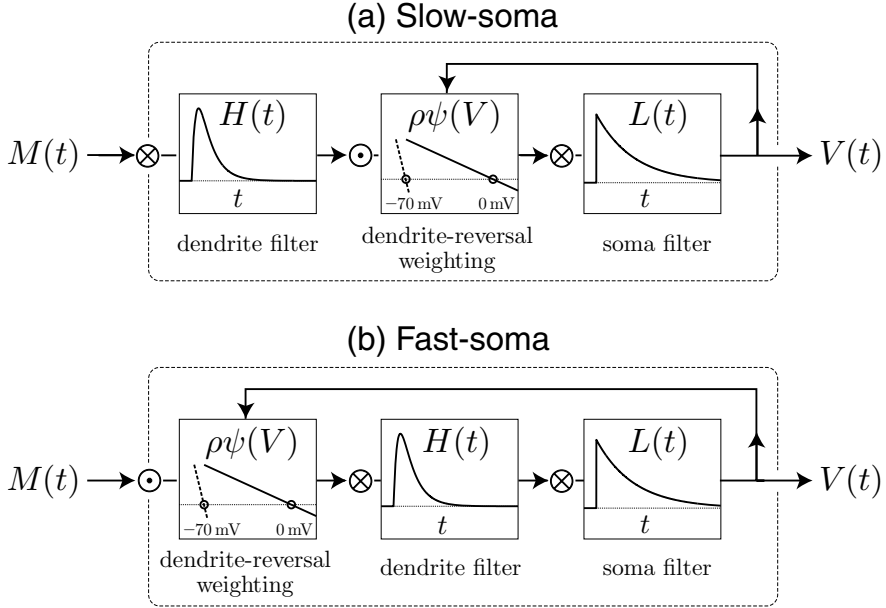


Fig. 12.1 Dendrite-to-soma flow diagrams for (a) slow-soma and (b) fast-soma cortical models. M is the average spike-rate input arriving at the dendrite via chemical synapses; V is the resulting voltage perturbation at the soma (for simplicity, we ignore the constant V^{rest} offset here); ρ is the synaptic strength; H and L are respectively the dendrite and soma impulse-response functions. The ψ reversal-potential weighting function provides immediate feedback from soma to dendrite. Symbols \otimes and \odot represent convolution and product operations respectively. (a) For slow-soma, the input flux is modulated by the reversal-weighting *after* integration at the dendrite filter, while for fast-soma, the reversal modulation occurs *prior* to dendrite integration.

$$\Phi_{ab}(t) = H_{ab}(t) \otimes M_{ab}(t), \quad (12.10)$$

and where

$$M_{ab}(t) = N_{ab}^{\alpha} \phi_{ab}^{\alpha}(t) + N_{ab}^{\beta} \phi_{ab}^{\beta}(t) + N_{eb}^{\text{sc}} \phi_{eb}^{\text{sc}}, \quad (12.11)$$

with subcortical inputs

$$\phi_{eb}^{\text{sc}} = s \cdot Q_e^{\text{max}}, \quad \phi_{ib}^{\text{sc}} = 0. \quad (12.12)$$

Equation (12.11) defines M , the total input flux of type a (e or i) entering neurons of type b (e or i). The superscript labels α , β , sc indicate long-range, short-range, and subcortical chemical-synaptic inputs respectively. The $N^{\alpha,\beta,\text{sc}}$ are the number of synaptic connections; the $\phi^{\alpha,\beta,\text{sc}}$ are per-synapse flux rates. The $\phi^{\alpha,\beta}$ fluxes obey wave equations detailed below in Eqs (12.18, 12.20). The long-range and subcortical inputs are excitatory only, so $N_{ie}^{\alpha} = N_{ii}^{\alpha} = N_{ie}^{\text{sc}} = N_{ii}^{\text{sc}} = 0$, and $\phi_{ie}^{\text{sc}} = \phi_{ii}^{\text{sc}} = 0$. Here s is a subcortical scaling parameter whose value can range between 0 and 1. Since $Q_e^{\text{max}} = 100 \text{ s}^{-1}$ (see Table 12.4 in the Appendix), choosing $s = 0.1$ will ensure

compatibility with the earlier modeling work by Rennie *et al.* [20] in which the default level for per-synapse subcortical drive was set at $\phi^{\text{sc}} = 10 \text{ s}^{-1}$.

The slow-soma integral equations of Eq. (12.9) are equivalent to a pair of first-order differential equations of motion for soma voltage,

$$\tau_b \frac{dV_b(t)}{dt} = V_b^{\text{rest}} - V_b(t) + \rho_e \psi_{eb}(t) \Phi_{eb}(t) + \rho_i \psi_{ib}(t) \Phi_{ib}(t). \quad (12.13)$$

Taking the biexponential form of the Eq. (12.1) postsynaptic potential, the four dendrite convolutions of Eq. (12.10) can be rewritten as four second-order ODEs in $\Phi(t)$,

$$\left(\frac{d}{dt} + \alpha_{ab} \right) \left(\frac{d}{dt} + \beta_{ab} \right) \Phi_{ab}(t) = \alpha_{ab} \beta_{ab} M_{ab}(t). \quad (12.14)$$

12.2.1.2 Fast-soma limit

For the fast-soma version of the cortical model, we use the revised form of the postsynaptic rate given in Eq. (12.8),

$$U_{ab}(t) = H_{ab}(t) \otimes [\psi_{ab}(t) \cdot M_{ab}(t)], \quad (12.15)$$

leading to two fast-soma differential equations for the V_b ($b = e, i$) neuron voltage (cf. Eq. (12.13)),

$$\tau_b \frac{dV_b(t)}{dt} = V_b^{\text{rest}} - V_b(t) + \rho_e U_{eb}(t) + \rho_i U_{ib}(t), \quad (12.16)$$

with dendrite ODEs (cf. Eq. (12.14)),

$$\left(\frac{d}{dt} + \alpha_{ab} \right) \left(\frac{d}{dt} + \beta_{ab} \right) U_{ab}(t) = \alpha_{ab} \beta_{ab} \psi_{ab}(t) \cdot M_{ab}(t). \quad (12.17)$$

Comparing Eqs (12.16, 12.17) with (12.13, 12.14), we see that, in the fast-soma limit, the ψ_{ab} reversal-potential weights are applied *directly* to the incoming M_{ab} synaptic flux, with the product being integrated at the dendrite to give the (weighted) dendritic flux U_{ab} ; whereas in the slow-soma model, the ψ_{ab} weights are applied *after* the M_{ab} input flux has been integrated at the dendrite.

12.2.1.3 Wave equations

The axonal wave equations described here apply equally to both the slow- and fast-soma cortical models. Following Robinson *et al.* [21], we assume that the ϕ^α long-

range excitatory fluxes obey a pair of 2-D damped wave equations generated by excitatory sources $Q_e(\mathbf{r}, t)$,

$$\left[\left(\frac{\partial}{\partial t} + v^\alpha \Lambda_{eb}^\alpha \right)^2 - (v^\alpha)^2 \nabla^2 \right] \phi_{eb}^\alpha(\mathbf{r}, t) = (v^\alpha \Lambda_{eb}^\alpha)^2 Q_e(\mathbf{r}, t), \quad b \in \{e, i\} \quad (12.18)$$

where Λ^α is the inverse-length scale for axonal connections [cm^{-1}], and v^α is the axonal conduction speed [cm/s]. Q is the sigmoidal mapping from soma voltage to neuronal firing rate,

$$Q_a(\mathbf{r}, t) = \frac{Q_a^{\max}}{1 + \exp[-C(V_a(\mathbf{r}, t) - \theta_a)/\sigma_a]}, \quad a \in \{e, i\} \quad (12.19)$$

with $C = \pi/\sqrt{3}$. Here, θ_a is the population-average threshold for firing, σ_a is its standard deviation, and Q_a^{\max} is the maximum firing rate.

In previous work [23–25, 32–34], we have assumed that the short-range axonal signals propagate instantaneously, allowing local spike-rate fluxes ϕ_{ab}^β to be replaced by their sources Q_a . In this chapter we allow for finite propagation speeds by writing four wave equations for the *short-range fluxes* ϕ_{ab}^β traveling on unmyelinated axons,

$$\left[\left(\frac{\partial}{\partial t} + v^\beta \Lambda_{ab}^\beta \right)^2 - (v^\beta)^2 \nabla^2 \right] \phi_{ab}^\beta(\mathbf{r}, t) = (v^\beta \Lambda_{ab}^\beta)^2 Q_a(\mathbf{r}, t). \quad (12.20)$$

12.2.1.4 Subcortical inputs

Equation (12.12) is applicable when the subcortical drive is a fixed constant. To allow noise to enter the cortex, we replace (12.12) with the stochastic form

$$\phi_{eb}^{\text{sc}}(\mathbf{r}, t) = s Q_e^{\max} + \gamma \sqrt{s Q_e^{\max}} \xi_m(\mathbf{r}, t), \quad m = 1, 2 \quad (12.21)$$

where γ is a constant noise scale-factor, and the ξ_m are a pair of Gaussian-distributed, zero-mean, spatiotemporal white-noise sources that are delta-correlated in time and space,

$$\langle \xi_m(\mathbf{r}, t) \rangle = 0, \quad (12.22)$$

$$\langle \xi_m(\mathbf{r}, t) \xi_n(\mathbf{r}', t') \rangle = \delta_{mn} \delta(t - t') \delta(\mathbf{r} - \mathbf{r}'). \quad (12.23)$$

In the grid simulations presented in Sect. 12.3, we specify s , the subcortical drive (e.g., $s = 0.1$), and initialize the 2-D sheet of cortical tissue at the homogeneous steady-state corresponding to this level of subcortical stimulation. Then, using Eq. (12.21), we distribute spatially-independent small-amplitude random perturbations across the cortical grid to allow the model to explore its proximal state space. If

the homogeneous equilibrium is unstable, these small-scale deviations from homogeneity can organize and grow into large-scale Turing structures, Hopf oscillations, and gamma-band standing-wave patterns.

12.2.2 Input from electrical synapses

The existence of gap junctions in the mammalian brain has been known for decades, but only recently has their electrophysiological significance with respect to neuron coupling and rhythm synchronization become apparent. According to the review article by Bennett and Zukin [2], electrical transmission (via gap junctions) between neurons is “likely to be found wherever it is useful,” with its selective advantage being the communication of subthreshold potentials that facilitate synchronization. The reported abundances of gap-junction connections in brain tissue are increasing as detection methods become more sensitive and discriminating. These connections are found between pairs of inhibitory interneurons (via connexin Cx36 channels [10]), between interneurons and their supporting glial cells (via Cx32), and between the glial cells themselves (via Cx42). The neuron-to-glia and glia-to-glia connections have been detected in all layers of the rat cerebral cortex [17]. These findings support the notion of a diffusively-intercoupled continuous scaffolding that links networks of active (neuronal) and passive (glial) cells.

Fukuda *et al.* [10] reported that, on average, each L-type inhibitory interneuron in the cat visual cortex was coupled to $N_{ii}^{\text{gap}} = 60 \pm 12$ other L-type interneurons via Cx36 connexin channels, that the connections were randomly and uniformly distributed over a disk of radius $\sim 200 \mu\text{m}$ centered on a given neuron, and that L-type abundance was $\sim 400 \text{ mm}^{-2}$, implying a connection density of $\sim 24\,000$ gap-junctions per mm^2 . The Fukuda measurements were specific to Cx36 connexins only, so total interneuron gap-junction interconnectivity could be considerably higher. In [26], we used the Fukuda measurements to construct a theoretical 2-D lattice of square “Fukuda cells”, with each lattice cell representing the effective area u of diffusive influence for a single L-type interneuron; see Fig. 12.2.

We assume that the neuron at the lattice center has resting voltage V^{rest} , capacitance C , membrane resistance R_m , and receives diffusive current along four resistive arms, each of resistance $R = R^{\text{gap}} / \frac{1}{4} N_{ii}^{\text{gap}} = R^{\text{gap}} / 15$, where R^{gap} is the resistance of a single Cx36 gap junction. The total current to “ground” (i.e., to the extracellular space) is the sum of membrane current $(V - V^{\text{rest}}) / R_m$ plus capacitive current $C dV/dt$, and this must match the addition of chemical synaptic currents I^{syn} (not shown) plus gap-junction diffusive currents I^{gap} ,

$$(V - V^{\text{rest}}) / R_m + C \frac{\partial V}{\partial t} = I^{\text{syn}} + I^{\text{gap}}, \quad (12.24)$$

where

$$I^{\text{gap}} = \frac{u}{R} \nabla^2 V. \quad (12.25)$$

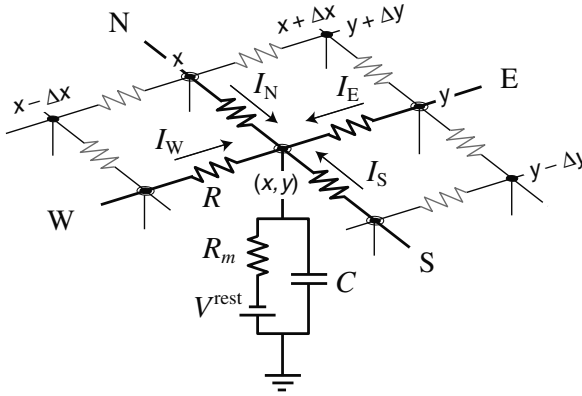


Fig. 12.2 Equivalent electrical circuit for nearest-neighbor gap-junction connections between neurons in a 2-D cortex. Diffusion currents I_N, I_S, I_E, I_W enter the neuron at the central node from the four neighboring nodes via gap-junction resistances R shown in bold. For clarity, chemical-synaptic currents are not shown. (Figure reproduced from [26].)

The effective area u of the Fukuda cell depends on the underlying connectivity assumption. If the connectivity is uniform across the $200\text{-}\mu\text{m}$ radius disk, we circumscribe the circle with a square of side 0.4 mm , giving an upper bound of $u \approx 0.16\text{ mm}^2$. More realistically, we can fit a Gaussian distribution to the Fukuda data (see Appendix A of [26] for details), leading to $u \approx 0.03\text{ mm}^2$, a factor of five smaller. We note that these estimates for u will *increase* if subsequent determinations of gap-junction dendritic extent and distal abundance are found to be larger than those reported by Fukuda *et al.*

Rearranging Eq. (12.24), we obtain the differential equation for soma voltage,

$$\tau \frac{\partial V}{\partial t} = (V^{\text{rest}} - V) + I^{\text{syn}} R_m + D_{ii} \nabla^2 V, \quad (12.26)$$

where $\tau = R_m C$ is the membrane time-constant, $I^{\text{syn}} R_m$ is the voltage contribution at the soma arising from chemical-synaptic currents, and D_{ii} is the diffusive coupling strength for inhibitory-to-inhibitory gap-junction currents,

$$D_{ii} = u \frac{R_m}{R} = \frac{u N_{ii}^{\text{gap}}}{4} \frac{R_m}{R^{\text{gap}}}. \quad (12.27)$$

In [26] we estimated $R_m \approx 7100\text{ M}\Omega$, $R^{\text{gap}} \approx 290\text{ M}\Omega$ (corresponding to a gap-junction in its fully-open configuration). Setting $u = 0.16\text{ mm}^2$ gives $D_{ii} \approx 0.6\text{ cm}^2$, but we emphasize that all four components ($u, N_{ii}, R_m, R^{\text{gap}}$) in the D_{ii} expression are uncertain, and are likely to vary with time, neuromodulatory state, and stage of development: it is not implausible that the “true” value for D_{ii} might lie within an uncertainty band that extends from $+1$ to -2 orders-of-magnitude above and below the nominal value quoted here.

12.2.2.1 Slow-soma limit with gap junctions

We incorporate the effect of gap-junction diffusion currents entering a slow-soma neuron by combining Eq. (12.26) with the slow-soma membrane Eqs (12.14) to give

$$\tau_b \frac{\partial V_b(\mathbf{r}, t)}{\partial t} = V_b^{\text{rest}} - V_b(\mathbf{r}, t) + [\rho_e \Psi_{eb}(\mathbf{r}, t) \Phi_{eb}(\mathbf{r}, t) + \rho_i \Psi_{ib}(\mathbf{r}, t) \Phi_{ib}(\mathbf{r}, t)] + D_{bb} \nabla^2 V_b(\mathbf{r}, t), \quad (12.28)$$

where the terms in square brackets [...] are the contributions from chemical-synaptic flux entering the *slow-soma* model. Here, $D_{bb} = D_{ee}$ for excitatory-to-excitatory diffusion, and $D_{bb} = D_{ii}$ for inhibitory-to-inhibitory diffusion.² We note that direct electrical connections between pairs of same-family inhibitory interneurons are common, but apparently gap junctions between pairs of excitatory neurons are rare, so in the work below we set D_{ee} to be a small (but non-zero) fraction of D_{ii} , with $D_{ee} = D_{ii}/100$.

12.2.2.2 Fast-soma limit with gap junctions

Combining Eq. (12.26) with the fast-soma membrane Eqs (12.17) give the pair ($b = e, i$) of partial differential equations for the diffusion-enhanced *fast-soma* model,

$$\tau_b \frac{\partial V_b(\mathbf{r}, t)}{\partial t} = V_b^{\text{rest}} - V_b(\mathbf{r}, t) + [\rho_e U_{eb}(\mathbf{r}, t) + \rho_i U_{ib}(\mathbf{r}, t)] + D_{bb} \nabla^2 V_b(\mathbf{r}, t), \quad (12.29)$$

12.3 Results

12.3.1 Stability predictions

The stability characteristics of the slow- and fast-soma cortical models are referenced to a homogeneous steady-state corresponding to a given value of subcortical drive s . This reference state is determined by zeroing the ξ_m noise terms in Eq. (12.21), and removing all time- and space-dependence by setting $d/dt = \nabla^2 = 0$ in *either* the slow-soma differential equations (12.28, 12.14, 12.18, 12.20), *or* the fast-soma equations (12.29, 12.17, 12.18, 12.20). Either procedure gives a set of nonlinear simultaneous equations that we solve numerically to locate the steady-state soma voltage (V_e^0, V_i^0) and firing rate (Q_e^0, Q_i^0). Note that, at steady state, the distinction between the slow-soma $U = \psi \cdot [H \otimes M]$ (Eq. (12.7)) and fast-soma $U = H \otimes [\psi \cdot M]$ (Eq. (12.8)) convolution forms vanishes, and therefore the *equilibrium* states for the slow-soma and fast-soma cortical models are *identical*. Yet despite their shared equilibria, we will show that the *dynamical* properties of the

² Later we simplify the subscripting notation for diffusion so that $(D_{ee}, D_{ii}) \equiv (D_1, D_2)$.

Table 12.1 State variables for slow- and fast-soma models

Variable	Symbol	Unit	Equations
Soma voltage	V_e, V_i	mV	(12.13, 12.28) ^a ; (12.16, 12.29) ^b
Dendritic flux response	$U_{ee}, U_{ei}, U_{ie}, U_{ii}$	s ⁻¹	(12.14) ^a ; (12.17) ^b
Long-range flux input	$\phi_{ee}^\alpha, \phi_{ei}^\alpha$	s ⁻¹	(12.18)
Short-range flux input	$\phi_{ee}^\beta, \phi_{ei}^\beta, \phi_{ie}^\beta, \phi_{ii}^\beta$	s ⁻¹	(12.20)

a Slow soma: $U_{ab} = \psi_{ab} \cdot \Phi_{ab} = \psi_{ab} \cdot [H_{ab} \otimes M_{ab}]$

b Fast soma: $U_{ab} = H_{ab} \otimes [\psi_{ab} \cdot M_{ab}]$

two models—as predicted by linear eigenvalue analysis, and confirmed by nonlinear grid simulations—are very different.

The 12 state variables for the slow-soma and fast-soma models are listed in Table 12.1; related system variables appear in Table 12.2. The 12 state variables are governed by two first-order (V_e, V_i) and 10 second-order ($U_{ab}, \phi_{eb}^\alpha, \phi_{ab}^\beta$) differential equations, so are equivalent to 22 coupled first-order DEs, and therefore, after linearization about homogeneous steady state, own 22 eigenvalues.

The linearization proceeds by expressing each of the 22 first-order variables (12 state variables plus 10 auxiliaries) as its homogeneous equilibrium value plus a fluctuating component. For example, the excitatory soma voltage is written

$$V_e(\mathbf{r}, t) = V_e^0 + \delta V_e(\mathbf{r}, t), \quad (12.30)$$

where \mathbf{r} is the 2-D position vector, and δV_e , the fluctuation about equilibrium V_e^0 , has spatial Fourier transform

$$\widetilde{\delta V_e}(\mathbf{q}, t) = \int_{-\infty}^{\infty} \delta V_e(\mathbf{r}, t) e^{-i\mathbf{q} \cdot \mathbf{r}} d\mathbf{r} \quad (12.31)$$

with \mathbf{q} being the 2-D wave vector. This is equivalent to assuming that the voltage perturbation can be expressed as a spatiotemporal mode of the form,

$$\delta V_e(\mathbf{r}, t) = \delta V_e(\mathbf{r}, 0) e^{\Lambda t} e^{i\mathbf{q} \cdot \mathbf{r}} \quad (12.32)$$

where Λ is its (complex) eigenvalue. If Λ has a positive real part, the perturbation will grow, indicating that the equilibrium state is unstable.

After linearizing the 22 first-order DEs about homogeneous equilibrium, then Fourier transforming in space, we compute numerically the 22 eigenvalues of the Jacobian matrix for a range of finely-spaced wavenumbers, $q = |\mathbf{q}|$. Arguing that the stability behavior of the cortical model will be dominated by the eigenvalue Λ whose real part is least negative (or most positive), we plot the distribution of dominant eigenvalues as a function of wavenumber, looking for regions for which $\text{Re}[\Lambda(q)] > 0$, indicating the presence of spatial modes that can destabilize the homogeneous rest state.

Table 12.2 Other system variables for cortical model

Variable	Symbol	Unit	Equation
Total flux input	M_{ab}	s^{-1}	(12.11)
Firing rate	Q_a	s^{-1}	(12.19)
Dendrite filter impulse	H_{ab}	s^{-1}	(12.1)
Soma filter impulse	L_b	s^{-1}	(12.4)
Reversal-potential weight	ψ_{ab}	—	(12.3)
Subcortical flux	ϕ_{eb}^{sc}	s^{-1}	(12.21)

12.3.2 Slow-soma stability

Figure 12.3 shows that, in the slow-soma limit, the homogeneous steady state can be destabilized either by *increasing* the inhibitory diffusion D_2 (Fig. 12.3(a)), or by *decreasing* the level of subcortical drive s (Fig. 12.3(b)). Instability at a given wavenumber q is predicted when its dominant eigenvalue crosses the zero-axis, changing sign from negative (decaying mode) to positive (exponentially-growing mode). For the case $D_2 = 4 \text{ cm}^2$, $s = 0.1$ (top curve of Fig. 12.3(a)), all wavenumbers in the range $0.24 \lesssim q/2\pi \lesssim 0.7 \text{ cm}^{-1}$ support growing modes, with strongest growth predicted at $q/2\pi \approx 0.4 \text{ cm}^{-1}$ (i.e., wavelength $\approx 2.5 \text{ cm}$), at the peak of the dispersion curve. At this wavenumber, the eigenvalue has a zero imaginary part, so the final pattern is expected to be a *stationary* periodic pattern in space—a Turing structure of intermixed regions of high- and low-firing cortical activity. This spontaneous Turing emergence is similar to that reported in [26] for an earlier version of the slow-soma model that had access to three homogeneous steady states (two stable, one unstable).

12.3.3 Fast-soma stability

Figure 12.4(a) shows the set of dispersion curves obtained for the *fast-soma* version of the model operating at the same $s = 0.1$ level of subcortical excitation used in Fig. 12.3(a). In marked contrast to the slow-soma case, maximum instability is obtained in the limit of *zero* inhibitory diffusion ($D_2 = 0$), with the instability being promptly damped out as the diffusion increases. We see that small increases in diffusion serve to narrow the range of spatial frequencies able to destabilize the equilibrium state. Thus, when $D_2 = 0$, the instability is distributed across the broad range $0.35 < q/2\pi < 3.48 \text{ cm}^{-1}$ (this upper value is not shown on the Fig. 12.4 graph), shrinking to $0.40\text{--}0.67 \text{ cm}^{-1}$ when $D_2 = 0.04 \text{ cm}^2$, and vanishing completely for $D_2 \geq 0.06 \text{ cm}^2$.

The fact that the dominant eigenvalue has a nonzero imaginary part indicates that these fast-soma spatial instabilities will tend to oscillate in time: for spatial frequency $q/2\pi = 0.5 \text{ cm}^{-1}$, the predicted temporal frequency is $\sim 29 \text{ Hz}$, at the lower end of the gamma band. Writing $\omega = \text{Im}[\Lambda]$, the slope of the ω -vs- q graph is nearly flat (thin curves of Fig. 12.4(a)), implying that these wave instabilities

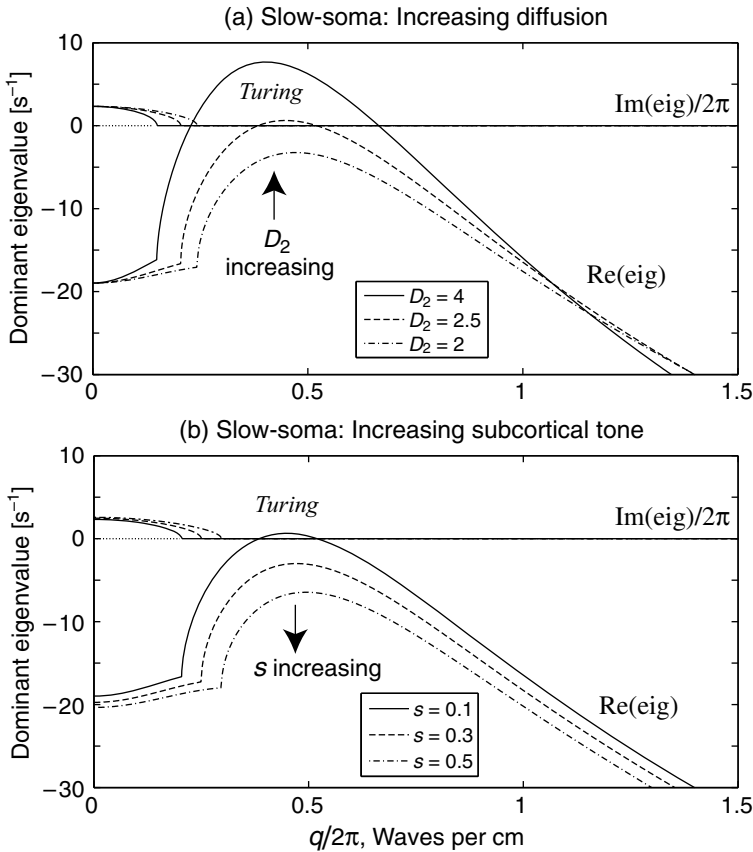


Fig. 12.3 *Slow-soma* dispersion curves for (a) increasing inhibitory diffusion D_2 and (b) increasing subcortical drive s . (a) Imaginary (upper thin traces) and real (lower thick traces) parts of dominant eigenvalue plotted as a function of scaled wavenumber $q/2\pi$ for three values of diffusion strength, $D_2 = [2.0, 2.5, 4.0] \text{ cm}^2$; excitatory diffusion strength is set at 1% of inhibitory strength: $D_1 = D_2/100$. Subcortical drive is fixed at $s = 0.1$ (i.e., $\phi^{\text{sc}} = 10 \text{ s}^{-1}$), corresponding to homogeneous steady state $(V_e^0, V_i^0) = (-59.41, -59.41) \text{ mV}$; $(Q_e^0, Q_i^0) = (6.37, 12.74) \text{ s}^{-1}$. (See Table 12.4 for parameter values.) The homogeneous state is predicted to be unstable at all spatial frequencies for which the real part of the dispersion curve is positive. Stationary Turing patterns are predicted at $q/2\pi \approx 0.45 \text{ cm}^{-1}$ for $D_2 \gtrsim 2.5 \text{ cm}^2$, and are *enhanced* by increases in D_2 coupling strength. (b) Eigenvalue distribution for three values of subcortical drive, $s = [0.1, 0.3, 0.5]$; excitatory and inhibitory diffusion strength are fixed at $(D_1, D_2) = (0.025, 2.5) \text{ cm}^2$. The *slow-soma* Turing instability at $q/2\pi = 0.45 \text{ cm}^{-1}$ is *damped out* by increases in subcortical tone, restoring stability to the homogenous steady-state. (Figure reproduced from Ref. [27].)

will propagate but slowly. For example, with $D_2 = 0.04 \text{ cm}^2$, the group velocity at $q/2\pi = 0.5 \text{ cm}^{-1}$ is $d\omega/dq = 3.8 \text{ cm/s}$, so that, over the timescale a single gamma oscillation ($\sim 0.034 \text{ s}$), the wave will travel only 1.3 mm. This suggests that the

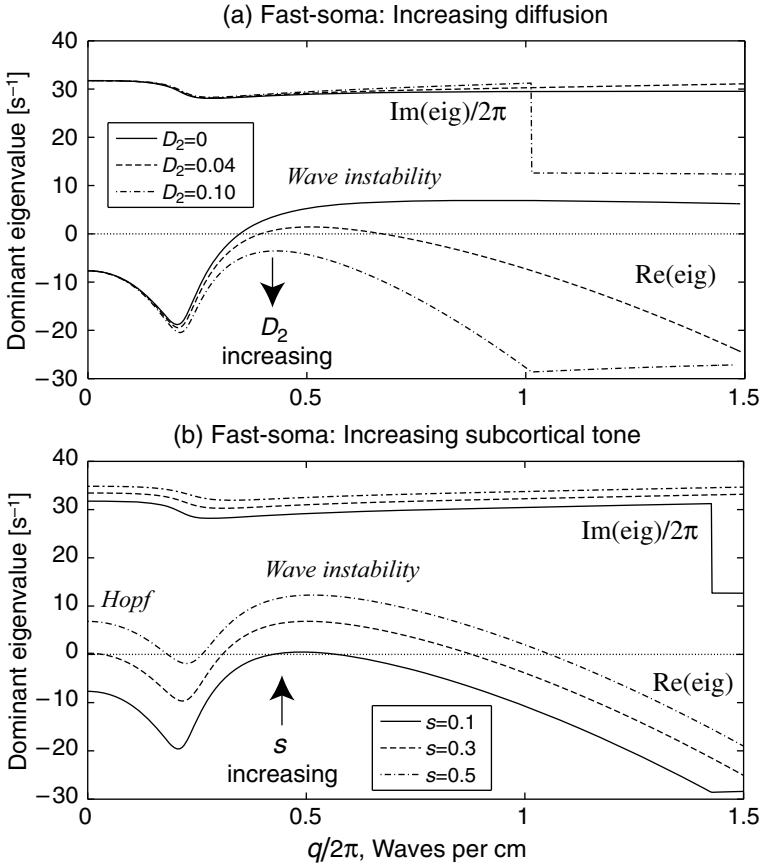


Fig. 12.4 *Fast-soma* dispersion curves for (a) increasing inhibitory diffusion D_2 and (b) increasing subcortical tone s . (a) The three pairs of eigenvalue curves correspond to three values for inhibitory diffusion, $D_2 = [0.0, 0.04, 0.10] \text{ cm}^2$, with subcortical drive kept fixed at $s = 0.1$. Wave instabilities of temporal frequency $\sim 29\text{-Hz}$ and spatial frequency 0.5 waves/cm are expected when $D_2 \lesssim 0.04 \text{ cm}^2$. (b) Eigenvalue distribution for three values of subcortical drive, $s = [0.1, 0.3, 0.5]$, corresponding to subcortical flux rates of $\phi^{\text{sc}} = [10, 30, 50] \text{ s}^{-1}$, giving homogeneous steady-state firing rates $Q_0^0 = [6.37, 7.28, 8.10] \text{ s}^{-1}$; excitatory and inhibitory diffusion is fixed at $(D_1, D_2) = (0.0005, 0.05) \text{ cm}^2$. For $s = 0.1$ (lower-thick and upper-thin solid curves), homogeneous steady-state destabilizes in favor of 29-Hz traveling waves of spatial frequency 0.49 cm^{-1} . Increasing subcortical tone to 0.3 and 0.5 strengthens the wave instability, and raises its frequency slightly to 31 and 32.5 Hz respectively. For $s = 0.5$, the peak at $q/2\pi = 0$ indicates that the wave pattern will be modulated by a whole-cortex Hopf instability of frequency 35 Hz. (Figure reproduced from Ref. [27].)

instability will manifest as a slowly-drifting standing-wave pattern of 29-Hz gamma oscillations, with wavelength ~ 2 cm.

Another surprising difference in the behavior of the slow- and fast-soma models is their contrasting stability response to alterations in the level of subcortical drive s . If the subcortical tone is stepped, say, from $s = 0.1$ to 0.3 to 0.5 , the homogeneous excitatory firing rate for both models increases slightly, from $Q_e^0 = 6.4$ to 7.3 to 8.1 spikes/s (not shown here). In Fig. 12.4(b) we observe that this boost in subcortical tone tends to *destabilize* the fast-soma cortex, increasing both the strength and the frequency of the gamma wave instability—but for the slow-soma cortex (Fig. 12.3(b)) the effect is precisely opposite, acting to *damp out* Turing instabilities, encouraging restoration of the homogeneous equilibrium state. We will argue in Sect. 12.4 that these interesting divergences in the slow- and fast-cortical responses are consistent with the notion that the slow-soma could describe the idling or default background state of the conscious brain, while the fast-soma could describe the genesis of gamma resonances that characterize the active, cognitive state.

12.3.4 Grid simulations

To test the linear-stability predictions of Turing and traveling-wave activity in the cortical model, we ran a series of numerical simulations of the full nonlinear slow-soma (12.28, 12.14, 12.18, 12.20) and fast-soma (12.29, 12.17, 12.18, 12.20) cortical equations. The substrate was a 240×240 square grid, of side-length 6 cm, joined at the edges to provide toroidal boundaries. We used a forward-time, centered-space Euler algorithm custom-written in MATLAB 7.6, with the diffusion and wave-equation ∇^2 Laplacians implemented as wrap-around (toroidal) convolutions³ of the 3×3 second-difference mask against the grid variables holding $V_{e,i}(\mathbf{r}, t)$, the excitatory and inhibitory membrane voltages. The grid was initialized at the homogeneous steady state corresponding to a specified value of subcortical drive s , then driven continuously by two independent sources of small-amplitude unfiltered spatiotemporal white noise representing unstructured subcortical tone $\phi_{ee,ei}^{sc}$ (see Eq. (12.21)). The timestep was set sufficiently small to ensure numerical stability, ranging from $\Delta t = 100 \mu\text{s}$ for the fast-soma (weak diffusion) runs, down to $1 \mu\text{s}$ for the slow-soma runs with strongest inhibitory diffusion (i.e., $D_2 = 6 \text{ cm}^2$).

The upper bound for the slow-soma timestep was obtained by recognizing that D_2/τ_i , the ratio of diffusive strength to membrane relaxation time, defines a diffusion coefficient [units: cm/s] for inhibitory voltage change, so in time Δt , a voltage

³ The 2-D circular convolution algorithm was written by David Young, Department of Informatics, University of Sussex, UK. His `convolve2()` MATLAB function can be downloaded from The MathWorks File Exchange, www.mathworks.com/matlabcentral/fileexchange.

perturbation is expected to diffuse through an rms distance $d^{\text{rms}} = \sqrt{4D_2\Delta t/\tau_i}$. Setting $d^{\text{rms}} = \Delta x = \Delta y$, the lattice spacing, and solving for Δt , gives

$$\Delta t = \frac{1}{4}(\Delta x)^2\tau_i/D_2, \quad (12.33)$$

which we replaced with the more conservative

$$(\Delta t)^{\text{max}} = \frac{1}{5}(\Delta x)^2\tau_i/D_2, \quad (12.34)$$

to ensure that, on average, the diffusive front would have propagated by less than one lattice spacing between consecutive timesteps.

For the fast-soma case, diffusion values are very weak, so it is the long-distance wave equations (12.18) that set the upper bound for the timestep. The Courant stability condition for the 2-D explicit-difference method requires $v^\alpha\Delta t/\Delta x \leq 1/\sqrt{2}$. Setting $\Delta x = L_x/240 = 0.025$ cm, and $v^\alpha = 140$ cm/s (see Table 12.4) gives $\Delta t \leq 126$ μs , so the fast-soma choice of $\Delta t = 100$ μs is safely conservative.

12.3.5 Slow-soma simulations

Figure 12.5 shows a sequence of snapshots of the firing-rate patterns that evolve spontaneously in the *slow-soma* model when the inhibitory diffusion is sufficiently strong, here set at $D_2 = 4$ cm^2 . Starting from the homogeneous steady-state

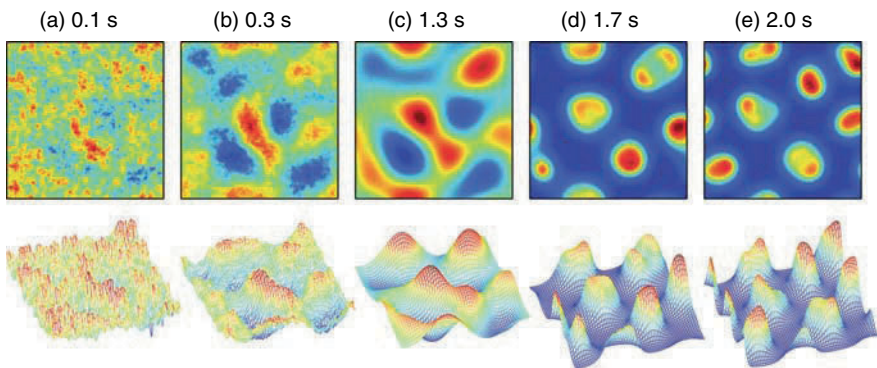


Fig. 12.5 [Color plate] Grid simulation for *slow-soma* cortical model with subcortical drive $s = 0.1$ (i.e., $\phi^{\text{sc}} = 10$ s^{-1}), inhibitory diffusion $D_2 = 4.0$ cm^2 . Cortex is a 240×240 square grid of side-length 6 cm, with toroidal boundaries, initialized at its homogeneous steady-state firing-rate $(Q_e^0, Q_i^0) = (6.37, 12.74)$ s^{-1} , driven continuously with small-amplitude spatiotemporal white noise. Snapshots show the spatial and temporal evolution of Q_e as bird's-eye (top row) and mesh (bottom row) perspectives. Consistent with Fig. 12.3(a), cortical sheet spontaneously organizes into stationary Turing patterns of wavelength ~ 2.5 cm. Turing structures grow strongly with time; see Fig. 12.6. Grid resolution $\Delta x = \Delta y = 0.25$ mm; timestep $\Delta t = 1.5$ μs . (Reproduced from Ref. [27].)

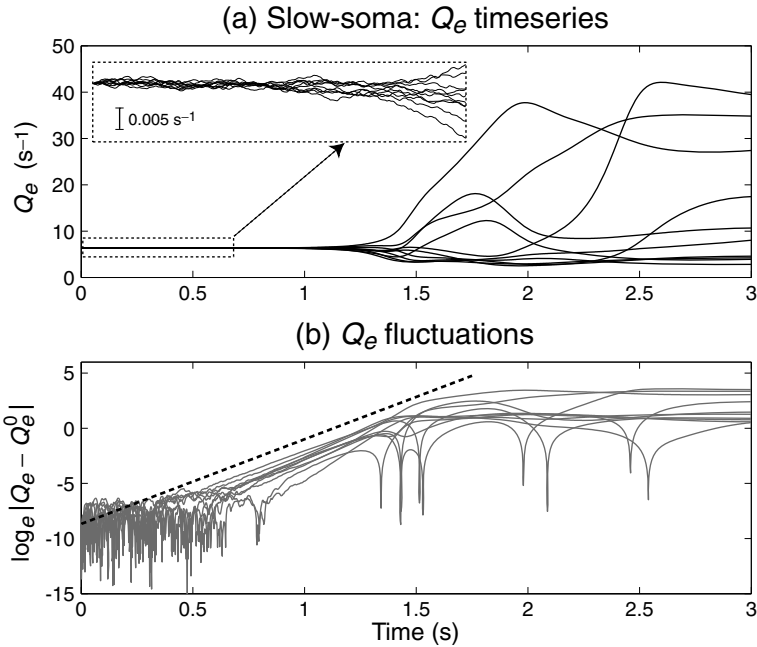


Fig. 12.6 Time-series showing formation of Fig. 12.5 Turing patterns. (a) Q_e vs time for 10 sample points distributed down the middle of the cortical sheet. *Inset:* Zoomed view of the first 0.7 s of evolution from the homogeneous equilibrium firing rate $Q_e^0 = 6.3677 s^{-1}$; scale bar = $0.005 s^{-1}$. Spatial patterns are fully developed after about 2 s. (b) Growth of fluctuations (deviations from equilibrium firing-rate) plotted on a log-scale. Dashed line shows that, for the first 1.5 s, fluctuations grow exponentially; the slope is $7.7 s^{-1}$, consistent with the Fig. 12.3(a) slow-soma prediction.

corresponding to a subcortical stimulation rate of $\phi^{sc} = 10 s^{-1}$ (i.e., $s = 0.1$), small-amplitude white-noise perturbations (with zero mean) destabilize the uniform equilibrium in favor of a spatially-organized stationary state consisting of intermixed regions of high-firing and low-firing cortical activity. The pattern wavelength of ~ 2.5 cm is consistent with the Fig. 12.3(a) prediction of maximum instability at wavenumber $q/2\pi \approx 0.4 cm^{-1}$. As is evident from Fig. 12.6, the patterns evolve promptly, with fluctuations obeying an exponential growth law $\sim e^{\alpha t}$ with $\alpha \approx 7.7 s^{-1}$, matching the Fig. 12.3(a) prediction for the dominant eigenvalue. The Turing structures are fully formed after ~ 2 s, evolving on much slower time-scales thereafter.

The Fig.-12.9 gallery of 24 snapshot images explores the sensitivity of the slow-soma cortex to changes in subcortical stimulus intensity (s increasing from left-to-right across the page), and to inhibitory diffusion (D_2 increasing from top-to-bottom). In Sect. 12.3.7 we compare and contrast these slow-soma patterns with the corresponding Fig.-12.10 gallery of images for the fast-soma case.

12.3.6 *Fast-soma simulations*

Figure 12.7 illustrates the response of the unstable *fast-soma* cortex to the imposition of continuous low-level spatiotemporal white noise. With excitatory and inhibitory diffusion strengths set at $(D_1, D_2) = (0.0005, 0.05) \text{ cm}^2$, and background subcortical flux at $\phi^{\text{sc}} = 30 \text{ s}^{-1}$, the cortical sheet organizes itself into a dynamic pattern of standing oscillations of temporal frequency $\sim 31 \text{ Hz}$ and wavelength $\sim 2 \text{ cm}$, consistent with the peak in the $(s = 0.3, D_2 = 0.05 \text{ cm}^2)$ dispersion curve of Fig. 12.4(b). The dominant spatial mode grows at the expense of higher- and lower-frequency spatial modes that either decay with time (have dominant eigenvalues whose real part is negative) or that grow more slowly than the favored mode.

The semilog plot of fluctuation amplitude vs time in Fig. 12.8(b) reveals an exponential growth law $\sim e^{\alpha t}$ with $\alpha \approx 3.9 \text{ s}^{-1}$ that persists until the onset of saturation effects at $t \approx 2.2 \text{ s}$. We note that the growth rate for the standing-wave instability is about a factor of two slower than the eigenvalue prediction of Fig. 12.4(b); the reason for this anomalous slowing has not been investigated. Nevertheless, the quantitative confirmation, via nonlinear simulation, of gamma-frequency wave activity emerging at the expected spatial and temporal frequencies, is most encouraging.

12.3.7 *Response to inhibitory diffusion and subcortical excitation*

The bold arrows labeled “ D_2 increasing” and “ s increasing” in Figs 12.3 and 12.4 highlight the fact that increases in inhibitory diffusion and in subcortical driving are predicted to act in contrary directions with respect to breaking or maintaining spatial symmetry across the cortical sheet. For the slow-soma cortex, increased *ii* diffusion D_2 encourages formation of Turing patterns, while increased subcortical stimulation s acts to restore stability to the uniform, unstructured state. These counteracting slow-soma tendencies—predicted by Fig. 12.3—are illustrated in Fig. 12.9. This figure presents a 6×4 gallery of excitatory firing-rate images captured after 2 s of continuous white-noise stimulation. A vertical top-to-bottom traverse shows that for constant subcortical drive (e.g., $s = 0.01$, left-most column), Turing formation is enhanced as D_2 diffusion is strengthened. In contrast, if diffusion is held constant (e.g., at $D_2 = 2.5 \text{ cm}^2$, second row), a horizontal scan from left-to-right shows the Turing structures losing contrast, tending to wash out as subcortical drive is increased.

For the fast-soma cortex, Fig. 12.4 indicates that these counteracting tendencies are reversed: gamma-wave instability should be enhanced as subcortical drive is boosted, but suppressed when inhibitory diffusion is increased. These theoretical claims are verified in the Fig. 12.10 gallery of fast-soma snapshots of the gamma-band standing-wave patterns. The patterns have maximum contrast in the top-right corner where subcortical drive is strong and inhibitory diffusion is minimal. A vertical downwards traverse shows the patterns broadening and weakening as diffusion is increased, eventually disappearing altogether when diffusion is set to moderate levels. It is evident that the fast-soma gamma instabilities are vastly more sensitive to small changes in inhibitory diffusion than are the slow-soma Turing instabilities.

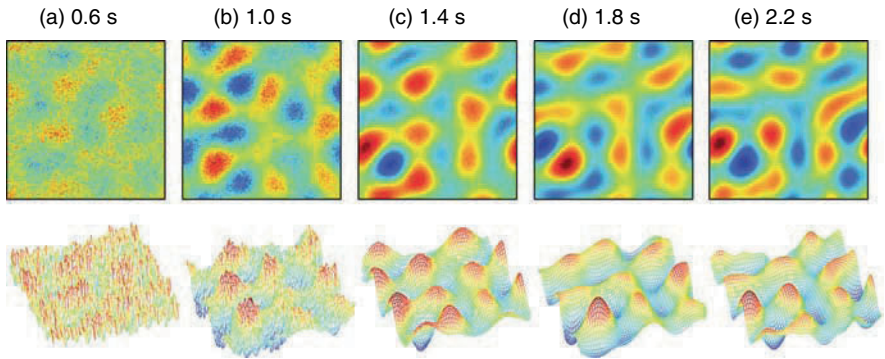


Fig. 12.7 [Color plate] Grid simulation for *fast-soma* cortical model with subcortical drive $s = 0.3$ (i.e., $\phi^{sc} = 30 \text{ s}^{-1}$), inhibitory diffusion $D_2 = 0.05 \text{ cm}^2$. Cortex was initialized at homogeneous steady-state $(Q_e^0, Q_i^0) = (7.28, 14.55) \text{ s}^{-1}$, and driven continuously with small-amplitude spatiotemporal white noise. Snapshots show the spatial and temporal evolution of Q_e at 0.4-s intervals. Consistent with Fig. 12.4(b), grid evolves into a 31-Hz standing-wave pattern of wavelength $\sim 2.0 \text{ cm}$. Red = high-firing; blue = low-firing. The fluctuations grow strongly with time, with successive panels displaying amplitude excursions, in s^{-1} , of (a) ± 0.005 ; (b) ± 0.01 ; (c) ± 0.05 ; (d) ± 0.2 ; (e) ± 1.3 about the 7.28-s^{-1} steady-state. Timestep $\Delta t = 100 \mu\text{s}$; other settings as for Fig. 12.5. (Figure reproduced from Ref. [27].)

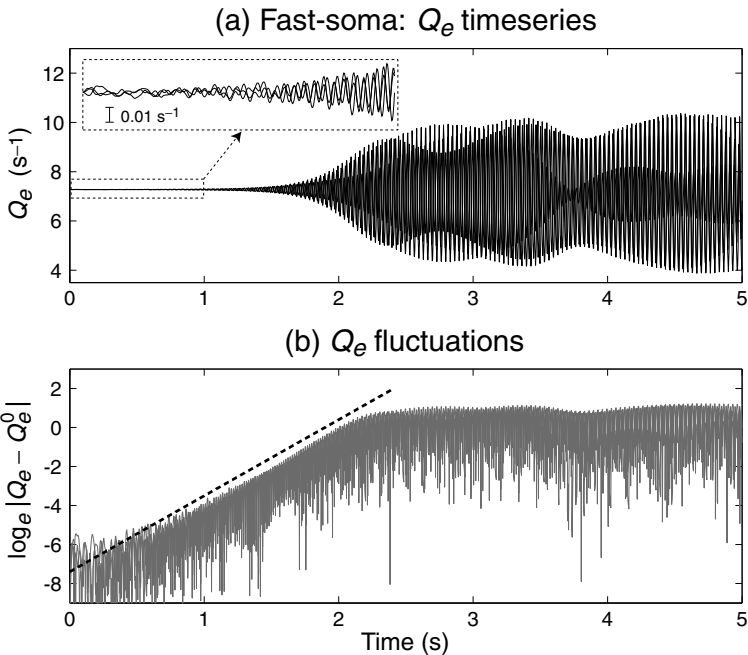


Fig. 12.8 Time-series showing formation of gamma-band standing-waves of Fig. 12.7. (a) Q_e vs time for three sample points in the cortical sheet. *Inset*: Zoomed view of the first 1 s of evolution from homogeneous equilibrium firing rate $Q_e^0 = 7.2762 \text{ s}^{-1}$; scale bar = 0.01 s^{-1} . Spatial patterns are fully developed after about 2.5 s. (b) Fluctuation growth plotted on a log-scale. Dashed line shows that, for the first 2 s, fluctuation amplitudes increase with an exponential growth rate of $\sim 3.9 \text{ s}^{-1}$.

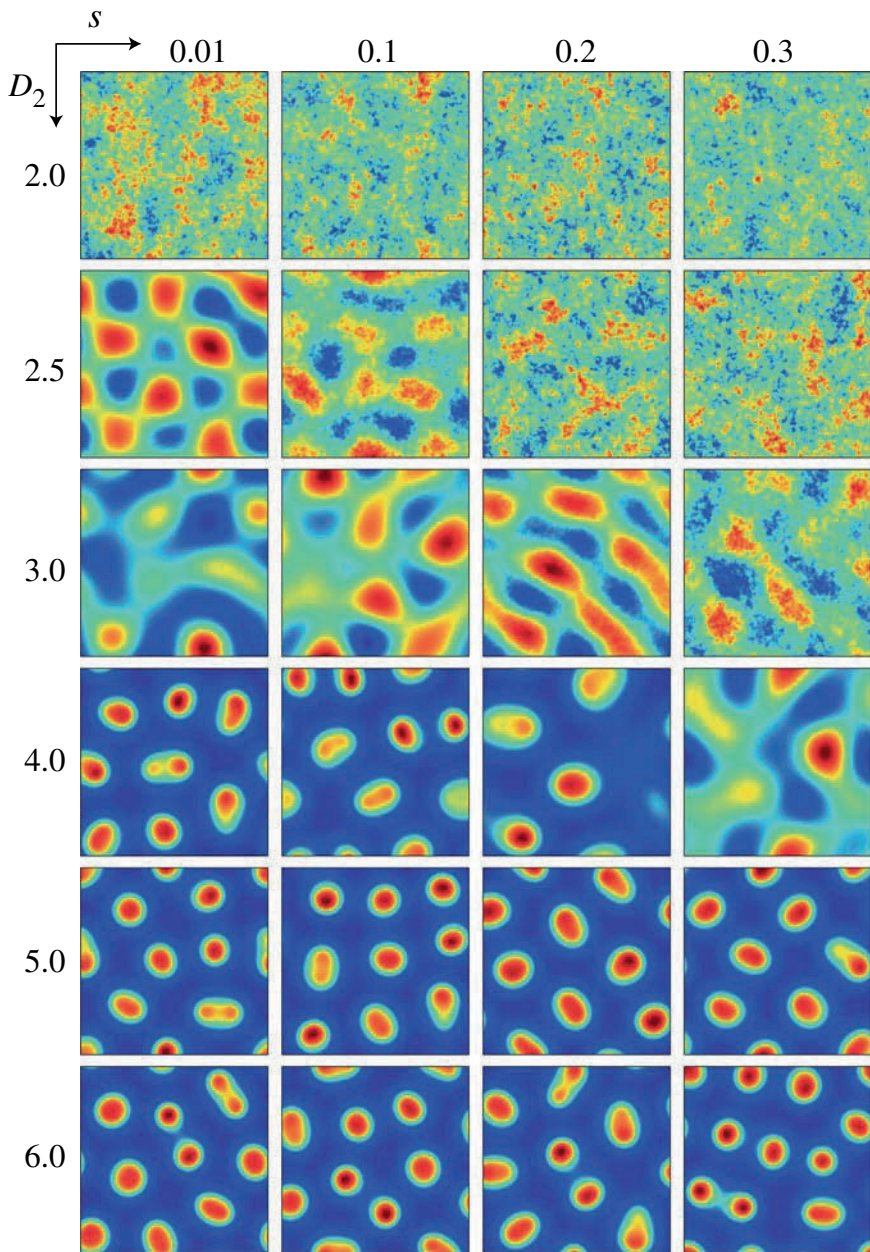


Fig. 12.9 [Color plate] Gallery of *slow-soma* Turing patterns for four values of subcortical drive s (horizontal axis), six values of inhibitory diffusion D_2 (vertical axis). Cortical sheet is initialized at homogeneous equilibrium, driven continuously by low-level continuous spatiotemporal white noise, and iterated for 2 s. Red = high-firing; blue = low-firing. Settings: $L_x = L_y = 6.0$ cm; $\Delta x = \Delta y = 0.25$ mm. Larger D_2 values required smaller timestepping: from top-to-bottom, timestep was set at $\Delta t = [3, 2, 2, 1.5, 1, 1]$ μ s.

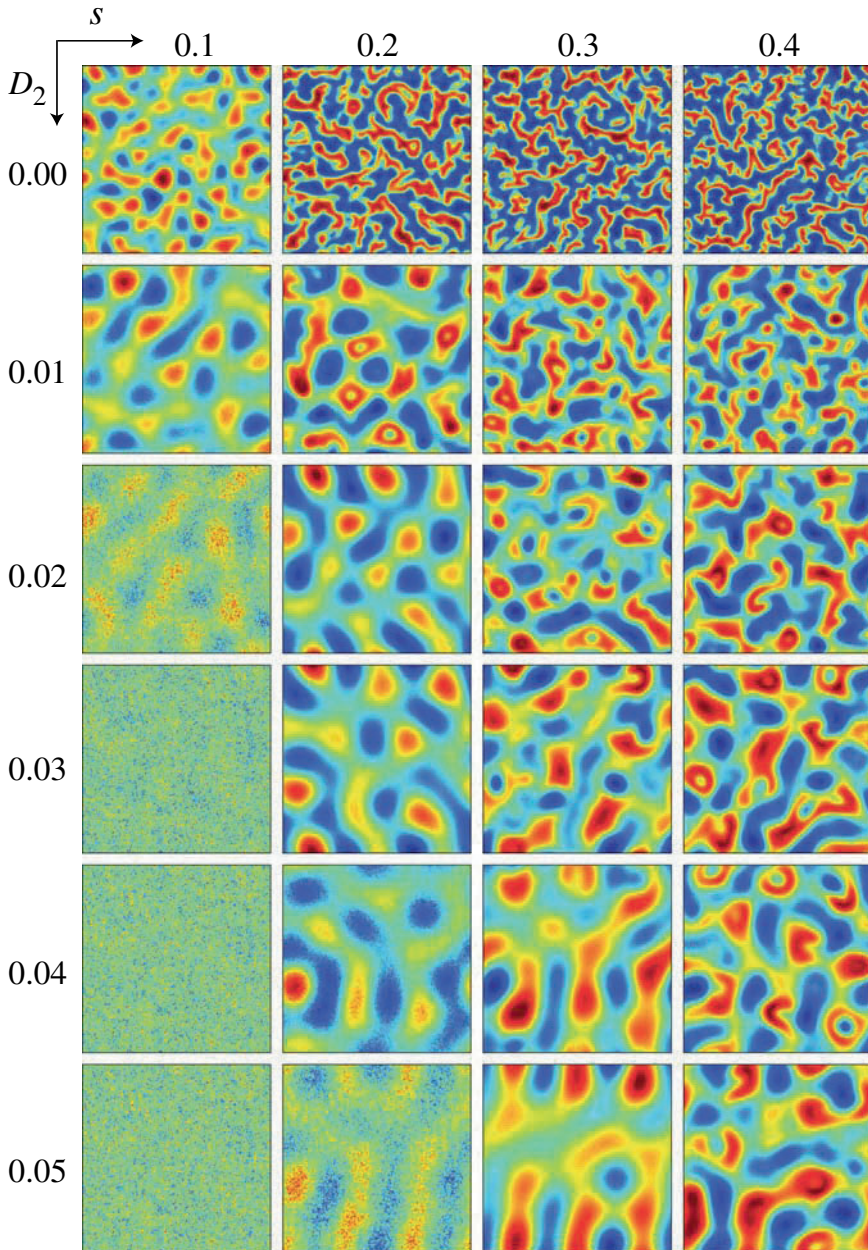


Fig. 12.10 [Color plate] Gallery of *fast-soma* standing-wave patterns for four values of subcortical drive s (horizontal axis), six values of inhibitory diffusion D_2 (vertical axis). Wave instabilities can emerge if diffusion is weak and subcortical stimulation is sufficiently strong. The wave patterns oscillate in place at ~ 30 Hz, with red and blue extrema exchanging position every half-cycle. Timestep $\Delta t = 100 \mu\text{s}$; other settings as for Fig. 12.9. (Figure reproduced from Ref. [27].)

12.4 Discussion

The crucial distinction between the slow-soma and fast-soma models is the temporal ordering of *reversal-potential weighting* (the scaling by ψ) relative to *Dendritic integration* (convolution with H): the slow-soma model assumes that the input flux M is integrated at the dendrite, then modulated by the reversal function, while the fast-soma model applies the reversal function directly to the input flux, then integrates at the dendrite. In both cases, the dendrite-filtered output flux is integrated at the soma, via convolution with the soma impulse response L , to give $\Delta V = V - V^{\text{rest}}$, the voltage perturbation from rest.

Ignoring the contribution from gap-junction diffusion, the voltage perturbation equations have the general forms,

$$V - V^{\text{rest}} = L \otimes [\rho \psi \cdot (H \otimes M)], \quad (\text{slow soma}), \quad (12.35)$$

$$V - V^{\text{rest}} = L \otimes [H \otimes (\rho \psi \cdot M)], \quad (\text{fast soma}), \quad (12.36)$$

where ρ is the (constant) synaptic strength at resting voltage. These general forms are illustrated as flow-diagrams in Fig. 12.1, with the instantaneous voltage-feedback—from soma to the dendritic reversal-weighting function—shown explicitly. As we have demonstrated in this chapter, swapping the order of dendrite filtering and reversal-potential weighting makes qualitative changes to the dynamical properties of the cortex.

For the slow-soma configuration (Fig. 12.1(a)), stationary Turing patterns can form if inhibitory diffusion is sufficiently strong; in addition, a low-frequency (~ 2 -Hz) whole-of-cortex Hopf oscillation emerges if the inhibitory PSP decay time-constant is sufficiently prolonged (not shown here; refer to [28] for details). But no evidence of higher-frequency rhythms—such as gamma oscillations—has been predicted or observed in the slow-soma stability analysis or its numerical simulation.

In contrast, the fast-soma model (Fig. 12.1(b)) supports ~ 30 -Hz gamma rhythms as cortical standing waves distributed across the 2-D cortex, with simulation behaviors being consistent with linear eigenvalue prediction. For this configuration, gamma oscillations only arise when the inhibitory diffusion is weak or non-existent. The contrary behaviors of the slow- and fast-soma mean-field models indicate that a primary determinant of cortically-generated rhythms is the nature and timeliness of the feedback from soma to dendrite.

Earlier mean-field work by Rennie *et al.* [20] also predicted gamma oscillations. However, although we have chosen our model constants (and corresponding steady-states) to be closely similar to theirs, the underlying convolution structures are very different. Translating the Rennie *et al.* symbols to match those used here by way of Table 12.3, their formulation for soma voltage reads,

$$V - V^{\text{rest}} = H \otimes [(\rho \psi \otimes L) \cdot M], \quad (\text{Rennie } et al.). \quad (12.37)$$

In the Rennie form, the incoming flux M is multiplied by a composite filter formed by convolving the reversal-potential weight ψ against the soma filter L ; the soma

Table 12.3 Correspondence between major symbols used here and those used by Rennie *et al.* [20]. Note that in our work, double-subscripts are read left-to-right, thus ab implies $a \rightarrow b$; in Rennie *et al.*, subscripts are read right-to-left.

Quantity	Symbol used here	Symbol used by Rennie <i>et al.</i>
Synaptic strength	ρ_a	S_{ba}
Reversal-potential weight	ψ_{ab}	R_{ba}
Dendritic response function	H_{ab}	S_{ba}
Soma response function	L_{ab}	H_{ba}

voltage is then obtained by integrating the product at the *dendrite* filter H . The flow-diagram for this sequence of operations (not shown) is difficult to interpret.

What is the biological significance of these slow- and fast-soma modeling predictions? In [26] we identified the slow-soma Turing patterns with the default noncognitive background state of the cortex that manifests when there is little subcortical stimulation, and in [28] we demonstrated that these patterns can be made to oscillate in place at ~ 1 Hz with a reduction in γ_{ie} , the rate-constant for the inhibitory post-synaptic potential (equivalent to rate-constant β_{ie} in this chapter). We argue that these slow patterned oscillations might relate to the even slower hemodynamic oscillations observed in the BOLD (blood-oxygen-level dependent) signals detected in fMRI (functional magnetic resonance imaging) measurements recorded from relaxed, non-engaged human subjects [7, 8].

Increases in the level of subcortical activation tend to wash out the slow-soma Turing patterns. Therefore any spatial patterns of firing activity observed during times of *elevated* subcortical stimulation—for example, during active cognition—cannot be explained using a slow-soma (with its implicit slow soma-to-dendrite feedback) limit. Instead, we replace the slow feedback assumption with the prompt feedback interactions implicit in the ordering of the convolutions adopted for the fast-soma case. In this limit, increases in subcortical drive favor emergence of traveling-wave instabilities of temporal frequency ~ 30 Hz, in the low-gamma band. Our grid simulations show that these gamma oscillations are coherent over distances of several centimeters, synchronized by an underlying standing-wave modulation of neuronal firing rates that provides a basis for the “instantaneous” action-at-a-distance observed in cognitive EEG experiments [22].

The contrasting sensitivity to inhibitory gap-junction diffusion predicted by the slow- and fast-soma models finds clinical support in brain-activity measurements from schizophrenic patients. The brains of schizophrenics carry excess concentration of the neuromodulator dopamine [30]. Dopamine is known to have a number of physiological impacts, one of these being a tendency to block neuronal gap junctions [12]. For the slow-soma model, the closure of gap junctions will reduce the inhibitory diffusion D_2 , and therefore, for a given value of subcortical drive s , will reduce the likelihood of forming Turing pattern spatial coherences during the default noncognitive state—e.g., consider a bottom-to-top traverse of the Fig. 12.9 slow-soma gallery. This degraded ability to form default-mode Turing structures

leads to the prediction that schizophrenics should exhibit impairments in the function of their default networks, and this is confirmed in two recent clinical studies [3, 19].

Applying excess dopamine to the fast-soma model, coherent gamma activity is predicted to emerge in the cognitively-active schizophrenic brain (see upper-right panel of Fig. 12.10), but these patterns will be “spindly” and less spatially generalized than those observed in a normal brain with lower dopamine levels and therefore stronger inhibitory diffusion (e.g., bottom-right panel of Fig. 12.10). This is consonant with gamma-band EEG measurements captured during cognitive tasks: compared with healthy controls, schizophrenics exhibited diminished levels of long-range phase synchrony in their gamma activity [29].

We acknowledge that, although the slow- and fast-soma models share identical steady states, the two models, as presently constructed, are not strictly comparable. This mismatch is evident in two respects. First, we find that the fast-soma model is about two orders of magnitude more sensitive to variations in inhibitory diffusion D_2 than is the slow-soma. Second, in order to bring the instability peaks in the respective dispersion graphs (Figs 12.3 and 12.4) into rough alignment at similar wavenumbers, we found it necessary to set the spatial decay-rate Λ_{eb}^α for the slow-soma wave-equation four times larger (i.e., axonal connectivity drops off four times faster) than for the fast-soma case (see Table 12.4). Despite these somewhat arbitrary model adjustments, we consider that the qualitative findings presented in this chapter are robust, namely that:

- delayed soma-to-dendrite feedback, via membrane reversal potentials, supports stationary or slowly fluctuating spatial firing-rate patterns
- prompt feedback from soma to dendrite enhances spatially-coherent gamma oscillations
- gap-junction diffusion has a strong influence on the stability and spatial extent of neural pattern coherence.

Acknowledgments We thank Chris Rennie for helpful discussions on convolution formulations for the cortex. This research was supported by the Royal Society of New Zealand Marsden Fund, contract 07-UOW-037.

Appendix

Table 12.4 Standard values for the neural model. Subscript label b means destination cell can be either of type e (excitatory) or i (inhibitory). Most of these values are drawn from Rennie *et al.* [20].

Symbol	Description	Value	Unit
$\tau_{e,i}$	soma time constant	0.050, 0.050	s
$V_{e,i}^{\text{rev}}$	reversal potential for AMPA, GABA channels	0, -70	mV
$V_{e,i}^{\text{rest}}$	resting potential	-60, -60	mV
$\rho_{e,i}$	synaptic gain at resting potential	$(2.4, -5.9) \times 10^{-3}$	mV·s
β_{ee}, β_{ie}	PSP rise-rate in excitatory neurons	500, 500	s^{-1}
β_{ei}, β_{ii}	PSP rise-rate in inhibitory neurons	500, 500	s^{-1}
α_{ee}	EPSP decay-rate in excitatory neurons	68	s^{-1}
α_{ei}	EPSP decay-rate in inhibitory neurons	176	s^{-1}
α_{ie}	IPSP decay-rate in excitatory neurons	47	s^{-1}
α_{ii}	IPSP decay-rate in inhibitory neurons	82	s^{-1}
N_{eb}^{α}	long-range $e \rightarrow b$ axonal connectivity	3710	-
$N_{eb}^{\beta}, N_{ib}^{\beta}$	local $e \rightarrow b, i \rightarrow b$ axonal connectivity	410, 800	-
$N_{eb}^{\text{sc}}, N_{ib}^{\text{sc}}$	subcortical $e \rightarrow b, i \rightarrow b$ axonal connectivity	80, 0	-
s	control parameter for subcortical synaptic flux	0.1	-
v_{eb}^{α}	long-range $e \rightarrow b$ axonal speed	140	cm s^{-1}
$v_{eb}^{\beta}, v_{ib}^{\beta}$	local $e \rightarrow b, i \rightarrow b$ axonal speed	20, 20	cm s^{-1}
Λ_{eb}^{α}	inverse-length scale for long-range $e \rightarrow b$ axons (slow-soma)	4	cm^{-1}
Λ_{eb}^{β}	inverse-length scale for long-range $e \rightarrow b$ axons (fast-soma)	1	cm^{-1}
$\Lambda_{eb}^{\beta}, \Lambda_{ib}^{\beta}$	inverse-length scale for local $e \rightarrow b, i \rightarrow b$ axons	50, 50	cm^{-1}
$Q_{e,i}^{\text{max}}$	maximum firing rate	100, 200	s^{-1}
$\theta_{e,i}$	threshold voltage for firing	-52, -52	mV
$\sigma_{e,i}$	standard deviation for threshold	5, 5	mV
$L_{x,y}$	length, width of cortical sheet	6, 6	cm

References

1. Alvarez-Maubecin, V., García-Hernández, F., Williams, J.T., Van Bockstaele, E.J.: Functional coupling between neurons and glia. *J. Neurosci.* **20**, 4091–4098 (2000)
2. Bennett, M.V., Zukin, R.S.: Electrical coupling and neuronal synchronization in the mammalian brain. *Neuron* **41**, 495–511 (2004)
3. Bluhm, R.L., Miller, J., Lanius, R.A., Osuch, E.A., Boksman, K., Neufeld, R.W.J., Théberge, J., Schaefer, B., Williamson, P.: Spontaneous low frequency fluctuations in the BOLD signal in schizophrenic patients: Anomalies in the default network. *Schizophrenia Bulletin* **33**(4), 1004–1012 (2007)
4. Bressloff, P.C.: New mechanism for neural pattern formation. *Phys. Rev. Lett.* **76**(24), 4644–4647 (1996), doi:10.1103/PhysRevLett.76.4644
5. Coombes, S., Lord, G.J., Owen, M.R.: Waves and bumps in neuronal networks with axo-dendritic synaptic interactions. *Physica D* **178**, 219–241 (2003), doi:10.1016/S0167-2789(03)00002-2
6. Ermentrout, G.B., Cowan, J.D.: Temporal oscillations in neuronal nets. *Journal of Mathematical Biology* **7**, 265–280 (1979)
7. Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., van Essen, D.C., Raichle, M.E.: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. USA* **102**(27), 9673–9678 (2005), doi:10.1073/pnas.0504136102
8. Fransson, P.: Human spontaneous low-frequency BOLD signal fluctuations: An fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum. Brain Mapp.* **26**, 15–29 (2005), doi:10.1002/hbm.20113
9. Freeman, W.J.: *Mass Action in the Nervous System*. Academic Press, New York (1975)
10. Fukuda, T., Kosaka, T., Singer, W., Galuske, R.A.W.: Gap junctions among dendrites of cortical GABAergic neurons establish a dense and widespread intercolumnar network. *J. Neurosci.* **26**, 3434–3443 (2006)
11. Haken, H.: *Brain Dynamics: Synchronization and Activity Patterns in Pulse-Coupled Neural Nets with Delays and Noise*. Springer, Berlin (2002)
12. Hampson, E.C.G.M., Vaney, D.I., Weile, R.: Dopaminergic modulation of gap junction permeability between amacrine cells in mammalian retina. *J. Neurosci.* **12**, 4911–4922 (1992)
13. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (Lond.)* **117**, 500–544 (1952)
14. Hutt, A., Bestehorn, M., Wennekers, T.: Pattern formation in intracortical neuronal fields. *Network: Computation in Neural Systems* **14**, 351–368 (2003)
15. Laing, C.R., Troy, W.C., Gutkins, B., Ermentrout, G.B.: Multiple bumps in a neuronal model of working memory. *SIAM J. Appl. Math.* **63**(1), 62–97 (2002), doi:10.1137/S0036139901389495
16. Liley, D.T.J., Cadusch, P.J., Wright, J.J.: A continuum theory of electro-cortical activity. *Neurocomputing* **26–27**, 795–800 (1999)
17. Nadarajah, B., Thomaidou, D., Evans, W.H., Parnavelas, J.G.: Gap junctions in the adult cerebral cortex; Regional differences in their distribution and cellular expression of connexins. *Journal of Comparative Neurology* **376**, 326–342 (1996)
18. Nunez, P.L.: The brain wave function: A model for the EEG. *Mathematical Biosciences* **21**, 279–297 (1974)
19. Ouyang, L., Deng, W., Zeng, L., Li, D., Gao, Q., Jiang, L., Zou, L., Cui, L., Ma, X., Huang, X.: Decreased spontaneous low-frequency BOLD signal fluctuation in first-episode treatment-naïve schizophrenia. *Int. J. Magn. Reson. Imaging* **1**(1), 61–64 (2007)
20. Rennie, C.J., Wright, J.J., Robinson, P.A.: Mechanisms for cortical electrical activity and emergence of gamma rhythm. *J. Theor. Biol.* **205**, 17–35 (2000)
21. Robinson, P.A., Rennie, C.J., Wright, J.J.: Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* **56**, 826–840 (1997)

22. Rodriguez, E., George, N., Lachaux, J.P., Martinerie, J., Renault, B., Varela, F.J.: Perception's shadow: long-distance synchronization of human brain activity. *Nature* **397**, 430–433 (1999)
23. Steyn-Ross, D.A., Steyn-Ross, M.L., Sleigh, J.W., Wilson, M.T., Gillies, I.P., Wright, J.J.: The sleep cycle modelled as a cortical phase transition. *Journal of Biological Physics* **31**, 547–569 (2005)
24. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W.: Modelling general anaesthesia as a first-order phase transition in the cortex. *Progress in Biophysics and Molecular Biology* **85**, 369–385 (2004)
25. Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W., Liley, D.T.J.: Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: Evidence for a general anesthetic-induced phase transition. *Phys. Rev. E* **60**, 7299–7311 (1999)
26. Steyn-Ross, M.L., Steyn-Ross, D.A., Wilson, M.T., Sleigh, J.W.: Gap junctions mediate large-scale Turing structures in a mean-field cortex driven by subcortical noise. *Phys. Rev. E* **76**, 011916 (2007), doi:10.1103/PhysRevE.76.011916
27. Steyn-Ross, M.L., Steyn-Ross, D.A., Wilson, M.T., Sleigh, J.W.: Modeling brain activation patterns for the default and cognitive states. *NeuroImage* **45**, 298–311 (2009), doi:10.1016/j.neuroimage.2008.11.036
28. Steyn-Ross, M.L., Steyn-Ross, D.A., Wilson, M.T., Sleigh, J.W.: Interacting Turing and Hopf instabilities drive pattern formation in a noise-driven model cortex. In: R. Wang, F. Gu, E. Shen (eds.), *Advances in Cognitive Neurodynamics ICCN 2007*, chap. 40, pp. 227–232, Springer (2008)
29. Uhlhaas, P.J., Linden, D.E.J., Singer, W., Haenschel, C., Lindner, M., Maurer, K., Rodriguez, E.: Dysfunctional long-range coordination of neural activity during gestalt perception in schizophrenia. *J. Neurosci.* **26**, 8168–8175 (2006)
30. Uhlhaas, P.J., Singer, W.: Neural synchrony in brain disorders: Relevance for cognitive dysfunctions and pathophysiology. *Neuron* **52**(1), 155–168 (2006), doi:10.1016/j.neuron.2006.09.020
31. Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973)
32. Wilson, M.T., Sleigh, J.W., Steyn-Ross, D.A., Steyn-Ross, M.L.: General anesthetic-induced seizures can be explained by a mean-field model of cortical dynamics. *Anesthesiology* **104**, 588–593 (2006)
33. Wilson, M.T., Steyn-Ross, D.A., Sleigh, J.W., Steyn-Ross, M.L., Wilcocks, L.C., Gillies, I.P.: The slow oscillation and K-complex in terms of a mean-field cortical model. *Journal of Computational Neuroscience* **21**, 243–257 (2006)
34. Wilson, M.T., Steyn-Ross, M.L., Steyn-Ross, D.A., Sleigh, J.W.: Predictions and simulations of cortical dynamics during natural sleep using a continuum approach. *Phys. Rev. E* **72**, 051910 (2005)
35. Wright, J.J., Liley, D.T.J.: A millimetric-scale simulation of electrocortical wave dynamics based on anatomical estimates of cortical synaptic density. *Network: Comput. Neural Syst.* **5**, 191–202 (1994), doi:10.1088/0954-898X/5/2/005
36. Wright, J.J., Robinson, P.A., Rennie, C.J., Gordon, E., Bourke, P.D., Chapman, C.L., Hawthorn, N., Lees, G.J., Alexander, D.: Toward an integrated continuum model of cerebral dynamics: the cerebral rhythms, synchronous oscillation and cortical stability. *BioSystems* **73**, 71–88 (2001)

Index

- 5-HT (serotonin), 155

- AAS (ascending arousal system), 194–198
- Acetylcholine, see ACh
- ACh (acetylcholine), 17, 20, 152, 155, 208, 226
- Activity
 - aminergic, 206
 - cholinergic, 206–208, 210
 - neuromodulator, 206, 210
 - spindle, 215
- Adenosine, 17, 20, 195, 206, 226
- Adiabatic approximation, 74
- AHP (after-hyperpolarization) current, 159
- Alpha oscillation, 90, 118–141, 190–193
 - genesis, 120–121
 - historical overview, 118
 - in EEG, 121
 - in thalamus, 120–121
 - modeling, 121–122
 - mu rhythm, 119
 - properties, 119–120, 138–140
 - tau rhythm, 119
- Aminergic activity, 206
- AMPA
 - receptor, 273
 - reversal potential, 274
- Anesthesia, 11–15, 41–43
 - biphasic paradox, 2, 11, 13
 - EEG fluctuation power, 15
 - effect on bioluminescence, 11–12
 - Guedel’s stages for induction, 11
 - hysteresis, 13
 - numerical model, 13–15
 - phase transition, 2, 13, 167
- ARMA (autoregressive moving-average) models, 30–35

- Asymptotic stability, 68
- Avalanche, 100

- Benzodiazepine, 121, 133
- Beta oscillation, 122, 156
- BFGS quasi-Newton algorithm, 39
- Bifurcation
 - Bogdanov–Takens, 131, 133
 - homoclinic, 131
 - Hopf, 8, 83–85, 131, 133–134, 211, 227–229
 - pitchfork, 74
 - route to chaos, 133
 - saddle–node, 9, 19, 22, 83, 131, 133, 182, 191, 193, 196, 237
 - Shilnikov, 131, 133
 - Turing, 61–63, 71–77, 272–274, 276, 284–294
- Bioluminescence
 - quantification of anesthetic potency, 12
- Blood–brain barrier, 13
- Bogdanov–Takens bifurcation, 131, 133
- BOLD (blood-oxygen-level dependent) signal, 90, 295
 - ultra-slow oscillation, 87
- Brain resting-state
 - characteristics, 94–96
 - correlated and anticorrelated networks, 87
 - dynamics, 92, 94
 - instability, 86–87
 - philosophical implications, 96
 - primate brain simulation, 86
 - spatiotemporal analysis, 87
 - transient thought, 94
- Breathing mode, 230–233
- Brown noise, 256
- Brownian motion, 7, 10

- Chaos, 92, 133
 Chloroform, 11
 Cholinergic activity, 206–208, 210
 Circadian sleep-drive, 195
 Circular causality, 73
 Coherence, 104, 156, 190, 193
 Coherent Infomax theory, 259
 Computational methods, 149–150
 Connections
 cortico-cortical, 62
 intracortical inhibitory, 62
 Connectivity kernel, 272
 exponential, 272
 Macdonald-function, 272
 Mexican hat, 272
 Continuum models, 271–272
 Correlated brain networks
 “task-negative” regions, 87
 “task-positive” regions, 87
 Critical fluctuations, 7–9, 70–71
 scaling law, 9
 Critical modes, 72
 Critical slowing, 5, 9, 19, 21
 Cx36 (connexin-36) gap junction, 273, 280
 Cyclopropane, 11

 Default (noncognitive) mode, 287, 295
 Default network, 86
 Delta oscillation, 211, 272
 Dendrite, 272–274, 276, 278, 281
 Dendrite filter response, 275
 Diffusive coupling, 231, 271, 273–274, 281
 Dopamine, 295
 Dynamic repertoire, 2, 19, 83, 85

 ECoG (electrocorticogram), 16, 148, 205, 209, 255
 ECT (electroconvulsive therapy), 163–167
 EEG (electroencephalogram), 15, 101, 104, 119, 122, 136–138, 148, 162, 166, 223, 229, 235–236, 271, 295
 artifacts, 89
 dynamical invariants, 89–94
 mutual information, 90–92
 nonlinearity, 137–138
 spatiotemporal analysis, 93–94
 spectrum, 190
 time-series analysis, 90–93
 Eigenvalue, 8, 10, 15, 32–33, 83, 94, 135, 211, 228, 283–286
 Electrical stimulation, 162
 Epilepsy, 101, 110
 as phase transition, 103
 limbic, 103
 seizure, 45–48
 spreading, 100–113
 temporal lobe, 103
 EPSP (excitatory PSP), 204, 206–208, 210, 225
 Equilibrium, *see* Steady state
 Ether, 11

 Feed-forward networks, 54
 FitzHugh–Nagumo neuron, 84–86, 150
 Fluctuation
 critical, 7–9, 83
 scaling law at critical point, 9
 spectrum, 10
 subthreshold, 5–7
 variance, 9–10
 fMRI (functional magnetic resonance imaging), 295
 Fokker–Planck equation, 68, 71, 75
 Fourier expansion, 57–59, 61, 63–65, 67, 69–70, 72
 Fourier transform, 89–90, 94, 127, 283
 Frankenhaeuser–Huxley neuron, 150, 167
 Freeman, Walter, 2, 6, 152, 162–163, 243
 Fukuda cell, 280

 GABA
 receptor, 101, 273
 reversal potential, 274
 Gamma
 oscillation, 133, 155, 245–254, 272, 285–287
 correlation with alpha activity, 122
 effect on cortex, 257–260
 Gap-junction
 connectivity, 273, 280
 diffusion, 272–274
 resistance, 281
 synapse, 273
 GARCH modeling, 36–40
 recommendations, 39
 state prediction, 36–37
 Gaussian distribution, 28, 56, 67–69
 Ghost, 23
 Glial cell, 273, 280
 Green’s function, 54, 69
 Group velocity, 286

 Halothane, 11
 Hartman–Grobman theorem, 128
 Heteroscedasticity, 36
 Hilbert transform, 94, 255
 Hippocampus, 101, 214
 information processing, 103

- memory formation, 103
 - Hodgkin–Huxley neuron, 3, 81, 84, 150, 158, 273
 - Homeostatic sleep–drive, 195
 - Homoclinic bifurcation, 131
 - Hopf bifurcation, 8, 83–85, 131, 133–134, 211, 227–229, 286
 - subcritical, 84
 - supercritical, 84
 - Hopf oscillator, 85
 - Hurst exponent, 92
 - Hypnic jerk, 20–23
 - Hysteresis, 13, 22, 193, 196–197

 - Ichogenesis, 101
 - IDE (integro-differential equation), 57, 272
 - Impulse response
 - alpha-function, 274
 - biexponential, 274, 278
 - Inhibitory diffusion, 284
 - Instability, *see* also Bifurcation
 - non-oscillatory, 70
 - of brain resting-state, 86–87
 - oscillatory, 63–66
 - spatiotemporal, 54, 77
 - wave, 273, 285–286, 290, 295
 - Integrator neuron, 9
 - subthreshold “resonance at dc”, 10
 - Interneurons
 - L-type, 280
 - Ion-channel density, 168–169
 - Ionic conductance, 3, 159, 165
 - IPSP (inhibitory PSP), 215, 225, 229
 - IS (intermediate sleep), 210, 215
 - IS–REM sleep transition, 212

 - Jacobian matrix, 283

 - K-complex, 237–238
 - Kalman filter, 30, 36, 38, 40
 - Kernel function, 57–59, 61–65, 67, 70
 - Ketamine, 209
 - Kuramoto–Sivashinsky equation, 57

 - LFP (local field potential), 150, 160
 - Limbic system, 103
 - Linear stochastic null hypothesis, 136
 - Luciferase, 11
 - Lyapunov exponent, 58, 68–70, 91–92, 131

 - Macrocolumn, 55, 59, 224
 - Mean-field model, 122–136, 204–207, 213–215, 223–226
 - construction, 124
 - dynamics, 130–136
 - EEG spectrum, 190
 - extended Liley model, 124–128
 - limitations of, 223
 - linear instability, 191
 - linearization, 128–129, 189
 - macrocolumn, 224
 - nonlinear instability, 191, 194
 - of neuronal activity, 181–183, 187
 - parameters, 186
 - physiological plausibility, 129–130
 - rationale, 180
 - steady states, 187
- Mean-square stability, 68–70
 - MEG (magnetoencephalogram), 148, 271
 - Membrane
 - capacitance, 159
 - resistance, 280
 - time-constant, 281
 - Mesoscopic brain dynamics, 148–150
 - Metastability, 140–141
 - Hebbian perspective, 140
 - Mexican hat, 161, 256
 - Microsleeps, 196
 - Monte Carlo spreading simulation, 109
 - Mu rhythm, 119
 - Mutual information, 90–92

 - Narcolepsy, 196
 - Natural (internally-induced) phase transitions, 150–162
 - Nelder–Mead simplex algorithm, 39
 - Neocortical network model, 157–160
 - Network models, 99–112
 - cat cortex, 102
 - chronic limbic epilepsy, 104–105
 - hierarchical, 101–103, 105–112
 - neocortical, 157–160
 - paleocortical, 151–153
 - random, 106, 108, 111
 - scale-free, 101
 - self-organized, 100–101
 - small-world, 106, 108–110, 112
 - topology, 100, 106, 108, 111
 - Neuromodulator activity, 206, 210
 - Neuromodulator-induced phase transition, 155–156
 - Neuron models
 - classification, 5
 - FitzHugh–Nagumo, 84–86, 150
 - Frankenhaeuser–Huxley, 150, 167
 - Hodgkin–Huxley, 3, 81, 84, 150, 158, 273
 - phase transition, 2–10
 - type-I (integrator), 5, 9

- Neuron models (*cont.*)
 - type-II (resonator), 5, 9
 - Wilson spiking model, 3–5
- Neuronal connectivity
 - anatomical, 81
 - functional, 82
- Neurotransmitter
 - AMPA, 225
 - GABA, 225
 - NMDA, 259, 263
- Noise covariance, 36
- Noise-induced phase transition, 77, 150–151, 153–155
- Nonlinear time-series analysis, 137–138
- Nonstationarity, 36
- NSF (nonspecific flux), 245–254
- Null spike, 255–256
 - as marker for phase transition, 256
- Nyquist frequency, 34

- Olfactory cortex, 162–163
- Order-parameter, 75
- Orexin, 195
- Ornstein–Uhlenbeck (Brownian motion)
 - process, 7, 10, 75

- Pacemakers, 190
- Paleocortical network model, 151–153
- PCA (principal component analysis), 87, 94
- PDE (partial differential equation), 57, 272
- Phase cone, 254–255
- Phase slip, 256
- Phase synchronization, 93
- Phase transition, 179–181, 190, 193, 198, 227, 255–256
 - anesthetic, 2, 13, 167
 - artificial (externally induced), 162–167
 - attention-induced, 156–157
 - ECT-induced seizure, 163–167
 - in single neuron, 2–10
 - intermediate to REM sleep, 212
 - natural (internally induced), 150–162
 - neuromodulator-induced, 155–156
 - noise-induced, 150–151, 153–155
 - SWS to REM sleep, 15–16
 - wake–sleep, 20–23
- Phase-space trajectory, 238
- Pitchfork bifurcation, 74
- Power spectra, 248
 - centimetric scale, 248
 - macrocolumnar scale, 248, 251
- Probability density function, 55, 58, 71, 75
- Propagation speed, 56–58, 66
- Propofol, 13

- PseudoECoG, 210
- PSP (postsynaptic potential), 13, 54, 224, 272
 - excitatory, 54, 120, 204, 206–208, 210, 275
 - inhibitory, 54, 131, 215, 275

- Random fluctuations, 66–69, 71, 74, 76
- Recurrent network, 54
- REM (rapid-eye-movement) sleep, 188, 207, 210, 214–215, 227
- Resonator neuron, 9
 - subthreshold ringing, 5, 10
- Reversal potential, 3, 272–275

- Saddle–node bifurcation, 9, 19, 22, 83, 131, 133, 182, 191, 193, 196, 237
- Schizophrenia, 295
- SDE (stochastic differential equation), 67
- Seizure, 101, 104, 110
 - epilepsy, 240
 - ictogenesis, 101
 - spreading, 100–113
 - status epilepticus, 104
- Self-organized criticality, 100–101
- Serotonin, 155
- Shilnikov bifurcation, 131, 133
- Short-range flux, 279
- Sigmoidal function, 56, 225
- Slaving principle, 73
- Sleep
 - in fetal sheep, 43–45
 - intermediate (IS), 212
 - rapid-eye-movement (REM), 227
 - slow-wave (SWS), 227
- Sleep–wake cycle
 - AAS (ascending arousal system), 194–198
 - circadian drive, 195
 - homeostatic drive, 195
 - microsleeps, 196
 - physiological basis, 194
 - spectral characteristics, 190
- Small-world network, *see* Network models, small-world
- Soma
 - “fast-soma” model, 274–278, 282–296
 - “slow-soma” model, 274–278, 282–296
 - impulse-response, 275, 277
 - time-constant, 275
 - voltage, 274–276, 278, 281–283
- Soma potential, 205, 210
- Somnogen, 206
- Spatial diffusion, 231
- Spatial interactions
 - local excitation–lateral inhibition (Mexican hat), 60, 65

- local inhibition–lateral excitation (inverse Mexican hat), 60
 - topological inhibition, 112
- Spatial mode, 68–71, 77
- Spindle, 215
- Spiral wave, 240
- Stability
 - asymptotic, 68, 70
 - linear, 53
 - mean-square, 68–70
 - stochastic, 68–70
- Stable modes, 72
- Standing wave, 272, 276, 287
- State-space modeling, 30–35
 - modal representation, 32–33
- Stationary state, 58, 63, 68–74, 226–227, 282, 287
- STLR (spatiotemporal learning rule), 259
- Stochastic
 - analysis, 74, 76
 - center manifold, 74–76
 - exploration, 3
 - stability, 68–70
 - volatility modeling, 36
- Subcortical drive, 282, 287
- Swift–Hohenberg equation, 57
- SWS (slow-wave sleep), 206, 210, 227
- SWS–REM sleep transition, 15–16, 206, 210, 214
- Synchronization, 157
- Synchronous oscillation, 251–252
- Tau rhythm, 119
- TCF (transcortical flux), 245–254
- Thalamus, 184, 187, 190
- Theta oscillation, 155, 211, 214
- Time-series modeling
 - innovation, 28
 - Markov process, 28
 - maximum-likelihood estimation, 28–29
 - white-noise, 28
- Topological inhibition, 112
- Traveling wave, 251–252, 254–255
- Turing bifurcation, 61–63, 71–77, 272–274, 276, 284–292
- Unmyelinated axon, 279
- Visual attention, 160–162
- Volume conduction, 184
- Wake–sleep transition, 20–23
- Wave equation, 278–279
- White noise, 3, 14, 28, 86, 92, 210, 224, 251, 279, 290
- Wiener process, 67
- Wilson neuron, 3–5
 - stochastic simulation, 5
- Xylazine, 209
- Zero-lag synchrony, 252