

frontiers

RESEARCH TOPICS

NEURAL MASSES AND FIELDS: MODELLING THE DYNAMICS OF BRAIN ACTIVITY

Topic Editors

Dimitris Pinotsis, Peter Robinson,
Peter beim Graben and Karl Friston



frontiers in
COMPUTATIONAL NEUROSCIENCE



frontiers

FRONTIERS COPYRIGHT STATEMENT

© Copyright 2007-2015
Frontiers Media SA.
All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

Cover image provided by lbbl sarl, Lausanne CH

ISSN 1664-8714

ISBN 978-2-88919-427-8

DOI 10.3389/978-2-88919-427-8

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

NEURAL MASSES AND FIELDS: MODELLING THE DYNAMICS OF BRAIN ACTIVITY

Topic Editors:

Dimitris Pinotsis, University College London, United Kingdom

Peter Robinson, University of Sydney, Australia

Peter beim Graben, Humboldt-Universität zu Berlin, Germany

Karl Friston, University College London, United Kingdom

Biophysical modelling of brain activity has a long and illustrious history and has recently profited from technological advances that furnish neuroimaging data at an unprecedented spatiotemporal resolution. Neuronal modelling is a very active area of research, with applications ranging from the characterization of neurobiological and cognitive processes, to constructing artificial brains in silico and building brain-machine interface and neuroprosthetic devices. Biophysical modelling has always benefited from interdisciplinary interactions between different and seemingly distant fields; ranging from mathematics and engineering to linguistics and psychology. This Research Topic aims to promote such interactions by promoting papers that contribute to a deeper understanding of neural activity as measured by fMRI or electrophysiology.

In general, mean field models of neural activity can be divided into two classes: neural mass and neural field models. The main difference between these classes is that field models prescribe how a quantity characterizing neural activity (such as average depolarization of a neural population) evolves over both space and time as opposed to mass models, which characterize activity over time only; by assuming that all neurons in a population are located at (approximately) the same point. This Research Topic focusses on both classes of models and considers several aspects and their relative merits that: span from synapses to the whole brain; comparisons of their predictions with EEG and MEG spectra of spontaneous brain activity; evoked responses, seizures, and fitting data - to infer brain states and map physiological parameters.

Table of Contents

05	<i>Neural Masses and Fields: Modeling the Dynamics of Brain Activity</i>	Dimitris Pinotsis, Peter Robinson, Peter beim Graben and Karl Friston
08	<i>Traveling Pulses in a Stochastic Neural Field Model of Direction Selectivity</i>	Paul C. Bressloff and Jeremy Wilkerson
22	<i>A Biophysical Observation Model for Field Potentials of Networks of Leaky Integrate-and-Fire Neurons</i>	Peter beim Graben and Serafim Rodrigues
35	<i>Spatiotemporal Imaging of Complexity</i>	Stephen E. Robinson, Arnold J. Mandell and Richard Coppola
49	<i>The Anesthetic Propofol Shifts the Frequency of Maximum Spectral Power in EEG During General Anesthesia: Analytical Insights From a Linear Model</i>	Axel Hutt
59	<i>On the Dynamics of Cortical Development: Synchrony and Synaptic Self-Organization</i>	James Joseph Wright and Paul David Bourke
76	<i>How Adaptation Shapes Spike Rate Oscillations in Recurrent Neuronal Networks</i>	Moritz Augustin, Josef Ladenbauer and Klaus Obermayer
87	<i>Inferring Network Properties of Cortical Neurons with Synaptic Coupling and Parameter Dispersion</i>	Dipanjan Roy and Viktor Jirsa
107	<i>Ketamine, Propofol, and the EEG: A Neural Field Analysis of HCN1-Mediated Interactions</i>	Ingo Bojak, Harry C. Day and David T. J. Liley
121	<i>Neural Mass Modeling of Power-Line Magnetic Fields Effects on Brain Activity</i>	J. Modolo, A. W. Thomas and A. Legros
136	<i>Stability Constraints on Large-Scale Structural Brain Networks</i>	Richard T. Gray and Peter A. Robinson
149	<i>Top-Down Influences on Local Networks: Basic Theory with Experimental Implications</i>	Ramesh Srinivasan, Samuel Thorpe and Paul L. Nunez
164	<i>Cortical Information Flow in Parkinson's Disease: A Composite Network/Field Model</i>	Cliff C. Kerr, Sacha J. Van Albada, Samuel A. Neymotin, George L. Chadderdon, P. A. Robinson and William W. Lytton

- 178** *The Mesoscopic Modeling of Burst Suppression During Anesthesia*
David T. J. Liley and Matthew Walsh
- 190** *Neural Masses and Fields in Dynamic Causal Modeling*
Rosalyn Moran, Dimitris A. Pinotsis and Karl Friston
- 202** *Implementing the Cellular Mechanisms of Synaptic Transmission in a Neural Mass Model of the Thalamo-Cortical Circuitry*
Basabdatta S. Bhattacharya
- 213** *Modulation of Epileptic Activity by Deep Brain Stimulation: A Model-Based Study of Frequency-Dependent Effects*
Faten Mina, Pascal Benquet, Anca Pasnicu, Arnaud Biraben and Fabrice Wendling
- 229** *On Conductance-Based Neural Field Models*
Dimitris A. Pinotsis, Marco Leite and Karl J. Friston



Neural masses and fields: modeling the dynamics of brain activity

Dimitris Pinotsis^{1*}, Peter Robinson^{2,3,4}, Peter beim Graben⁵ and Karl Friston¹

¹ The Wellcome Trust Centre for Neuroimaging, University College London, London, UK

² School of Physics, University of Sydney, NSW, Australia

³ Center for Integrative Brain Function, University of Sydney, NSW, Australia

⁴ Brain Dynamics Centre, Westmead Millennium Institute, Westmead, NSW, Australia

⁵ Institut für deutsche Sprache und Linguistik, Humboldt-Universität zu Berlin, Berlin, Germany

*Correspondence: d.pinotsis@ucl.ac.uk

Edited and reviewed by:

Misha Tsodyks, Weizmann Institute of Science, Israel

Keywords: integro-differential equations, neural field theory, neural masses, anaesthesia, electrophysiology, electroencephalogram, oscillations, neural disorders

Biophysical modeling of brain activity has a long and illustrious history (Ermentrout, 1998; Deco et al., 2008; Coombes, 2010) and has recently profited from technological advances that furnish neuroimaging data at an unprecedented spatiotemporal resolution (Guillory and Bujarski, 2014; Sporns, 2014). Neuronal modeling is a very active area of research, with applications ranging from the characterization of neurobiological and cognitive processes, (Jirsa, 2004b,a; Bojak and Liley, 2005; Phillips and Robinson, 2009; Rolls and Treves, 2011) to constructing artificial brains *in silico* and building brain-machine interface and neuro-prosthetic devices, e.g., Einevoll et al., 2013; Whalen et al., 2013. Biophysical modeling has always benefited from interdisciplinary interactions between different and seemingly distant fields; ranging from mathematics and engineering to linguistics and psychology. This Research Topic aims to promote such interactions by promoting papers that contribute to a deeper understanding of neural activity as measured by fMRI or electrophysiology.

In general, mean field models of neural activity can be divided into two classes: neural mass and neural field models. The main difference between these classes is that field models prescribe how a quantity characterizing neural activity (such as average depolarization of a neural population) evolves over *both space and time* as opposed to mass models, which characterize activity over time only; by assuming that all neurons in a population are located at (approximately) the same point. This Research Topic focusses on both classes of models and considers several aspects and their relative merits that: span from synapses to the whole brain; comparisons of their predictions with EEG and MEG spectra of spontaneous brain activity; evoked responses, seizures, and fitting data—to infer brain states and map physiological parameters.

EXTENSIONS OF MEAN FIELD MODELS AND MODELING OF ANAESTHETIC ACTION

Some of the contributions consider extensions of neural mass and field models and their relation with other classes of models, with a particular focus on modeling the action of anesthetics:

Liley and Walsh (2013) hypothesize that fast-slow dynamics, as exhibited in individual neuron bursting, dynamically underpins electroencephalographic bursting. They are able to modify a

well-known mean field model of the electroencephalogram by adding slow variables. This can be seen as a metaphor for anesthetic action, and allows them to produce a wide variety of burst-like activities. Bojak et al. (2013) look at quantitative modulations of EEG activity resulting from manipulating the anesthetics ketamine and propofol. They are able to determine parameter ranges that produce observed modulations in alpha peak frequency, and predict antagonistic drug interactions. The action of anesthetics, in the context of mean field models, is also discussed by Hutt (2013). The author considers a linear neural population model and presents an analytic derivation of the power spectrum that depends on propofol concentration. He then explains the anesthetic-induced power increase in neural activity as a result of an oscillatory instability and derives conditions under which the power peak shifts to higher frequencies, as observed experimentally in EEG.

The roles of neural mass, conductance based, and neural field models in dynamic causal modeling (DCM) are reviewed and explored by Moran et al. (2013). These authors show that such models can reproduce the characteristics of spectra and evoked responses observed empirically, with conductance based models having a richer repertoire of dynamics than neural mass models. Neural field models are able to capture lateral interactions and allow detailed analysis of structure-function relationships in the cortex.

Modolo et al. (2013) discuss neural masses designed to study the interaction between power-line magnetic fields and brain activity. They demonstrate that EEG alpha power could be modulated by weak membrane depolarization induced by the exposure to power-line magnetic fields and explore the role of input noise on EEG power modulation. A different use of neural fields is presented in Wright and Bourke (2013). These authors propose that both synchronous firing of neurons—and their competition for limited metabolic resources during neural development—lead to ultra-small-world neural networks. These networks then exhibit Möbius strip-like topologies that putatively reflect structure in striatal visual cortex.

The contribution of Pinotsis et al. (2013) introduces a conductance-based neural field model combining biologically

realistic synaptic dynamics with neural field equations. These authors demonstrate that both the evoked responses and induced responses show qualitative differences depending on the chosen model, either neural mass or neural field.

EXPLAINING ACTIVITY OBSERVED IN NEUROLOGICAL DISORDERS AND COGNITIVE TASKS

Other articles in this Research Topic relate to the use of field models to explain aberrant neural activity and dynamics recorded during cognitive tasks: Kerr et al. (2013) integrate field and network models in a multiscale model. This allows the authors to reveal alterations in cortical information flow between normal subjects and Parkinsonian patients, quantified by a decrease in Spectral Granger Causality between cortical layers in the beta frequency.

Frequency-dependent effects in deep brain stimulation in epileptic patients are studied using computational modeling and intracerebral EEG data in Mina et al. (2013). This paper describes the biophysics of direct stimulation of the thalamic compartment of an established thalamocortical model at the cellular level. It also demonstrates that low-frequency and high-frequency stimulation are beneficial for suppressing epileptic seizures, but that intermediate frequencies favor thalamic oscillations and entrain epileptic dynamics, rather than suppressing them.

Bhattacharya's paper (Bhattacharya, 2013) also focuses on explaining brain oscillations in sickness and health. The author replaces the "alpha function" approximation for synaptic transmission by a kinetic framework of neurotransmitter and receptor dynamics. The results are compared with experimental studies and shown to be consistent; they also lead to an order of magnitude improvement in simulation times compared to the alpha function approach commonly adopted in neural mass models.

In Srinivasan et al. (2013) the authors study an important phenomenon observed in EEG data, called phase-amplitude coupling, and show how it can be modeled using classical Wilson and Cowan equations. This is not only a mathematical exercise; it allows for a description of important top-down influences on local networks as a result of behavioral (e.g., attentional) or pharmacological manipulations—and fits well with results from the animal and human literature.

In another paper, Robinson et al. (2012) explore the functional neuroimaging measurements required to characterize neocortical activity. In particular, they show that some state changes can occur independently of changes in average amplitude, power, or metabolic indexes. They then introduce a new measure of complexity that can uncover the corresponding dynamical structure inherent in cortical activity, which would otherwise be difficult or impossible to detect.

Finally, beim Graben and Rodrigues (2012) reduce a simplified 3-compartment neuron model into a leaky integrate-and-fire (LIF) model describing spiking dynamics and derive an observation model for dendritic dipole currents in extracellular space that contributes to the local field potential (LFP) of a neural population. They introduce a new way to predict LFPs in network simulations involving only single-compartment neurons and compare their method with the results of an earlier approach (Mazzoni et al., 2008).

THEORY OF MEAN FIELD MODELS

In addition to papers focussing on applications, this Research Topic includes theoretical papers studying the mathematical aspects of mean field theory: Bressloff and Wilkerson (2012) study rigorous aspects of field models using an off-centered connectivity kernel that can serve as a model for direction selectivity. They prove the existence and stability of stimulus-induced activity pulses assuming a Heaviside firing rate function and including spatiotemporal noise. These authors conclude that freely moving pulses are more sensitive to multiplicative noise than stimulus-locked pulses.

In Gray and Robinson (2013), the authors address an important issue in the literature on neural networks; that is, what are the effects of time delays and dendritic time constants on the stability constraints of the network dynamics. They approach this question from the perspective of their prior work, in particular the Robinson, Rennie Wright model (RRW). Within this framework, they introduce a constant time delay and then systematically analyze the stability of a network state as a function of time delays and other parameters.

Roy and Jirsa (2013) show how a novel neurocomputational unit model qualitatively captures the complex dynamics exhibited by a full network of parabolic bursting neurons. The reduced representation is mathematically tractable and allows the authors to derive appropriate boundary conditions for various dynamical regimes. This approach sheds light on the role of slow oscillations for determining the global behavior of brain networks. Finally, Augustin et al. (2013) examine how the dynamics of adaptation currents contribute to spike rate oscillations in recurrent neural networks. They find frequency-dependent effects that can have roles in generation of specific frequencies and selective signal propagation.

The above anthology of papers provides illustrative examples of recent advances in biophysical modeling. This line of work speaks to the hope that such models may help explain neural dynamics that underpin disorders like epilepsy or Parkinson's disease as well as normal functions like attention or working memory; an endeavor we hope the articles in this volume will progress.

REFERENCES

- Augustin, M., Ladenbauer, J., and Obermayer, K. (2013). How adaptation shapes spike rate oscillations in recurrent neuronal networks. *Front. Comput. Neurosci.* 7:9. doi: 10.3389/fncom.2013.00009
- beim Graben, P., and Rodrigues, S. (2012). A biophysical observation model for field potentials of networks of leaky integrate-and-fire neurons. *Front. Comput. Neurosci.* 6:100. doi: 10.3389/fncom.2012.00100
- Bhattacharya, B. S. (2013). Implementing the cellular mechanisms of synaptic transmission in a neural mass model of the thalamo-cortical circuitry. *Front. Comput. Neurosci.* 7:81. doi: 10.3389/fncom.2013.00081
- Bojak, I., Day, H. C., and Liley, D. T. (2013). Ketamine, propofol, and the EEG: a neural field analysis of HCN1-mediated interactions. *Front. Comput. Neurosci.* 7:22. doi: 10.3389/fncom.2013.00022
- Bojak, I., and Liley, D. T. (2005). Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* 71:041902. doi: 10.1103/PhysRevE.71.041902
- Bressloff, P. C., and Wilkerson, J. (2012). Traveling pulses in a stochastic neural field model of direction selectivity. *Front. Comput. Neurosci.* 6:90. doi: 10.3389/fncom.2012.00090
- Coombes, S. (2010). Large-scale neural dynamics: simple and complex. *Neuroimage* 52, 731–739. doi: 10.1016/j.neuroimage.2010.01.045

- Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., and Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* 4:e1000092. doi: 10.1371/journal.pcbi.1000092
- Einevoll, G. T., Kayser, C., Logothetis, N. K., and Panzeri, S. (2013). Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nat. Rev. Neurosci.* 14, 770–785. doi: 10.1038/nrn3599
- Ermentrout, B. (1998). Neural networks as spatio-temporal pattern-forming systems. *Rep. Prog. Phys.* 61, 353. doi: 10.1088/0034-4885/61/4/002
- Gray, R. T., and Robinson, P. A. (2013). Stability constraints on large-scale structural brain networks. *Front. Comput. Neurosci.* 7:31. doi: 10.3389/fncom.2013.00031
- Guillory, S. A., and Bujarski, K. A. (2014). Exploring emotions using invasive methods: review of 60 years of human intracranial electrophysiology. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nsu002. [Epub ahead of print].
- Hutt, A. (2013). The anesthetic propofol shifts the frequency of maximum spectral power in EEG during general anesthesia: analytical insights from a linear model. *Front. Comput. Neurosci.* 7:2. doi: 10.3389/fncom.2013.00002
- Jirsa, V. K. (2004a). Connectivity and dynamics of neural information processing. *Neuroinformatics* 2, 183–204. doi: 10.1385/NI:2:2:183
- Jirsa, V. K. (2004b). Perceptual-cognitive control as a special case in equivalent multisensory-sensorimotor interactions. *J. Mot. Behav.* 36, 385–386.
- Kerr, C. C., Van Albada, S. J., Neymotin, S. A., Chadderdon, G. L., Robinson, P. A., and Lytton, W. W. (2013). Cortical information flow in Parkinson's disease: a composite network/field model. *Front. Comput. Neurosci.* 7:39. doi: 10.3389/fncom.2013.00039
- Liley, D. T., and Walsh, M. (2013). The mesoscopic modeling of burst suppression during anesthesia. *Front. Comput. Neurosci.* 7:46. doi: 10.3389/fncom.2013.00046
- Mazzoni, A., Panzeri, S., Logothetis, N. K., and Brunel, N. (2008). Encoding of naturalistic stimuli by local field potential spectra in networks of excitatory and inhibitory neurons. *PLoS Comput. Biol.* 4:e1000239. doi: 10.1371/journal.pcbi.1000239
- Mina, F., Benquet, P., Pasnicu, A., Biraben, A., and Wendling, F. (2013). Modulation of epileptic activity by deep brain stimulation: a model-based study of frequency-dependent effects. *Front. Comput. Neurosci.* 7:94. doi: 10.3389/fncom.2013.00094
- Modolo, J., Thomas, A. W., and Legros, A. (2013). Neural mass modeling of power-line magnetic fields effects on brain activity. *Front. Comput. Neurosci.* 7:34. doi: 10.3389/fncom.2013.00034
- Moran, R., Pinotsis, D. A., and Friston, K. (2013). Neural masses and fields in dynamic causal modeling. *Front. Comput. Neurosci.* 7:57. doi: 10.3389/fncom.2013.00057
- Phillips, A. J. K., and Robinson, P. A. (2009). Potential formulation of sleep dynamics. *Phys. Rev. E* 79:021913. doi: 10.1103/PhysRevE.79.021913
- Pinotsis, D., Leite, M., and Friston, K. (2013). On conductance-based neural field models. *Front. Comput. Neurosci.* 7:158. doi: 10.3389/fncom.2013.00158
- Robinson, S. E., Mandell, A. J., and Coppola, R. (2012). Spatiotemporal imaging of complexity. *Front. Comput. Neurosci.* 6:101. doi: 10.3389/fncom.2012.00101
- Rolls, E. T., and Treves, A. (2011). The neuronal encoding of information in the brain. *Prog. Neurobiol.* 95, 448–490. doi: 10.1016/j.pneurobio.2011.08.002
- Roy, D., and Jirsa, V. (2013). Inferring network properties of cortical neurons with synaptic coupling and parameter dispersion. *Front. Comput. Neurosci.* 7:20. doi: 10.3389/fncom.2013.00020
- Sporns, O. (2014). Contributions and challenges for network models in cognitive neuroscience. *Nat. Neurosci.* 17, 652–660. doi: 10.1038/nn.3690
- Srinivasan, R., Thorpe, S., and Nunez, P. L. (2013). Top-down influences on local networks: basic theory with experimental implications. *Front. Comput. Neurosci.* 7:29. doi: 10.3389/fncom.2013.00029
- Whalen, A. J., Brennan, S. N., Sauer, T. D., and Schiff, S. J. (2013). Observability and controllability of neuronal network motifs. *arXiv preprint arXiv.1307.5478*.
- Wright, J. J., and Bourke, P. D. (2013). On the dynamics of cortical development: synchrony and synaptic self-organization. *Front. Comput. Neurosci.* 7:4. doi: 10.3389/fncom.2013.00004

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 October 2014; accepted: 30 October 2014; published online: 17 November 2014.

Citation: Pinotsis D, Robinson P, beim Graben P and Friston K (2014) Neural masses and fields: modeling the dynamics of brain activity. *Front. Comput. Neurosci.* 8:149. doi: 10.3389/fncom.2014.00149

This article was submitted to the journal *Frontiers in Computational Neuroscience*. Copyright © 2014 Pinotsis, Robinson, beim Graben and Friston. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Traveling pulses in a stochastic neural field model of direction selectivity

Paul C. Bressloff* and Jeremy Wilkerson

Department of Mathematics, University of Utah, Salt Lake City, UT, USA

Edited by:

Dimitris Pinotsis, University College London, UK

Reviewed by:

Axel Hutt, INRIA CR Nancy, France
Carlo Laing, Massey University, New Zealand

***Correspondence:**

Paul C. Bressloff, Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, UT 84112, USA.
e-mail: bressloff@math.utah.edu

We analyze the effects of extrinsic noise on traveling pulses in a neural field model of direction selectivity. The model consists of a one-dimensional scalar neural field with an asymmetric weight distribution consisting of an offset Mexican hat function. We first show how, in the absence of any noise, the system supports spontaneously propagating traveling pulses that can lock to externally moving stimuli. Using a separation of time-scales and perturbation methods previously developed for stochastic reaction-diffusion equations, we then show how extrinsic noise in the activity variables leads to a diffusive-like displacement (wandering) of the wave from its uniformly translating position at long time-scales, and fluctuations in the wave profile around its instantaneous position at short time-scales. In the case of freely propagating pulses, the wandering is characterized by pure Brownian motion, whereas in the case of stimulus-locked pulses, it is given by an Ornstein-Uhlenbeck process. This establishes that stimulus-locked pulses are more robust to noise.

Keywords: stochastic processes, traveling waves, neural field theory, direction selectivity, stimulus-driven

INTRODUCTION

Continuum neural field models represent the large-scale dynamics of spatially structured networks of neurons in terms of non-linear integro-differential equations, whose associated integral kernels represent the spatial distribution of neuronal synaptic connections (Wilson and Cowan, 1972, 1973; Amari, 1977). As in the case of non-linear partial differential equation (PDE) models of diffusively coupled excitable systems (Keener, 1981; Kuramoto, 1984), non-local neural fields can exhibit a diverse range of spatiotemporal dynamics, including solitary traveling fronts and pulses, stationary pulses, and spatially localized oscillations (breathers), spiral waves, and Turing-like patterns. See, for example, the reviews Ermentrout (1998), Coombes (2005), and Bressloff (2012). In recent years, neural fields have been used to model a wide range of neurobiological phenomena, including wave propagation in cortical slices (Pinto and Ermentrout, 2001; Richardson et al., 2005) and *in vivo* (Huang et al., 2004), geometric visual hallucinations (Ermentrout and Cowan, 1979; Bressloff et al., 2001), EEG rhythms (Nunez, 1995; Robinson et al., 2001; Liley et al., 2002; Steyn-Ross et al., 2003), orientation tuning in primary visual cortex (V1) (Ben-Yishai et al., 1995; Somers et al., 1995), short term working memory (Camperi and Wang, 1998; Laing and Chow, 2002), control of head direction (Zhang, 1996), direction selectivity (Xie and Giese, 2002), motion perception (Giese, 1999), and binocular rivalry waves (Bressloff and Webber, 2012a). One particularly useful feature of neural fields is that analytical techniques for solving these integro-differential equations can be adapted from previous studies of non-linear PDEs. These include regular and singular perturbation methods, weakly non-linear analysis and pattern formation, symmetric bifurcation theory, Evans functions and wave stability, and

homogenization theory (Bressloff, 2012). In particular, we have recently shown how perturbation methods for studying fluctuating fronts in reaction-diffusion PDEs (Schimansky-Geier et al., 1983; de Pasquale et al., 1992; Armero et al., 1998; Sagues et al., 2007) can be extended to the problem of front propagation in stochastic neural fields (Bressloff and Webber, 2012b), and have used this to investigate the effects of noise on binocular rivalry waves (Webber and Bressloff, submitted). Such methods exploit a separation of time-scales in which there is a diffusive-like displacement (wandering) of the front from its uniformly translating position at long time-scales, and fluctuations in the front profile around its instantaneous position at short time-scales.

In this paper, we extend our theory of wave propagation in stochastic neural fields to the case of a neural field that supports traveling pulses rather than fronts. A typical mechanism for generating traveling pulses in an excitatory network is to include some form of slow adaptation, such as spike frequency adaptation (Pinto and Ermentrout, 2001; Coombes and Owen, 2005) or synaptic depression (Kilpatrick and Bressloff, 2010a,b), which suppresses the trailing edge of the wave. One of the motivations for considering excitatory neural fields is that traveling pulses are observed in *in vitro* cortical slices that have been disinhibited. Here we consider an alternative mechanism for generating pulses, based on asymmetric excitatory/inhibitory synaptic connections. Such a network architecture supports freely propagating pulses without any adaptation, and has been proposed as a simple recurrent mechanism for generating direction selectivity in a network driven by moving stimuli (Mineiro and Zipser, 1998; Xie and Giese, 2002). Most classical models for the direction selectivity of cortical neurons are based on feedforward mechanisms, namely, the linear or non-linear spatiotemporal filtering

of afferent thalamo-cortical inputs (Reichardt, 1961; Adelson and Bergen, 1985; Koch and Poggio, 1985; van Santen and Sperling, 1985). Some of these models also involve a combination of lagged (time-delayed) and non-lagged inputs (Saul and Humphrey, 1990; Baker and Bair, 2012). However, there is now considerable experimental data demonstrating that the response of cortical cells is strongly influenced by intracortical circuitry. This has motivated a number of modeling studies that show how direction selectivity can be reproduced by recurrent neural network models with asymmetric lateral excitatory or inhibitory connections and non-direction-selective inputs (Suarez et al., 1995; Maex and Urban, 1996; Mineiro and Zipser, 1998; Xie and Giese, 2002). In this paper, we base our investigation of stochastic traveling pulses on the particular version introduced by Xie and Giese (2002).

The main results of the paper are as follows. We first analyze freely propagating pulses and stimulus-locked pulses in the deterministic case, expanding the analysis of Xie and Giese (2002). In particular, we construct a stability diagram showing the existence and stability of stimulus-locked pulses as a function of stimulus velocity and amplitude. We then turn to a corresponding stochastic version of the model. We show how extrinsic noise in the activity variable leads to a diffusive-like displacement (wandering) of the wave from its uniformly translating position at long time-scales, and fluctuations in the wave profile around its instantaneous position at short time-scales. In the case of freely propagating pulses, the wandering is characterized by pure Brownian motion, whereas in the case of stimulus-locked pulses, it is given by an Ornstein–Uhlenbeck process. This establishes that stimulus-locked pulses are more robust to noise. One major difference between pulses and fronts is that, in principle, noise could significantly affect both the location (center-of-mass) and width of the pulse. We find that fluctuations in the width can be neglected in the case of freely propagating pulses, whereas the saturation of the mean-square displacement of the center-of-mass of the pulse for stimulus-locked pulses means that fluctuations in pulse width can no longer be ignored.

MATERIALS AND METHODS

NEURAL FIELD MODEL OF DIRECTION SELECTIVITY

In this paper we consider a scalar neural field equation of the form

$$\tau \frac{\partial u(x, t)}{\partial t} = -u(x, t) + \int_{-\infty}^{\infty} w(x - x') F(u(x', t)) dx' + h(x, t) \quad (1)$$

Here $u(x, t)$ is a measure of activity (current or voltage) within a local population of excitatory and inhibitory neurons at position $x \in \mathbb{R}$ and time t , τ is a membrane time constant (of order 10 msec), $w(x)$ denotes the spatial distribution of synaptic connections between local populations, $F(u)$ is a non-linear firing rate function and $h(x, t)$ is an external input. (We fix the time-scale by setting $\tau = 1$). F is usually taken to be a sigmoid function

$$F(u) = \frac{1}{1 + e^{-\gamma(u-\kappa)}} \quad (2)$$

with gain γ and threshold κ . In the high-gain limit $\gamma \rightarrow \infty$, this reduces to the Heaviside function

$$F(u) \rightarrow H(u - \kappa) = \begin{cases} 1 & \text{if } u > \kappa \\ 0 & \text{if } u \leq \kappa. \end{cases} \quad (3)$$

The function $w(x - x')$ represents the distribution of synaptic weights from the local population at x' to the population at x . Usually, w is taken to be a symmetric or even function such that $w(x) = w(-x)$. A common choice for the weight distribution is a “Mexican hat” function, with a center excitatory region surrounded by flanking inhibitory regions. As originally shown by Amari (1977), symmetric Mexican hat functions tend to support stationary activity “bumps.” Following Xie and Giese (2002), however, we will use an asymmetric Mexican hat function whose maximum is offset by an amount x_0 , that is $w(x - x_0) = w(-[x - x_0])$; the resulting neural field then supports freely propagating pulses that depend on the degree of offset. Note that such a choice should be contrasted with a symmetric function w with peaks offset from zero see e.g., (Hutt and Atay, 2005). In the case of exponential functions, w takes the form (see **Figure 1**)

$$w(x) = a_e e^{-\sigma_e |x - x_0|} - a_i e^{-\sigma_i |x - x_0|}, \quad (4)$$

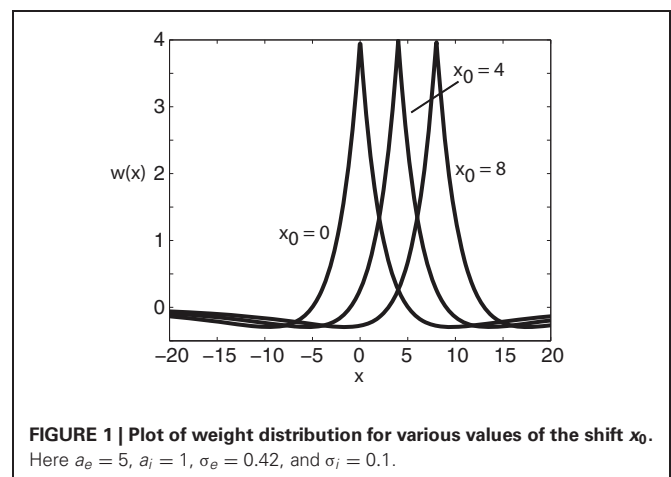
where $a_e > a_i$ and $\sigma_e > \sigma_i$. Setting $x_0 = 0$ recovers the standard Mexican hat function. Note that one could equally use other functions such as a difference-of-Gaussians without changing the main results of the paper; the advantage of exponentials is that one can carry out explicit calculations.

Finally, the external input $h(x, t)$ consists of two components:

$$h(x, t) = I(x - vt) + \sqrt{\epsilon} g(u(x, t)) \xi(x, t). \quad (5)$$

Here $I(x - vt)$ represents an external pulse-like stimulus moving with constant speed v and amplitude I_0 , whereas the second term represents an extrinsic, multiplicative noise source. In particular, $\xi(x, t)$ is a Gaussian process with zero mean and two-point correlations

$$\langle \eta(x, t) \eta(x', t') \rangle = 2C(|x - x'|/\lambda) \delta(t - t'). \quad (6)$$



Thus the noise is white in time and colored in space with correlation length λ . Formally speaking, $\eta(x, t)dt = dW(x, t)$ where $dW(x, t)$ is a corresponding Wiener process. The amplitude of the noise is determined by the parameter ϵ , and the function $g(u)$ incorporates any activity-dependence. Note that (Xie and Giese, 2002) only considered the deterministic case ($\epsilon = 0$). They showed how the deterministic neural field supports freely propagating pulses of fixed speed c when $I_0 = 0$. This then provides a mechanism for direction selectivity, since these pulses can lock to a moving stimulus of speed v provided that $|c - v|$ is sufficiently small; the range of locking depends on the amplitude I_0 . In this paper, we develop a more systematic analysis of stimulus-locking in the absence of noise, and then investigate the effects of noise on both freely propagating and stimulus-locked pulses.

RESULTS

DETERMINISTIC NEURAL FIELD

We begin by analyzing traveling pulse solutions of the neural field Equation (1) in the absence of noise ($\epsilon = 0$). Following the original formulation of Amari (1977), we investigate the existence and stability of traveling pulses by setting the firing rate function to be the Heaviside (Equation 3).

Freely propagating pulses

For the moment, suppose that there are no external inputs so that $h(x, t) = 0$ in Equation (1). A traveling pulse of velocity c is then defined according to $u(x, t) = U(\xi)$, with $\xi = x - ct$ a traveling wave coordinate such that $\lim_{\xi \rightarrow \pm\infty} U(\xi) = 0$. Moreover, the wave profile is restricted to be super threshold in a connected interval of width d . Since the neural field is equivariant with respect to uniform translations (in the absence of external stimuli), we choose the two threshold crossing points to be

$$U(0) = \kappa, \quad U(d) = \kappa. \quad (7)$$

Thus, $U(\xi) > \kappa$ for $0 < \xi < d$, $U(\xi) < \kappa$ for $\xi < 0$, and $\xi > d$. It turns out the wave travels in the same direction as the offset so we restrict ourselves to the case $x_0 > 0$ and $c > 0$. Substituting the traveling pulse solution into Equation (1) gives

$$-c \frac{\partial U(\xi)}{\partial \xi} = -U(\xi) + \int_0^d w(\xi - \xi') d\xi' \quad (8)$$

Multiplying both sides by $e^{-\xi/c}$ and integrating gives the following equation for the wave solution:

$$U(\xi) = \frac{e^{\xi/c}}{c} \int_{\xi}^{\infty} W(\xi') e^{-\xi'/c} d\xi', \quad (9)$$

where

$$W(\xi) \equiv \int_{\xi-d}^{\xi} w(x) dx.$$

It is convenient to express the weight function in piecewise form as follows:

$$w(x) = \begin{cases} a_e e^{-\sigma_e(x-x_0)} - a_i e^{-\sigma_i(x-x_0)} \equiv w_1(x), & \text{if } x \geq x_0 \\ a_e e^{\sigma_e(x-x_0)} - a_i e^{\sigma_i(x-x_0)} \equiv w_2(x), & \text{if } x \leq x_0 \end{cases} \quad (10)$$

We then obtain a piecewise expression for $W(\xi)$ of the form

$$W(\xi) = \begin{cases} W_3(\xi) \equiv \int_{\xi-d}^{\xi} w_2(x) dx, & \text{if } \xi \leq x_0 \\ W_2(\xi) \equiv \int_{\xi-d}^{x_0} w_2(x) dx + \int_{x_0}^{\xi} w_1(x) dx, & \text{if } x_0 \leq \xi \leq x_0 + d \\ W_1(\xi) \equiv \int_{\xi-d}^{\xi} w_1(x) dx, & \text{if } \xi \geq x_0 + d \end{cases} \quad (11)$$

We then have

$$U(\xi) = \begin{cases} \frac{1}{c} e^{\xi/c} (M_3(\xi) + M_1(x_0 + d) + M_2(x_0)), & \text{if } \xi \leq x_0 \\ \frac{1}{c} e^{\xi/c} (M_2(\xi) + M_1(x_0 + d)), & \text{if } x_0 \leq \xi \leq x_0 + d \\ \frac{1}{c} e^{\xi/c} M_1(\xi), & \text{if } \xi \geq x_0 + d \end{cases}$$

where

$$M_n(\xi) = \int_{\xi}^{\xi_n} W_n(\xi') e^{-\xi'/c} d\xi' \quad (12)$$

with $\xi_1 = \infty$, $\xi_2 = x_0 + d$ and $\xi_3 = x_0$.

Having obtained the piecewise wave profile $U(\xi)$, the threshold conditions (Equation 7) can now be used to determine the pulse speed c and width d ; the resulting transcendental equations have to be solved numerically. **Figure 2** shows solutions for the pulse speed and width as functions of the threshold. It turns out that the solution with slower speed (and larger width)

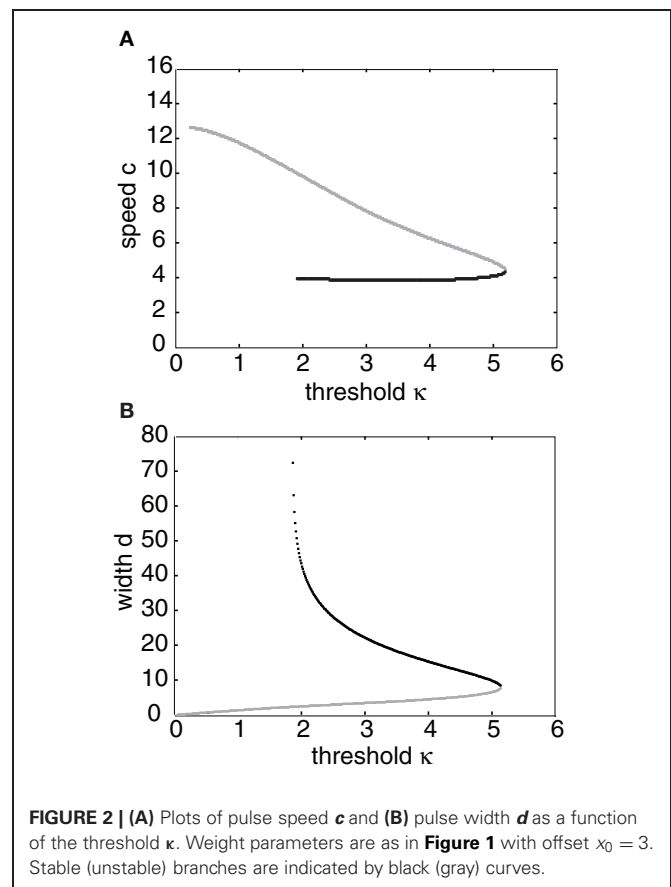


FIGURE 2 | (A) Plots of pulse speed c and **(B)** pulse width d as a function of the threshold κ . Weight parameters are as in **Figure 1** with offset $x_0 = 3$. Stable (unstable) branches are indicated by black (gray) curves.

is stable (see below). This differs from traveling pulse solutions found in adaptive neural fields, where the faster wave (with larger width) tends to be stable (Pinto and Ermentrout, 2001; Kilpatrick and Bressloff, 2010a). **Figure 3A** shows a typical pulse waveform and **Figure 3B** shows a numerical simulation of the neural field Equation (1) using the wave solution as the initial condition. The pulse propagates at the predicted speed without changing shape significantly. This occurs because the parameters were chosen to make the pulse solution linearly stable.

Stability

In order to determine the linear stability of a traveling pulse solution $U(\xi)$ in the moving frame, we linearize Equation (1) with $h(x, t) = 0$ by setting

$$U(\xi, t) = U(\xi) + \varphi(\xi, t),$$

and Taylor expanding to first order in φ . This gives

$$\begin{aligned} \frac{\partial \varphi(\xi, t)}{\partial t} = & c \frac{\partial \varphi(\xi, t)}{\partial \xi} - \varphi(\xi, t) + \int_{-\infty}^{\infty} w(\xi - y) \\ & \times F'(U(y)) \varphi(y, t) dy. \end{aligned} \quad (13)$$

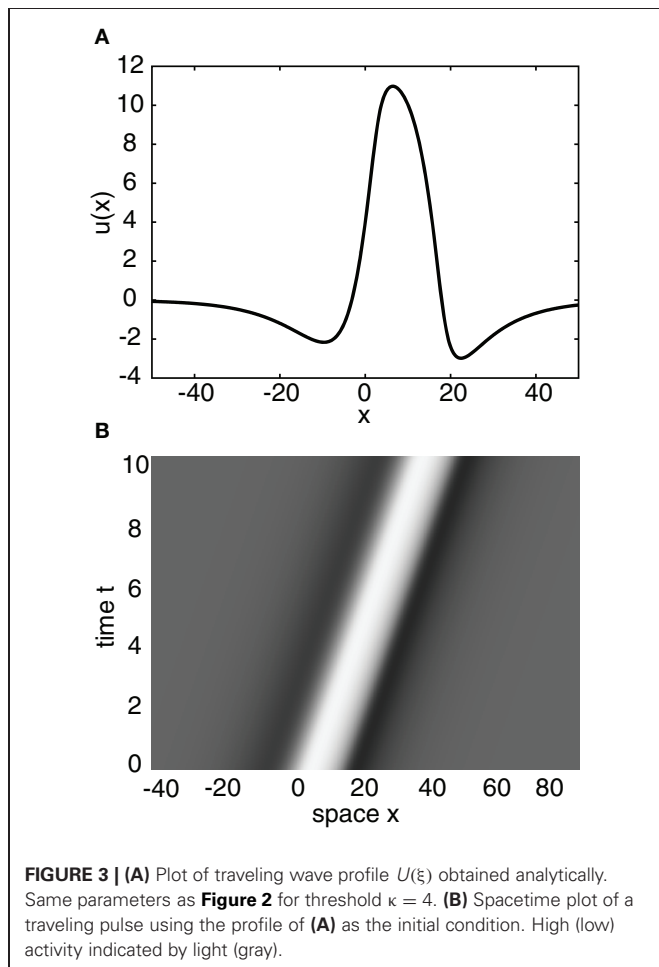


FIGURE 3 | (A) Plot of traveling wave profile $U(\xi)$ obtained analytically. Same parameters as **Figure 2** for threshold $\kappa = 4$. **(B)** Spacetime plot of a traveling pulse using the profile of **(A)** as the initial condition. High (low) activity indicated by light (gray).

In the case of the Heaviside rate function (Equation 3), we have

$$F'(U(\xi)) = \frac{\delta(\xi)}{|U'(0)|} + \frac{\delta(\xi - d)}{|U'(d)|}. \quad (14)$$

Moreover, differentiating Equation (9) with respect to ξ shows that

$$U'(\xi) = \frac{1}{c}(U(\xi) - W(\xi)).$$

Substituting the previous two results into Equation (13) gives

$$\begin{aligned} \frac{\partial \varphi(\xi, t)}{\partial t} = & \mathcal{L}\varphi(\xi, t) \\ \equiv & c \frac{\partial \varphi(\xi, t)}{\partial \xi} - \varphi(\xi, t) + \frac{c\varphi(0, t)}{|\kappa - W(0)|}w(\xi) \\ & + \frac{c\varphi(d, t)}{|\kappa - W(d)|}w(\xi - d). \end{aligned} \quad (15)$$

where κ is the threshold. Looking for solutions of the form

$$\varphi(\xi, t) = e^{\lambda t} \varphi(\xi). \quad (16)$$

then leads to the spectral problem

$$\mathcal{L}\varphi(\xi, t) = \lambda \varphi(\xi, t). \quad (17)$$

We take the linear operator \mathcal{L} to act on a Banach space \mathcal{B} of continuous, bounded functions $\psi(\xi)$ that are defined for $\xi \in \mathbb{R}$, and that decay exponentially as $\xi \rightarrow \pm\infty$. Let $\sigma(\mathcal{L})$ denote the spectrum of the linear operator \mathcal{L} , and define the associated resolvent operator according to $\mathcal{R}_\lambda \equiv (\mathcal{L} - \lambda I)^{-1}$, where I is the identity operator. The spectrum can be defined as those values of λ for which $\mathcal{T}_\lambda \equiv \mathcal{L} - \lambda I$ is not bijective. The spectrum is composed of three disjoint sets, the point or discrete spectrum, the residual spectrum, and the continuous spectrum. The point spectrum is defined as the values of λ (eigenvalues) for which the resolvent does not exist. The residual spectrum are the spectral values for which the resolvent exists but is not defined on a dense subset of \mathcal{B} . The continuous spectrum are the spectral values for which the resolvent exists and is densely defined but is unbounded (Kreyszig, 1978). Given these definitions, the traveling pulse is said to be linearly stable if (1) $\text{Re}(\lambda) < 0$ for all $\lambda \in \sigma(\mathcal{L})$, $\lambda \neq 0$ and (2) the zero eigenvalue is simple. The existence of a zero eigenvalue with corresponding eigenfunction $\varphi(\xi, t) = U'(\xi)$ reflects translation invariance, and immediately follows from differentiating (Equation 9) with respect to ξ .

We first consider the discrete spectrum by solving the eigenvalue equation

$$\frac{d\psi(\xi)}{d\xi} - \frac{\lambda + 1}{c}\psi(\xi) + K_0\psi(0)w(\xi) + K_d\psi(d)w(\xi - d) = 0,$$

where we have introduced the constants

$$K_0 = \frac{1}{|\kappa - W(0)|} \text{ and } K_d = \frac{1}{|\kappa - W(d)|}.$$

Multiplying both sides by $e^{-(\lambda+1)\xi/c}$ and integrating gives

$$\begin{aligned}\psi(\xi) = & K_0\psi(0) \int_{\xi}^{\infty} w(y)e^{(\lambda+1)(\xi-y)/c} dy \\ & + K_d\psi(d) \int_{\xi-d}^{\infty} w(y)e^{(\lambda+1)(\xi-d-y)/c} dy,\end{aligned}$$

which can be rewritten in the more compact form

$$\psi(\xi) = K_0\psi(0)(w \times P_{\lambda})(\xi) + K_d\psi(d)(w \times P_{\lambda})(\xi - d) \quad (18)$$

with

$$(w \times P_{\lambda})(\xi) = \int_{-\infty}^{\infty} w(y)P_{\lambda}(\xi - y) dy, \quad P_{\lambda}(\xi) = H(-\xi)e^{(\lambda+1)\xi/c} \quad (19)$$

The eigenvalues are now determined by imposing self-consistency at $\xi = 0$ and $\xi = d$. Setting $\xi = 0$ and $\xi = d$ in Equation (18) leads to the vector equation

$$\begin{bmatrix} K_0(w \times P_{\lambda})(0) - 1 & K_d(w \times P_{\lambda})(-d) \\ K_0(w \times P_{\lambda})(d) & K_d(w \times P_{\lambda})(0) - 1 \end{bmatrix} \begin{bmatrix} \psi(0) \\ \psi(d) \end{bmatrix} = 0 \quad (20)$$

This has a non-trivial solution if and only if the determinant of the matrix is zero. The determinant expressed as a function of λ , $\mathcal{E}(\lambda)$, is a complex analytic function known as the Evans function:

$$\begin{aligned}\mathcal{E}(\lambda) = & [K_0(w \times P_{\lambda})(0) - 1] [K_d(w \times P_{\lambda})(0) - 1] \\ & - K_0K_d(w \times P_{\lambda})(d)(w \times P_{\lambda})(-d).\end{aligned} \quad (21)$$

Thus, the zeros of the Evans function determine the discrete spectrum of the linear operator formed by linearizing the neural field equation about the pulse solution. Evans functions were originally introduced within the context of the stability of solitary pulses in diffusive Hodgkin–Huxley type equations describing action potential propagation in nerve axons (Evans, 1975). Since then the Evans function construction has been extended to a wide range of PDEs, see the review (Sandstede, 2002). It has also recently been applied to neural field equations (Zhang, 2003; Coombes and Owen, 2004; Rubin, 2004; Folias and Bressloff, 2005; Pinto et al., 2005; Sandstede, 2007) and more general non-local problems (Kapitula et al., 2004). An example plot of the real and imaginary parts of $\mathcal{E}(\lambda) = 0$ on the complex plane is shown in **Figure 4**. It can be seen that there is a zero eigenvalue and one negative real eigenvalue, indicating that the corresponding traveling pulse is linearly stable.

To find the essential spectrum, which is the union of the residual and continuous spectra, we will derive an explicit expression for the resolvent \mathcal{R}_{λ} . We start by writing an inhomogeneous equation of the form $\mathcal{T}_{\lambda}\psi(\xi) = h(\xi)$, where $h(\xi)$ represents a general function from the Banach space \mathcal{B} . This can be manipulated as before to give

$$\begin{aligned}\psi(\xi) = & -(h \times P_{\lambda})(\xi) + K_0\psi(0)(w \times P_{\lambda})(\xi) \\ & + K_d\psi(d)(w \times P_{\lambda})(\xi - d)\end{aligned} \quad (22)$$

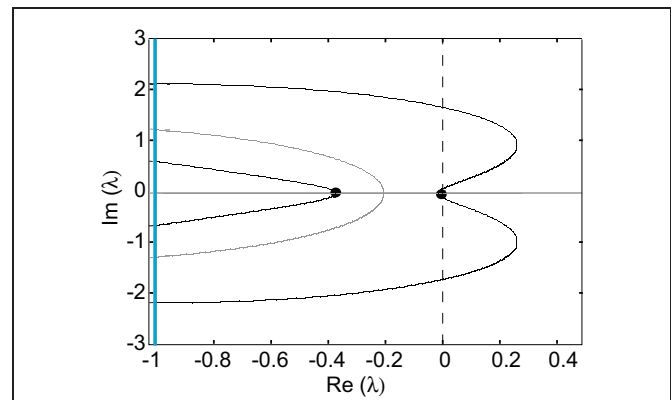


FIGURE 4 | Graphs of the zero sets of the real (dark curves) and imaginary (light curves) parts of the Evans function determining the stability of a freely propagating pulse; intersection points (filled circles) indicate eigenvalues. The line $\text{Im } \lambda = -1$ indicates the essential spectrum. Same parameter values as **Figure 3**.

By evaluating at $\xi = 0$ and $\xi = d$ as before, we arrive at the following vector equation, which differs from (Equation 20) only on the right-hand side:

$$\begin{bmatrix} K_0(w \times P_{\lambda})(0) - 1 & K_d(w \times P_{\lambda})(-d) \\ K_0(w \times P_{\lambda})(d) & K_d(w \times P_{\lambda})(0) - 1 \end{bmatrix} \begin{bmatrix} \psi(0) \\ \psi(d) \end{bmatrix} = \begin{bmatrix} (h \times P_{\lambda})(0) \\ (h \times P_{\lambda})(d) \end{bmatrix}.$$

Since we are looking for spectral values outside the discrete spectrum, the determinant of the matrix satisfies $\mathcal{E}(\lambda) \neq 0$. Therefore, multiplying both sides by the inverse matrix yields expressions for $\psi(0)$ and $\psi(d)$ in terms of h :

$$\psi(0) = \frac{S_0 h}{\mathcal{E}(\lambda)}, \quad \psi(d) = \frac{S_d h}{\mathcal{E}(\lambda)}$$

where

$$\begin{aligned}S_0 h = & (K_d(w \times P_{\lambda})(0) - 1) (h \times P_{\lambda})(0) \\ & - K_d(w \times P_{\lambda})(-d) (h \times P_{\lambda})(d)\end{aligned}$$

and

$$\begin{aligned}S_d h = & -K_0(w \times P_{\lambda})(d) (h \times P_{\lambda})(0) \\ & + (K_0(w \times P_{\lambda})(0) - 1) (h \times P_{\lambda})(d).\end{aligned}$$

Substituting into Equation (22) gives the following expression for the resolvent operator, of the form $\mathcal{R}_{\lambda} h = \varphi$:

$$\begin{aligned}- (h \times P_{\lambda})(\xi) + & \frac{K_0 S_0 h}{\mathcal{E}(\lambda)} (w \times P_{\lambda})(\xi) \\ & + \frac{K_d S_d h}{\mathcal{E}(\lambda)} (w \times P_{\lambda})(\xi - d) = \varphi(\xi)\end{aligned} \quad (23)$$

The resolvent is well-defined for all h in \mathcal{B} , so the residual spectrum of \mathcal{L} is empty. To find the continuous spectrum, we Fourier transform Equation (23):

$$-\hat{h}(k)\hat{P}_\lambda(k) + \frac{K_0 S_0 h}{\varepsilon(\lambda)} \hat{w}(k)\hat{P}_\lambda(k) + \frac{K_d S_d h}{\varepsilon(\lambda)} \hat{w}(k)\hat{P}_\lambda(k) e^{-2\pi i d k} = \hat{\psi}(k) \quad (24)$$

It follows that the resolvent operator is unbounded when \hat{P}_λ is unbounded. Equation (19) implies that

$$\hat{P}_\lambda(k) = \int_{-\infty}^{\infty} P_\lambda(\xi) e^{2\pi i k \xi} d\xi = \frac{1}{\frac{\lambda+1}{c} + 2\pi i k}. \quad (25)$$

Hence, \hat{P}_λ is unbounded for $\lambda = -1 - 2\pi i k c$ so that the continuous spectrum of \mathcal{L} is a vertical line in the complex plane at $\text{Re}(\lambda) = -1$. Since $\text{Re}(\lambda) < 0$, the continuous spectrum will not make any pulse solution of our model unstable.

Stimulus-locked pulses

Now suppose that the neural field is driven by a moving external pulse stimulus of speed v so that Equation (1) becomes

$$\frac{\partial u(x, t)}{\partial t} = -u(x, t) + \int_{-\infty}^{\infty} w(x-y) H(u(y, t) - \kappa) dy + I(x - vt). \quad (26)$$

In order to study the existence of stimulus-locked pulses, we will define a “stimulus coordinate” $\xi = x - vt$ and look for pulse solutions that move at the same speed as the stimulus, that is, $u(x, t) = U(\xi)$ with

$$-v \frac{\partial U(\xi)}{\partial \xi} = -U(\xi) + \int_{-\infty}^{\infty} w(\xi - y) H(U(y) - \kappa) dy + I(\xi). \quad (27)$$

For concreteness, the stimulus will be represented by a rectangular wave of amplitude I_0 and width d , defined formally as

$$I(\xi) = \begin{cases} I_0, & \text{if } 0 \leq \xi \leq d \\ 0, & \text{if } \xi < 0 \text{ or } \xi > d. \end{cases}$$

Since translation invariance no longer holds, it is necessary to determine both threshold crossing points, which we denote by $\xi = d_1$ and $\xi = d_2$. Proceeding in a similar fashion to the case of freely propagating pulses, we find that

$$U(\xi) = \frac{e^{\xi/v}}{v} \int_{\xi}^{z_0} e^{-y/v} W(y) dy + \frac{e^{\xi/v}}{v} \int_{\xi}^{z_0} e^{-y/v} I(y) dy,$$

where $z_0 = \infty$ if $v > 0$, $z_0 = -\infty$ if $v < 0$, and

$$W(\xi) \equiv \int_{\xi-d_2}^{\xi-d_1} w(x) dx.$$

The latter can be expressed in the piecewise form

$$W(\xi) = \begin{cases} W_3(\xi) \equiv \int_{\xi-d_2}^{\xi-d_1} w_2(x) dx, & \text{if } \xi \leq x_0 + d_1 \\ W_2(\xi) \equiv \int_{\xi-d_2}^{x_0} w_2(x) dx \\ + \int_{x_0}^{\xi-d_1} w_1(x) dx, & \text{if } x_0 + d_1 \leq \xi \leq x_0 + d_2 \\ W_1(\xi) \equiv \int_{\xi-d_2}^{\xi-d_1} w_1(x) dx, & \text{if } \xi \geq x_0 + d_2 \end{cases} \quad (28)$$

where w_1 and w_2 are defined as in Equation (10). After evaluating the integrals along similar lines to section “Neural Field Model of Direction Selectivity,” we obtain the following expressions for the pulse solution, defined independently for positive and negative stimulus directions:

$v > 0$:

$$U(\xi) = \begin{cases} U_3(\xi), & \text{if } \xi \leq x_0 + d_1 \\ U_2(\xi), & \text{if } x_0 + d_1 \leq \xi \leq x_0 + d_2, \\ U_1(\xi), & \text{if } \xi \geq x_0 + d_2, \end{cases}$$

with

$$U_3(\xi) = \frac{1}{v} e^{\xi/v} (M_3(\xi) + M_1(x_0 + d_2) + M_2(x_0 + d_1)) + Z(\xi)$$

$$U_2(\xi) = \frac{1}{v} e^{\xi/v} (M_2(\xi) + M_1(x_0 + d_2)) + Z(\xi)$$

$$U_1(\xi) = \frac{1}{v} e^{\xi/v} M_1(\xi) + Z(\xi),$$

$$M_n(\xi) = \int_{\xi}^{\xi_n} W_n(\xi') e^{-\xi'/v} d\xi'$$

for $\xi_1 = \infty$, $\xi_2 = x_0 + d_2$, $\xi_3 = x_0 + d_1$, and

$$Z(\xi) = \begin{cases} (e^{\xi/v} - e^{(\xi-d)/v}) I_0, & \text{if } \xi < 0 \\ (1 - e^{(\xi-d)/v}) I_0, & \text{if } 0 \leq \xi \leq d. \\ 0, & \text{if } \xi > d \end{cases}$$

$v < 0$:

$$U(\xi) = \begin{cases} U_3(\xi), & \text{if } \xi \leq x_0 + d_1 \\ U_2(\xi), & \text{if } x_0 + d_1 \leq \xi \leq x_0 + d_2, \\ U_1(\xi), & \text{if } \xi \geq x_0 + d_2 \end{cases}$$

with

$$U_3(\xi) = -\frac{1}{v} e^{\xi/v} N_3(\xi) + Z(\xi)$$

$$U_2(\xi) = -\frac{1}{v} e^{\xi/v} (N_2(\xi) + N_3(x_0 + d_1)) + Z(\xi)$$

$$U_1(\xi) = -\frac{1}{v} e^{\xi/v} (N_1(\xi) + N_2(x_0 + d_1) + N_3(x_0 + d_1)) + Z(\xi)$$

$$N_n(\xi) = \int_{\xi_n}^{\xi} W_n(\xi') e^{-\xi'/v} d\xi'$$

for $\xi_3 = -\infty$, $\xi_2 = x_0 + d_1$, $\xi_1 = x_0 + d_2$, and

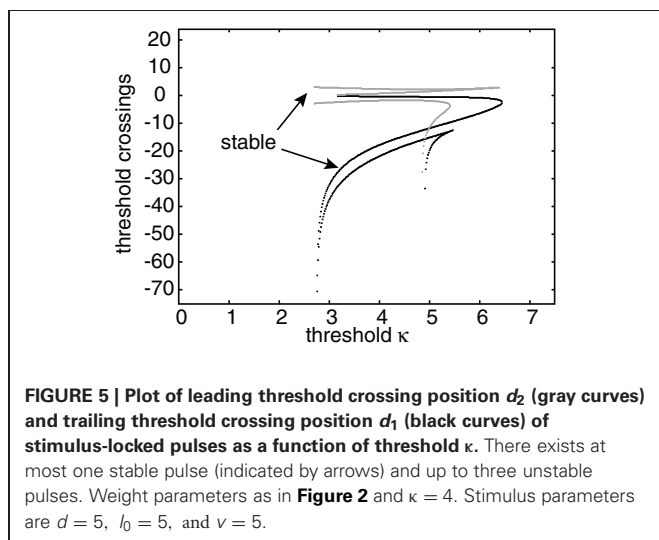
$$Z(\xi) = \begin{cases} 0, & \text{if } \xi < 0 \\ (1 - e^{\xi/\nu}) I_0, & \text{if } 0 \leq \xi \leq d \\ (e^{(\xi-d)/\nu} - e^{\xi/\nu}) I_0, & \text{if } \xi > d \end{cases}$$

The threshold crossing points (d_1 and d_2) are determined in the same way the pulse speed and width were determined in the no-stimulus case, which is by numerically solving a system of two transcendental equations. The first equation is given by $U_3(d_1) = \kappa$. The second equation is $U_3(d_2) = \kappa$ if $d_2 < x_0 + d_1$, else it is given by $U_2(d_2) = \kappa$. **Figure 5** shows a plot of d_1 (black curves) and d_2 (gray curves) vs. the threshold κ . It can be seen that for a certain range of thresholds there exists more than one stable/unstable pair of pulses. **Figure 6** shows the linear stability and the number of solutions for different combinations of stimulus speed (ν) and strength (I_0). The offset $x_0 = 3$ and the corresponding spontaneous wave speed is $c = 4$. (Note that for smaller offsets x_0 and thus smaller wave speeds c , one finds stimulus-locked waves for negative values of ν). The stability of solutions in the presence of a stimulus is determined in much the same way as without a stimulus. We again define $u(x, t) = U(\xi) + \varphi(\xi, t)$ and look at the behavior of the perturbations described by $\varphi(\xi, t)$. Substituting into Equation (26), the stimulus term drops out when we perform the linearization, so that

$$\frac{\partial \varphi(\xi, t)}{\partial t} = \nu \frac{\partial \varphi(\xi, t)}{\partial \xi} - \varphi(\xi, t) + \int_{-\infty}^{\infty} w(\xi - y) \times F'(U(y)) \varphi(y, t) dy.$$

Setting $\varphi(\xi, t) = e^{\lambda t} \varphi$ ultimately yields the spectral problem

$$\begin{aligned} \lambda \varphi(\xi) &\equiv \mathcal{L} \varphi(\xi) \\ &= \nu \frac{\partial \varphi(\xi, t)}{\partial \xi} - \varphi(\xi, t) \end{aligned} \quad (29)$$



$$\begin{aligned} &+ \frac{|v| \varphi(d_1)}{|\kappa - W(d_1) - I(d_1)|} w(\xi - d_1) \\ &+ \frac{|v| \varphi(d_2)}{|\kappa - W(d_2) - I(d_2)|} w(\xi - d_2). \end{aligned}$$

The corresponding Evans function is now

$$\mathcal{E}(\lambda) = [K_1(w \times P_\lambda)(0) - 1] [K_2(w \times P_\lambda)(0) - 1] - K_1 K_2(w \times P_\lambda)(d_2 - d_1)(w \times P_\lambda)(d_1 - d_2), \quad (30)$$

where

$$K_n = \frac{\text{sgn}(\nu)}{|\kappa - W(d_n) - I(d_n)|}, \quad n = 1, 2,$$

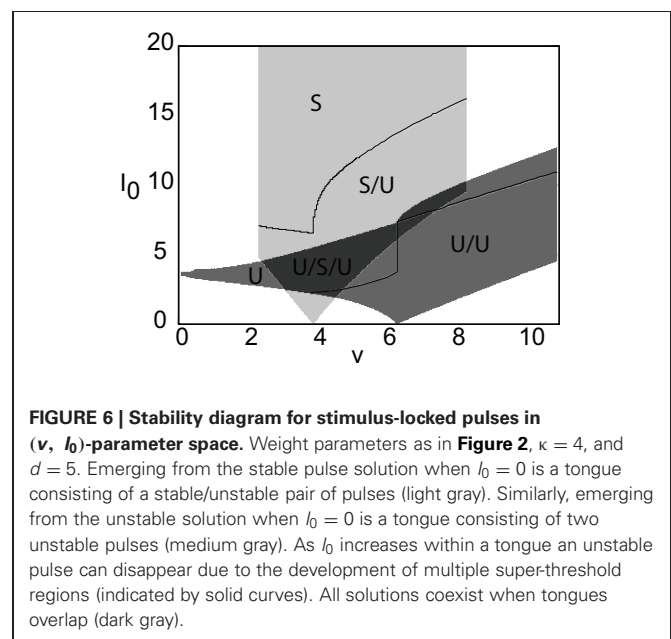
and

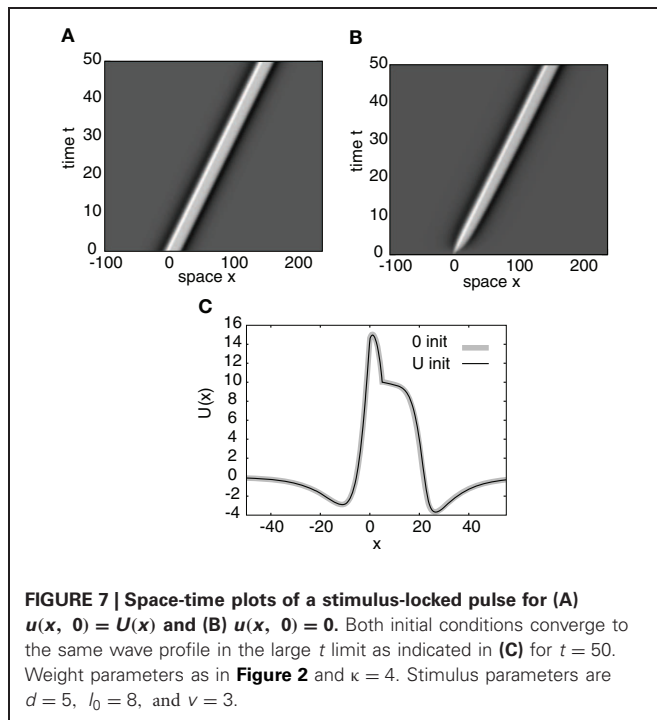
$$P_\lambda(\xi) = H(-\text{sgn}(\nu)\xi) e^{(\lambda+1)\xi/\nu}.$$

It is easy to establish as before that the residual spectrum is empty and that the continuous spectrum consists of a vertical line in the complex plane at $\text{Re}(\lambda) = -1$. So the stability is again determined only by the discrete spectrum, which consists of the zeros of the Evans function. **Figure 7A** shows an example of a numerical simulation of a stable stimulus-locked pulse solution of Equation (26) with the analytical pulse solution $U(x)$ as an initial condition. **Figure 7B** shows the same simulation except with the zero initial condition $u(x, 0) = 0$. It can be seen from **Figure 7C** that both initial conditions converge to the same pulse profile.

STOCHASTIC NEURAL FIELD

Several recent studies have considered stochastic versions of neural field equations that are based on a corresponding Langevin





equation formulation (Brackley and Turner, 2007; Hutt et al., 2008; Faugeras et al., 2009; Bressloff and Webber, 2012b). Motivated by these examples, we consider the following Langevin equation (or stochastic PDE) for the stochastic activity variable $U(x, t)$, which is a rewriting of Equation (1) with $h(x, t)$ given by Equation (5) for $l_0 = 0$ and $\epsilon > 0$:

$$dU(x, t) = \left[-U(x, t) + \int_{-\infty}^{\infty} w(x-y)F(U(y, t))dy \right] dt + \epsilon^{1/2}g(U(x, t))dW(x, t), \quad (31)$$

where $dW(x, t)$ is an independent Wiener process with zero mean and correlation given by

$$\langle dW(x, t)dW(x', t') \rangle = 2C(|x - x'|/\lambda)\delta(t - t')dt dt'. \quad (32)$$

Here λ is the spatial correlation length of the noise such that $C(x/\lambda) \rightarrow \delta(x)$ in the limit $\lambda \rightarrow 0$, and ϵ determines the strength of the noise, which is assumed to be weak. For the sake of generality, we take the noise to be multiplicative rather than additive; however, the main results of the paper hold for both. Following standard formulations of Langevin equations (Gardiner, 2009), the multiplicative noise term is taken to be of Stratonovich form in the case of extrinsic noise. Note, however, that an alternative formulation of stochastic neural field theory has been developed in terms of a neural master equation (Buice and Cowan, 2007; Bressloff, 2009, 2010; Buice et al., 2010), in which the underlying deterministic equations are recovered in the thermodynamic limit $N \rightarrow \infty$, where N is a measure of the system size of each local population. In the case of large but finite N , a Kramers-Moyal expansion of the master equation yields a Langevin neural

field equation with multiplicative noise of the Ito form Bressloff (2009, 2010). Multiplicative noise in the Stratonovich sense causes a shift in the speed and width of the pulse. This happens because $\langle g(U)dW \rangle \neq 0$, even though $\langle dW \rangle = 0$. We can use Novikov's theorem (Novikov, 1965) to calculate the former average:

$$\epsilon^{1/2}\langle g(U)dW \rangle = \epsilon C(0)\langle g'(U)g(U) \rangle dt.$$

The average can also be calculated by Fourier transforming Equation (31) and taking averages using the corresponding Fokker-Planck equation (Armero et al., 1998; Bressloff and Webber, 2012b). In the limit that λ approaches 0, we set $C(0) \rightarrow 1/\Delta x$, where Δx is a lattice cut-off that can be identified with the spatial discretization step size in numerical simulations (Bressloff and Webber, 2012b). Following Ref. Armero et al. (1998), we rewrite Equation (31) so the fluctuating term has zero mean:

$$dU(x, t) = \left[h(U(x, t)) + \int_{-\infty}^{\infty} w(x-y)F(U(y, t))dy \right] dt + \epsilon^{1/2}dR(U, x, t), \quad (33)$$

where

$$h(U) = -U + \epsilon C(0)g'(U)g(U) \quad (34)$$

and

$$dR(U, x, t) = g(U)dW(x, t) - \epsilon^{1/2}C(0)g'(U)g(U)dt. \quad (35)$$

The stochastic process R has zero mean and correlation

$$\begin{aligned} \langle dR(U, x, t)dR(U, x', t') \rangle \\ = \langle g(U(x, t))dW(x, t)g(U(x', t'))dW(x', t') \rangle + \mathcal{O}(\epsilon^{1/2}). \end{aligned} \quad (36)$$

Separation of time-scales

The effects of additive or multiplicative extrinsic noise on traveling waves can be analyzed using multiple time-scale methods originally developed for reaction-diffusion equations (Schimansky-Geier et al., 1983; de Pasquale et al., 1992; Armero et al., 1998; Sagues et al., 2007), which were recently extended to neural field equations in Ref. Bressloff and Webber (2012b). The main idea is to assume that the fluctuating term generates two distinct phenomena that occur on different time-scales: a diffusive-like displacement of the traveling wave from its uniformly translating position at long time-scales, and fluctuations in the wave profile around its instantaneous position at short time-scales. It is important to point out that, in contrast to traveling front solutions of scalar neural field equations (Bressloff and Webber, 2012b), we are now considering traveling pulse solutions. Thus in addition to the center-of-mass of the traveling pulse wave, which moves with speed c in the absence of noise, there is an additional degree of freedom corresponding to the “width” of the pulse. (In the case of a Heaviside rate function, the width Δ is determined by the threshold crossing points). For simplicity, we assume that the width of the wave is only weakly affected by the noise; this is consistent with what is found

numerically. We now express the solution U of Equation (33) as a combination of a fixed wave profile U_0 that is displaced by an amount $\Delta(t)$ from its uniformly translating position $\xi = x - c_\epsilon t$, where c_ϵ is a noise-dependent speed, and a time-dependent fluctuation Φ in the wave shape about its instantaneous position:

$$U(x, t) = U_0(\xi - \Delta(t)) + \epsilon^{1/2} \Phi(\xi - \Delta(t), t). \quad (37)$$

The wave profile U_0 and associated wave speed/width c_ϵ , Δ_ϵ are obtained by solving the modified deterministic equation

$$-c_\epsilon \frac{dU_0(\xi)}{d\xi} - h(U_0(\xi)) = \int_{-\infty}^{\infty} w(\xi - \xi') F(U_0(\xi')) d\xi'. \quad (38)$$

The results depend on ϵ due to the ϵ -dependence of h . Equation (38) is chosen so that to leading order, the stochastic variable $\Delta(t)$ undergoes unbiased Brownian motion with a diffusion coefficient $D(\epsilon) = \mathcal{O}(\epsilon)$ (see below). The next step is to substitute the decomposition Equation (37) into (33) and expand to first order in $\mathcal{O}(\epsilon^{1/2})$:

$$\begin{aligned} & -[c_\epsilon + \dot{\Delta}]U'_0(\xi_\Delta)dt + \epsilon^{1/2} [d\Phi(\xi_\Delta, t) - [c_\epsilon + \dot{\Delta}]\Phi'(\xi_\Delta, t)dt] \\ & = h(U_0(\xi_\Delta))dt + \epsilon^{1/2} h'(U_0(\xi_\Delta))\Phi(\xi_\Delta, t)dt \\ & + \int_{-\infty}^{\infty} w(\xi - \xi') F(U_0(\xi'_\Delta)) d\xi' dt \\ & + \epsilon^{1/2} \int_{-\infty}^{\infty} w(\xi - \xi') F'(U_0(\xi'_\Delta)) \Phi(\xi'_\Delta, t) d\xi' dt \\ & + \epsilon^{1/2} dR(U_0(\xi_\Delta), \xi, t) + \mathcal{O}(\epsilon). \end{aligned}$$

where we have set $\xi_\Delta = \xi - \Delta(t)$ and $\xi'_\Delta = \xi' - \Delta(t)$. We now use Equation (38) for U_0 , after shifting $\xi \rightarrow \xi - \Delta(t)$, to eliminate terms and then divide through by $\sqrt{\epsilon}$. This gives the inhomogeneous equation to $\mathcal{O}(\epsilon^{1/2})$

$$\begin{aligned} d\Phi(\xi_\Delta, t) - \widehat{L}\Phi(\xi_\Delta, t)dt &= \epsilon^{-\frac{1}{2}} U'_0(\xi_\Delta) d\Delta(t) \\ &+ dR(U_0(\xi_\Delta), \xi, t) \end{aligned} \quad (39)$$

where the non-self-adjoint linear operator

$$\begin{aligned} \widehat{L}A(\xi) &\equiv c_\epsilon \frac{dA(\xi)}{d\xi} + h'(U_0(\xi))A(\xi) \\ &+ \int_{-\infty}^{\infty} w(\xi - \xi') F'(U_0(\xi')) A(\xi') d\xi' \end{aligned} \quad (40)$$

is defined for all functions $A(\xi)$ in $\mathcal{L}_2(\mathbb{R})$. Note that for all terms in Equation (40) to be of the same order we have taken $\Delta(t) = \mathcal{O}(\epsilon^{1/2})$. It then follows that $U_0(\xi - \Delta(t)) = U_0(\xi) + \mathcal{O}(\epsilon^{1/2})$ etc., and Equation (39) reduces to

$$d\Phi(\xi, t) - \widehat{L}\Phi(\xi, t)dt = \epsilon^{-\frac{1}{2}} U'_0(\xi) d\Delta(t) + dR_u(U_0(\xi), \xi, t) \quad (41)$$

If $U_0(\xi)$ were a traveling front solution of a neural field model with a symmetric, excitatory weight distribution w , then it could

be proven that the operator \widehat{L} has a 1D null space spanned by $U'_0(\xi)$ (Ermentrout and McLeod, 1993). We will assume that such a result carries over to traveling pulse solutions of a neural field with w given by an asymmetric Mexican hat function; the fact that $U'_0(\xi)$ belongs to the null space follows immediately from differentiating Equation (38) with respect to ξ . We then have the solvability condition for the existence of a non-trivial bounded solution of Equation (41), namely, that the inhomogeneous part is orthogonal to all elements of the null space of the adjoint operator \widehat{L}^* . The latter is defined with respect to the inner product

$$\int_{-\infty}^{\infty} B(\xi) \widehat{L}A(\xi) d\xi = \int_{-\infty}^{\infty} [\widehat{L}^*B(\xi)] A(\xi) d\xi.$$

Integrating by parts and using (Equation 14) leads to

$$\begin{aligned} \widehat{L}^*B(\xi) &= -c_\epsilon \frac{dB(\xi)}{d\xi} + h'(U_0(\xi))B(\xi) \\ &+ F'(U_0(\xi)) \int_{-\infty}^{\infty} w(\xi' - \xi) B(\xi') d\xi'. \end{aligned} \quad (42)$$

We will assume that the null space of the adjoint operator \widehat{L}^* is also one-dimensional and is spanned by some yet to be determined function $\mathcal{V}(\xi)$. (In the case of a Heaviside firing function, we will determine the null space explicitly). Hence, we can write the solvability condition as

$$\int_{-\infty}^{\infty} \mathcal{V}(\xi) [U'_0(\xi) d\Delta(t) + \epsilon^{1/2} dR(U_0(\xi), \xi, t)] d\xi = 0.$$

which leads directly to the stochastic differential equation

$$d\Delta(t) = -\epsilon^{1/2} \frac{\int_{-\infty}^{\infty} \mathcal{V}(\xi) dR(U_0, \xi, t) d\xi}{\int_{-\infty}^{\infty} \mathcal{V}(\xi) U'_0(\xi) d\xi}.$$

Using the lowest order approximations $dR(U_0, \xi, t) = g(U_0(\xi))dW(\xi, t)$, we deduce that [for $\Delta(0) = 0$]

$$\langle \Delta(t) \rangle = 0, \quad \langle \Delta(t)^2 \rangle = 2D(\epsilon)t, \quad (43)$$

where $D(\epsilon)$ is the effective diffusivity

$$D(\epsilon) = \epsilon \frac{\int_{-\infty}^{\infty} \mathcal{V}^2(\xi) g(U_0(\xi))^2 d\xi}{[\int_{-\infty}^{\infty} \mathcal{V}(\xi) U'_0(\xi) d\xi]^2}. \quad (44)$$

Explicit results for Heaviside rate function

In order to illustrate the above analysis, we consider a particular example where the mean speed c_ϵ and diffusion coefficient $D(\epsilon)$ can be calculated explicitly. That is, set $g(U) = g_0 U$ for the multiplicative noise term and take $F(U) = H(u - \kappa)$. (The constant g_0 has units of $\sqrt{\text{length/time}}$). Note that the choice for $g(U)$ can be interpreted physiologically in terms of an effective modification in the membrane time constant of neurons due to stochastic background synaptic activity (Bernander et al., 1991; Rapp et al., 1992; Bressloff, 1994). The deterministic Equation (38) for U_0

then reduces to

$$-\frac{dU_0(\xi)}{d\xi} + \Gamma(\epsilon)U_0(\xi) = \frac{1}{c_\epsilon} \int_{-\infty}^{\infty} w(\xi - \xi')H(U_0(\xi') - \kappa)d\xi', \quad (45)$$

where $\Gamma(\epsilon) = (1 - \epsilon C(0)g_0^2)/c_\epsilon$. Hence,

$$U(\xi) = \frac{e^{\Gamma\xi}}{c_\epsilon} \int_{-\infty}^{\infty} W(\xi')e^{-\Gamma\xi'} d\xi'. \quad (46)$$

The deterministic pulse profile can be evaluated along identical lines to section “Neural Field Model of Direction Selectivity.” In order to calculate the diffusion coefficient, it is first necessary to determine the null vector $\mathcal{V}(\epsilon)$ of the adjoint linear operator \widehat{L}^* . Substituting $F(U) = H(U - \kappa)$ and $g(U) = g_0U$ into Equation (42) shows that

$$\begin{aligned} \frac{d\mathcal{V}(\xi)}{d\xi} + \Gamma(\epsilon)\mathcal{V}(\xi) &= \frac{\delta(\xi)}{c|U'_0(0)|} \int_{-\infty}^{\infty} w(z)\mathcal{V}(z)dz \\ &+ \frac{\delta(\xi - \Delta)}{c|U'_0(\Delta)|} \int_{-\infty}^{\infty} w(z - \Delta)\mathcal{V}(z)dz. \end{aligned} \quad (47)$$

Proceeding along similar lines to Bressloff (2001) and Kilpatrick et al. (2008), we make the ansatz that

$$\mathcal{V}(\xi) = AH(\xi)e^{-\Gamma\xi} + BH(\xi - \Delta)e^{-\Gamma(\xi - \Delta)}. \quad (48)$$

Substituting into Equation (47) shows that

$$A = \frac{1}{|U'_0(0)|} [Ab(0) + Bb(\Delta)], \quad B = \frac{1}{|U'_0(\Delta)|} [Ab(-\Delta) + Bb(0)]$$

where

$$b(z) \equiv \frac{1}{c} \int_z^{\infty} e^{-\Gamma(\xi' - z)} w(\xi') d\xi'. \quad (49)$$

Differentiating Equation (46) shows that $U'(\xi) = b(\xi) - b(\xi - \Delta)$, so that we obtain the vector equation

$$\begin{bmatrix} \frac{b(0)}{b(0) - b(-\Delta)} - 1 & \frac{b(\Delta)}{b(0) - b(-\Delta)} \\ \frac{b(-\Delta)}{b(0) - b(\Delta)} & \frac{b(0)}{b(0) - b(\Delta)} - 1 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = 0$$

The matrix has rank 1, confirming that the linear operator \widehat{L}^* has a 1D null-space. The latter is spanned by the function

$$\mathcal{V}(\xi) = b(\Delta)H(\xi)e^{-\Gamma\xi} - b(-\Delta)H(\xi - \Delta)e^{-\Gamma(\xi - \Delta)}. \quad (50)$$

In **Figure 8** we show the temporal evolution of a freely propagating stochastic traveling pulse, which is obtained by numerically solving the Langevin Equation (31) for $F(U) = H(U - \kappa)$, $g(U) = U$ and the asymmetric difference-of-exponentials (Equation 4). Note that the location of the stochastic wave appears to coincide with the underlying mean solution. However, over longer time-scales the wandering of the pulse about its

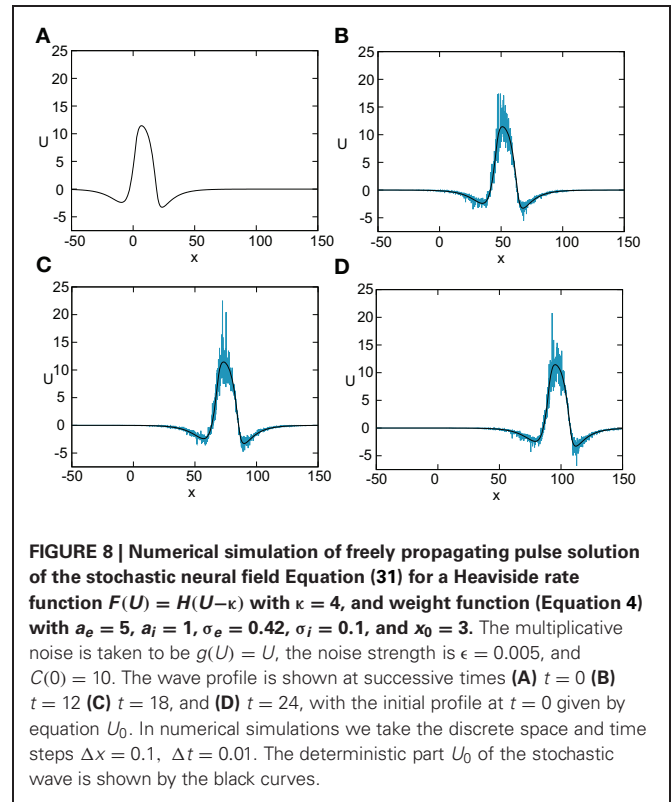


FIGURE 8 | Numerical simulation of freely propagating pulse solution of the stochastic neural field Equation (31) for a Heaviside rate function $F(U) = H(U - \kappa)$ with $\kappa = 4$, and weight function (Equation 4) with $a_e = 5$, $a_i = 1$, $\sigma_e = 0.42$, $\sigma_i = 0.1$, and $x_0 = 3$. The multiplicative noise is taken to be $g(U) = U$, the noise strength is $\epsilon = 0.005$, and $C(0) = 10$. The wave profile is shown at successive times (A) $t = 0$ (B) $t = 12$ (C) $t = 18$, and (D) $t = 24$, with the initial profile at $t = 0$ given by equation U_0 . In numerical simulations we take the discrete space and time steps $\Delta x = 0.1$, $\Delta t = 0.01$. The deterministic part U_0 of the stochastic wave is shown by the black curves.

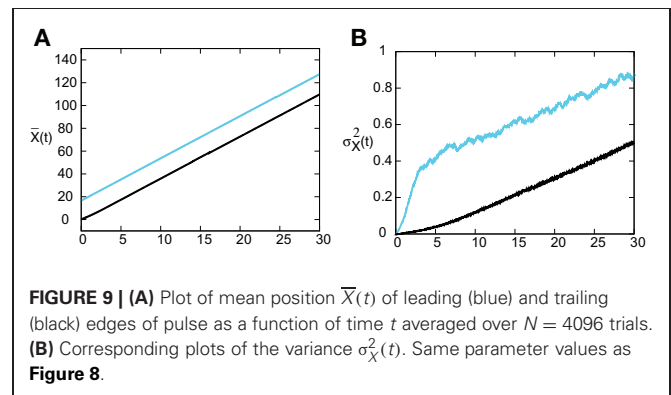


FIGURE 9 | (A) Plot of mean position $\bar{X}(t)$ of leading (blue) and trailing (black) edges of pulse as a function of time t averaged over $N = 4096$ trials. **(B)** Corresponding plots of the variance $\sigma_X^2(t)$. Same parameter values as **Figure 8**.

mean position would be seen. In **Figure 9** we plot the mean position $\bar{X}(t)$ and variance $\sigma_X^2(t)$ of the leading and trailing edges of the pulse as a function of t . It can be seen that they all vary linearly with t , consistent with the assumption that there is a diffusive-like displacement of the center-of-mass of the pulse from its uniformly translating position at long time-scales. The slopes of these curves then determine the effective wave speed and diffusion coefficient according to $\bar{X}(t) \sim c_\epsilon t$ and $\sigma_X^2(t) \sim 2D(\epsilon)t$. Both the leading and trailing edges exhibit the same speeds and diffusivities (after a transient phase). The transients are caused by fluctuations in the mean width of the pulse which can be neglected for large t , where the difference in the size of fluctuations of the leading and trailing edges can be neglected.

In order to find the mean location of the leading or trailing edge of the pulse as a function of time, we numerically carry out

a large number of level set position measurements. That is, we determine the positions $X_a(t)$ such that $U(X_a(t), t) = a$, for various level set values a and then define the mean location to be $\bar{X}(t) = \mathbb{E}[X_a(t)]$, where the expectation is first taken with respect to the sampled values a and then averaged over N trials. The corresponding variance is given by $\sigma_X^2(t) = \mathbb{E}[(X_a(t) - \bar{X}(t))^2]$. In order to compare the numerical results with our theoretical analysis, we assume that $X_a(t) = \Delta(t) + c_\epsilon t + X_a(0)$ for each a on either the leading or trailing edge. It then follows that $\bar{X}(t) = c_\epsilon t + X_a(0)$ and $\sigma_X^2(t) = \langle \Delta(t)^2 \rangle$. In **Figure 10** we plot the numerically estimated diffusion coefficient for various values of the threshold κ and compare these to the corresponding theoretical curves obtained using the above analysis. It can be seen that there is excellent agreement with our theoretical predictions. Finally, note that we can also use the level set data to estimate fluctuations in the width of the pulse. Suppose that $X_d(t)$ and $Y_d(t)$ denote the threshold crossing points of the leading and trailing edges of the pulse at time t . Then the stochastic width of the pulse can be defined according to $D(t) = X_d(t) - Y_d(t)$. We find that after a transient phase, $\langle D(t)^2 \rangle - \langle D(t) \rangle^2 \ll \sigma_X^2(t)$.

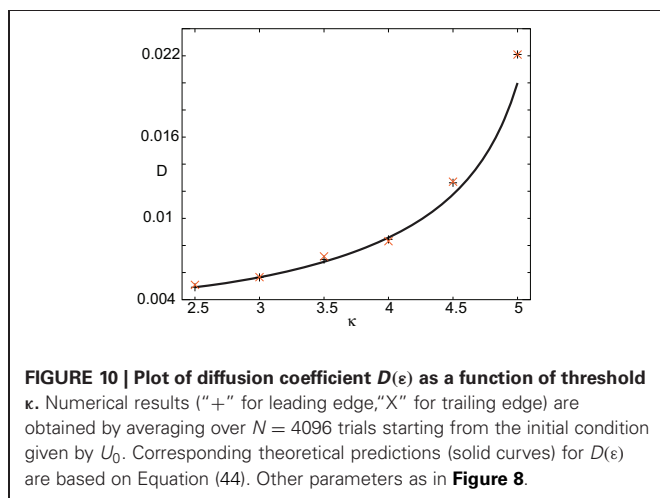
Stimulus-locked pulses

We now add a stimulus term I to the stochastic neural field Equation (33), that is

$$dU(x, t) = \left[h(U(x, t)) + \int_{-\infty}^{\infty} w(x - y) F(U(y, t)) dy \right] dt + I(x - vt)dt + \epsilon^{1/2} dR(U, x, t), \quad (51)$$

where the stimulus is again a rectangular wave of amplitude I_0 and width d , moving with speed v . Here h and dR are defined by Equations (34) and (35). The stochastic activity variable is now decomposed according to Equation (37), with $\xi = x - vt$, and the modified deterministic equation

$$-v \frac{dU_0}{d\xi} - h(U_0(\xi)) - I(\xi) = \int_{-\infty}^{\infty} w(\xi - \xi') F(U_0(\xi')) d\xi'. \quad (52)$$



Through a similar process as in the previous section, we expand to $\mathcal{O}(\epsilon^{1/2})$ to obtain the inhomogeneous equation

$$d\Phi(\xi, t) - \hat{L}\Phi(\xi, t)dt = -\epsilon^{-1/2} [U'_0(\xi) - I'(\xi)] d\Delta(t) - dR(U_0(\xi), \xi, t) + \mathcal{O}(\epsilon^{1/2}), \quad (53)$$

where \hat{L} is defined as in Equation (40) but with $c_\epsilon \rightarrow v$. The solvability condition is now

$$\int_{-\infty}^{\infty} \mathcal{V}(\xi) [U'_0(\xi) d\Delta(t) + I'(\xi) d\Delta(t) + \epsilon^{1/2} dR(U_0(\xi), \xi, t)] d\xi = 0. \quad (54)$$

This can be manipulated to give, to leading order, the Ornstein–Uhlenbeck equation (Gardiner, 2009):

$$d\Delta(t) + A\Delta(t)dt = d\hat{W}(t), \quad (55)$$

where

$$A = \frac{\int_{-\infty}^{\infty} \mathcal{V}(\xi) I'(\xi) d\xi}{\int_{-\infty}^{\infty} \mathcal{V}(\xi) U'_0(\xi) d\xi},$$

and

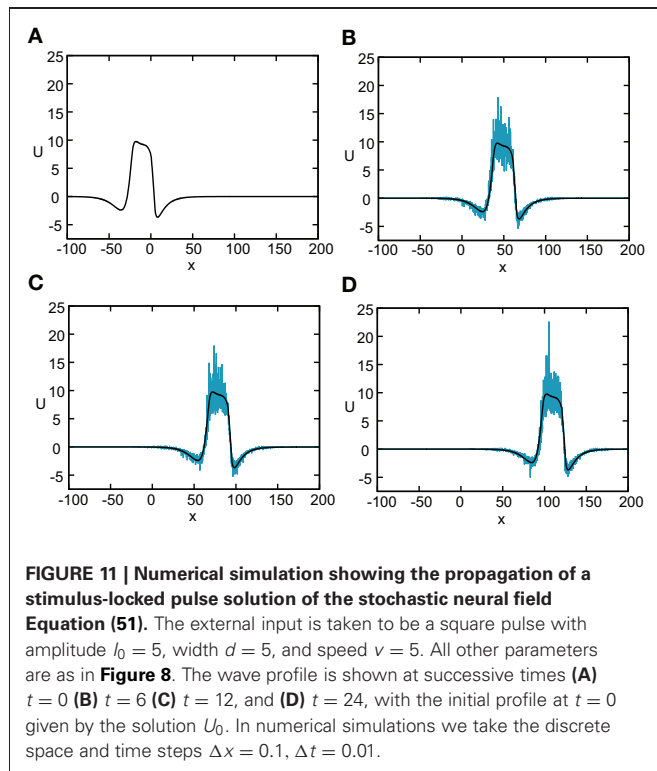
$$d\hat{W}(t) = -\epsilon^{1/2} g_0 \frac{\int_{-\infty}^{\infty} \mathcal{V}(\xi) U_0(\xi) dW(\xi, t) d\xi}{\int_{-\infty}^{\infty} \mathcal{V}(\xi) U'_0(\xi) d\xi}.$$

Solving the stochastic differential equation in Equation (55) and taking averages shows that $\langle \Delta(t) \rangle = \Delta(0)e^{-At}$ and

$$\langle \Delta(t)^2 \rangle - \langle \Delta(t) \rangle^2 \approx \frac{D(\epsilon)}{A} [1 - e^{-2At}], \quad (56)$$

where $D(\epsilon)$ is given by Equation (44) except for a modified null vector $\mathcal{V}(\xi)$. Thus the variance of $\Delta(t)$ approaches $D(\epsilon)/A$ in the large t limit.

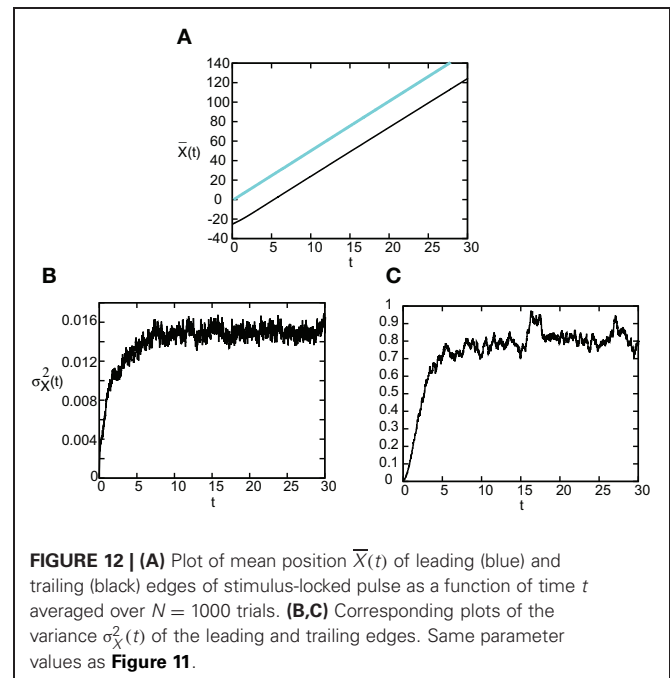
As in the case of freely propagating pulses, we can explicitly solve for $\mathcal{V}(\epsilon)$ and thus calculate the diffusion coefficient $D(\epsilon)$ when $F(U) = H(U - \kappa)$ and $g(U) = U$. Since the steps are similar to the previous case, we simply present our results here. In **Figure 11** we show the temporal evolution of a single stimulus-locked front, which is obtained by numerically solving the Langevin Equation (51) for $F(U) = H(U - \kappa)$, $g(U) = U$ and the weight distribution (Equation 4). The external input is taken to be a square pulse of amplitude $I_0 = 5$, width $d = 5$, and speed $v = 5$. Next we determine the mean $\bar{X}(t)$ and variance $\sigma_X^2(t)$ of the position of the leading and trailing edges by averaging over level sets along identical lines to the freely-propagating case. The results are shown in **Figure 12**. It can be seen that, as predicted by the analysis, $\bar{X}(t)$ varies linearly with t with a slope equal to the stimulus speed $v = 5$. Moreover, the variance $\sigma_X^2(t)$ approaches a constant value as $t \rightarrow \infty$ for both the trailing and leading edges. Thus, we find that stimulus-locked pulses are much more robust to noise than freely propagating pulses, since the variance of the mean position of the leading and



trailing edges saturate as $t \rightarrow \infty$. Consequently, stimulus locking persists in the presence of noise over most of the parameter range for which stimulus locking is predicted to occur. However, the trailing edge has an asymptotic variance that is at least an order of magnitude larger than the leading edge, which implies that fluctuations in the width of the pulse can no longer be neglected.

DISCUSSION

In this paper we have explored the effects of extrinsic noise on propagating pulses in a one-dimensional scalar neural field with asymmetric weights. Such a network has previously been proposed as a continuum model of direction selectivity. We have shown that the effects of noise on the wandering of the mean front



position depends on properties of the underlying deterministic pulse. In the case of a freely propagating pulse, we find diffusive wandering with the mean square displacement growing linearly with time t . Moreover, in the large time limit, fluctuations in the width of the pulse can be neglected. On the other hand, if the pulse is locked to a moving pulse-like stimulus, then the wandering is described by an Ornstein–Uhlenbeck process and the mean square displacement saturates in the long time limit. However, we find that fluctuations in the pulse width can no longer be ignored.

In summary, this paper further illustrates how methods developed for studying wave propagation in stochastic PDEs can be adapted to study wave propagation in stochastic neural fields. As we have previously found for fronts, stimulus-locked waves are more robust to noise, which is a desirable property of a network performing some form of stimulus-processing such as direction selectivity.

REFERENCES

- Adelson, E. H., and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am.* 2, 284–299.
- Amari, S. (1977). Dynamics of pattern formation in lateral inhibition type neural fields. *Biol. Cybern.* 27, 77–87.
- Armero, J., Casademunt, J., Ramirez-Piscina, L., and Sancho, J. M. (1998). Ballistic and diffusive corrections to front propagation in the presence of multiplicative noise. *Phys. Rev. E* 58, 5494–5500.
- Baker, P. M., and Bair, W. (2012). Inter-neuronal correlation distinguishes mechanisms of direction selectivity in cortical circuit models. *J. Neurosci.* 32, 8800–8816.
- Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3844–3848.
- Bernander, O., Douglas, R. J., Martin, K. A. C., and Koch, C. (1991). Synaptic background activity influences spatiotemporal integration in single pyramidal cells. *Proc. Natl. Acad. Sci. U.S.A.* 88, 11569–11573.
- Brackley, C. A., and Turner, M. S. (2007). Random fluctuations of the firing rate function in a continuum neural field model. *Phys. Rev. E* 75, 041913.
- Bressloff, P. C. (1994). A green's function approach to analysing the effects of random synaptic background activity in a model neural network. *J. Phys. A* 27, 4097–4113.
- Bressloff, P. C. (2001). Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Physica D* 155, 83–100.
- Bressloff, P. C. (2009). Stochastic neural field theory and the system-size expansion. *SIAM J. Appl. Math.* 70, 1488–1521.
- Bressloff, P. C. (2010). Metastable states and quasicycles in a stochastic Wilson-Cowan model of neuronal population dynamics. *Phys. Rev. E* 85, 051903.
- Bressloff, P. C. (2012). Spatiotemporal dynamics of continuum neural fields. *J. Phys. A* 45, 033001.
- Bressloff, P. C., Cowan, J. D., Golubitsky, M., Thomas, P. J.,

- and Wiener, M. (2001). Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. *Phil. Trans. R. Soc. Lond. B* 356, 299–330.
- Bressloff, P. C., and Webber, M. A. (2012a). Neural field model of binocular rivalry waves. *J. Comput. Neurosci.* 32, 233–252.
- Bressloff, P. C., and Webber, M. A. (2012b). Front propagation in stochastic neural fields. *SIAM J. Appl. Dyn. Syst.* 11, 708–740.
- Buice, M., and Cowan, J. D. (2007). Field-theoretic approach to fluctuation effects in neural networks. *Phys. Rev. E* 75, 051919.
- Buice, M., Cowan, J. D., and Chow, C. C. (2010). Systematic fluctuation expansion for neural network activity equations. *Neural Comput.* 22, 377–426.
- Camperi, M., and Wang, X.-J. (1998). A model of visuospatial short-term memory in prefrontal cortex: recurrent network and cellular bistability. *J. Comput. Neurosci.* 5, 383–405.
- Coombes, S. (2005). Waves, bumps and patterns in neural field theories. *Biol. Cybern.* 93, 91–108.
- Coombes, S., and Owen, M. R. (2004). Evans functions for integral neural field equations with Heaviside firing rate function. *SIAM J. Appl. Dyn. Syst.* 4, 574–600.
- Coombes, S., and Owen, M. R. (2005). Bumps, breathers, and waves in a neural network with spike frequency adaptation. *Phys. Rev. Lett.* 94, 148102.
- de Pasquale, F., Gorecki, J., and Poielawski, J. (1992). On the stochastic correlations in a randomly perturbed chemical front. *J. Phys. A* 25, 433.
- Ermentrout, G. B. (1998). Neural networks as spatio-temporal pattern-forming systems. *Rep. Prog. Phys.* 61, 353–430.
- Ermentrout, G. B., and Cowan, J. (1979). A mathematical theory of visual hallucination patterns. *Bio. Cybern.* 34, 137–150.
- Ermentrout, G. B., and McLeod, J. B. (1993). Existence and uniqueness of travelling waves for a neural network. *Proc. R. Soc. Edinburgh A* 123, 461–478.
- Evans, J. (1975). Nerve axon equations iv: the stable and unstable impulse. *Indiana Univ. Math. J.* 24, 1169–1190.
- Faugeras, O., Touboul, J., and Cessac, B. (2009). A constructive mean-field analysis of multi-population neural networks with random synaptic weights and stochastic inputs. *Front. Comput. Neurosci.* 3:1. doi: 10.3389/fnro.10.001. 2009
- Folias, S. E., and Bressloff, P. C. (2005). Stimulus-locked traveling pulses and breathers in an excitatory neural network. *SIAM J. Appl. Math.* 65, 2067–2092.
- Gardiner, C. W. (2009). *Handbook of Stochastic Methods, 4th Edn.* Berlin: Springer.
- Giese, M. A. (1999). *Neural Field Theory for Motion Perception.* Dordrecht: Kluwer Academic Publishers.
- Huang, X., Troy, W. C., Yang, Q., Ma, H., Laing, C. R., Schiff, S. J., et al. (2004). Spiral waves in disinhibited mammalian neocortex. *J. Neurosci.* 24, 9897–9902.
- Hutt, A., and Atay, F. (2005). Analysis of nonlocal neural fields for both general and gamma-distributed connectivities. *Physica D* 203, 30–54.
- Hutt, A., Longtin, A., and Schimansky-Geier, L. (2008). Additive noise-induces Turing transitions in spatial systems with application to neural fields and the Swift-Hohenberg equation. *Physica D* 237, 755–773.
- Kapitula, T., Kutz, N., and Sandstede, B. (2004). The Evans function for non-local equations. *Indiana Univ. Math. J.* 53, 1095–1126.
- Keener, J. P. (1981). Waves in excitable media. *SIAM J. Appl. Math.* 39, 528–548.
- Kilpatrick, Z. P., and Bressloff, P. C. (2010a). Effects of synaptic depression and adaptation on spatiotemporal dynamics of an excitatory neuronal network. *Physica D* 239, 547–560.
- Kilpatrick, Z. P., and Bressloff, P. C. (2010b). Spatially structured oscillations in a two-dimensional neuronal network with synaptic depression. *J. Comput. Neurosci.* 28, 193–209.
- Kilpatrick, Z. P., Folias, S. E., and Bressloff, P. C. (2008). Traveling pulses and wave propagation failure in inhomogeneous neural media. *SIAM J. Appl. Dyn. Syst.* 7, 161–185.
- Koch, C., and Poggio, T. (1985). “Orientation selectivity in the visual cortex?” in *Models of the Visual Cortex*, eds D. Rose and V. G. Dobson (New York, NY: Wiley), 408–419.
- Kreyszig, E. (1978). *Introductory Functional Analysis with Applications.* New York, NY: Wiley.
- Kuramoto, Y. (1984). *Chemical Oscillations, Waves and Turbulence.* New York, NY: Springer-Verlag.
- Laing, C. R., and Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *J. Comput. Neurosci.* 12, 39–53.
- Liley, D. J. T., Cadusch, P. J., and Dafilis, M. P. (2002). A spatially continuous mean field theory of electrocortical activity. *Network* 13, 67–113.
- Maex, R., and Urban, G. A. (1996). Model circuit of spiking neurons generating directional selectivity in simple cells. *J. Neurophysiol.* 75, 1515–1545.
- Mineiro, P., and Zipser, D. (1998). Analysis of direction selectivity arising from recurrent cortical interactions. *Neural Comput.* 10, 353–371.
- Novikov, E. A. (1965). Functionals and the random-force method in turbulence theory. *Sov. Phys. JETP* 20, 1290.
- Nunez, P. I. (1995). *Neocortical Dynamics and Human EEG Rhythms.* New York, NY: Oxford University Press.
- Pinto, D., and Ermentrout, G. B. (2001). Spatially structured activity in synaptically coupled neuronal networks: I. Traveling fronts and pulses. *SIAM J. Appl. Math.* 62, 206–225.
- Pinto, D., Jackson, R. K., and Wayne, C. E. (2005). Existence and stability of traveling pulses in a continuous neuronal network. *SIAM J. Appl. Dyn. Syst.* 4, 954–984.
- Rapp, M., Yarom, Y., and Segev, I. (1992). The impact of parallel fiber background activity on the cable properties of cerebellar Purkinje cells. *Neural Comput.* 4, 518–533.
- Reichardt, W. (1961). “Autocorrelation, a principle for the evaluation of sensory information by the central nervous system,” in *Sensory Communication*, ed A. Rosenblith (Cambridge, MA: MIT Press), 303–317.
- Richardson, K. A., Schiff, S. J., and Gluckman, B. J. (2005). Control of traveling waves in the mammalian cortex. *Phys. Rev. Lett.* 94, 028103.
- Robinson, P. A., Rennie, C. J., Wright, J. J., Bahramali, H., Gordon, E., and Rowe, D. I. (2001). Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E* 63, 021903.
- Rubin, J. E. (2004). A nonlocal eigenvalue problem for the stability of a traveling wave in a neuronal medium. *Discrete Contin. Dyn. Syst.* 10, 925–940.
- Sagundes, E., Sancho, J. M., and Garcia-Ojalvo, J. (2007). Spatiotemporal order out of noise. *Rev. Mod. Phys.* 79, 829–882.
- Sandstede, B. (2002). “Stability of travelling waves,” in *Handbook of Dynamical Systems, Vol. 2*, ed B. Fiedler (Amsterdam: North-Holland), 983–1055.
- Sandstede, B. (2007). Evans functions and nonlinear stability of traveling waves in neuronal network models. *Int. J. Bifurcation Chaos Appl. Sci. Eng.* 17, 2693–2704.
- Saul, A. B., and Humphrey, A. L. (1990). Spatial and temporal response properties of lagged and nonlagged cells in cat lateral geniculate nucleus. *J. Neurophysiol.* 64, 206–224.
- Schimansky-Geier, L., Mikhailov, A. S., and Ebeling, W. (1983). Effects of fluctuations on plane front propagation in bistable nonequilibrium systems. *Ann. Phys.* 40, 277.
- Somers, D. C., Nelson, S., and Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.* 15, 5448–5465.
- Steyn-Ross, M. L., Steyn-Ross, D. A., Sleigh, J. W., and Whiting, D. R. (2003). Theoretical predictions for spatial covariance of the electroencephalographic signal during the anesthetic-induced phase transition: increased correlation length and emergence of spatial self-organization. *Phys. Rev. E* 68, 021902.
- Suarez, H., Koch, C., and Douglas, R. J. (1995). Modeling direction selectivity of simple cells in striate visual cortex using the canonical microcircuit. *J. Neurosci.* 15, 6700–6719.
- van Santen, J. P., and Sperling, G. (1985). Elaborated Reichardt detectors. *J. Opt. Soc. Am. A* 256, 300–320.
- Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–23.
- Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic

- nervous tissue. *Kybernetik* 13, 55–80.
- Xie, X., and Giese, M. A. (2002). Nonlinear dynamics of direction-selective recurrent neural media. *Phys. Rev. E* 65, 051904.
- Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J. Neurosci.* 16, 2112–2126.
- Zhang, L. (2003). On the stability of traveling wave solutions in synaptically coupled neuronal networks. *Diff. Integral Equat.* 16, 513–536.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 24 September 2012; accepted: 10 October 2012; published online: 29 October 2012.
- Citation: Bressloff PC and Wilkerson J (2012) Traveling pulses in a stochastic neural field model of direction selectivity. *Front. Comput. Neurosci.* 6:90. doi: 10.3389/fncom.2012.00090
- Copyright © 2012 Bressloff and Wilkerson. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



A biophysical observation model for field potentials of networks of leaky integrate-and-fire neurons

Peter beim Graben^{1,2*} and Serafim Rodrigues³

¹ Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany

² Department of German Language and Linguistics, Humboldt-Universität zu Berlin, Berlin, Germany

³ Centre for Robotics and Neural Systems, School of Computing and Mathematics, University of Plymouth, Plymouth, UK

Edited by:

Dimitris Pinotsis, University College London, UK

Reviewed by:

Gaute T. Einevoll, Norwegian University of Life Sciences, Norway
Julien Modolo, Lawson Health Research Institute, Canada

*Correspondence:

Peter beim Graben, Department of German Language and Linguistics, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099, Berlin, Germany.
e-mail: peter.beim.graben@hu-berlin.de

We present a biophysical approach for the coupling of neural network activity as resulting from proper dipole currents of cortical pyramidal neurons to the electric field in extracellular fluid. Starting from a reduced three-compartment model of a single pyramidal neuron, we derive an observation model for dendritic dipole currents in extracellular space and thereby for the dendritic field potential (DFP) that contributes to the local field potential (LFP) of a neural population. This work aligns and satisfies the widespread dipole assumption that is motivated by the “open-field” configuration of the DFP around cortical pyramidal cells. Our reduced three-compartment scheme allows to derive networks of leaky integrate-and-fire (LIF) models, which facilitates comparison with existing neural network and observation models. In particular, by means of numerical simulations we compare our approach with an *ad hoc* model by Mazzoni et al. (2008), and conclude that our biophysically motivated approach yields substantial improvement.

Keywords: biophysics, neural networks, leaky integrate-and-fire neuron, current dipoles, extracellular medium, field potentials

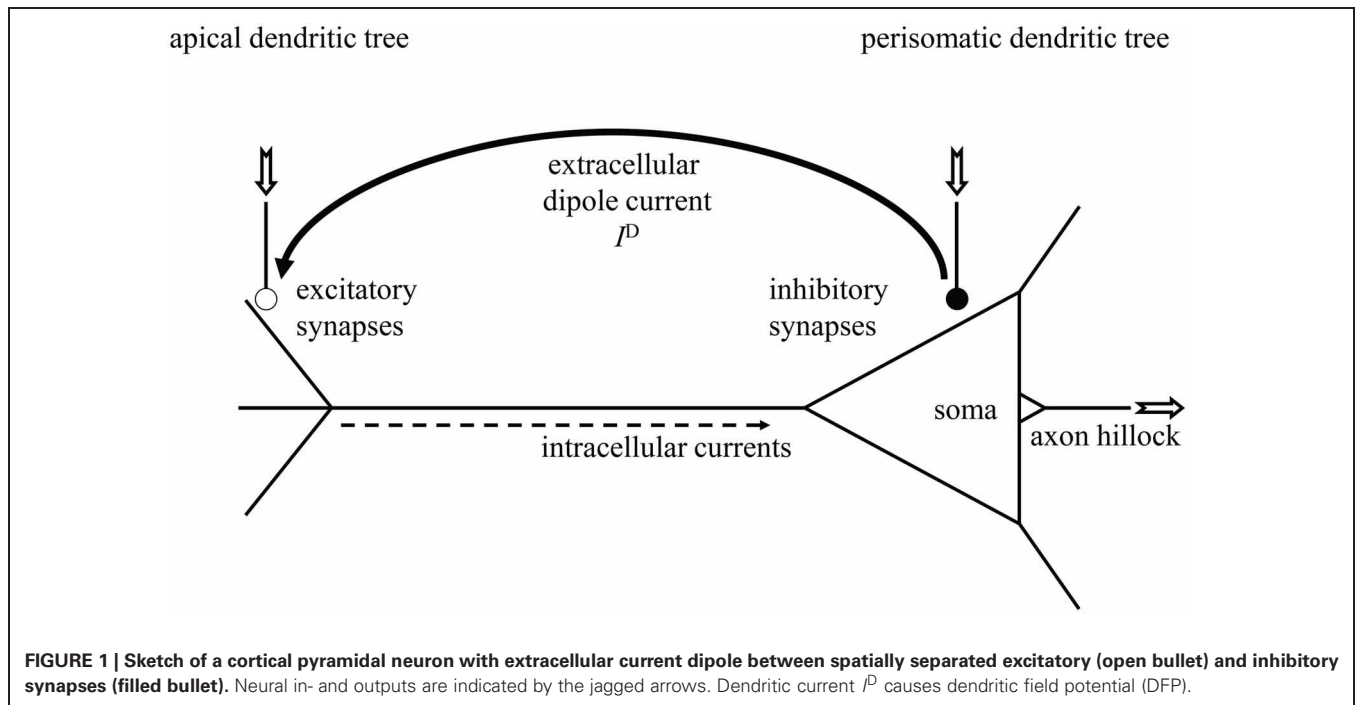
1. INTRODUCTION

Since Hans Berger's 1924 discovery of the human *electroencephalogram* (EEG) (Berger, 1929), neuroscientists achieved much progress in clarifying its neural generators (Creutzfeldt et al., 1966a,b; Nunez and Srinivasan, 2006; Schomer and Lopes da Silva, 2011). These are the cortical pyramidal neurons, as sketched in **Figure 1**, that possess a long dendritic trunk separating mainly excitatory synapses at the apical dendritic tree from mainly inhibitory synapses at the soma and at the perisomatic basal dendritic tree (Creutzfeldt et al., 1966a; Spruston, 2008). In addition, they exhibit an axial symmetry and are aligned in parallel to each other, perpendicular to the cortex' surface, thus forming a palisade of cell bodies and dendritic trunks. When both kinds of synapses are simultaneously active, inhibitory synapses generate current sources and excitatory synapses current sinks in extracellular space, hence causing the pyramidal cell to behave as a microscopic dipole surrounded by its characteristic electrical field, the *dendritic field potential* (DFP). The densely packed pyramidal cells form then a dipole layer whose superimposed currents give rise to the *local field potential* (LFP) of neural masses and eventually to the EEG (Nunez and Srinivasan, 2006; Lindén et al., 2010; Lindén et al., 2011; Schomer and Lopes da Silva, 2011).

Despite of the progress from experimental neuroscience, theoretically understanding the coupling of complex neural network dynamics to the electromagnetic field in the extracellular space poses challenging problems; some of them have been addressed to some extent by Bédard et al. (2004); Bédard and Destexhe (2009), and Bédard and Destexhe (2012).

In computer simulation studies, neural mass potentials, such as LFP and EEG are most realistically simulated by means of multicompartmental models (Protopapas et al., 1998; Sargsyan et al., 2001; Lindén et al., 2010; Lindén et al., 2011). Lindén et al. (2010) calculated the current dipole momentum of the DFP for single pyramidal and stellate cells, based on several hundreds compartments of the dendritic trees. Their results were in compliance with the standard dipole approximation of the electrostatic multipole expansion in the far-field (more than 1 mm remote from the dendritic trunk), but they found rather poor agreement with that approximation in the vicinity of the cell body. For comparison they also computed a “two-monopole” model of one synaptic current and its counterpart, the somatic return current, estimated from the current dipole momentum of the whole dendritic tree. This “two-monopole” model, which corresponds to an electrically equivalent single dipole model, obtained from the decomposition of the dendrite into two compartments, better approximates the true current dipole momentum in the vicinity of the pyramidal neuron. By superimposing the DFPs of pyramidal cells to the ensemble LFP, Lindén et al. (2011) found that LFP properties cannot be attributed to the far-field dipole approximation.

However, realistic multicompartmental models are computationally too expensive for large-scale neural network simulations. Therefore, various techniques have been proposed and employed to overcome computational complexity. These include networks of point models (i.e., devoid from any spatial representation), based on conductance models (Hodgkin and Huxley, 1952; Mazzoni et al., 2008), population density models



(Omurtag et al., 2000), or firing rate models (Wilson and Cowan, 1972), which can be seen as a sub class of population density models, with uniform density distribution (Chizhov et al., 2007). In these kinds of models, mass potentials such as LFP or EEG are conventionally described as averaged membrane potential. A different class of models are neural mass models (Jansen and Rit, 1995; Wendling et al., 2000; David and Friston, 2003; Rodrigues et al., 2010), where mass potentials are estimated either through sums (or actually differences) of excitatory postsynaptic potentials (EPSP) (David and Friston, 2003) or of excitatory postsynaptic currents (EPSC) (Mazzoni et al., 2008).

In particular, the model of Mazzoni et al. (2008) which is based on Brunel and Wang (2003), recently led to a series of follow-up studies (Mazzoni et al., 2010, 2011) addressing the correlations between numerically simulated and experimentally measured LFP/EEG with spike rates by means of statistical modeling and information theoretic measures. In all of the above point models and their extension to population models, it is assumed that the extracellular space is iso-potential and the majority of studies thereby neglect the effect of extracellular resistance. That is, the extracellular space constitutes a different and isolated domain with no effect on neuronal dynamics.

In this article we extend the *ad hoc* model of Mazzoni et al. (2008) toward a biophysically better justified approach, taking the dipole character of extracellular currents and fields into account. Basically, our model corresponds to the “two-monopole,” or, equivalent dipole model of Lindén et al. (2010) which gave a good fit of the DFP close to the cell body of a cortical pyramidal neuron. However, we aim to keep the simplicity of the Mazzoni et al. (2008) model in terms of computational complexity, by endowing the extracellular space with resistance and by keeping point-like neuronal circuits. That is,

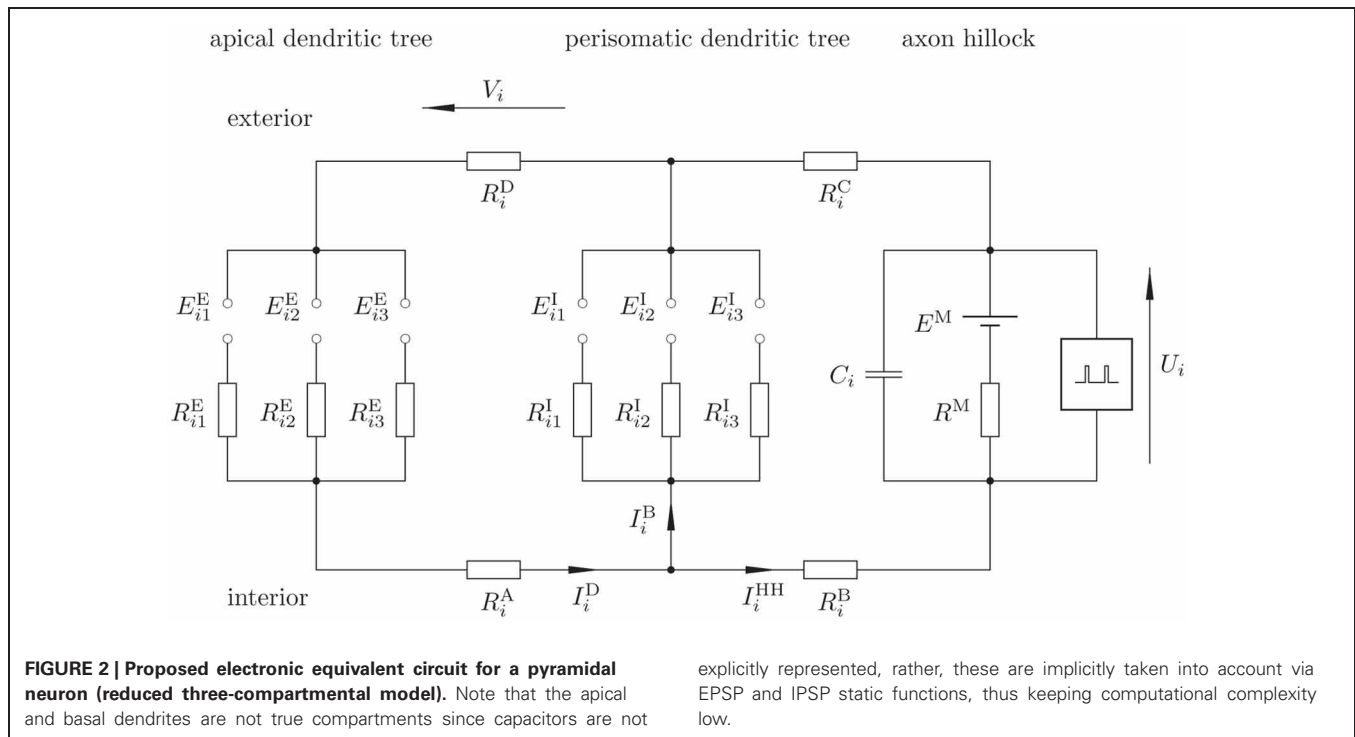
in our case we do not quite consider point neurons, nor spatially extended models with detailed compartmental morphology, yet an intermediate level of description is achieved. To this end we propose a *reduced three-compartmental model* of a single pyramidal neuron (Destexhe, 2001; Wang et al., 2004; beim Graben, 2008), and derive an observation model for the dendritic dipole currents in the extracellular space and thereby for the DFP that contributes to the LFP of a neural population. Interestingly, our reduced three-compartmental model enables us to derive a leaky integrate-and-fire (LIF) mechanism [as for a point model (Mazzoni et al., 2008)], with additional observation equations for the DFP, which all together allows to study the relationship between spike rates and LFP. Our derivations also nicely map realistic electrotonic parameters to phenomenological parameters considered in Mazzoni et al. (2008).

2. MATERIALS AND METHODS

Mazzoni et al. (2008) consider three populations of neurons, namely excitatory cortical pyramidal cells (population 1), inhibitory cortical interneurons (population 2), and excitatory thalamic relay neurons (population 3), passing sensory input to the cortex that is simulated by a random (Erdős–Rényi) graph of $K = 4000$ pyramidal and $L = 1000$ interneurons with connection probability $P = 0.2$.

2.1. THEORY

We describe the i th cortical pyramidal neuron (Figure 1) from population 1 via the electronic equivalent (reduced) three-compartment model (Figure 2) (Destexhe, 2001; Wang et al., 2004; beim Graben, 2008), which is parsimonious to derive our observation model: one compartment for the apical dendritic



tree, another one for soma and perisomatic basal dendritic tree (Lindén et al., 2010), and the third—actually a LIF unit—for the axon hillock where membrane potential is converted into spike trains by means of an integrate-and-fire mechanism.

Excitatory synapses are represented by the left-most branch, where EPSP at a synapse between a neuron j from population 1 or 3 and neuron i act as electromotoric forces E_{ij}^E . These potentials drive EPSC I_{ij}^E , essentially consisting of sodium ions, through the cell plasma with resistance R_{ij}^E from the synapse toward the axon hillock.

The middle branch describes the inhibitory synapses between a neuron k from population 2 and neuron i . Here, inhibitory postsynaptic potentials (IPSP) E_{ik}^I provide a shortcut between the excitatory branch and the trigger zone, where inhibitory postsynaptic currents (IPSC) I_{ik}^I (essentially chloride ions) close the loop between the apical and perisomatic dendritic trees. The resistivity of the current paths along the cell plasma is given by R_{ik}^I .

The cell membrane at the axon hillock itself is represented by the branch at the right hand side. Here, a capacitor C_i reflects the temporary storage capacity of the membrane. The serial circuit consisting of a battery E^M and a resistor R^M denotes the Nernst resting potential and the leakage conductance of the membrane, respectively (Johnston and Wu, 1997). Finally, a spike generator (Hodgkin and Huxley, 1952; Mazzoni et al., 2008) (indicated by a “black box”) is regarded of having infinite input impedance. Both, EPSP and IPSP result from the interaction of postsynaptic receptor kinetics with dendritic low-pass filtering in compartments one and two, respectively (Destexhe et al., 1998; Lindén et al., 2010). Hence the required capacitances, omitted in **Figure 2**, are already taken into account by E_{ij}^E , E_{ik}^I . Therefore, we refer to our model as to a “reduced compartment model” here.

The three compartments are coupled through longitudinal resistors, R_i^A , R_i^B , R_i^C , and R_i^D where R_i^A , R_i^B denote the resistivity of the cell plasma and R_i^C , R_i^D that of extracellular space (Holt and Koch, 1999).

Finally, the membrane voltage at the axon hillock U_i (the dynamical state variable) and the DFP V_i , which measures the drop in electrical potential along the extracellular resistor R_i^D are indicated. For the aim of calculation, the mesh currents I_i^D (the dendritic current), I_i^B (the basal current), and I_i^{IF} (the integrate-and-fire current) are indicated.

The circuit in **Figure 2** obeys the following equations:

$$I_i^D = \sum_{j=1}^p I_{ij}^E \quad (1)$$

$$I_i^B = \sum_{k=1}^q I_{ik}^I \quad (2)$$

$$I_i^{IF} = I_i^D - I_i^B \quad (3)$$

$$I_i^{IF} = C_i \frac{dU_i}{dt} + \frac{U_i - E^M}{R^M} \quad (4)$$

$$E_{ij}^E = R_{ij}^E I_{ij}^E + (R_i^A + R_i^D) I_i^D + (R_i^B + R_i^C) I_i^{IF} + U_i, 1 \leq j \leq p \quad (5)$$

$$E_{ik}^I = R_{ik}^I I_{ik}^I + (R_i^B + R_i^C) I_i^{IF} + U_i, 1 \leq k \leq q \quad (6)$$

$$V_i = R_i^D I_i^D. \quad (7)$$

Here, p is the number of excitatory and q is the number of inhibitory synapses connected to neuron i .

The circuit described by Equations (1–7) shows that the neuron i is likely to fire when the excitatory synapses are activated. Then, the integrate-and-fire current I_i^{IF} equals the dendritic current I_i^{D} . If, by contrast, also the inhibitory synapses are active, the dendritic current I_i^{D} is shunted between the apical and perisomatic basal dendritic trees and only a portion could evoke spikes at the trigger zone (Equation 4). On the other hand, the large dendritic current I_i^{D} flowing through the extracellular space of resistance R_i^{D} , gives rise to a large DFP V_i .

In order to simplify the following derivations, we gauge the resting potential (Equation 4) to $E^{\text{M}} = 0$, yielding

$$I_i^{\text{IF}} = C_i \frac{dU_i}{dt} + \frac{U_i}{R^{\text{M}}}. \quad (8)$$

From Equation (5) we obtain the individual EPSC's as

$$I_{ij}^{\text{E}} = \frac{1}{R_{ij}^{\text{E}}} [E_{ij}^{\text{E}} - (R_i^{\text{A}} + R_i^{\text{D}}) I_i^{\text{D}} - (R_i^{\text{B}} + R_i^{\text{C}}) I_i^{\text{IF}} - U_i]. \quad (9)$$

And accordingly, the individual IPSC's from Equation (6)

$$I_{ik}^{\text{I}} = \frac{1}{R_{ik}^{\text{I}}} [E_{ik}^{\text{I}} - (R_i^{\text{B}} + R_i^{\text{C}}) I_i^{\text{IF}} - U_i]. \quad (10)$$

Inserting Equation (9) into Equation (1) yields the excitatory dendritic current

$$I_i^{\text{D}} = \sum_{j=1}^p \frac{1}{R_{ij}^{\text{E}}} E_{ij}^{\text{E}} - g_i^{\text{E}} [(R_i^{\text{A}} + R_i^{\text{D}}) I_i^{\text{D}} + (R_i^{\text{B}} + R_i^{\text{C}}) I_i^{\text{IF}} + U_i], \quad (11)$$

where we have introduced the excitatory dendritic conductivity

$$g_i^{\text{E}} = \sum_{j=1}^p \frac{1}{R_{ij}^{\text{E}}}. \quad (12)$$

Likewise we obtain the inhibitory dendritic currents from Equations (2) and (10) as

$$I_i^{\text{B}} = \sum_{k=1}^q \frac{1}{R_{ik}^{\text{I}}} E_{ik}^{\text{I}} - g_i^{\text{I}} [(R_i^{\text{B}} + R_i^{\text{C}}) I_i^{\text{IF}} + U_i], \quad (13)$$

with the inhibitory dendritic conductivity

$$g_i^{\text{I}} = \sum_{k=1}^q \frac{1}{R_{ik}^{\text{I}}}. \quad (14)$$

With these results, we obtain an interface equation for an observation model as follows. Rearranging Equation (11) yields

$$I_i^{\text{D}} [1 + g_i^{\text{E}} (R_i^{\text{A}} + R_i^{\text{D}})] = \sum_{j=1}^p \frac{1}{R_{ij}^{\text{E}}} E_{ij}^{\text{E}} - g_i^{\text{E}} [(R_i^{\text{B}} + R_i^{\text{C}}) I_i^{\text{IF}} + U_i] \quad (15)$$

Next, we eliminate I_i^{IF} through Equation (8):

$$I_i^{\text{D}} [1 + g_i^{\text{E}} (R_i^{\text{A}} + R_i^{\text{D}})] = \sum_{j=1}^p \frac{1}{R_{ij}^{\text{E}}} E_{ij}^{\text{E}} - g_i^{\text{E}} \times \left[C_i (R_i^{\text{B}} + R_i^{\text{C}}) \frac{dU_i}{dt} + U_i \left(1 + \frac{R_i^{\text{B}} + R_i^{\text{C}}}{R^{\text{M}}} \right) \right].$$

Division by $1 + g_i^{\text{E}} (R_i^{\text{A}} + R_i^{\text{D}})$ gives the desired expression for the extracellular dendritic dipole current:

$$I_i^{\text{D}} = \sum_{j=1}^p \alpha_{ij} E_{ij}^{\text{E}} - \beta_i \frac{dU_i}{dt} - \gamma_i U_i, \quad (16)$$

with the following electrotonic parameters

$$\alpha_{ij} = \frac{1}{R_{ij}^{\text{E}} [1 + g_i^{\text{E}} (R_i^{\text{A}} + R_i^{\text{D}})]} \quad (17)$$

$$\beta_i = \frac{C_i g_i^{\text{E}} (R_i^{\text{B}} + R_i^{\text{C}})}{1 + g_i^{\text{E}} (R_i^{\text{A}} + R_i^{\text{D}})} \quad (18)$$

$$\gamma_i = \frac{g_i^{\text{E}} (R^{\text{M}} + R_i^{\text{B}} + R_i^{\text{C}})}{R^{\text{M}} [1 + g_i^{\text{E}} (R_i^{\text{A}} + R_i^{\text{D}})]}. \quad (19)$$

In order to derive the evolution equation we consider the integrate-and-fire current I_i^{IF} that is given through Equation (3). The individual EPSCs and IPSCs have already been obtained in Equations (9) and (10), respectively. Inserting Equation (13) into Equation (3) yields

$$I_i^{\text{IF}} [1 - g_i^{\text{I}} (R_i^{\text{B}} + R_i^{\text{C}})] - g_i^{\text{I}} U_i = I_i^{\text{D}} - \sum_{k=1}^q \frac{1}{R_{ik}^{\text{I}}} E_{ik}^{\text{I}}.$$

Next we insert our interface equation Equation (16) and also Equation (8):

$$\begin{aligned} & \left[C_i \frac{dU_i}{dt} + \frac{U_i}{R^{\text{M}}} \right] [1 - g_i^{\text{I}} (R_i^{\text{B}} + R_i^{\text{C}})] - g_i^{\text{I}} U_i \\ &= \sum_{j=1}^p \alpha_{ij} E_{ij}^{\text{E}} - \beta_i \frac{dU_i}{dt} - \gamma_i U_i - \sum_{k=1}^q \frac{1}{R_{ik}^{\text{I}}} E_{ik}^{\text{I}} \end{aligned}$$

and obtain after some rearrangements

$$\begin{aligned} & \{C_i [1 - g_i^{\text{I}} (R_i^{\text{B}} + R_i^{\text{C}})] + \beta_i\} \frac{dU_i}{dt} \\ &+ \frac{1 - g_i^{\text{I}} (R_i^{\text{B}} + R_i^{\text{C}} + R^{\text{M}}) + R^{\text{M}} \gamma_i}{R^{\text{M}}} U_i = \sum_{j=1}^p \alpha_{ij} E_{ij}^{\text{E}} - \sum_{k=1}^q \frac{1}{R_{ik}^{\text{I}}} E_{ik}^{\text{I}} \end{aligned}$$

and after multiplication with

$$r_i = \frac{R^{\text{M}}}{1 - g_i^{\text{I}} (R_i^{\text{B}} + R_i^{\text{C}} + R^{\text{M}}) + R^{\text{M}} \gamma_i}$$

the dynamical law for the membrane potential at axon hillock:

$$\tau_i \frac{dU_i}{dt} + U_i = \sum_{j=1}^p w_{ij}^E E_{ij}^E - \sum_{k=1}^q w_{ik}^I E_{ik}^I, \quad (20)$$

where we have introduced the following parameters:

- *time constants*

$$\tau_i = r_i \{C_i [1 - g_i^I (R_i^B + R_i^C)] + \beta_i\} \quad (21)$$

- *excitatory synaptic weights*

$$w_{ij}^E = r_i \alpha_{ij} \quad (22)$$

- *inhibitory synaptic weights*

$$w_{ik}^I = \frac{r_i}{R_{ik}^I}. \quad (23)$$

Using the result Equation (20), we can also eliminate the temporal derivative in the interface equation Equation (16) through

$$\frac{dU_i}{dt} = \frac{1}{\tau_i} \left[\sum_{j=1}^p w_{ij}^E E_{ij}^E - \sum_{k=1}^q w_{ik}^I E_{ik}^I - U_i \right] \quad (24)$$

which yields

$$I_i^D = \sum_{j=1}^p \left(\alpha_{ij} - \frac{\beta_i}{\tau_i} w_{ij}^E \right) E_{ij}^E + \sum_{k=1}^q \frac{\beta_i}{\tau_i} w_{ik}^I E_{ik}^I + \left(\frac{\beta_i}{\tau_i} - \gamma_i \right) U_i.$$

And eventually, by virtue of Equation (7) after multiplication with R_i^D the DFP

$$V_i = \sum_{j=1}^p \tilde{w}_{ij}^E E_{ij}^E + \sum_{k=1}^q \tilde{w}_{ik}^I E_{ik}^I + \xi_i U_i, \quad (25)$$

with parameters

$$\tilde{w}_{ij}^E = R_i^D w_{ij}^E \left(\frac{1}{r_i} - \frac{\beta_i}{\tau_i} \right) \quad (26)$$

$$\tilde{w}_{ik}^I = R_i^D w_{ik}^I \frac{\beta_i}{\tau_i} \quad (27)$$

$$\xi_i = R_i^D \left(\frac{\beta_i}{\tau_i} - \gamma_i \right). \quad (28)$$

The change in sign of the inhibitory contribution from Equation (20) to Equation (25) has an obvious physical interpretation: In Equation (20), the change of membrane potential U_i and therefore the spike rate is enhanced by EPSPs but diminished by IPSPs. On the other hand, the dendritic shunting current I_i^D in Equation (25) is large for both, large EPSPs and large IPSPs.

From Equation (20) we eventually obtain the neural network's dynamics by taking into account that postsynaptic potentials

are obtained from presynaptic spike trains through temporal convolution with postsynaptic impulse response functions, i.e.,

$$E_{ij}^{E|I}(t) = \int_{-\infty}^t s_i^{E|I}(t-t') R_j(t') dt' \quad (29)$$

where $s_i^{E|I}(t)$ are excitatory and inhibitory synaptic impulse response functions, respectively, and R_j is the spike train

$$R_j(t) = \sum_{t_v} \delta(t - t_v - \tau_L) \quad (30)$$

coming from presynaptic neuron j , when spikes were emitted at times t_v . The additional time constant τ_L is attributed to synaptic transmission delay (Mazzoni et al., 2008). These events are obtained by integrating Equation (20) with initial condition

$$U_i(t_v) = E. \quad (31)$$

where E is some steady-state potential (Mazzoni et al., 2008). If at time $t = t_v$ the membrane reaches a threshold

$$U_i(t) \geq \theta_i(t) \quad (32)$$

[with possibly a time-dependent activation threshold $\theta_i(t)$] from below $\frac{dU_i(t)}{dt} > 0$ then an output spike $\delta(t - t_v)$ is generated, which is then followed by a potential resetting as follows

$$U_i(t_{v+1}) \leftarrow E. \quad (33)$$

Additionally, the integration of the dynamical law is restarted at time $t = t_{v+1} + \tau_{rp}$ after interrupting the dynamics for a refractory period τ_{rp} .

Inserting Equation (29) into Equation (20) entails the evolution equation of the neural network

$$\tau_i \frac{dU_i}{dt} + U_i = \sum_{j=1}^p w_{ij}^E s_i^E(t) * R_j(t) + \sum_{k=1}^q w_{ik}^I s_i^I(t) * R_k(t), \quad (34)$$

where the signs had been absorbed by the synaptic weights, such that $w_{ij}^E > 0$ for excitatory synapses and $w_{ik}^I < 0$ for inhibitory synapses, respectively.

Following Mazzoni et al. (2008) an individual postsynaptic current $I_{ij}^{E|I}$ at a synapse between neurons i and j obeys

$$\tau_d^{E|I} \frac{dI_{ij}^{E|I}}{dt} + I_{ij}^{E|I} = x_{ij}^{E|I} \quad (35)$$

$$\tau_r^{E|I} \frac{dx_{ij}^{E|I}}{dt} + x_{ij}^{E|I} = F_{ij}^{E|I}, \quad (36)$$

where $\tau_d^{E|I}$ are decay time constants and $\tau_r^{E|I}$ are rise time constants of EPSC and IPSC, respectively. Auxiliary variables are denoted by

x_{ij}^{EI} , while F_{ij}^{EI} prescribes presynaptic forcing

$$F_{ij}^{\text{EI}} = \tau_i J_{ij} R_j(t) \quad (37)$$

with spike train Equation (30). Here, $J_{ij} = \nu w_{ij}^{\text{EI}}$ denotes synaptic gain with $\nu = 1$ mV as voltage unit.

Note that Equation (37) is essentially a weighted sum of delta functions, such that a single spike can be assumed as particular forcing

$$F = F_0 \delta(t), \quad (38)$$

with some constant F_0 .

Derivating Equation (35) and eliminating x_{ij}^{EI} transforms Equations (35, 36) into a linear second-order differential equation with constant coefficients

$$\tau_d^{\text{EI}} \tau_r^{\text{EI}} \frac{d^2 I_{ij}^{\text{EI}}}{dt^2} + (\tau_d^{\text{EI}} + \tau_r^{\text{EI}}) \frac{d I_{ij}^{\text{EI}}}{dt} + I_{ij}^{\text{EI}} = F_{ij}^{\text{EI}}. \quad (39)$$

Equation (39) with the particular forcing Equation (38) is solved by a Green's function $s_i^{\text{EI}}(t)$ such that the general solution of Equation (39) is obtained as the temporal convolution

$$I_{ij}^{\text{EI}}(t) = \int_{-\infty}^t s_i^{\text{EI}}(t - t') F_{ij}^{\text{EI}}(t) dt'. \quad (40)$$

For $t \neq 0$, Equation (39) assumes its homogeneous form and is easily solved by means of the associated characteristic polynomial

$$\tau_d^{\text{EI}} \tau_r^{\text{EI}} \lambda^2 + (\tau_d^{\text{EI}} + \tau_r^{\text{EI}}) \lambda + 1 = 0 \quad (41)$$

with roots $\lambda_1 = -1/\tau_d^{\text{EI}}$ and $\lambda_2 = -1/\tau_r^{\text{EI}}$, entailing the Green's functions

$$s_i^{\text{EI}}(t) = (A^{\text{EI}} e^{t/\tau_r^{\text{EI}}} - B^{\text{EI}} e^{t/\tau_d^{\text{EI}}}) \Theta(t) \quad (42)$$

with the Heaviside step function $\Theta(t)$.

The constants $A^{\text{EI}}, B^{\text{EI}} > 0$ are obtained from the initial conditions $s_i^{\text{EI}}(t) = 0$, reflecting causality, and a suitable normalization

$$\int_0^\infty s_i^{\text{EI}}(t) dt = 1.$$

The initial condition yields $A^{\text{EI}} = B^{\text{EI}} \equiv S^{\text{EI}}$, while the remaining constant

$$S^{\text{EI}} = \frac{1}{\tau_d^{\text{EI}} - \tau_r^{\text{EI}}},$$

due to normalization. Therefore, the normalized Green's functions are those of Brunel and Wang (2003)

$$s_i^{\text{EI}}(t) = \nu \frac{\tau_i}{\tau_d^{\text{EI}} - \tau_r^{\text{EI}}} (e^{t/\tau_r^{\text{EI}}} - e^{t/\tau_d^{\text{EI}}}) \Theta(t). \quad (43)$$

Now, we are able to compare our DFP V_i (Equation 25) with the estimate of Mazzoni et al. (2008) which is given (in our notation) as the sums of the moduli of excitatory and inhibitory synaptic currents, i.e.,

$$V_i^{\text{MPLB}} = \sum_j |I_{ij}^{\text{E}}| + \sum_k |I_{ik}^{\text{I}}| \quad (44)$$

where “MPLB” refers to the authors Mazzoni et al. (2008).

From Equations (25) and (44), respectively, we compute two models of the LFP. First, by summing DFP across all pyramidal neurons (beim Graben and Kurths, 2008; Mazzoni et al., 2008), and, second by taking the DFP average (Nunez and Srinivasan, 2006), which yields

$$L_1 = \sum_i V_i^{\text{MPLB}} \quad (45)$$

$$L_2 = \frac{1}{K} \sum_i V_i^{\text{MPLB}} \quad (46)$$

$$L_3 = \sum_i V_i \quad (47)$$

$$L_4 = \frac{1}{K} \sum_i V_i, \quad (48)$$

where K is number of pyramidal neurons.

2.2. PARAMETER ESTIMATION

Next, we relate the electrotonic parameters of our model to the phenomenological parameters of Mazzoni et al. (2008). To this end, we first report their synaptic efficacies in **Table 1**.

From these, we compute the synaptic weights through

$$w_{ij}^{\text{E}} = J_{ij}^{\text{E}}/\nu = \begin{cases} 0.42 & \text{if } j \text{ “cortical”} \\ 0.55 & \text{if } j \text{ “thalamic”} \end{cases} \quad (49)$$

and

$$w_{ik}^{\text{I}} = J_{ik}^{\text{I}}/\nu = 1.7$$

Next, we determine the factors r_i by virtue of Equation (23) through

$$r_i = \frac{w_{ik}^{\text{I}}}{\bar{g}_{\text{GABA}}} = \frac{1.7}{1 \text{ nS}} = 1.7 \text{ G}\Omega$$

using the inhibitory synaptic conductivity $\bar{g}_{\text{GABA}} = 1$ nS, correspondingly, Equation (22) allows us to express α_{ij} in terms of the

Table 1 | Parameters laid as in Mazzoni et al. (2008).

Synaptic efficacies/mV	On interneurons	On pyramidal neurons
GABA	2.7	1.7
Recurrent cortical AMPA	0.7	0.42
External thalamic AMPA	0.95	0.55

excitatory synaptic weights through

$$\alpha_{ij} = \frac{w_{ij}^E}{r_i} = \begin{cases} 0.25 \text{ nS} & \text{if } j \text{ "cortical"} \\ 0.32 \text{ nS} & \text{if } j \text{ "thalamic"} \end{cases}$$

From α_{ij} we can determine the total excitatory synaptic conductivities g_i^E according to Equation (17) through

$$\begin{aligned} \alpha_{ij} &= \frac{1}{R_{ij}^E [1 + g_i^E (R_i^A + R_i^D)]} \\ g_i^E \left[1 - (R_i^A + R_i^D) \sum_{j=1}^p \alpha_{ij} \right] &= \sum_{j=1}^p \alpha_{ij} \\ g_i^E &= \frac{\sum_{j=1}^p \alpha_{ij}}{1 - (R_i^A + R_i^D) \sum_{j=1}^p \alpha_{ij}} \end{aligned} \quad (50)$$

and hence

$$R_{ij}^E = \frac{1}{\alpha_{ij} [1 + g_i^E (R_i^A + R_i^D)]} \quad (51)$$

Inserting next Equation (18) into Equation (21) yields

$$\tau_i = r_i C_i \frac{1 + g_i^E (R_i^A + R_i^D) + (R_i^B + R_i^C) \{g_i^E - g_i^I [1 + g_i^E (R_i^A + R_i^D)]\}}{1 + g_i^E (R_i^A + R_i^D)} \quad (52)$$

Equation (52) could constraint the choice of the membrane capacitance C_i by choosing $\tau_i = 20 \text{ ms}$ (Mazzoni et al., 2008).

In order to also determine the DFP parameters Equations (26–28), we finally compute the ratios

$$\frac{\beta_i}{\tau_i} = \frac{g_i^E (R_i^B + R_i^C)}{r_i \{1 + g_i^E (R_i^A + R_i^D) + (R_i^B + R_i^C) \{g_i^E - g_i^I [1 + g_i^E (R_i^A + R_i^D)]\}\}}.$$

The remaining electrotonic parameters R_i^M , R_i^A , R_i^B , R_i^C , and R_i^D are estimated from cell geometries as follows. The resistance R of a volume conductor is proportional to its length ℓ and reciprocally proportional to its cross-section A , i.e.,

$$R = \rho \frac{\ell}{A} \quad (53)$$

where ρ is the (specific) resistivity of the medium. **Table 2** shows the resistivities of the three kinds of interest which then allows to evaluate the various volume conductor resistances according to Equation (53).

We consider a total dendritic length of $2\ell = 20 \mu\text{m}$ and a dendritic radius of $a = 7 \mu\text{m}$, that are generally subjected to variation. Equally, parameters that were allowed to vary are the length and radius of the axon hillock, yet herein we consider a length of $2\ell = 20 \mu\text{m}$ and radius of $a = 0.5 \mu\text{m}$ (Mainen et al., 1995; Destexhe, 2001; Kole and Stuart, 2012). To evaluate the intracellular (R_A , R_B) and extracellular (R_D , R_C) resistances, respectively, according to Equation (53), we consider a simple implementation

Table 2 | Resistivities of cell membrane, cell plasma and extracellular space.

Medium	$\rho/\Omega\text{cm}$
Cell membrane (at axon hillock)	5×10^7
Cell plasma (cytoplasm)	200
Extracellular space	333

Parameters from Rall (1977); Mainen et al. (1995); Kole and Stuart (2012), and Gold et al. (2007). Note that the resistivity of the cell membrane has to be related to the constant membrane thickness ($\approx 10 \text{ nm}$).

where the length ℓ is half of the dendritic length (i.e., basal and apical length are symmetrical, but this can be broken). However, the cross sectional area for the cytoplasm is simply $A = \pi a^2$. Finally, the area of the axon hillock is simply the surface area of a cylinder.

In order to also determine the cross-section of extracellular space between dendritic trunks we make the following approximations. We assume that dendritic trunks are parallel aligned cylinders of radius a and length ℓ that are hexagonally dense packed. Then the centers of three adjacent trunks form an equilateral triangle with side length $2a$ and hence area $2\sqrt{3}a^2$. The enclosed space is then given by the difference of the triangle area and the area of three sixth circle sectors, therefore

$$A_{\text{space}} = 2\sqrt{3}a^2 - \frac{3}{6}\pi a^2 = \left(2\sqrt{3} - \frac{1}{2}\pi\right)a^2.$$

Hence, the cross-section of extracellular space surrounding one trunk is

$$A = 6A_{\text{space}} = \left(12\sqrt{3} - 3\pi\right)a^2. \quad (54)$$

2.3. SIMULATIONS

Subsequently, we implement an identical network to the one considered by Mazzoni et al. (2008) with *Brian Simulator*, that is a Python-based environment (Goodman and Brette, 2009). However, the derivations from the previous section enables the possibility of setting a dipole observable that measures the local DFP on each pyramidal neurons, given by Equation (25). This allows then to define a mesoscopic LFP observable, which can be equated either as averaged DFP or simply given as the sum of DFP, given by Equations (45–48). Primarily, we compare our LFP measure L_4 , proposed as the average of DFP, with the Mazzoni et al. LFP L_1 which is defined as the sum of absolute values of GABA and AMPA currents (Equation 44). Additionally, we also compare all possible measures, namely, mean membrane potential $\frac{1}{K} \sum_i U_i$, Mazzoni et al. LFP L_1 , average of Mazzoni et al. DFP L_2 , sum of DFP L_3 , and the average of DFP L_4 .

For completeness, we briefly summarize the description of the network [we refer the reader to Mazzoni et al. (2008) for details]. The network models a cortical tissue with LIF neurons, composed of 1000 inhibitory interneurons and 4000 pyramidal neurons, which are described by the evolution Equation (34). The threshold crossings given by Equation (32) is considered static with

$\theta_i = 18$ mV and the reset potential $E = 11$ mV. The refractory period for excitatory neurons is $\tau_{rp} = 2$ ms while for inhibitory neurons it is $\tau_{rp} = 1$ ms. The network connectivity is random and sparse with a 0.2 probability of directed connection between any pair of neurons. The evolution of synaptic currents, fast GABA (inhibitory) and AMPA (excitatory) are described via the second order evolution Equations (35, 36), which are activated by incoming presynaptic spikes represented by Equation (30). The latency of the postsynaptic currents is set to $\tau_L = 1$ ms and the rise and decay times are given by **Table 3**.

Moreover, synaptic efficacies, J_{ij}^{EI} , for simulation were presented in **Table 1**. Note that Relation (49) then allows to determine the synaptic weights. Additionally, all neurons receive external thalamic excitatory inputs, that is, via AMPA-type synapses, which are activated by random Poisson spike trains, with a time varying rate that is identical for all neurons. Specifically, the thalamic inputs are the only source of noise, which attempts to account for both cortical heterogeneity and spontaneous activity. This is achieved by modeling a two level noise, where the first level is an Ornstein–Uhlenbeck process superimposed with a constant signal and the second level is a time varying inhomogeneous Poisson process. Thus, we have the following time varying rate, $\lambda(t)$, that feeds into inhomogeneous Poisson process:

$$\tau_n \frac{dn(t)}{dt} = -n(t) + \sigma_n \sqrt{\frac{2}{\tau_n}} \eta(t) \quad (55)$$

$$\lambda(t) = [c_0 + n(t)]_+ \quad (56)$$

where $\eta(t)$ represents Gaussian white noise, c_0 represents a constant signal (but equally could be periodic or other), and the operation $[\cdot]_+$ is the threshold-linear function, $[x]_+ = x$ if $x > 0$, $[x]_+ = 0$ otherwise, which circumvents negative rates. The constant signal c_0 can range between 1.2 and 2.6 spikes/ms. The parameters of the Ornstein–Uhlenbeck process are $\tau_n = 16$ ms and the standard deviation $\sigma_n = 0.4$ spikes/ms.

For complete exposition, we note that from an implementation viewpoint (within the Brian simulator), a copy of the postsynaptic impulse response function (Equation 29) has to be evaluated to calculate the DFP (Equation 25) with weights \tilde{w}_{ij}^{EI} . This implies evaluating the second order process (Equations 35, 36) with a different forcing term. Specifically, starting from $I_{ij}^{EI}(t) \equiv w_{ij}^{EI} E_{ij}^{EI}(t) = s_i^{EI}(t) * F_{ij}^{EI}$ and pre-multiplying both sides with \tilde{w}_{ij}^{EI} and subsequently re-arranging we obtain the desired forcing term $\tilde{F}_{ij}^{EI} = \tilde{w}_{ij}^{EI} F_{ij}^{EI} / w_{ij}^{EI}$. Note

further that by expanding the term F_{ij}^{EI} with Equation (37) and using Relation (49) we finally obtain $\tilde{F}_{ij}^{EI} = \tilde{w}_{ij}^{EI} \tau_i \nu R_j(t)$.

3. RESULTS

Following Mazzoni et al. (2008), the network simulations are run for 2 s with three different noise levels, specifically, receiving a constant signal with three different rates 1.2, 1.6, and 2.4 spikes/ms as depicted in **Figure 3**. Note that these input rates do not mean that a single neuron fires at these high rates. Rather, it can be obtained from multiple neurons that jointly fire with slower, yet desynchronized, rates converging at the same postsynaptic cell. The Poisson process ensures that this is well represented.

The focus is to compare our proposed measure L_4 , defined as mean of the DFP (Equation 48), with the Mazzoni et al. LFP L_1 from Equation (45). In **Figure 3** one sees two main striking differences between the two measures, namely in frequency and in amplitude. Specifically, L_1 responds instantaneously to the spiking network activity by means of high frequency oscillations. Moreover, L_1 also exhibits a large amplitude. In contrast, our mean DFP L_4 measures comparably to experimental LFP, that is, in the order of millivolts, and although it responds to population activity, it has a relatively smoother response. Actually one can realize that our LFP estimate represents low-pass filtered thalamic input.

The physiological relevance of this is not yet clear in our work. However, recent work (Poulet et al., 2012) shows that desynchronized cortical state during active behavior is driven by a centrally generated increase in thalamic action potential firing (i.e., thalamic firing controls cortical states). Thus, it seems that cortical synchronous activity is suppressed when thalamic input increases, thereby suggesting that cortical desynchronized states to be related to sensory processing. This work further quantifies these observations by applying Fast Fourier Transform (FFT) to cortical EEG and subsequently comparing with thalamic firing rate by means of Pearson correlation coefficient. Unfortunately they do not quantify the amount of thalamic oscillations contained within the cortical EEG.

Yet, to keep a comparable comparison between measures, we also compute the average of the Mazzoni et al. DFP L_2 (Equation 48) and additionally the mean membrane potential (the standard considered in the neuroscientific literature). These are shown in **Figure 4**.

Clearly, in terms of time profile, the summed and averaged observables are similar within the same class of LFP measures. However, in all cases the Mazzoni et al. LFP L_1 exhibits a significantly larger order of magnitude, which diverges substantially from experimental LFP amplitudes, typically varying between 0.5 and 2 mV (Lakatos et al., 2005; Niedermeyer, 2005). In contrast, although the mean DFP is not contained within the interval from 0.5 to 2 mV it arguably performs better. However, we do concede further work is required. Some gains in improving the different LFP measures can be achieved by applying for example a weighted average, which would mimic the distance of an electrode to a particular neuron by means of a lead field kernel (Nunez and Srinivasan, 2006). For example, a convolution of either L_1

Table 3 | Synaptic rise (τ_r) and decay times (τ_d).

Synaptic times	τ_r /ms	τ_d /ms
GABA	0.25	5
AMPA on interneurons	0.2	1
AMPA on pyramidal neurons	0.4	2

Parameters laid as in Mazzoni et al. (2008).

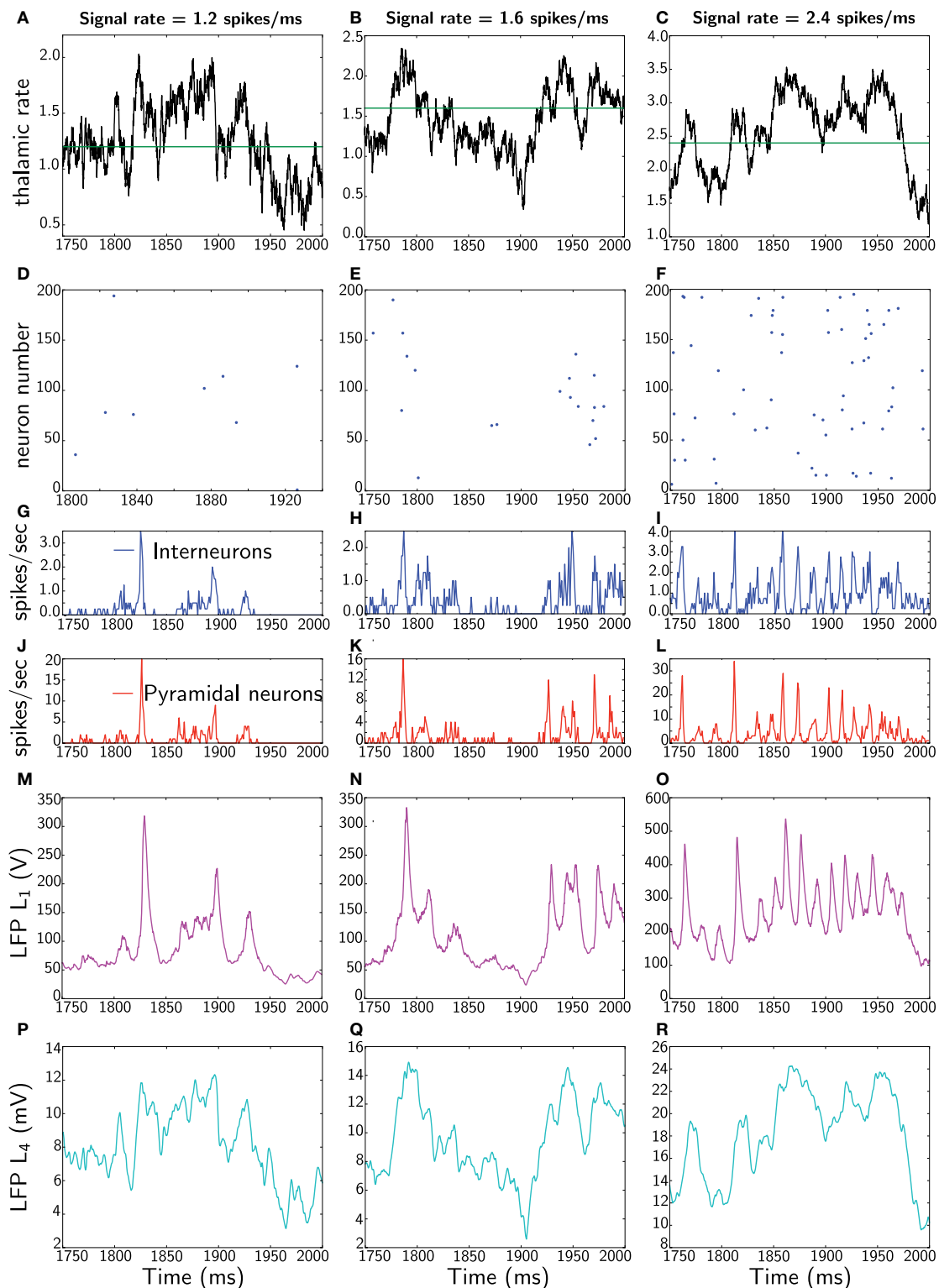


FIGURE 3 | Dynamics of the network and LFP comparisons: the three columns represent different runs of the network for three different rates, 1.2, 1.6, and 2.4 spikes/ms. In each column, all panels show the same 250 ms (extracted from 2 s simulations). The first panels (A–C) represent thalamic inputs with the different rates. The second panels (D–F) corresponds to a raster plot of the activity of 200 pyramidal neurons. The

third panels (G–I) depict average instantaneous firing rate (computed on a 1 ms bin) of interneurons (blue) and fourth panels (J–L) correspond to average instantaneous firing rate of pyramidal neurons. The fifth panels (M–O) show the Mazzoni et al. LFP L_1 from Equation (45). Finally, the last panels (P–R) depict our proposed LFP measure L_4 , which is the average of dendritic field potential (DFP) (Equation 48).

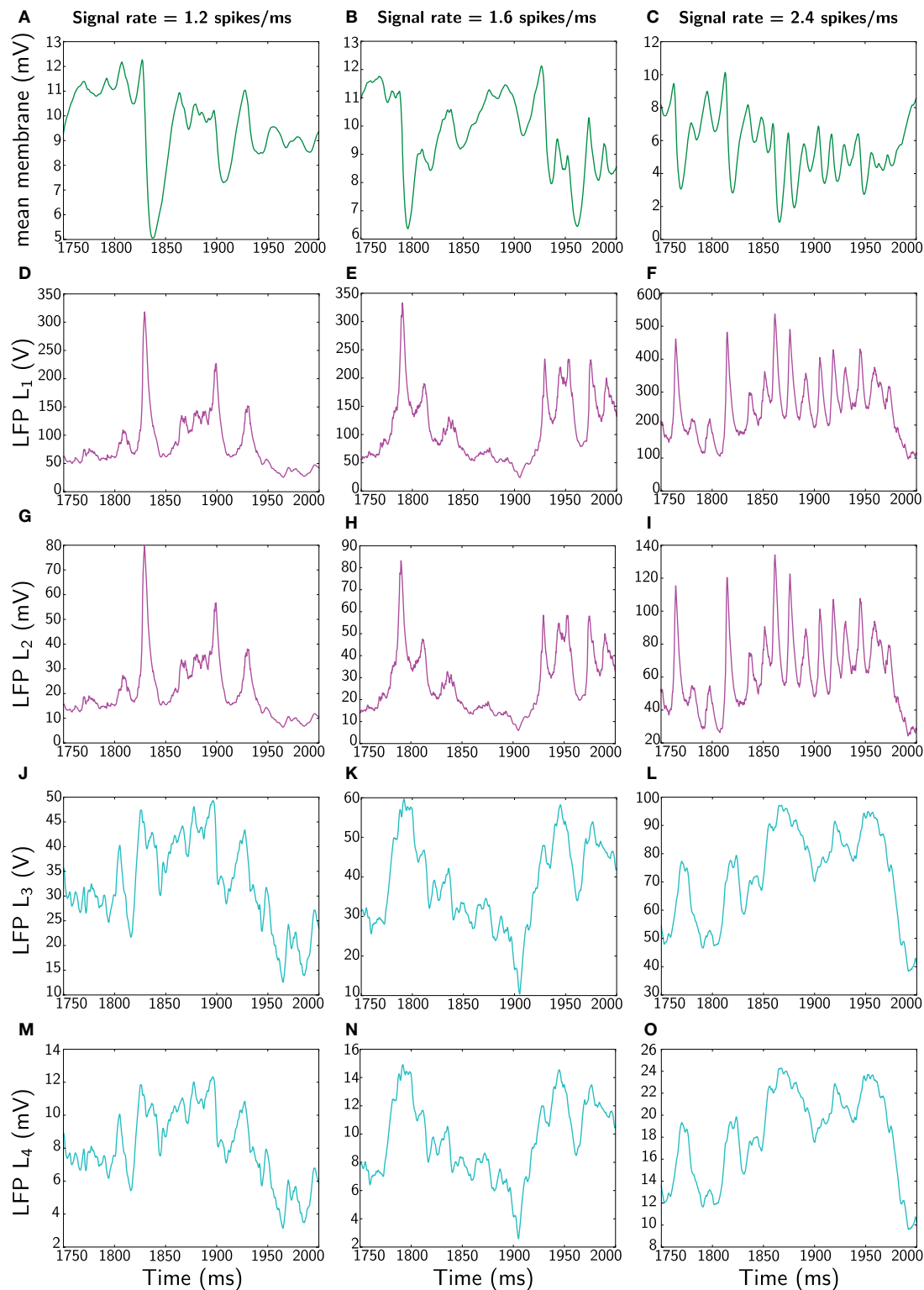


FIGURE 4 | Comparison of different LFP measures when the network receives constant signal with three different rates (1.2, 1.6, and 2.4 spikes/ms). Again, only 250 ms is represented (extracted from 2 s simulation). The first panels (A–C) corresponding to the different rates shows the most widespread LFP measure used in the literature, namely average membrane potential $\frac{1}{N} \sum_i U_i$. The second

panels (D–F) shows the Mazzoni et al. LFP L_1 from Equation (45). The third panels (G–I) displays the average of the Mazzoni et al. DFP L_2 (Equation 46). Similarly, the fourth panels (J–L) shows the total, L_3 , (Equation 47) and the last panels (M–O) depicts the averaged, L_4 , (Equation 48) LFP measure. Note the different amplitude scales between measures.

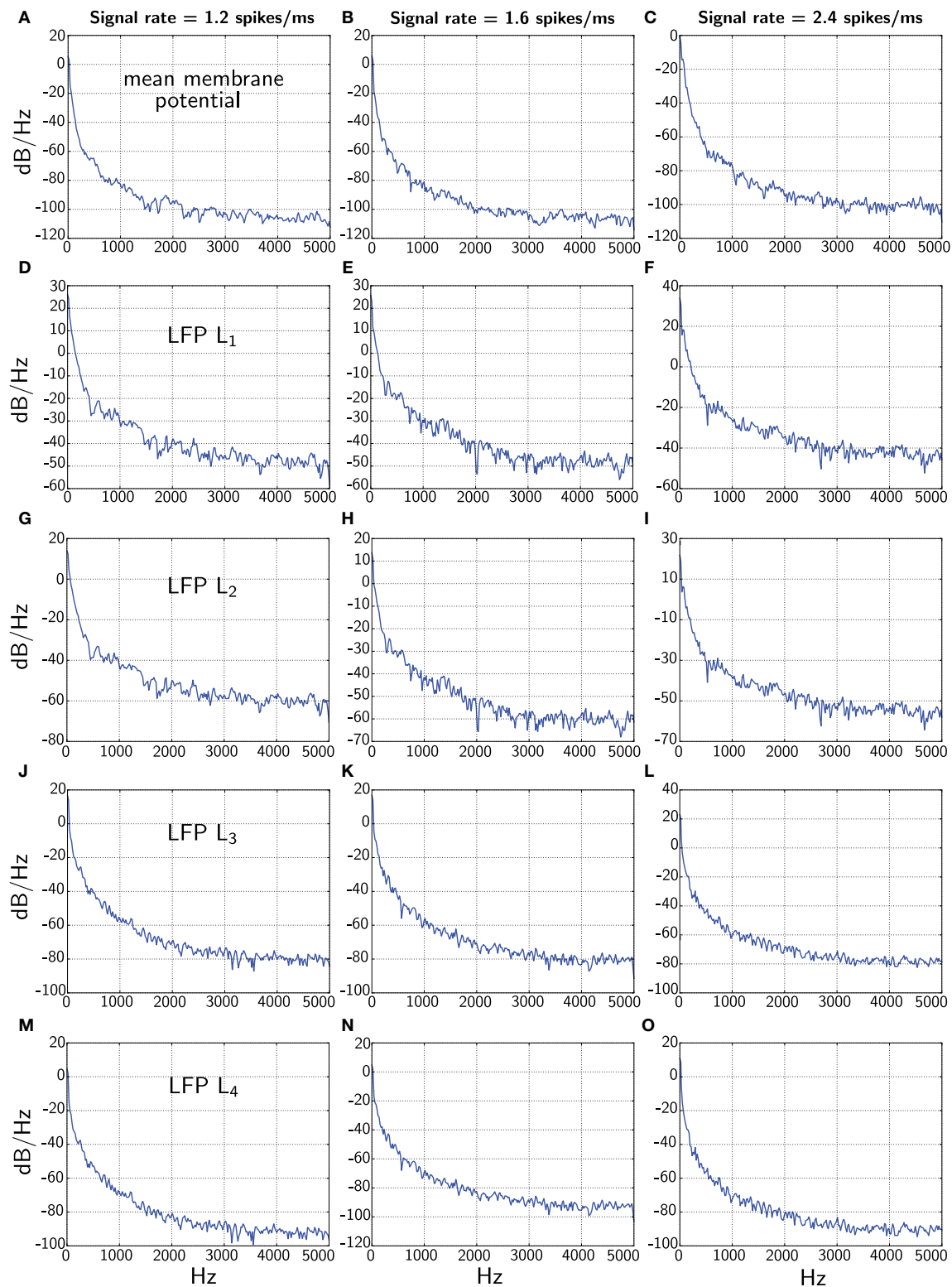


FIGURE 5 | Comparison of power spectra of the various LFP measures when the network receives constant signal with three different rates (1.2, 1.6, and 2.4 spikes/ms). The first panels (A–C) corresponding to the different rates shows the power spectrum of the average membrane potential $\frac{1}{N} \sum_i U_i$. The second panels (D–F) and third panels (G–I) show

power spectra of the total and average of L_1 and L_2 corresponding to Mazzoni et al. (2008), respectively. The fourth panels (J–L) and the last panels (M–O) display power spectra of the L_3 and L_4 measures from our model, respectively. Note we show the full spectrum up to 5 kHz only for convenience due to the fine sample rate.

or L_2 with a Gaussian kernel (representing the distance to a neuron), would yield a measure that captures better the LFP or better the DFP of the nearest neurons. However, further work will be required to properly quantify the gain when space is taken into account.

In **Figure 5** we finally contrast the power spectra of the different LFP measures.

One interesting feature is that the power spectrum of the Mazzoni et al. LFP measures decays much more slowly than the average membrane potential for higher frequencies. This observation is true for both, L_1 and L_2 . In contrast, our LFP measures L_3 and L_4 fare better, and in particular, L_4 decays at an approximately similar rate as the average membrane potential.

4. DISCUSSION

In this article we derived a model for cortical dipole fields, such as DFP/LFP from biophysical principles. To that aim we decomposed a cortical pyramidal cell, the putative generator of those potentials, into three compartments: the apical dendritic tree as the place of mainly excitatory (AMPA) synapses, the soma and the perisomatic dendritic tree as the place of mainly inhibitory (GABA) synapses, and the axon hillock as the place of wave-to-spike conversion by means of an integrate-and-fire mechanism. From Kirchhoff's laws governing an electronic equivalent circuit of our model, we were then able to derive the evolution equation for neural network activity (Equation 34) and, in addition, an observation equation (25) for the dendritic dipole potential contributing to the LFP of a cortical population.

In order to compare our approach with another model discussed in the recent literature (Mazzoni et al., 2008, 2010, 2011) we aligned the parameters of our model with the model of Mazzoni et al. (2008) who approximated DFP as the sum of moduli of excitatory and inhibitory synaptic currents (Equation 44). From both approaches, we computed four different LFP estimates: L_1 , the sum of Mazzoni et al. DFP, L_2 , the population average of Mazzoni et al. DFP, L_3 the sum of our dipole DFP, and L_4 the population average of our dipole DFP (Equations 45–48).

Our results indicate two main effects between our dipole LFP measures and those of Mazzoni et al. Firstly, the measures based on Mazzoni et al. (2008) systematically overestimate LFP amplitude by almost one order of magnitude. One reason for that could be attributed to the direct conversion of synaptic current into voltage without taking extracellular conductivity into account, as properly done in our approach. Yet, another, even more crucial reason is disclosed by our equivalent circuit (**Figure 2**). In our approach there is just *one* extracellular current I^D flowing from the perisomatic to the apical dendritic tree. In the model of Mazzoni et al. (2008), however, two synaptic currents that might be of the same order of magnitude are superimposed to the DFP. Secondly, the measures based on Mazzoni et al. (2008) also systematically overestimate LFP frequencies. This could probably be attributed partly to spurious higher harmonics introduced by computing absolute values. Moreover, taking the power spectrum shows that the Mazzoni et al. (2008) measure decays much more slowly than the average membrane potential, which is at variance with experimental data.

However, at the current stage, both models, that of Mazzoni et al. (2008) and our own, agree with respect to the polarity of DFP and LFP. The measures based on Mazzoni et al. (2008) have positive polarity simply due to the moduli. On the other hand, also the direction of current dipoles in our model is constrained by the construction of the equivalent circuit (**Figure 2**) where current sources are situated at the perisomatic and current sinks are situated at apical dendritic tree. Taking this polarity as positive also entails positive DFP and LFP that could only change in strength. However, it is well known from brain anatomy that pyramidal cells appear in at least two layers, III and VI, of neocortex. This is reflected in experiments when an electrode traverses different layers by LFP polarity reversals, and, of course, by the fact that LFP and EEG oscillate between positive and negative polarity. Adapting our model to this situation could be straightforwardly accomplished in the framework of neural field theory by fully representing space and simulating layered neural fields (Amari, 1977; Jirsa and Haken, 1996; beim Graben, 2008). By contrast such a generalization is impossible at all with the model of Mazzoni et al. (2008) due to the presence of absolute values.

On these grounds we have good indication that our measure is an improvement to the Mazzoni et al. LFP measures, and, quite importantly, it is biophysically better motivated than the *ad hoc* model of Mazzoni et al. (2008). However, much considerable effort is still required to underpin all the relevant LFP mechanisms and to better represent experimental LFP/EEG dynamics.

Finally, our work provides a new framework where DFPs and the relationship between firing rates and local fields can be explored without the extreme demand on computational complexity involved in multicompartmental modeling (Protopapas et al., 1998; Sargsyan et al., 2001; Lindén et al., 2010; Lindén et al., 2011) by adopting reduced compartment circuits. For example, we envisage to extend our recent work which maps firing rate model (derived from LIF models) to population density models (Chizhov et al., 2007), but now incorporating our observational DFP model. In addition, our framework is analytically amenable and thus can be applied to any linear differential equation, for instance, GIF (Gif-sur-Yvette Integrate Fire) models, which are improvements to the LIF models and compute more accurately spike activations (Rudolph-Lilith et al., 2012). Also resonant membranes (mediated by Ca^{2+} and a Ca^{2+} -activated K^+ ionic currents) that describe sub-threshold oscillations and which can be easily expressed by linear equations (Mauro et al., 1970) can be incorporated in our derivations. We note, however, that our framework can be applied to non-linear equations, with Hodgkin and Huxley (1952) type activation, but it will fall short from explicit and analytical observation equations.

ACKNOWLEDGMENTS

We thank Michelle Lilith, Claude Bédard, Alain Destexhe, and Jürgen Kurths for fruitful discussion. In addition, we would like to thank Samantha Adams for providing help with Brian Simulator installations and initial discussions of Brian usage. This research was supported by a DFG Heisenberg grant awarded to PbG (GR 3711/1-1).

REFERENCES

- Amari, S.-I. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87.
- Bédard, C., and Destexhe, A. (2009). Macroscopic models of local field potentials and the apparent 1/f noise in brain activity. *Biophys. J.* 96, 2589–2603.
- Bédard, C., and Destexhe, A. (2012). “Modeling local field potentials and their interaction with the extracellular medium,” in *Handbook of Neural Activity Measurement*, eds R. Brette and A. Destexhe (Cambridge, MA: Cambridge University Press), 136–191.
- Bédard, C., Kröger, H., and Destexhe, A. (2004). Modeling extracellular field potentials and the frequency-filtering properties of extracellular space. *Biophys. J.* 86, 1829–1842.
- beim Graben, P. (2008). “Foundations of neurophysics,” in *Lectures in Supercomputational Neuroscience: Dynamics in Complex Brain Networks*, Springer Complexity Series, eds P. b. Graben, C. Zhou, M. Thiel, and J. Kurths (Berlin: Springer), 3–48.
- beim Graben, P., and Kurths, J. (2008). Simulating global properties of electroencephalograms with minimal random neural networks. *Neurocomputing* 71, 999–1007.
- Berger, H. (1929). Über das Elektroencephalogramm des Menschen. *Archiv für Psychiatrie* 87, 527–570.
- Brunel, N., and Wang, X.-J. (2003). What determines the frequency of fast network oscillations with irregular neural discharges? I. synaptic dynamics and excitation-inhibition balance. *J. Neurophysiol.* 90, 415–430.
- Chizhov, A. V., Rodrigues, S., and Terry, J. R. (2007). A comparative analysis of a firing-rate model and conductance-based neural population model. *Phys. Lett. A* 369, 31–36.
- Creutzfeldt, O. D., Watanabe, S., and Lux, H. D. (1966a). Relations between EEG phenomena and potentials of single cortical cells. I. evoked responses after thalamic and epicortical stimulation. *Electroencephalogr. Clin. Neurophysiol.* 20, 1–18.
- Creutzfeldt, O. D., Watanabe, S., and Lux, H. D. (1966b). Relations between EEG phenomena and potentials of single cortical cells. II. spontaneous and convulsoid activity. *Electroencephalogr. Clin. Neurophysiol.* 20, 19–37.
- David, O., and Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage* 20, 1743–1755.
- Destexhe, A. (2001). Simplified models of neocortical pyramidal cells preserving somatodendritic voltage attenuation. *Neurocomputing* 38–40, 167–173.
- Destexhe, A., Mainen, F., and Sejnowski, T. J. (1998). “Kinetic models of synaptic transmission,” in *Methods in Neuronal Modelling. From Ions to Networks*, eds C. Koch and I. Segev (Cambridge, MA: MIT Press), 1–25.
- Gold, C., Henze, D., and Koch, C. (2007). Using extracellular action potential recordings to constrain compartmental models. *J. Comput. Neurosci.* 23, 39–58.
- Goodman, D., and Brette, R. (2009). The Brian simulator. *Front. Neurosci.* 3, 192–197. doi: 10.3389/neuro.01.026.2009
- Hodgkin, A. L., and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* 117, 500–544.
- Holt, G. R., and Koch, C. (1999). Electrical interactions via the extracellular potential near cell bodies. *J. Comput. Neurosci.* 6, 169–184.
- Jansen, B. H., and Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.* 73, 357–366.
- Jirsa, V. K., and Haken, H. (1996). Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* 77, 960–963.
- Johnston, D., and Wu, S. M.-S. (1997). *Foundations of Cellular Neurophysiology*. Cambridge, MA: MIT Press.
- Koch, C., and Segev, I. (eds.). (1998). *Methods in Neuronal Modelling. From Ions to Networks*, 2nd Edn., Computational Neuroscience. Cambridge, MA: MIT Press.
- Kole, M., and Stuart, G. (2012). Signal processing in the axon initial segment. *Neuron* 73, 235–247.
- Lakatos, P., Shah, A., Knuth, K., Ulbert, I., Karmos, G., and Schroeder, C. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J. Neurophysiol.* 94, 1904–1911.
- Lindén, H., Pettersen, K., and Einevoll, G. (2010). Intrinsic dendritic filtering gives low-pass power spectra of local field potentials. *J. Comput. Neurosci.* 29, 423–444.
- Lindén, H., Tetzlaff, T., Potjans, T. C., Pettersen, K. H., Grün, S., Diesmann, M., et al. (2011). Modeling the spatial reach of the LFP. *Neuron* 72, 859–872.
- Mainen, Z., Joerges, J., Huguenard, J., and Sejnowski, T. (1995). A model of spike initiation in neocortical pyramidal neurons. *Neuron* 15, 1427–1439.
- Mauro, A., Conti, F., Dodge, F., and Schor, R. (1970). Subthreshold behavior and phenomenological impedance of the squid giant axon. *J. Gen. Physiol.* 55, 497–532.
- Mazzoni, A., Brunel, N., Cavallari, S., Logothetis, N. K., and Panzeri, S. (2011). Cortical dynamics during naturalistic sensory stimulations: experiments and models. *J. Physiol.* 105, 2–15.
- Mazzoni, A., Panzeri, S., Logothetis, N. K., and Brunel, N. (2008). Encoding of naturalistic stimuli by local field potential spectra in networks of excitatory and inhibitory neurons. *PLoS Comput. Biol.* 4:e1000239. doi: 10.1371/journal.pcbi.1000239
- Mazzoni, A., Whittingstall, K., Brunel, N., Logothetis, N. K., and Panzeri, S. (2010). Understanding the relationships between spike rate and delta/gamma frequency bands of LFPs and EEGs using a local cortical network model. *Neuroimage* 52, 956–972.
- Niedermeyer, E. (2005). “The normal EEG of the waking adult,” in *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 5th Edn, eds E. Niedermeyer and F. L. D. Silva (Philadelphia, PA: Lippincott Williams & Wilkins), 167–192.
- Nunez, P. L., and Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. 2nd Edn. New York, NY: Oxford University Press.
- Omurtag, A., Knight, B., and Sirovich, L. (2000). On the simulation of large populations of neurons. *J. Comput. Neurosci.* 8, 51–63.
- Poulet, J. F., Fernandez, L. M., Crochet, S., and Petersen, C. C. (2012). Thalamic control of cortical states. *Nat. Neurosci.* 15, 370–372.
- Protopapas, A., Vanier, M., and Bower, J. M. (1998). “Simulating large networks of neurons,” in *Methods in Neuronal Modelling. From Ions to Networks*, eds C. Koch and I. Segev (Cambridge, MA: MIT Press), 461–498.
- Rall, W. (1977). “Core conductor theory and cable properties of neurons,” in *Handbook of Physiology – The Nervous System, Cellular Biology of Neurons*, Vol. 1, ed E. R. Kandel (Bethesda, MD: American Physiological Society), 39–97.
- Rodrigues, S., Chizhov, A., Marten, F., and Terry, J. (2010). Mappings between a macroscopic neural-mass model and a reduced conductance-based model. *Biol. Cybern.* 102, 361–371.
- Rudolph-Lilith, M., Dubois, M., and Destexhe, A. (2012). Analytical integrate-and-fire neuron models with conductance-based dynamics and realistic postsynaptic potential time course for event-driven simulation strategies. *Neural Comput.* 34, 1426–1461.
- Sargsyan, A. R., Papatheodoropoulos, C., and Kostopoulos, G. K. (2001). Modeling of evoked field potentials in hippocampal CA1 area describes their dependence on NMDA and GABA receptors. *J. Neurosci. Methods* 104, 143–153.
- Schomer, D. L., and Lopes da Silva, F. H. (eds.). (2011). *Niedermeyer’s Electroencephalography. Basic Principles, Clinical Applications, and Related Fields*. 6th Edn. Philadelphia, PA: Lippincott Williams & Wilkins.
- Spruston, N. (2008). Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.* 9, 206–221.
- Wang, X.-J., Tegnér, J., Constantinidis, C., and Goldman-Rakic, P. S. (2004). Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. *Proc. Natl. Acad. Sci. U.S.A.* 101, 1368–1373.
- Wendling, F., Bellanger, J. J., Bartolomei, F., and Chauvel, P. (2000). Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals. *Biol. Cybern.* 83, 367–378.
- Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 September 2012; accepted: 14 December 2012; published online: 04 January 2013.

Citation: beim Graben P and Rodrigues S (2013) A biophysical observation model for field potentials of networks of leaky integrate-and-fire neurons. *Front. Comput. Neurosci.* 6:100. doi: 10.3389/fncom.2012.00100

Copyright © 2013 beim Graben and Rodrigues. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Spatiotemporal imaging of complexity

Stephen E. Robinson^{1*}, Arnold J. Mandell² and Richard Coppola¹

¹ MEG Core Facility, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA

² Department of Psychiatry, University California, San Diego, CA, USA

Edited by:

Peter Robinson, The University of Sydney, Australia

Reviewed by:

G. Brad Ermentrout, University of Pittsburgh, USA

J. Michael Herrmann, The University of Edinburgh, UK

*Correspondence:

Stephen E. Robinson, MEG Core Group, National Institute of Mental Health, 10 Center Drive, Bethesda, MD 20814, USA.
e-mail: robinsonse@mail.nih.gov

What are the functional neuroimaging measurements required for more fully characterizing the events and locations of neocortical activity? A prime assumption has been that modulation of cortical activity will inevitably be reflected in changes in energy utilization (for the most part) changes of glucose and oxygen consumption. Are such measures complete and sufficient? More direct measures of cortical electrophysiological activity show event or task-related modulation of amplitude or band-limited oscillatory power. Using magnetoencephalography (MEG), these measures have been shown to correlate well with energy utilization sensitive BOLD fMRI. In this paper, we explore the existence of state changes in electrophysiological cortical activity that can occur independently of changes in averaged amplitude, source power or indices of metabolic rates. In addition, we demonstrate that such state changes can be described by applying a new measure of complexity, rank vector entropy (RVE), to source waveform estimates from beamformer-processed MEG. RVE is a non-parametric symbolic dynamic informational entropy measure that accommodates the wide dynamic range of measured brain signals while resolving its temporal variations. By representing the measurements by their rank values, RVE overcomes the problem of defining embedding space partitions without resorting to signal compression. This renders RVE-independent of absolute signal amplitude. In addition, this approach is robust, being relatively free of tunable parameters. We present examples of task-free and task-dependent MEG demonstrating that RVE provides new information by uncovering hidden dynamical structure in the apparent turbulent (or chaotic) dynamics of spontaneous cortical activity.

Keywords: magnetoencephalography, neuroscience, cognitive, beamformer, complexity, nonlinear, turbulence, mixing

INTRODUCTION

ENTROPY AND COMPLEXITY

The term “entropy” is commonly defined as a measure of the order or disorder in a physical system. In the context of time-varying electrophysiological brain signals we use the term “complexity” instead of “entropy” in order to emphasize the temporal fluctuations of information rate rather than the total information of that signal. Signals having low complexity include synchronous events and oscillations. Signals with high complexity appear more chaotic and correspond to a higher information rate.

The topographic characteristics of spatiotemporal fluctuations in cortical electrophysiological activity are analogous to those of non-laminar, turbulent flow as visualized by optical imaging using voltage sensitive dyes (Cohen et al., 1978). On a macroscopic scale, the activity of individual neurons is hidden from external non-invasive measures such as magnetoencephalography (MEG) (Cohen, 1968) or electroencephalography (EEG) (Berger, 1989), the ensemble behavior of the underlying cortical neural network exhibits complex emergent and traveling fronts of excitation and inhibition that are supported by both short-range inter-neuron connections and “small world” longer-range connections. There have been a variety of experimental approaches that have been developed to characterize the dynamics of turbulent fluid

flow ranging from visualization of waves and eddies through the use of fluorescent dyes (Busse and Clever, 1979) and monitoring heat flux using cryogenic techniques (Swinney and Gollub, 1981) to laser Doppler techniques that allow measurement of local field velocities without perturbing the field significantly (Gollub and Steinman, 1981). The spatiotemporal patterns observed with changing parameter values include transitions from laminar (linear) flow via a Hopf bifurcation to periodic oscillations, followed by two or more simultaneous irrationally related periodic flows and finally the aperiodic oscillations of turbulent (chaotic) dynamics (Ruelle and Takens, 1971). The latter state is characterized by positive entropy generation (Eckmann and Ruelle, 1985) similar to that observed in the MEG record (Mandell et al., 2011a,b; Robinson et al., 2012).

Brain activity is most commonly modeled by narrow-band oscillatory regions that are coupled to one another via networks. Chaos theory is often used as a model to describe complex biological measurements for which linear theory is incomplete. The criteria for chaos modeling include sensitivity to initial conditions and topological mixing. Both of these conditions are satisfied by local measures of ongoing brain activity. This naturally leads to combining a measurement of local cortical signals such as a beamformer estimate derived from MEG to a sensitive and robust

measure of the broad bandwidth non-linear properties of those signals such as RVE.

The lack of suitable existing complexity measures for MEG is our motivation for developing rank vector entropy. First, we require a method for observing the spatiotemporal structure of cortical non-linear dynamics. Retaining temporal resolution allows for the study of the entropy change time-course that is needed to characterize event and task-related brain activity. The focus of current MEG complexity measures has been long-term properties having no temporal resolution such as Lempel-Ziv complexity (Fernandez et al., 2011), transfer complexity (Vakorin et al., 2010), and comparisons of multiple entropy/complexity measures (Bruna et al., 2012). Although there are sliding block methods for observing temporal changes in entropy (Adler and Marcus, 1979), this is computationally inefficient when applied to thousands of voxels for functional brain imaging. The estimated source time-series from beamformed MEG is efficiently transformed into an entropy time-series by RVE.

The *rank vector entropy* (RVE) algorithm is a non-parametric partial analog to metric (Kolmogorov) entropy (ME) (Kolmogorov, 1958; Crutchfield and Feldman, 2003). Both methods estimate the entropy of a one-dimensional series of measurements (e.g., a neurophysiological signal) computed on probability distributions of short sequence “states” encountered in the time series. A unique property of the RVE as a metric entropy lies in its initial encoding of the time series using the topological property of sequence order (Cornfield et al., 1982; Bandt and Pompe, 2002). In the more conventional ME, a short sequence of lagged measurement values from a one-dimensional time-series are mapped into an N -dimensional phase space (Ott, 1993). This embedding space is then partitioned into N -dimensional hypercubes. Each partition represents a state that the signal can manifest in its trajectory through N -dimensional space. The metric entropy of the signal is a measure on the probability distribution of trajectories passing through possible partitioned spaces. The number of partitions in ME can be arbitrary and must be sufficient in number, and with sufficient continuity to accommodate the dynamic range of the signal. In some implementations, dynamic range compression of the measurements is required in order to limit the number of partitions to a manageable number. The classical criteria of a “generating partition”—no more than one entry per partitioned space (Eckmann and Ruelle, 1985) is impractical in the context of real biological data. In contrast, in the RVE algorithm the measurement values of a short sequence of samples extracted from the entire time series are converted to their rank ordered values. The number of states, as partitions of the one-dimensional sliding window is quantized according to the number of elements in the sequence of measurements. Each sequence, with the topological property of relative “nearness,” is referred to as a “rank vector” which can also be thought of as a rank ordered one-dimensional embedding space. The RVE algorithm ignores the absolute signal amplitude in favor of its relative amplitude within the span of each sample window. A metric entropy is then computed on the probability distribution of the finite set of rank sequences. It will be shown that RVE is relatively free of arbitrarily tunable parameters (aside from selection of signal bandwidth, window length,

and a decay time constant). The more conventional ME requires the investigator to use a variety of schemes, many involving the Whitney embedding theorem (Milnor, 1997; Temin, 1997), to determine the number of dimensions and lags for the embedding space and the size and number of partitions in embedding space.

We measure the spatiotemporal complexity of brain activity by applying the RVE analysis to sensor mediated brain signals. This could be accomplished by analyzing the signals from individual sensors (Vakorin et al., 2010; Gomez et al., 2011), pairs of sensors (Mandell et al., 2011a,b), or source estimates from MEG using a scalar LCMV beamformer (Robinson and Black, 1990; Robinson and Rose, 1993; Van Veen et al., 1997; Vrba and Robinson, 2002). The latter completes the analogy to characterization of time-series of turbulent chaotic fluid flow, sampled at multiple spatial points within the flow.

MAGNETOENCEPHALOGRAPHY

Spontaneous MEG signals are on the order of 10^{-13} Tesla peak-to-peak. Despite its small signal strength, MEG has been made practical by larger DC-SQUID based sensor arrays covering the entire head in combination with excellent rejection of environmental magnetic interference (Fife et al., 2002). The application of RVE to MEG signals is a natural one. The magnetoencephalogram is the magnetic counterpart of the electroencephalogram. MEG is a measure of the magnetic field arising from primary (impressed) neural currents, whereas EEG measures the electrical potentials on the scalp that arise from the volume currents and is therefore dependent on tissue conductivity and its boundaries (Plonsey, 1981). The major contributor to the observed primary currents is the potential difference between the dendritic tree and soma of neocortical pyramidal neurons. Because MEG is less dependent on tissue conductivity its measurements can be modeled using simple analytic solutions (Sarvas, 1987). The accuracy of such analytic solutions for MEG enables the use of beamformers (detailed in section “Scalar LCMV Beamformer”) to estimate the source time series for any coordinate within the brain (Robinson, 1989). Application of the RVE transform (section “Rank Vector Entropy”) to any source time series yields a corresponding complexity time series of that activity. The relationship of spatiotemporal RVE to sensorimotor events and cognitive tasks can then be determined by signal averaging or by comparison of the RVE signals that have been parsed into active and control state time segments.

EXAMPLE MEG STUDIES

We have selected four different examples of MEG studies to explore the properties of spatiotemporal RVE analysis. These datasets were selected from an archive of normal control studies. The task-free (resting) MEG dataset was chosen to compare the spatial distribution and timing relationships of RVE complexity to power in a 4–150 Hz bandpass. The P300 study, also referred to as the mismatch negativity (MMN), demonstrates the properties of evoked signals to frequent and deviant auditory stimuli. In the EEG, comparison of the frequent to deviant tone signal averages is characterized by a slow wave after about 250 ms (Naatanen et al., 1978). The same protocol is also referred to as a P₃₀₀ study when a response is required for the deviant tones (Donchin, 1981).

The working memory (n-back) study is used to compare broadband RVE with power in an effortful short-term memory task (Kirchner, 1958). Lastly, we present a MEG study of self-paced voluntary finger movement to compare complexity with SCP. Self-paced voluntary movement is preceded by a slow “readiness” potential (Bereitschaftspotential) (Kornhuber and Deecke, 1964) and a corresponding “readiness” field (Bereitschaftsfield) (Deecke et al., 1982).

METHODS

RANK VECTOR ENTROPY

The RVE algorithm can be described as follow: first, consider a one-dimensional discretely sampled time series of length K samples: $\mathbf{X} = [x_1, x_2, \dots, x_K]$. Let the sample rate and low-pass “corner” frequency of \mathbf{X} be denoted by f_s and f_c , respectively. For any given low-pass frequency, \mathbf{X} is completely determined at $2f_c$ samples per second. It is unnecessary for the rank vector to represent every sample, sequentially. It is sufficient to define the lag ξ required to avoid oversampling as:

$$\xi = \frac{f_s}{2f_c}. \quad (1)$$

Thus, for any specified low-pass cutoff and sample rate we need only process every ξ th sample (ξ is rounded up to the nearest integer). Note the analogy of ξ to the sample lags that are used to define the dimensionality of the embedding space in conventional ME computations (Eckmann and Ruelle, 1985; Crutchfield and Feldman, 2003). The selection of lags in ME is usually based upon the mixing length reflected in the decay of the autocorrelation function (Walters, 1982). In this way it is biased in favor of the dominant signal and its ordering. These methods characteristically assume that the signal is stationary (which it is not) and it may also have the unfortunate side effect of aliasing information. Furthermore, it appears that with real biological data, some investigators have been occasionally arbitrary in their selection of lags (evaluating several different lags so as to obtain results more consistent with their expectations). In the RVE algorithm, the lags are rigorously defined by sample rate and low-pass corner frequency, thus not justifying any further modification. If the frequency band of interest is below the data acquisition bandwidth, the investigator can set the low-pass frequency accordingly, from which the lags are again automatically determined.

Let us select a sub-window of W samples (at integer ξ intervals) from \mathbf{X} , with the window beginning with the k th sample: $\mathbf{W}_k = [x_k, x_{k+\xi}, \dots, x_{k+(W-1)\xi}]$. The measured signal for each lagged sample within sub-window \mathbf{W}_k is initially converted to its integer rank values, forming a rank vector \mathbf{R}_k of length W : $\mathbf{R}_k = [\text{rank}_1, \text{rank}_2, \dots, \text{rank}_W]$. For a window of length W there are $W!$ (factorial) unique rank vectors (i.e., vectors of length W for which the ordered rank values do not repeat). Let there be a table \mathbf{S} to map rank vector sequences to symbols, where the symbol value is obtained from a “look up” table indexed for matching rank vectors. For example, let us consider a window length of $W = 5$ for which there are $5! = 120$ unique rank vectors from which we derive 120 state symbols, s_n . Counting and normalizing the number of occurrences of each unique symbol can then

generate a probability histogram. Counting these symbols in RVE is analogous to counting visits of the signal to the higher dimensional hypercube partitions in ME (Ott, 1993). For this example, we use $W = 5$, $f_s = 600$ Hz, $f_c = 100$ Hz, and $\xi = 3$. For example, let the measured values of x_k through $x_{k+4\xi}$ (i.e., \mathbf{W}_k) be (4.07, −3.12, 3.95, 8.51, −1.21). Its rank vector and symbol value are (2, 5, 3, 1, 4) and 45, respectively (a symbol value of 45 indicates its place in an ascending numerical order of rank vectors).

As the sub-window of W samples is advanced through \mathbf{X} one sample at a time, a new rank vector \mathbf{R}_k and new symbolic representation s_n is generated. The frequency of occurrence of each of these symbols is accumulated in a corresponding state histogram: $\mathbf{F}_k = [f_1(k), f_2(k), \dots, f_{W!}(k)]$. The resulting histogram contains the cumulative counts of each rank vector (state). Since our primary interest is in observing the time-dependent, event-related changes in the rank vector informational entropy, it is necessary to avoid the reduction in relative temporal sensitivity by saturation. We prevent this by introducing a time-constant determining the rate of decay of histogram counts with time. The integrator decay rate τ ($1/e$ time) is required in order to measure the fluctuations in entropy over time. The amplitude of the entropy fluctuations depends on the decay rate (longer τ yields smaller peak-to-peak fluctuations). The relative rather than the absolute changes in entropy are of interest for event or task-related functional imaging (including ICA or resting state MEG). Provided that τ is longer than the time required to completely fill all states, the entropy waveform will be independent of τ —except for its amplitude. As a practical matter, we select a time constant such that the $1/e$ time (in samples) corresponds to three times the number of states. This is implemented by defining constant α :

$$\alpha = e^{-1/(\tau f_s)}, \quad (2)$$

in which τ is the time for the counts to decay to $1/e$ of their initial values, such that for each time step:

$$\mathbf{F}_k = \alpha \mathbf{F}_{k-1}. \quad (3)$$

This constitutes what is termed a “leaky integrator.” The histogram count corresponding to the current state symbol is the incremented by one. Based upon the revised count frequencies, there will be a corresponding set of probabilities for each state:

$$\mathbf{P}_k = [p_1(k), p_2(k), \dots, p_{W!}(k)]. \quad (4)$$

If the cumulative entropy over all samples of \mathbf{X} is required, then $\alpha = 1$. Otherwise, the entropy is estimated as time-dependent. Finally, for each step k , we compute the Shannon entropy (Shannon, 1963) over all state probabilities greater than zero, as normalized by its maximum value, constraining entropy to the range 0–1:

$$h(k) = \frac{1}{\log_2 W!} \sum_{n=1}^{W!} -p_n(k) \log_2 p_n(k). \quad (5)$$

The steps in the algorithmic procedure for computing the time series of RVEs are as follows:

1. Initialize the state count histogram \mathbf{F} (set all f_n values to 1.0).
2. For each sample index k in time series \mathbf{X} , generate a length W rank vector \mathbf{R}_k with lags of ξ samples.
3. Look up state symbol index n corresponding to \mathbf{R}_k .
4. Multiply all elements of histogram \mathbf{F} by α , thus allowing the count histories to decay (Equation 3).
5. Increment the corresponding histogram count f_n by one (where n is the index in \mathbf{S} corresponding to \mathbf{R}_k).
6. Compute the probabilities of each state from the histogram of counts (Equation 4).
7. Compute the RVE metric entropy for this time step (Equation 5).
8. Advance sample index by one sample and continue to repeat steps 2 through 7 and in this way generating a new entropy value for each cycle of the algorithmic process.

SCALAR LCMV BEAMFORMER

Synthetic aperture magnetometry (SAM) is a scalar linearly constrained minimum variance (LCMV) beamformer estimating source activity from MEG signals for specified coordinates in the brain. The mathematics of the beamformer procedure can be traced to the minimum variance estimator first described by Gauss (1823). SAM minimizes the variance (power) of all correlated signals observed by an array of SQUID sensors, subject to a unity gain constraint for a specified coordinate. As such, the action of the SAM may be regarded as spatially selective noise reduction, where noise is defined as unwanted environmental or biological magnetic signals.

The computational procedure has been described in detail elsewhere but can be summarized briefly, as follows: consider measured MEG from an array of sensors. Let the signal space vector at time sample k be denoted by $\mathbf{M}(k)$. Given a sufficient number of time samples, we construct a source estimate $\hat{\mathbf{S}}_{\mathbf{r}}$ for coordinate \mathbf{r} as the weighted sum MEG measurements.

$$\hat{\mathbf{S}}_{\mathbf{r}}(k) = \mathbf{W}_{\mathbf{r}}^T \mathbf{M}(k). \quad (6)$$

To compute the weights using the method of Gauss, we use the quadratic form:

$$\left[\hat{\mathbf{S}}_{\mathbf{r}}(k) = \mathbf{W}_{\mathbf{r}}^T \mathbf{M}(k) \right]^2, \quad (7)$$

for which, after integrating over time, we obtain:

$$\hat{\mathbf{S}}_{\mathbf{r}}^2 = \mathbf{W}_{\mathbf{r}}^T \mathbf{C} \mathbf{W}_{\mathbf{r}}, \quad (8)$$

where \mathbf{C} is the covariance matrix computed over the integration time:

$$\mathbf{C} = \langle \mathbf{M} \mathbf{M}^T \rangle, \quad (9)$$

where $\langle . \rangle$ denotes the expectation value. We solve for \mathbf{W} by introducing a constraint such that $\hat{\mathbf{S}}_{\mathbf{r}}^2$ (power or variance) is

minimized subject to unit gain for a specified coordinate. One such constraint is:

$$\mathbf{W}_{\mathbf{r}}^T \mathbf{B}_{\mathbf{r}} = 1, \quad (10)$$

where $\mathbf{B}_{\mathbf{r}}$ is the *a priori* forward solution for the field observed by an array of M sensors generated by a current dipole source located at \mathbf{r} . That is:

$$\mathbf{B}_{\mathbf{r}} = \begin{bmatrix} b_1(\mathbf{r}) \\ b_2(\mathbf{r}) \\ \vdots \\ b_M(\mathbf{r}) \end{bmatrix}. \quad (11)$$

Suitable methods for computing forward solutions include the current dipole in a homogeneously conducting sphere model (Grynszpan and Geselowitz, 1973; Sarvas, 1987; Hari et al., 1988), multiple local spheres fitted to the surface of the head (Huang et al., 1999), boundary element methods (De Munck, 1992) or perturbative solutions derived from the spherical harmonic expansion of the conductive boundary (Nolte et al., 2004). However, to compute $\mathbf{B}_{\mathbf{r}}$ we must first estimate the dipole orientation. We do this by finding the dipole orientation yielding the highest signal-to-noise ratio (SNR) at location \mathbf{r} . We define $\mathbf{L}_{\mathbf{r}}$, the matrix of forward solutions for dipoles oriented in the three cardinal directions as:

$$\mathbf{L}_{\mathbf{r}} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ \vdots & \vdots & \vdots \\ b_{M1} & b_{M2} & b_{M3} \end{bmatrix}. \quad (12)$$

We also define a diagonal matrix Σ of uncorrelated sensor instrumental noise power σ_m^2 :

$$\Sigma = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \sigma_2^2 & \\ & & \ddots \\ 0 & & & \sigma_M^2 \end{bmatrix} \cong \bar{\sigma}^2 \mathbf{I}. \quad (13)$$

The source power \mathbf{S} in each of three cardinal directions is given by:

$$\mathbf{S}_{\mathbf{r}}^2 = \left[\mathbf{L}_{\mathbf{r}}^T \mathbf{C}^{-1} \mathbf{L}_{\mathbf{r}} \right]^{-1}, \quad (14)$$

and the noise power \mathbf{N} by:

$$\mathbf{N}_{\mathbf{r}}^2 = \left[\Sigma \mathbf{L}_{\mathbf{r}}^T \mathbf{C}^{-2} \mathbf{L}_{\mathbf{r}} \right]^{-1}. \quad (15)$$

The SNR is given by a generalized eigensystem of these two 3×3 matrices (Sekihara and Nagarajan, 2008). By assuming that the SQUID noise is nearly equal in all sensors, we can neglect Σ , as it represents a scalar ($\bar{\sigma}^2 \mathbf{I}$) that will not affect determination of the moment vector. The dipole orientation maximizing SNR is given by the eigenvector \mathbf{e}_{\max} corresponding to the maximum eigenvalue λ_{\max} of the generalized eigensystem:

$$\mathbf{L}_{\mathbf{r}}^T \mathbf{C}^{-2} \mathbf{L}_{\mathbf{r}} \mathbf{e}_k = \lambda_k \mathbf{L}_{\mathbf{r}}^T \mathbf{C}^{-1} \mathbf{L}_{\mathbf{r}} \mathbf{e}_k. \quad (16)$$

The forward solution for the dipole vector for the highest SNR is therefore:

$$\mathbf{B}_r = \mathbf{L}_r \mathbf{e}_{\max}. \quad (17)$$

Solving for the optimum scalar LCMV beamformer weights using Lagrange multipliers results in:

$$\mathbf{W}_r = \frac{\mathbf{C}^{-1} \mathbf{B}_r}{\mathbf{B}_r^T \mathbf{C}^{-1} \mathbf{B}_r}. \quad (18)$$

Substituting \mathbf{W}_r into Equation (6) yields an estimate of the source time series.

APPLICATION OF RVE TO MEG DATA

Subjects

MEG data from four healthy normal control subjects, one per example study, were randomly selected from a larger group of NIMH study subjects of both genders, mean age 27.6 years. All subjects gave written informed consent according to protocols approved by the NIH CNS Institutional Review Board.

Data acquisition

MEG data were acquired using a 275-channel whole head MEG (CTF Systems, Inc.) housed within a three layer magnetically shielded room (Vacuumschmelze AK-3). Three head localization coils were affixed to subjects at the nasion, right preauricular, and left preauricular points. The acquisition software energizes these coils with sinusoidal currents at three different frequencies before and after data acquisition in order to localize and establish each subject's head position relative to the MEG sensors. Data were sampled continuously, without breaking it into epochs or trials at 600 Hz in a bandpass of DC to 150 Hz, with the subjects in seated position. Stimulus and response trigger markers for each study were recorded with the data. Synthetic 3rd-gradient mode was used during data acquisition to obtain further reduction in magnetic noise (Vrba and Robinson, 2002). Raw data were archived on disk for subsequent analysis.

RVE analysis parameters for all studies were fixed at a 4–150 Hz bandpass, $W = 5$ (120 states), lags every two samples, and an integrator decay time constant of 0.6 s.

All subjects were given a T1-weighted volumetric MRI. Radiological markers were affixed to the identical fiducial points as were used for the head localization coils used during MEG acquisition. Markers were used to transform the MRI to the MEG head frame for subsequent processing, including segmentation of the cortical boundary, and coregistration of functional and anatomical data.

Data analysis

Analyses were applied to the unaveraged continuous data. We use the single-layer realistic head model (Nolte et al., 2001) to compute the forward solutions. The scalar beamformer processing includes the steps:

1. Coregister a T1-weighted MRI to the MEG frame and segment the MRI to extract subject's brain hull, using AFNI software (Cox, 1996).
2. Compute the outward-pointing normal vectors for the brain hull.
3. Estimate the MEG measurement covariance matrix for the required time segments and frequency bandpass (Equation 9).
4. Compute points (voxel coordinates) on a regular three-dimensional grid at 5 mm intervals within the head.
5. For each voxel within the hull boundary:
 - a. Compute the lead-field matrix (Equation 12) using the brain hull as a single-layer realistic head model (Nolte et al., 2001).
 - b. Compute the beamformer coefficients (Equation 18).
 - c. Estimate the source time series (Equation 6).

The RVE voxel time-series is computed for a 4–150 Hz bandpass (Equation 1–5). No additional filtering or smoothing is required as the RVE time-series is inherently smooth. Source power time-series are computed from the smoothed envelope of a Hilbert transform following bandpass filtering of the MEG data. The Hilbert envelope was smoothed using lowpass filter corresponding to the lowpass corner frequency of each selected bandpass.

The fluctuations in RVE are relatively small (on the order of 5–15% for $\alpha = 0.6$ s). RVE deviation relative to its statistical mode is used for analysis of the task-free data. For the remaining studies, we compute the Student's T-value for each latency in the RVE time, relative to a selected baseline. Static (3D) and spatiotemporal images (3D + time) were assembled and displayed using AFNI software (Cox, 1996).

Task-free (resting) protocol

Two hundred and forty seconds of task-free ("resting") MEG data were recorded from a normal subject (eyes opened), using data acquisition procedures outlined in section "Data Acquisition." Beamformer weights were computed as above for the entire duration at 5 mm voxel intervals on a three-dimensional grid occupying the entire head. The envelope of the source power S^2 and the RVE were then computed for each voxel and the results mapped as three dimensions plus time images at 50 ms intervals. Because RVE is a measure on temporally hierarchical brain signals, it is studied using a broad bandwidth of 4–150 Hz. In this way we avoid the potential reduction in complexity that would accompany our narrowing the bandwidth of observation. In comparing RVE, with the simultaneously studied envelopes of band width power, the latter was computed using a smoothed Hilbert transform for a sequence of bandwidths that included: 4–150 Hz, 4–8 Hz (theta), 8–13 Hz (alpha), 15–30 Hz (beta), 35–70 Hz (low gamma), and 70–150 Hz (high gamma). The RVE time series is inherently smooth and required no additional filtering.

Auditory P300 (Mismatched Negativity) protocol

The auditory P300 protocol consisted of random presentation of 200 frequent (1.0 kHz) and 50 rare (1.5 kHz) 50 ms duration tone bursts delivered binaurally via non-magnetic earphones (Etymotic), with a 1.0–1.5 s pseudo-random interstimulus interval (ISI). Subjects were instructed to respond to the rare tones by pressing a response button. MEG data were acquired as noted in

section “Data Acquisition” and stored to disk together with trigger markers for identifying the onset of the rare and frequent tone bursts, along with the button response.

Beamformer weights were computed from the *unaveraged* MEG data for each of three conditions: (1) both frequent and rare tones, (2) frequent tones, and (3) rare tones. The weights were computed in a 4–150 Hz bandpass. Three-dimensional images of RVE were then computed at 5 ms intervals over a time window from -0.2 to $+0.8$ s relative to the markers for each of the three conditions. Image maxima and minima coordinates were determined and additional beamformer weights computed for a DC to 100 Hz bandpass. These weights were then applied to the averaged RVE signal of the three conditions in order to show the time course of the entropy waveforms. Note that we compute the average of the entropy and *not* the entropy of the averaged signal. The RVE functional images and time series are displayed as Student’s *T*-values for each voxel, relative to the selected pre-stimulus baseline.

Working memory (N-back) protocol

The numerals 1 through 4 were randomly presented to the subject at 1.4-s intervals in 18 blocks of 11 trials each, using a DLP projector. Preceding each block the subject received instructions how to respond to via four numbered buttons, that corresponds to the numbers 1 through 4. For 0-back blocks, the subject simply pressed the button corresponding to the number presented. For 1 and 2-back conditions, the subject pressed the button corresponding to the numbers that were presented one or two trials back, respectively. MEG was recorded, along with trigger markers indicating the task (i.e., respond to 0, 1, or 2-back), which number was displayed and which response button was depressed, using the settings in section “Data acquisition.”

The RVE was applied to the continuous source time series of each voxel, resulting in an RVE time series. The RVE voxel time series was then parsed into 0.5-s segments (± 0.25 s relative to the button response) for the 0, 1, and 2-back conditions. The RVE was integrated over each segment. A Student’s *T*-test was used to compare the difference in integrated RVE signal for pairs of each condition (i.e., 2 vs. 0-back, 2 vs. 1-back) for each individual voxel. The results were displayed as *p*-values using AFNI.

In a similar manner, we computed and displayed the comparisons for beta-band (14–30 Hz) power using the same conditions.

Self-paced voluntary movement protocol

Subjects performed self-paced button presses at intervals of at least 10 s while continuous MEG was recorded together with time markers for each button press, as outlined in section “Data acquisition.” No cues were given to the subjects as to when to press the button.

Scalar beamformer weights were computed for the entire unaveraged dataset in a bandpass of 4–150 Hz. Weights were then applied to the measured MEG data yielding a source estimate time series for each voxel. The voxel time series was then transformed to a RVE time series. The mean entropy value over the interval from -3.0 to -2.5 s prior to the button press was designated as a baseline for comparison of later entropy changes; the RVE time-series was averaged relative to the voluntary button press, at

5 ms intervals from -3.0 to 2.0 s. The RVE averaged response is displayed as its Student’s *T*-value relative to the selected baseline. The voxel having the maximum relative RVE prior to the button press was also used to compute the averaged source strength time-series.

RESULTS

TASK-FREE (RESTING) PROTOCOL

Comparison of the spatial and temporal patterns in the 4–150 Hz bandpass reveals very little apparent correlation between RVE and the Hilbert envelope of power. The fluctuations in entropy are noticeably slower than that of power. In many cortical locations the RVE shows transient decreases from its modal value; spontaneous entropy increases are not as prominent in the resting condition (**Figure 1** top panel). Changes in power relative to its modal value (**Figure 1** center panel) are much more rapid than those of the RVE. The spatial distribution of RVE and power also are seen to differ (**Figure 1** bottom panel).

An expanded 30-s view of the time course of the smoothed Hilbert envelope of power and across multiple frequency bands shows no apparent correlation with RVE 4–150 Hz (**Figure 2**). Changes in entropy clearly show slower, longer wavelength features that are not evident in measures of power.

AUDITORY P300 TASK

In this characteristic example, we observed a reduction in entropy coinciding with the N_{100M} peak (**Figure 3B**) for frequent and combined frequent and rare trials (**Figure 3A**). The reduction in entropy at 100 ms was much smaller for the rare tones, alone (**Figure 3B**). The reduction in entropy was maximal in the vicinity of the left Heschl’s gyrus (**Figures 3C,D,E**), with a much smaller reduction in the right hemisphere. Although a flat baseline was seen in the pre-stimulus interval of -0.2 to 0.0 s in the averaged overlay of all sensors (**Figure 3B**), it is absent from the combined and frequent trial averages (**Figure 3A**). We also observed an increase in the entropy *T*-value after 100 ms for the rare tones (**Figure 3A**). At 700 ms this increase was maximal in anterior cingulate and left temporal cortex (**Figure 3F**). The increase does not appear in the sensor signal averaged overlay.

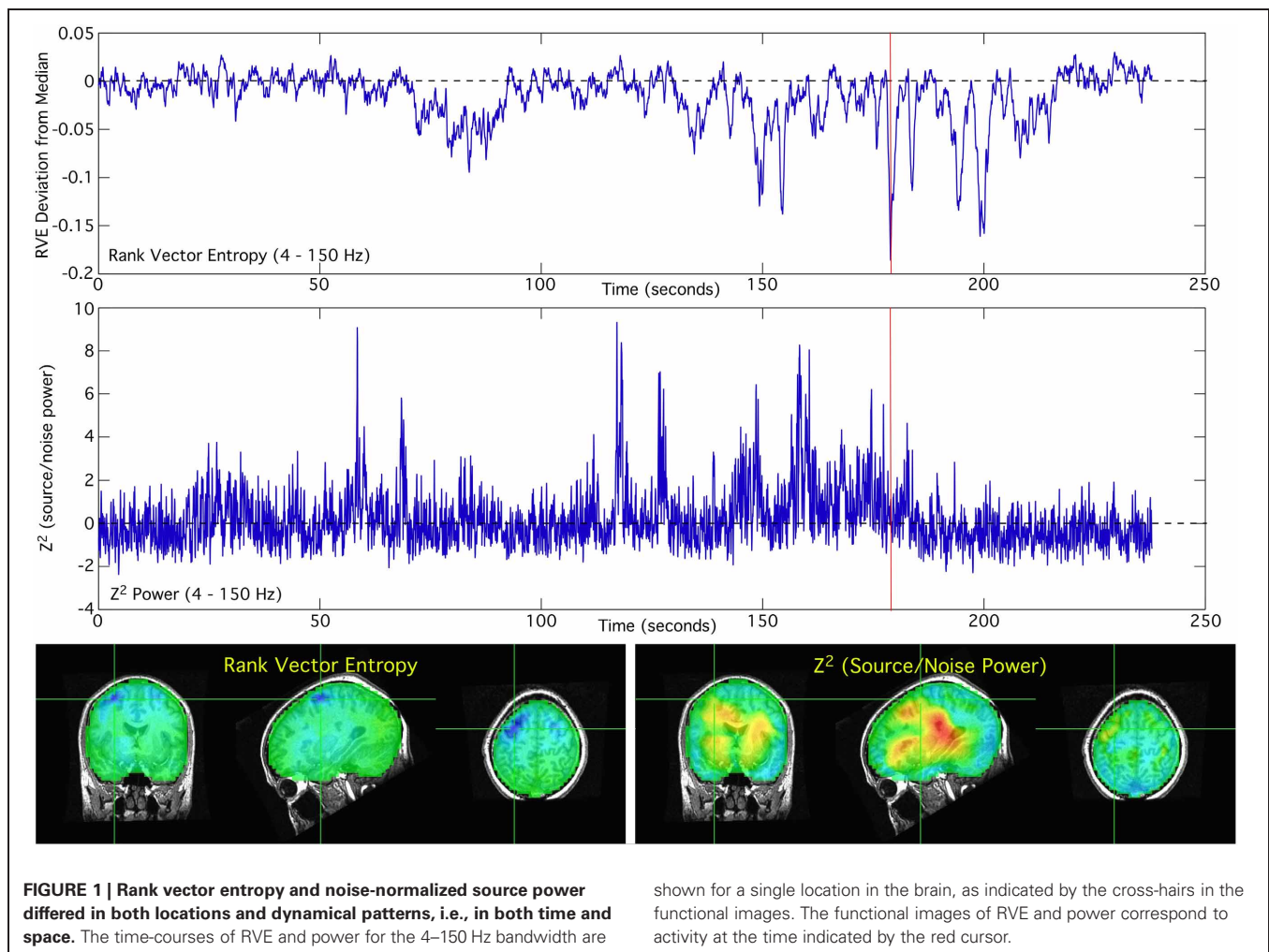
WORKING MEMORY TASK

In the 2 vs. 0-back comparison, we observe a reduction in beta-band power in the 2-back task relative to the 0-back task, ERD, in dorsolateral prefrontal cortex (DLPFC) and an increase in power, ERS, in inferior occipital cortex (**Figure 4A**). For the identical comparison, there is a significant increase in entropy ($p < 10^{-5}$) in the 2-back condition that is widely distributed throughout the brain, with the largest changes in the left hemisphere (**Figure 4B**).

In the 2 vs. 1-back comparison, beta-band ERD is seen in DLPFC and bilaterally near the inferior temporal pole (**Figure 5A**). The same comparison for RVE shows an increase in entropy in anterior cingulate cortex and a decrease in the right inferior temporal pole (**Figure 5B**).

SELF-PACED VOLUNTARY FINGER MOVEMENT

The Bereitschaftsfield (BF), average of 29 trials, appears as a slow low frequency rise in the source moment in premotor cortex



that occurs over the time span from -2.5 s to 0 s, relative to the button press (**Figure 6** top). We observed a peak dipole moment of about 10 nA \cdot m relative to a mean baseline over the interval -3.0 to -2.5 s. Signal-to-noise ratio for this source is low, requiring a lowpass filter of 5 Hz in order to see the slow changes. By contrast, the averaged RVE (4 – 150 Hz), displayed as a T-value relative to baseline, has excellent signal-to-noise and spans -1.75 to 0 s (**Figure 6** bottom). We refer to this wave as the “Bereitschaftskomplexität” (BK), to emphasize its relationship to the BF. Both the BF and BK waveforms are similar.

DISCUSSION

Our single most important finding is that the RVE measure of complexity adds new information about brain dynamics that was previously hidden within the apparent chaos of spontaneous cortical signals. Moreover, RVE, when combined with a scalar beamformer, reveals an underlying spatiotemporal complexity pattern that is modulated by stimuli and tasks. We have presented experimental evidence demonstrating that these patterns of complexity are decoupled from conventional measures of amplitude and oscillatory power. Thus, RVE reveals new information on the dynamics of brain activity.

Next, we will examine the experimental evidence regarding what directional changes in entropy signify. MEG measures such as band-limited oscillatory power or amplitude have been shown to be concordant with functional imaging by BOLD fMRI (Singh et al., 2002; Coppola et al., 2004). Induced changes in cortical activity, particularly for movement, are signaled by event-related desynchronization (ERD) of power (Taniguchi et al., 2000) and also by increases in the BOLD signal relative to a resting or control state. How does one interpret event-related changes in RVE? The answer is not straightforward. We show that a simple auditory stimulus induces a decrease in the RVE at a latency corresponding to the N_{100m} and location corresponding to primary auditory cortex in the left hemisphere. Comparable MEG studies also show ERD in the beta and alpha bands at the latency and locations corresponding to the N_{100m} , and fMRI shows a corresponding increase in the BOLD signal.

In our auditory P_{300} study the response to the rare tones elicited later increases in the RVE signal that persisted for over 1 s. The increase was seen broadly in anterior cingulate cortex and in left temporal cortex in a wide area centered on where the N_{100m} response appeared. These late responses are not visible in the averaged overlay of sensors and in the event-related changes

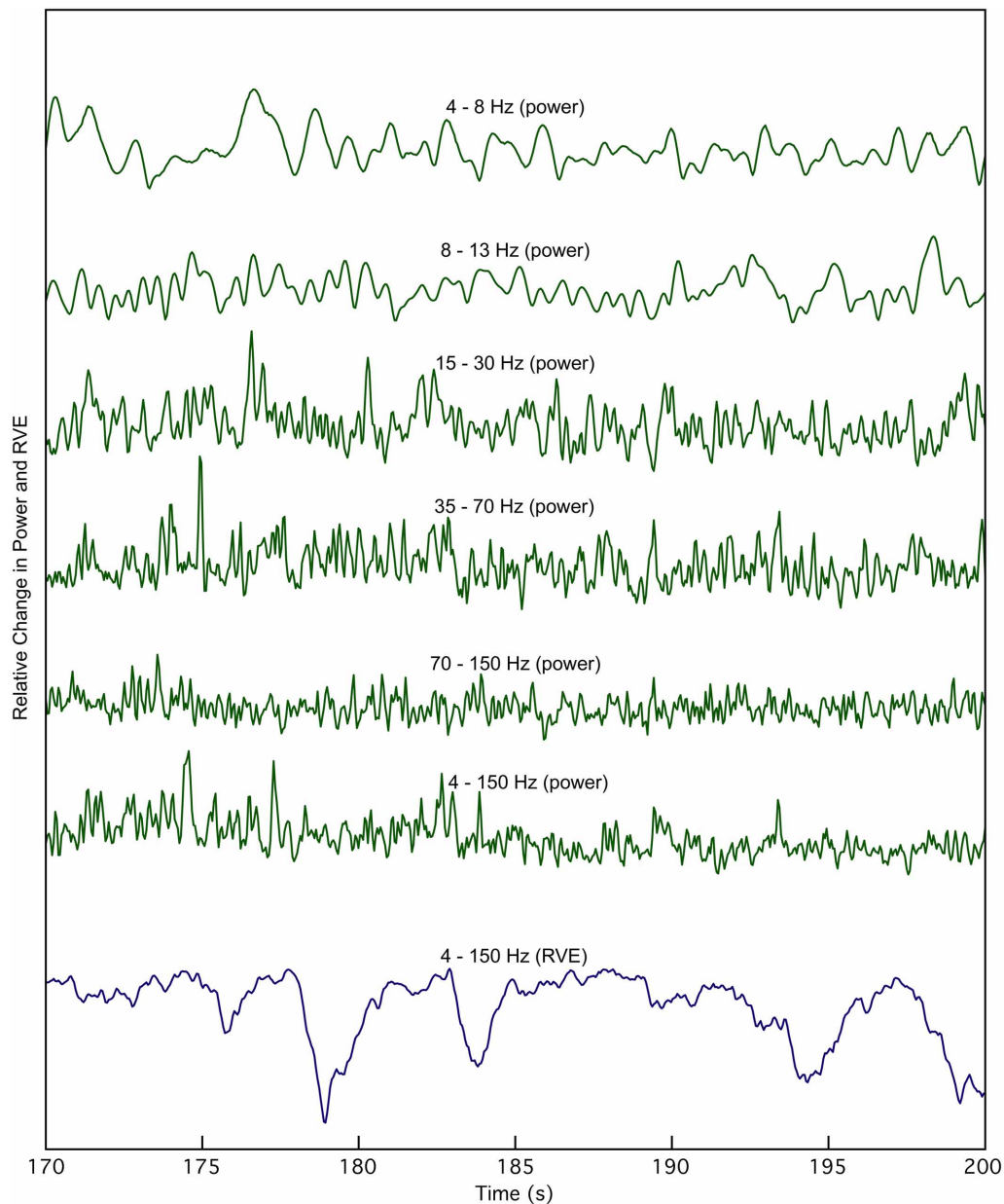


FIGURE 2 | The time course of rank vector entropy, $h(k)$ (blue trace), is apparently uncorrelated with changes in the envelope of source power, in any bandpass. The 4–150 Hz RVE is

compared with source power at the same location in multiple bandwidths for a 30 s expanded view derived from the waveforms shown in **Figure 1**.

in power. We see that the absence of a stable baseline for the frequent tones is a consequence of the slow return to baseline of the entropy increase evoked by the rare tones. That is, the 1.0–1.5 s ISI is not sufficient for the induced RVE increase to return to its resting value.

These P_{300} results demonstrate that stimuli and tasks can induce either an increase or a decrease in entropy relative to its resting value. The RVE of task-free MEG is characterized by transient decreases in entropy from its modal value of ~ 0.92 , lasting two or more seconds. Increases above the modal value are

not as prominent and are more rapid—suggesting that bidirectional modulation of RVE reflects the parabolic character of the entropy function and/or its governance by at least two different mechanisms. The N_{100} evoked component of frequent tones corresponds to a transient decrease in the RVE signal that is much shorter than the observed transients in the task-free MEG recordings. The localization of decrease was centered on the left auditory cortex, but included perisylvian areas. The broad distribution of the induced entropy changes confirm that the auditory N_{100} component arises from multiple regions and not simply primary

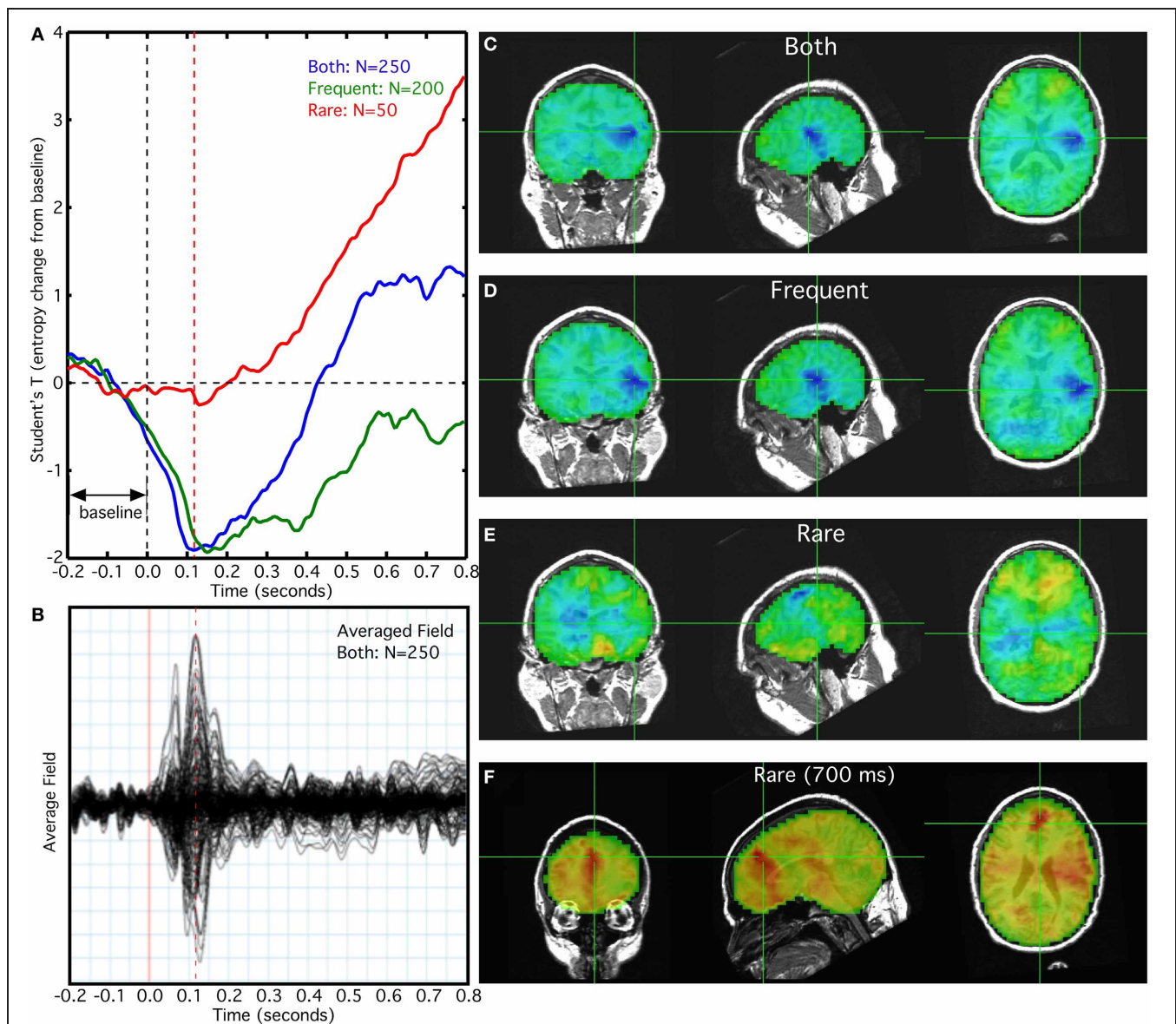


FIGURE 3 | Relationship of rank vector entropy to the auditory evoked response to rare and frequent tone bursts. (A) Average of entropy time series for left primary auditory cortex; **(B)** average overlay of 275 MEG

sensors; **(C)** RVE image of combined rare and frequent tones at 100 ms; **(D)** RVE image of frequent tones at 100 ms; **(E)** RVE image of rare tones at 100 ms; **(F)** RVE image of rare tones at 700 ms latency.

auditory cortex (Naatanen and Picton, 1987). Although the averaged evoked response to the rare tones showed only a small peak corresponding to the P₃₀₀ response, there was a profound slow increase in the RVE signal starting at about 250 ms. This slow increase is concordant with the mismatch negativity signal that is observed in similar protocols involving frequent and deviant tones. It appears that synchronized cortical activity, such as that leading to an averaged evoked response component, induces a transient decrease in entropy in the corresponding regions. By contrast, the slow and prolonged rise in entropy induced by the rare tones suggests an increase in asynchronous cortical activity that signifies attentional mechanisms. It should be noted that the frequent tone average (Figure 3A) shows what appears to be

activity where the pre-stimulus baseline should instead be flat. This is an artifact due to the long duration of the slow RVE component of the rare tones and the relatively short 1.5 s interstimulus interval.

RVE's relatively long time scales (in seconds) and the continuity and differentiability in its time-dependent changes suggests a relationship of this measure to the family of slow cortical potentials (SCPs), including the contingent negative variation (CNV) and the Bereitschaftspotential (Birbaumer et al., 1990). This is particularly evident when comparing BF and BK waveforms seen in Figure 6. The slow changes in entropy are only partially due to integration of symbolic state counts. The decay time constant incorporated into the integrator permits RVE to respond rapidly

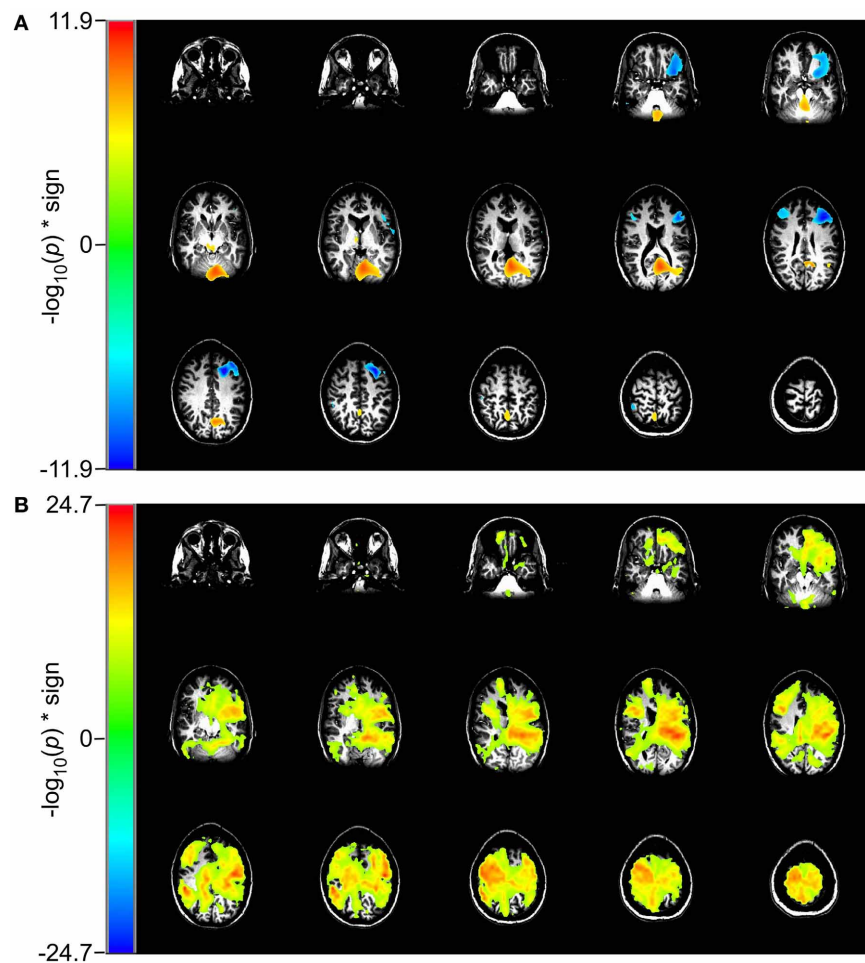


FIGURE 4 | Comparison of event-related changes in beta-band power and in rank vector entropy for 2 vs. 0-back working memory task centered on the response. Images are univariate p -values for a Student's T -test, with a threshold of 10^{-5} . **(A)** Event-related

desynchronization of beta band power (14–30 Hz) appears bilaterally in dorsolateral prefrontal cortex and in left supraorbital prefrontal cortex. **(B)** RVE (4–150 Hz) shows widely distributed increases that overlap with regions of beta-band ERD.

to changing complexity. The observed slow changes in RVE are significantly longer than the 0.6-s $1/e$ integration decay time constant used in these analyses. Thus, the slow shifts in entropy are not an artifact of the analysis but rather reflect cortical processes. Further studies will be needed to compare RVE with other SCP phenomena (e.g., the contingent negative variation paradigm and orienting responses).

The working memory task lends additional insight into how the RVE complexity measure is modulated by cognitive effort. The contrast for the 2-back to the 0-back condition for RVE (**Figure 4**) shows widespread and highly significant ($p < 10^{-5}$) increases in entropy throughout the brain. The 2-back task engages working memory whereas the 0-back task does not. Although the RVE spatial distribution is not uniform, it suggests greater and more significant changes than does event-related beta-band power. These same conditions viewed as beta-band power show the expected focal changes in DLPFC. The difference between RVE and beta-band power images suggests that there vast

tracts of cortex that have changed into an “up” state (i.e., activation of the thalamo-cortical attention circuit), without significant expenditure of energy. Is this arousal? Is it readiness? One might speculate that this widespread *increase* in entropy represents a decrease in the cortical excitability threshold—necessary for the efficient performance of the 2-back task. Such a mechanism would be concordant with our observations of other slow changes in RVE, such as that seen in the self-paced voluntary finger movement study and late P_{300} components for the rare tones. By contrast, the 2-back and 1-back conditions both engage working memory but with different levels of effort. For this comparison the RVE demonstrates a number of focal increases bilaterally in anterior cingulate gyri, supraorbital prefrontal cortex, and a focal decrease in the left inferior temporal pole ($p < 10^{-11}$) as shown in **Figure 5**. Thus, the contrast for increasing memory workload can result in both increases and decreases in RVE. The RVE results are qualitatively different from the same contrasts in task-related beta-band power (**Figures 4 and 5**). Beta-band changes appear

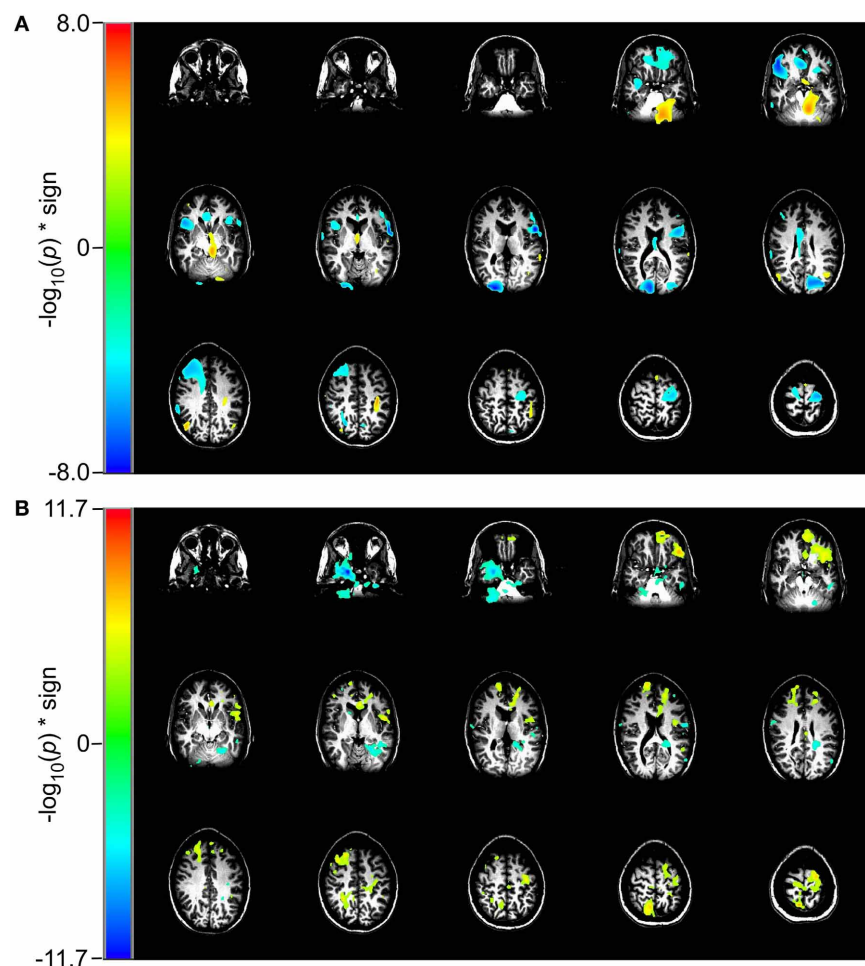


FIGURE 5 | Comparison of event-related changes in beta-band power and in entropy for 2 vs. 1-back working memory task centered on the response. Images are univariate p -values for a

Student's T -test, with a threshold of 10^{-3} . **(A)** Event-related desynchronization of beta band power (14–30 Hz) and **(B)** RVE (4–150 Hz).

mainly as ERD in dorsolateral prefrontal and parietal cortex for both 2 vs. 0-back and 2 vs. 1-back. The p -values for these regions are smaller than those for the RVE. Overall, the working memory task suggests the higher cognitive load is associated with increases in RVE.

The results obtained from these MEG studies provide evidence that RVE—specifically increases in complexity—are measures of attention and intention. The neurophysiological basis of attentional arises from the non-specific projections from the centromedian thalamic nucleus to the neocortex (Steriade, 1995). Activation of these projections depolarizes the apical dendrites of neocortical pyramidal neurons resulting in a decrease in the excitability threshold. An increase in asynchronous firing rate of these neurons will be reflected by an increase in complexity, as seen in the RVE signal. It should also be noted that this same mechanism is responsible for the low frequency negative signal appearing in the scalp EEG—the so-called SCP (Birbaumer et al., 1990). Detection of the EEG SCP signals requires careful removal of motion artifacts, eye-blinks, and higher frequency signals and

rhythms. By contrast the RVE signals corresponding to attention are readily measured in a broad bandwidth with little or no filtering. Thus RVE is a more sensitive measure of attentional mechanisms than is the SCP.

The complexity decrease associated with synchronous and evoked events such as the auditory N100 implies a *low information rate*. This calls to question the model that synchronous events represent transfer of information. Instead, we suggest a new model in which synchronous activity signifies degradation of the current cognitive context in a cortical region so that new information from sensory or associative sources can be incorporated into a new cognitive context. Exploration of this model is the subject of our ongoing research.

The RVE combines some properties of topological (relative nearness as non-numeric sequences) and metric (probability measure theoretic) entropies (Cornfield et al., 1982; Ornstein, 1989; Milnor, 1997). The log density of states (Equation 7) is a metric entropy. These states are a consequence of mapping topologically ordered sequences to vectorial states, then mapping these

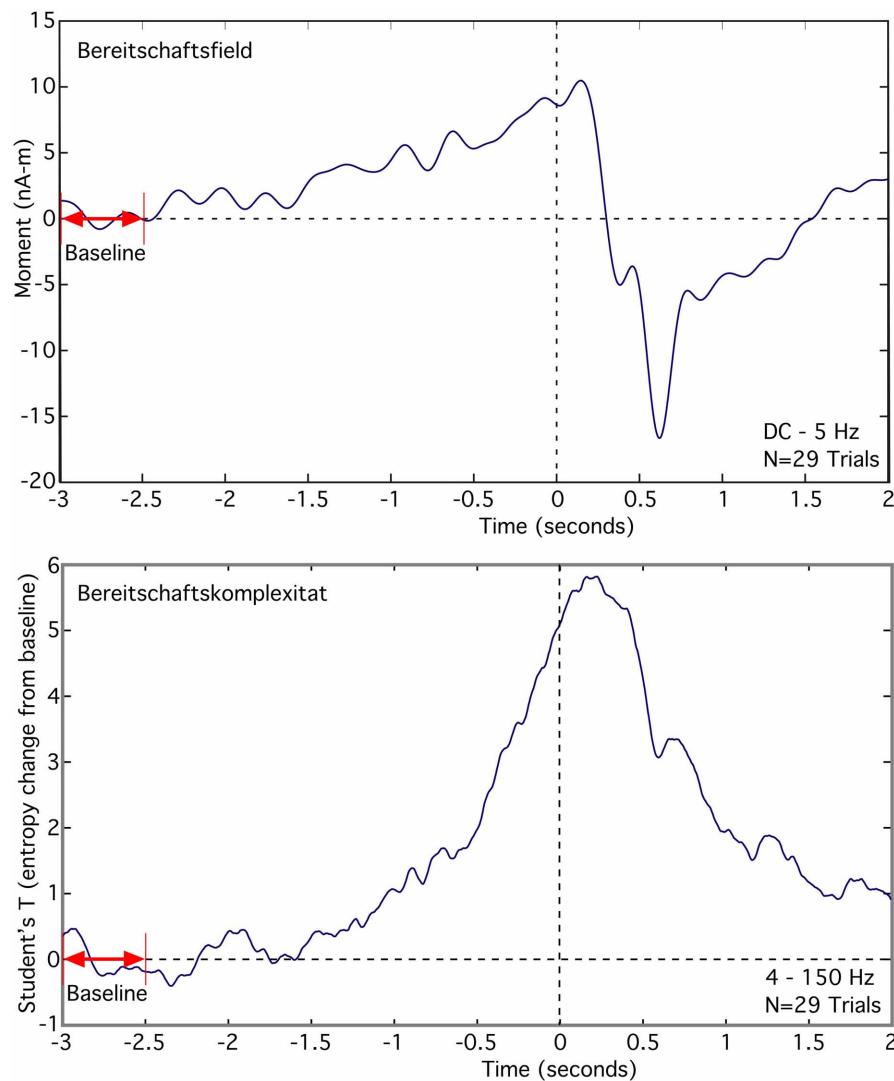


FIGURE 6 | Comparison of the averaged Bereitschaftsfield (DC to 5 Hz) to the averaged Bereitschaftskomplexität RVE (4–150 Hz) for a single voxel in premotor cortex contralateral to a self-paced voluntary finger movement (button press).

states into a symbolic dynamic topology, eventuating in the metric transformation of the density distribution of these symbolic states (Pollicott and Yuri, 1998).

The RVE is distinguished from ME in the order of operations. In ME, the states are defined by discrete partitions in the embedding space of *measured values* and the probabilities are derived from the number of counts in each n -dimensional partition. It can be difficult to define the boundaries of each partition—particularly when signals cover a wide dynamic range. In such cases, signal amplitude compression may be required. By contrast, RVE reduces the measurements to their rank values for each length W window. Each rank vector represents a state symbol (Adler et al., 1977) and partitions, as such, are no longer required; it only necessary to count the instances of each symbol. This step not only defines a finite number of available states but also renders the entropy measure-independent of absolute signal amplitude.

As a consequence of the integrator decay time constant, the mean of the entropy time series will be smaller than in the limit of $\alpha = 1$ for which integration takes place over all samples. A similar decrement from the maximum entropy was also found in older methods (Walters, 1982). This decay time constant is required in order to maintain sensitivity to entropy changes over time. As expected, RVE approaches its limiting value as the decay constant is increased. At long decay constants fluctuations will be very small due to the accruing “memory” of the state counts. For MEG data sampled at 600 Hz and $W = 5$, we choose a $1/e$ time constant 0.6 s—corresponding to three times the number of states, in samples. This is more efficient than the sliding block methods for which entropy is computed for the number of samples in each block, and the blocks are advanced one sample at a time (Adler and Marcus, 1979). It should be mentioned here that altering the decay time constant changes the peak-to-peak amplitude

of the entropy fluctuations, but not its waveform (thus the RVE measure)—provided the time constant is sufficiently long to allow the state counts to fully populate.

There is an inherent frequency bias in the RVE method. The duration of the sub-window from which the rank vector is obtained constitutes a highpass filter on the signal. To a measurable extent the contribution of the entropy of signals with wavelengths longer than this sub-window is reduced. This suggests future innovations of the RVE algorithm for the purpose of better conserving the range of broadband responses. For example, one could apply a 6 dB per octave increase in the measured signal for frequencies below the window cutoff to compensate for the decline in sensitivity.

The tunable parameters for RVE analysis are the rank vector length, signal lowpass frequency, and integration decay rate. We here note practical limitations of rank vector length. Since the number of possible states is exactly determined by the length of the rank vector, if $W = 4, 5$, and 6 there are 24, 120, and 720 possible states, respectively. As a compromise, $W = 5$ seems most attractive because it allows for a sufficient number of states to reflect the probabilities without too many or too few states to be impractical (i.e., to not fully represent the latent structure in the time-series). For example, when $W = 3$, there are only 6 allowable states and when $W = 7$ there are 5040 allowable states. The latter would require a very large number of samples to accurately reflect the probabilities of each state, while the former would provide very poor state resolution. The lags defining the phase space are determined by the signal lowpass frequency. Since signal complexity decreases as bandwidth is reduced, it defeats the purpose of complexity analysis to limit the bandwidth. Therefore, the

low pass should reflect the information content of the measured signal. Lastly, the integrator decay rate is selected so as to allow the state count histogram to be well-populated (a sufficient number symbol counts for obtaining a good estimate of the entropy) and have a short lead-in time (i.e., the time required for the entropy to approach its asymptotic value). For 120 states, and a sample rate of 600 Hz, a $1/e$ time of 0.6 s is sufficient. Note that, as indicated above, the integrator decay does not lowpass filter the entropy waveform. It only affects the relative amplitude of the entropy fluctuations around the asymptotic value.

The normal cerebral cortex is never quiescent—even when not engaged in specific tasks. We have shown that stimuli and tasks modulate the regional entropy of cortical activity, inducing either increases or decreases in our RVE measure. The question posed initially was whether RVE provides new information on brain activity that is not obvious when observing only changes in source power. We conclude that RVE is a sensitive measure of brain signal complexity that complements other MEG functional imaging techniques including event and task-related changes in power.

ACKNOWLEDGMENTS

The authors wish to thank Tom Holroyd for software for extracting the hull model from the MRI, and Frederick Carver and Judy Mitchell-Francis for access to the many archived MEG studies, and to Surjo Soekedar and Matthias Witkowski for the self-paced finger movement datasets. This work is supported by the Intramural Research Program of the National Institute of Mental Health, with additional support for Arnold J. Mandell from the Franklin-Fetzer Fund.

REFERENCES

- Adler, R. L., Goodwyn, L. W., and Weiss, B. (1977). Equivalence of topographic Markov shifts. *Israel. J. Math.* 27, 49–63.
- Adler, R. L., and Marcus, B. (1979). Topological entropy and equivalence of dynamical systems. *Mem. Am. Math. Soc.* 219, 114–126.
- Bandt, C., and Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.* 88:174102. doi: 10.1103/PhysRevLett.88.174102
- Berger, S. (1989). “On the electroencephalogram of man. The fourteen original reports on the human electroencephalogram,” in *Electroencephalography and Clinical Neurophysiology*, Trans. from the German and ed P. Gloor (1969) (New York, NY: Elsevier), xii + 350, (Suppl. 28), 151–171.
- Birbaumer, N., Elbert, T., Canavan, A. G., and Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiol. Rev.* 70, 1–41.
- Bruna, R., Poza, J., Gomez, C., Garcia, M., Fernandez, A., and Hornero, R. (2012). Analysis of spontaneous MEG activity in mild cognitive impairment and Alzheimer’s disease using spectral entropies and statistical complexity measures. *J. Neural Eng.* 9:036007. doi: 10.1088/1741-2560/9/3/036007
- Busse, F. H., and Clever, R. M. (1979). Instabilities of convection rolls in a fluid of moderate Prandtl number. *J. Fluid Mech.* 91, 319–327.
- Cohen, D. (1968). Magnetoencephalography: evidence of magnetic field produced by alpha rhythm currents. *Science* 161, 784–786.
- Cohen, L. B., Salzberg, B. M., and Grinvald, A. (1978). Optical methods for monitoring neuron activity. *Annu. Rev. Neurosci.* 1, 171–182.
- Coppola, R., Callicott, J. H., Holroyd, T., Verchinski, B. A., Sust, S., and Weinberger, D. R. (2004). “MEG activation comparison to fMRI BOLD for a working memory task,” in *Proceedings of the 14th International Conference on Biomagnetism*, eds E. Halgren, S. Ahlfors, M. Härmäläinen, and D. Cohen (Boston, MA), 7.
- Cornfield, B. R., Fomin, V., and Sinai, Y. (1982). *Ergodic Theory*. Berlin: Springer-Verlag.
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.
- Crutchfield, J. P., and Feldman, D. P. (2003). Regularities unseen, randomness observed: levels of entropy convergence. *Chaos* 13, 25–54.
- Deecke, L., Weinberg, H., and Brickett, P. (1982). Magnetic-fields of the human-brain accompanying voluntary movement – Bereitschaftsmagnetfeld. *Exp. Brain Res.* 48, 144–148.
- De Munck, J. C. (1992). A linear discretization of the volume conductor boundary integral equation using analytically integrated elements. *IEEE Trans. Biomed. Eng.* 39, 986–990.
- Donchin, E. (1981). Presidential address, 1980. Surprise!... Surprise? *Psychophysiology* 18, 493–513.
- Eckmann, J. P., and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* 57, 617–656.
- Fernandez, A., Lopez-Ibor, M. I., Turrero, A., Santos, J. M., Moron, M. D., Hornero, R., et al. (2011). Lempel-Ziv complexity in schizophrenia: a MEG study. *Clin. Neurophysiol.* 122, 2227–2235.
- Fife, A. A., Vrba, J., Haid, G., Hoang, T., Kubik, P. R., Lee, S., et al. (2002). “A 275-channel whole cortex MEG system,” in *Proceedings of the 13th International Conference on Biomagnetism*, eds H. Nowak, J. Hauesen, F. Gießler, and R. Huonker (Berlin: VDE Verlag GMBH), 912–914.
- Gauss, C. F. (1823). “Theoria combinationis observationum erroribus minimis obnoxiae,” Trans. Stewart, G. W. (1995). “Theory of the combination of observations least subject to errors,” in *Classics in Applied Mathematics, Chap. 4* (Philadelphia, PA: SIAM), 30.
- Gollub, J., and Steinman, J. (1981). Doppler imaging of the onset of turbulent convection. *Phys. Rev. Lett.* 47, 505–509.
- Gomez, C., Poza, J., Garcia, M., Fernandez, A., and Hornero, R. (2011). Regularity analysis of spontaneous MEG activity in Attention-Deficit/Hyperactivity

- Disorder. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2011, 1765–1768.
- Grynspan, F., and Geselowitz, D. B. (1973). Model studies of the magnetocardiogram. *Biophys. J.* 13, 911–925.
- Hari, R., Joutsiniemi, S. L., and Sarvas, J. (1988). Spatial resolution of neuromagnetic records: theoretical calculations in a spherical model. *Electroencephalogr. Clin. Neurophysiol.* 71, 64–72.
- Huang, M. X., Mosher, J. C., and Leahy, R. M. (1999). A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Phys. Med. Biol.* 44, 423–440.
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *J. Exp. Psychol.* 55, 352–358.
- Kolmogorov, A. N. (1958). A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces. *Dokl. Russ. Acad. Sci.* 119, 861–864.
- Kornhuber, H. H., and Deecke, L. (1964). Hirnpotentialänderungen beim Menschen vor und nach Willkurbewegungen dargestellt mit Magnetbandspeicherung und Ruckwartsanalyse. *Pflügers Arch. Eur. J. Physiol.* 281, 52.
- Mandell, A. J., Selz, K. A., Aven, J., Holroyd, T., and Coppola, R. (2011a). “Daydreaming, thought blocking and strudels in the taskless, resting human brain’s magnetic fields,” in *International Conference on Applications in Nonlinear Dynamics (ICAND 2010)*, eds V. In, P. Longhini, and A. Palacios (Lake Louise, Melville: American Institute of Physics), 7–24.
- Mandell, A. J., Selz, K. A., Holroyd, T., and Coppola, R. (2011b). “Intermittent vorticity, power spectral scaling and dynamical measures on resting brain magnetic field fluctuations,” in *Neuronal Variability and Its Functional Significance*, eds D. Glanzman and M. Ding (New York, NY: Oxford University Press), 328–390.
- Milnor, J. (1997). *Topology from the Differential Viewpoint*. Princeton, NJ: Princeton University Press.
- Naatanen, R., Gaillard, A. W., and Mantysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol. (Amst.)* 42, 313–329.
- Naatanen, R., and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425.
- Nolte, G., Fieseler, T., and Curio, G. (2001). Perturbative analytic solutions of the magnetic forward problem for realistic volume conductors. *J. Appl. Phys.* 89, 2360–2369.
- Nolte, G., Holroyd, T., Carver, F., Coppola, R., and Hallett, M. (2004). Localizing brain interactions from rhythmic EEG/MEG data. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2, 998–1001.
- Ornstein, D. S. (1989). Ergodic theory, randomness, and “chaos.” *Science* 243, 182–187.
- Ott, E. (1993). *Chaos in Dynamical Systems*. Cambridge: Cambridge University Press.
- Plonsey, R., (1981). “Generation of magnetic fields by the human body (theory),” in *Proceedings Third International Workshop on Biomagnetism*, eds S. N. Ern , H.-D. Halbohm, and H. L bbig (Berlin; New York: Walter de Gruyter), 177–205.
- Pollicott, M., and Yuri, M. (1998). *Dynamical Systems and Ergodic Theory*. Cambridge: Cambridge University Press.
- Robinson, S. E. (1989). “Theory and properties of lead field synthesis analysis,” in *Advances in Biomagnetism*, eds H. Manfred, S. J. Williamson, M. Kotani, and G. Stroink (New York, NY: Plenum Press), 599–602.
- Robinson, S. E., and Black, W. C. Jr. (1990). *Analysis of Biological Signals Using Data from Arrays of Sensors*. USA patent application 07/359,640. December 18, 1990.
- Robinson, S. E., Mandell, A. J., and Coppola, R. (2012). “Imaging complexity,” in *18th Annual Meeting of the Organization for Human Brain Mapping*. (Beijing).
- Robinson, S. E., and Rose, D. F. (1993). “Current source image estimation by spatially filtered MEG,” in *Biomagnetism: Clinical Aspects*, eds M. Hoke, S. N. Erne, Y. C. Okada, and G. L. Romani (Muenster, Germany: Elsevier Science Publishers, B.V.), 761–765.
- Ruelle, D., and Takens, F. (1971). On the nature of turbulence. *Commun. Math. Phys.* 82, 137–151.
- Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.* 32, 91–97.
- Sekihara, K., and Nagarajan, S. S. (2008). “Scalar adaptive spatial filter: deriving the optimum source orientation,” in *Adaptive Spatial Filters for Electromagnetic Brain Imaging*, ed J. H. Nagel (Berlin, Heidelberg: Springer-Verlag), 89–92.
- Shannon, C. E. (1963). *A Mathematical Theory of Communication*. Champaign-Urbana, IL: University of Illinois Press.
- Singh, K. D., Barnes, G. R., Hillebrand, A., Forde, E. M., and Williams, A. L. (2002). Task-related changes in cortical synchronization are spatially coincident with the hemodynamic response. *Neuroimage* 16, 103–114.
- Steriade, M. (1995). Neuromodulatory systems of thalamus and neocortex. *Sem. Neurosci.* 7, 361–370.
- Swinney, H., and Gollub, J. (eds.). (1981). *Hydrodynamic Instabilities and the Transition to Turbulence*. Berlin: Springer.
- Taniguchi, M., Kato, A., Fujita, N., Hirata, M., Tanaka, H., Kihara, T., et al. (2000). Movement-related desynchronization of the cerebral cortex studied with spatially filtered magnetoencephalography. *Neuroimage* 12, 298–306.
- Temin, R. (1997). *Infinite Dimensional Dynamical Systems in Mechanics and Physics*. Applied mathematical sciences, 2nd Edn. (Berlin: Springer), 648.
- Vakorin, V. A., Ross, B., Krakovska, O., Bardouille, T., Cheyne, D., and McIntosh, A. R. (2010). Complexity analysis of source activity underlying the neuromagnetic somatosensory steady-state response. *Neuroimage* 51, 83–90.
- Van Veen, B. D., Van Drongelen, W., Yuchtman, M., and Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* 44, 867–880.
- Vrba, J., and Robinson, S. E. (2002). SQUID sensor array configurations for magnetoencephalography applications. *Supercond. Sci. Technol.* 15, R51–R89.
- Walters, P. (1982). *An Introduction to Ergodic Theory*. Berlin: Springer.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 August 2012; accepted: 30 December 2012; published online: 24 January 2013.

Citation: Robinson SE, Mandell AJ and Coppola R (2013) Spatiotemporal imaging of complexity. *Front. Comput. Neurosci.* 6:101. doi: 10.3389/fncom.2012.00101

Copyright © 2013 Robinson, Mandell and Coppola. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



The anesthetic propofol shifts the frequency of maximum spectral power in EEG during general anesthesia: analytical insights from a linear model

Axel Hutt*

INRIA CR Nancy - Grand Est, Team CORTEX, Villers-les-Nancy, France

Edited by:

Peter Beim Graben,
Humboldt-Universität zu Berlin,
Germany

Reviewed by:

Moirá Steyn-Ross, University of
Waikato, New Zealand
Rikkert Hindriks, Universitat
Pompeu Fabra, Spain

***Correspondence:**

Axel Hutt, INRIA CR Nancy -
Grand Est, Team CORTEX, 615 rue
du Jardin Botanique, 54602
Villers-les-Nancy, France.
e-mail: axel.hutt@inria.fr

The work introduces a linear neural population model that allows to derive analytically the power spectrum subjected to the concentration of the anesthetic propofol. The analytical study of the power spectrum of the systems activity gives conditions on how the frequency of maximum power in experimental electroencephalographic (EEG) changes dependent on the propofol concentration. In this context, we explain the anesthetic-induced power increase in neural activity by an oscillatory instability and derive conditions under which the power peak shifts to larger frequencies as observed experimentally in EEG. Moreover the work predicts that the power increase only occurs while the frequency of maximum power increases. Numerically simulations of the systems activity complement the analytical results.

Keywords: general anesthesia, propofol, neural fields, power spectrum, EEG

1. INTRODUCTION

General anesthesia (GA) is an important medical application in today's hospital surgery. Although GA is omnipresent in recent medicine, its underlying neural interactions have been a long-standing mystery. In the last decades, the anesthetic phenomena have attracted theoreticians, e.g., (Steyn-Ross et al., 1999; Bojak and Liley, 2005; Hutt, 2011), who aim to describe mathematically some major experimental phenomena by population models (Steyn-Ross et al., 2004, 2012; Bojak and Liley, 2005; Hutt and Longtin, 2009; Hindriks and van Putten, 2012), or spiking-neuron models (McCarthy et al., 2008; Ching et al., 2010). Most theoretical studies aim to explain signal features of electroencephalographic (EEG) data observed during anesthesia. Such features comprise the diminution of α -activity accompanied by a subsequent enhancement of δ -activity while increasing anesthetic concentration (Gugino et al., 2001; Cimenser et al., 2011; Murphy et al., 2011) and the power enhancement of activity induced by some anesthetics (McCarthy et al., 2008; Ching et al., 2010). Another example is the increase of the frequency of maximum EEG-power to higher values as observed experimentally in several studies (Gugino et al., 2001; Ching et al., 2010; Murphy et al., 2011; Boly et al., 2012; Hindriks and van Putten, 2012). The current work focusses on the power enhancement and the frequency shift of maximum power while increasing the anesthetic concentration and gives insights into its origin by the analytical treatment of a linear neural field model.

One of the major objectives of this work is to answer the question whether it is possible to explain spectral EEG-features observed during GA by a low-dimensional linear model. The advantage of such a reduced model is the analytical tractability and an identification of underlying neural interactions or even the origin of the spectral feature. Here the difficulty is to find a simple

model, that, however, still involves important, i.e., realistic and neural interactions. We are convinced that such a model has been found in a previous work (Hutt and Longtin, 2009). The present work will simplify further this spatio-temporal model while taking into account the biophysical effects of the anesthetic propofol on synaptic receptors and hence retaining the neurobiological plausibility.

The simplicity of the model will allow to reveal the effect of different actions of the anesthetic propofol on synaptic receptors on the frequency of maximum spectral power. Moreover, the work gives criteria under which conditions the frequency of maximum power increases with increasing propofol concentration and when it may decrease. In the analytical treatment, we will see that the power enhancement for larger propofol concentrations may be explained by an oscillatory instability and we predict that it always occurs while the frequency of maximum power increases.

2. METHODS

2.1. THE MODEL

The neural field model under study (Hutt and Longtin, 2009) describes the evolution of the mean membrane potential of a neural population in a small spatial patch at spatial location x and at time t . Similar models have been derived and studied before (Wilson and Cowan, 1972; Amari, 1977; Ermentrout, 1998) and applied successfully to explain spatio-temporal neural activity observed experimentally (Ermentrout and Cowan, 1979; Huang et al., 2004; Angelucci and Bressloff, 2006; Schwabe et al., 2006). The population includes both excitatory and inhibitory neurons and takes into account excitatory and inhibitory synapses. Assuming that excitatory and inhibitory neurons exhibit identical effective membrane potentials, i.e., an identical difference between excitatory and inhibitory post-synaptic

potentials, the mean excitatory and inhibitory post-synaptic potentials $V_e(x, t)$ and $V_i(x, t)$, respectively, obey

$$\begin{aligned}\hat{L}_e V_e(x, t) &= a_e \int_{\mathcal{D}} K_e(x-y) S_e[V_e(y, t) - V_i(y, t)] dy + I(x, t) \\ \hat{L}_i V_i(x, t) &= a_i \int_{\mathcal{D}} K_i(x-y) S_i[V_e(y, t) - V_i(y, t)] dy\end{aligned}\quad (1)$$

with the circular spatial population domain \mathcal{D} of length L , i.e., assuming periodic boundary conditions. The model under study differs from some other previous models, e.g., by Liley and Bojak (2005), by the implementation of synaptic action, generation of action potentials or axonal connectivity [see also the work of Coombes et al. (2007) for a comparison of the current model and other models]. The functionals $S_e[\cdot]$ and $S_i[\cdot]$ are continuously increasing and represent the population firing rate of excitatory and inhibitory neurons, respectively. In the population the single neurons are connected by a complex system of axons from neuron somata to synapses. The kernels $K_e(x)$ and $K_i(x)$ are the probability density of such connections in the population. Here axonal transmission delay is neglected for simplicity although it is straightforward to include it in this type of model (Hutt and Longtin, 2009). In addition, the model considers excitatory and inhibitory synapses, $\hat{L}_e = \hat{L}_e(d/dt)$ and $\hat{L}_i = \hat{L}_i(d/dt)$ denote functional operators describing the corresponding temporal synaptic response phase and the factors a_e , a_i represent the corresponding synaptic efficacies.

Mathematically, the differential operators are the inverse of the integral operators in

$$V(t) = \int_{-\infty}^t h(t-\tau) P(\tau) d\tau \quad (2)$$

where $h(t)$ is the synaptic response function, or more precisely the electric current response in the synaptic receptor to an impact of binding neurotransmitters (Koch, 1999). The function $P(\tau) > 0$ is the mean pulse activity arriving at the synapses. In a reasonable approximation, the response function reads

$$h(t) = \frac{a}{\tau} e^{-t/\tau}$$

with the decay time constant τ and the synaptic efficacy $a > 0$. Then the response amplitude is $h(0) = a/\tau$ and the charge transferred in the receptor $\rho = a$ is the time integral over the current flow. The corresponding differential operator stipulates $\hat{L}V(t) = aP(t)$ leading to

$$\hat{L} \left(\frac{\partial}{\partial t} \right) = \tau \frac{\partial}{\partial t} + 1$$

and Equation (2) re-casts to

$$\tau \frac{\partial V(t)}{\partial t} + V(t) = aP(t).$$

These expressions hold for excitatory and inhibitory synapses.

The synaptic receptors are major targets of anesthetic agents. The present work considers the action of propofol on inhibitory synaptic and extra-synaptic GABA_A-receptors. The former receptor is supposed to be a major anesthetic target (Franks and Lieb, 1994) and there is growing evidence that extra-synaptic inhibitory receptors may play an important role in anesthesia as well (Orser, 2006; Hutt, 2012). The subsequent sections consider effects on synaptic receptors due to the well-established experimental evidence. Hence, the synaptic parameters of inhibitory synaptic receptors depend on the anesthetic concentration and are parameterized by the factor $p \geq 1$ (Steyn-Ross et al., 2001), i.e., the decay time of inhibitory synapses $\tau_2 = \tau_2(p)$ and the corresponding synaptic efficacy $a_i = a_i(p)$ depend on p .

The input in Equation (1) fluctuates randomly in space and time with $\xi(x, t)$ about a constant value $I_0 = \text{const}$, i.e., $I(x, t) = I_0 + \xi(x, t)$. The random fluctuations are independent in space and time and thus obey $\langle \xi(x, t) \rangle = 0$, $\langle \xi(x, t) \xi(y, T) \rangle = 2D\delta(t-T)\delta(x-y)$, where $\langle \cdot \rangle$ denotes the ensemble average.

Considering the latter definitions of synaptic properties, anesthetic action and external input, the final model equations read

$$\begin{aligned}\tau_1 \frac{\partial V_e(x, t)}{\partial t} &= -V_e(x, t) + a_e \int_{\mathcal{D}} K_e(x-y) S_e[V_e(y, t) \\ &\quad - V_i(y, t)] dy + I_0 + \xi(x, t) \\ \tau_2(p) \frac{\partial V_i(x, t)}{\partial t} &= -V_i(x, t) + a_i(p) \int_{\mathcal{D}} K_i(x-y) S_i[V_e(y, t) \\ &\quad - V_i(y, t)] dy\end{aligned}\quad (3)$$

with the decay time of excitatory synapses τ_1 .

Assuming that the random fluctuations are small and do not affect the stationary state [in contrast to recent results gained from non-linear systems (Hutt et al., 2007; Hutt, 2008)], the stationary state $V_e(x, t) = V_e^0 = \text{const}$, $V_i(x, t) = V_i^0 = \text{const}$ obeys $V_e^0 = a_e S_e[V_-] + I_0$, $V_i^0 = a_i(p) S_i[V_-]$ with $V_- = V_e^0 - V_i^0 = a_e S_e[V_-] - a_i(p) S_i[V_-] + I_0$ (Hutt and Longtin, 2009).

2.2. THEORETICAL POWER SPECTRUM

To compute the power spectrum, we employ the method of Greens function. Let us assume the activity variable vector $\mathbf{x}(t) \in \mathcal{R}^N$, the matrix \mathbf{A} , the external input vector $\xi(t) \in \mathcal{R}^N$, the Greens function matrix $\mathbf{G}(t) \in \mathcal{R}^{N \times N}$ and

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x} + \xi(t).$$

Then, for $t \rightarrow \infty$, the solution of the system is

$$\mathbf{x}(t) = \int_{-\infty}^{\infty} \mathbf{G}(t-\tau) \xi(\tau) d\tau. \quad (4)$$

and the Greens function obeys

$$\dot{\mathbf{G}} - \mathbf{A}\mathbf{G}(t) = \mathbf{1}\delta(t)$$

with the unitary matrix $\mathbf{1} \in \mathcal{R}^{N \times N}$. Applying the Fourier transform

$$\mathbf{G}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\mathbf{G}}(\omega) e^{i\omega t} d\omega. \quad (5)$$

yields

$$\tilde{\mathbf{G}}(\omega) = \frac{1}{\sqrt{2\pi}}(i\omega\mathbf{1} - \mathbf{A})^{-1} \quad (6)$$

$$= \frac{1}{\sqrt{2\pi}} \frac{\mathbf{F}(i\omega)}{P(i\omega)} \quad (7)$$

with the matrix $\mathbf{F}(i\omega)$ and the characteristic polynomial $P(i\omega)$. The matrix \mathbf{F} in Equation (7) includes the matrix elements of the inverse of $i\omega\mathbf{1} - \mathbf{A}$ and the characteristic polynomial P represents the corresponding matrix determinant.

Inserting Equation (6) into (5) allows to compute $\mathbf{G}(t)$ by the residue theorem in functional analysis

$$\mathbf{G}(t) = 2\pi i \sum_{n=1}^r \text{Res}(z_n, t) \Theta(t) \quad (8)$$

with the Heaviside function $\Theta(t)$ and the residues matrix $\text{Res}(z_n, t)$ of $\mathbf{F}(z)/P(z)$ at the roots z_n of the characteristic equation $P(z) = 0$. The condition $t > 0$ considered by the Heaviside function is the mathematical condition for the validity of Equation (8) while it also guarantees the causality of the system response. Equation (8) together with Equation (4) determines the time dependence of the solution x and is computed explicitly in section 3.4.

Finally, the power spectral density matrix $\mathbf{S}(\omega)$ of \mathbf{x} is the Fourier transform of the auto-correlation function matrix $\langle \mathbf{x}^t(t)\mathbf{x}(t-T) \rangle$ (Wiener-Khinchine Theorem) leading to

$$\mathbf{S}(\omega) = 2D\sqrt{2\pi}\tilde{\mathbf{G}}(\omega)\tilde{\mathbf{G}}^t(-\omega),$$

where the high index t denotes the transposed vector or matrix.

3. RESULTS

3.1. EFFECT OF PROPOFOL

The effect of the anesthetic propofol on neural properties is manifold (Alkire et al., 2008). It affects properties of membrane ion channels, synaptic receptors and extra-synaptic receptors, see Franks and Lieb (1994) for a review. Kitamura et al. (2002) have revealed in an experimental study how propofol affects post-synaptic phasic responses of inhibitory synapses to spontaneous neurotransmitter release. They have found that the decay time constant decreases with increasing anesthetic blood concentration, the charge transfer increases while the amplitude of the responses remains constant. Some previous studies (Hutt and Longtin, 2009; Hindriks and van Putten, 2012) have implemented these effects for a bi-exponential synaptic response function. The present work considers an exponential decay phase due to its mathematical simplicity, which nevertheless reflects the major anesthetic impact. To this end, the phasic response at inhibitory synapses is determined by the decay time constant and the response amplitude. Introducing the parameter $p = \tau_2(p)/\tau_2(1)$, $p = 1$ reflects the absence of anesthetic agents and increases of the anesthetic concentration yields an increase of p . In addition Kitamura et al. (2002) have shown that the amplitude in cortical neurons remains constant or changes slightly only, i.e.,

$h(0) = a/\tau_2(p) \approx \text{const}$ and the charge transfer $\rho(p) = a$ increase with increasing anesthetic concentration. In the case of constant the model

$$\tau_2(p) = \tau_2(1)p, \quad a_i(p) = H_0p. \quad (9)$$

with the constant $H_0 > 0$. These choices of anesthetic actions reflect different synaptic mechanism. The first relation reflects the increase of the time constant with increasing anesthetic concentration, and the second one both the constant amplitude and the resulting increasing charge transfer. However, synapses may have different properties in different brain areas, e.g., synapses in cortico-thalamic connections increase their amplitude with increasing propofol concentrations (Ying and Goldstein, 2005). Hence the relations (Equation 9) are specific properties of inhibitory synapses on cortical neurons only.

3.2. THE LINEAR MODEL

The following investigation considers the stability of stationary states and spectral properties of small deviations about them. These small deviations represent fluctuating currents on the dendritic trees of the neurons in the population generating an electric field on the scalp. They originate from fluctuations in the neuron membranes or from spontaneous neurotransmitter emission at synapses. The generated electric field is measured in terms of voltage differences between two spatial locations on the scalp which is the electroencephalogram (EEG) (Nunez and Srinivasan, 2006).

The small fluctuations about the stationary states $u_e(x, t) = V_e(x, t) - V_e^0$, $u_i(x, t) = V_i(x, t) - V_i^0$ obey

$$\begin{aligned} \tau_1 \frac{du_e(x, t)}{dt} &= -u_e(x, t) + a_e S'_e \int_{\mathcal{D}} K_e(x-y) (u_e(y, t) \\ &\quad - u_i(y, t)) dy + \xi(x, t) \\ \tau_2(p) \frac{du_i(x, t)}{dt} &= -u_i(x, t) + a_i(p) S'_i \int_{\mathcal{D}} K_i(x-y) (u_e(y, t) \\ &\quad - u_i(y, t)) dy \end{aligned}$$

with the somatic non-linear gain $S'_{e,i} = dS_{e,i}(x)/dx$ at $x = \{V_e^0, V_i^0\}$. By virtue of the finite spatial domain, the small deviations about the stationary state may be expanded into a discrete infinite Fourier series. The major contribution of neuronal activity to encephalographic activity on the scalp is modeled successfully by a spatially constant Fourier mode $u_e(x, t) = x(t)$, $u_i(x, t) = y(t)$ (Nunez and Srinivasan, 2006). Then $x(t)$, $y(t)$ obey

$$\begin{aligned} \tau_1 \frac{dx(t)}{dt} &= (-1 + N_1)x(t) - N_1y(t) + \gamma(t) \\ \tau_2(p) \frac{dy(t)}{dt} &= N_2x(t) + (-1 - N_2)y(t) \end{aligned} \quad (10)$$

with the synaptic non-linear gains $N_1 = a_e S'_e \tilde{K}_e(0)\sqrt{L}$, $N_2 = N_2(p) = a_i(p) S'_i \tilde{K}_i(0)\sqrt{L}$, the spatial Fourier transform of the kernels $\tilde{K}_e(k)$, $\tilde{K}_i(k)$ and the spatial Fourier transform of the external noise at zero wavenumber $\gamma(t)$. We point out that $\langle \gamma(t) \rangle = 0$, $\langle \gamma(t)\gamma(T) \rangle = 2D\delta(t-T)$.

3.3. STABILITY ANALYSIS

To study the dynamics about the stationary state, at first let us neglect the external input $\xi(t)$ since in a first approximation the stability of the linear system does not depend on the external input. Then the stationary state is asymptotically stable if the characteristic equation of Equation (10)

$$\lambda^2 - \lambda Tr + det = 0 \quad (11)$$

with $\lambda \in \mathbb{C}$ and

$$Tr = \frac{N_1 - 1}{\tau_1} - \frac{N_2 + 1}{\tau_2} \quad (12)$$

$$det = \frac{N_1 N_2}{\tau_1 \tau_2} - \frac{N_1 - 1}{\tau_1} \frac{N_2 + 1}{\tau_2}. \quad (13)$$

has solutions $Re(\lambda) < 0$. Here Tr and det are the trace and determinant of the linear matrix in Equation (10), respectively. In addition, the stationary state is a stable focus if $Im(\lambda) = \Omega \neq 0$, $\Omega \in \mathbb{R}$ and

$$\Omega^2 = \frac{N_1 N_2}{\tau_1 \tau_2} - \frac{1}{4} \left(\frac{N_1 - 1}{\tau_1} + \frac{N_2 + 1}{\tau_2} \right)^2.$$

We observe immediately from Equation (12), that the stable focus is asymptotically stable if $Tr < 0$ or $N_1 < 1$ for all other parameters and the system may lose stability only if $N_1 > 1$. Since the present work aims to give conditions for certain oscillation frequencies in the population, the subsequent part of the work considers stable foci only. In addition the inhibitory synaptic time scale $\tau_2 = \tau_2(p)$ depends on the anesthetic concentration, but not τ_1 . Thus τ_1 is treated as a constant.

The new variables $a = (N_1 - 1)^2$, $b = N_1 N_2 - N_1 + N_2 + 1$ and $c = (N_2 + 1)^2$ depend solely on N_1 , N_2 , and simplify the notation in the following analysis. Then for $N_1 > 1$, stable foci stipulate

$$Tr = \frac{\sqrt{a}}{\tau_1} - \frac{\sqrt{c}}{\tau_2} < 0 \quad \text{or} \quad \frac{\tau_2}{\tau_1} < \sqrt{\frac{c}{a}}, \quad (14)$$

$$\Omega = \frac{1}{2} \sqrt{\frac{1}{\tau_1} \left(\frac{2b}{\tau_2} - \frac{a}{\tau_1} - \frac{c\tau_1}{\tau_2^2} \right)} \in \mathbb{R}. \quad (15)$$

The last equation implies that the determinant is positive definite, i.e.,

$$d(\tau_2) = a\tau_2^2 - 2b\tau_1\tau_2 + c\tau_1^2 < 0.$$

A comparison to Equation (15) reveals that $d(\tau_2) = 0$ leads to $\Omega = 0$ which permits to compute the range of τ_2 for which the system exhibits stable foci:

$$\frac{\tau_2^-}{\tau_1} \leq \frac{\tau_2}{\tau_1} \leq \frac{\tau_2^+}{\tau_1} \quad (16)$$

$$\frac{\tau_2^-}{\tau_1} = \frac{b}{a} \left(1 + \sqrt{1 - \frac{ac}{b^2}} \right), \quad \frac{\tau_2^+}{\tau_1} = \frac{b}{a} \left(1 - \sqrt{1 - \frac{ac}{b^2}} \right)$$

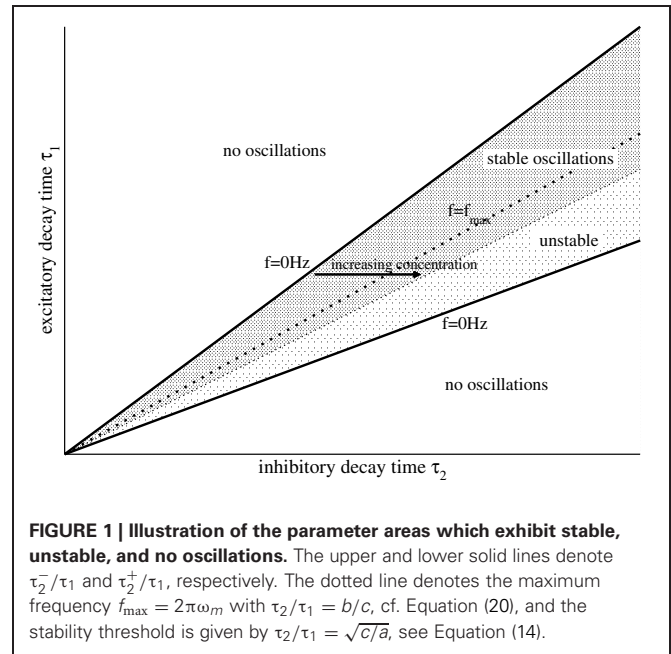


FIGURE 1 | Illustration of the parameter areas which exhibit stable, unstable, and no oscillations. The upper and lower solid lines denote τ_2^+/τ_1 and τ_2^-/τ_1 , respectively. The dotted line denotes the maximum frequency $f_{\max} = 2\pi\omega_m$ with $\tau_2/\tau_1 = b/c$, cf. Equation (20), and the stability threshold is given by $\tau_2/\tau_1 = \sqrt{c/a}$, see Equation (14).

implying $(b/a)^2 > c/a$. Condition (16) constrains the relation of both synaptic time scales τ_2/τ_1 by N_1 and N_2 , see **Figure 1** for the corresponding parameter space.

To learn more about the dynamics of the model, we consider parameters yielding strong oscillations with a predefined frequency, such as $f = 4$ Hz (δ -band), $f = 10$ Hz (α -band), or $f = 15$ Hz (β -band) as observed in experiments (Cimenser et al., 2011). To this end, we fix the frequency $\Omega = 2\pi f$ in Equation (15). Then inserting the threshold condition $\tau_2/\tau_1 = \sqrt{c/a}$ given in Equation (14) into (15) for a fixed τ_1 yields a relation between N_1 and N_2 , see **Figure 2A**. Similarly the condition $\Omega = 0$ given by Equation (16) determines the values of N_1 and N_2 at the threshold of oscillations. **Figures 2B–E** plot the parameter space $\tau_2/\tau_1 - N_2$ where the system exhibits stable oscillations. The figure also shows how the values change when increasing p (arrows in panels), i.e., increasing the propofol concentration, and it turns out that the system always approaches the stability threshold.

3.4. THE SPECTRAL POWER

To compute the power spectrum, we employ the method of Greens function. Equations (10) show that the solutions of the system obey

$$x(t) = \int_{-\infty}^{\infty} G_{11}(t - \tau) \gamma(\tau) d\tau, \quad y(t) = \int_{-\infty}^{\infty} G_{21}(t - \tau) \gamma(\tau) d\tau$$

with the Greens functions

$$G_{11}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{i\omega + (N_2 + 1)/\tau_2}{P(i\omega)} e^{i\omega t} d\omega \quad (17)$$

$$G_{21}(t) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{N_2/\tau_2}{P(i\omega)} e^{i\omega t} d\omega \quad (18)$$

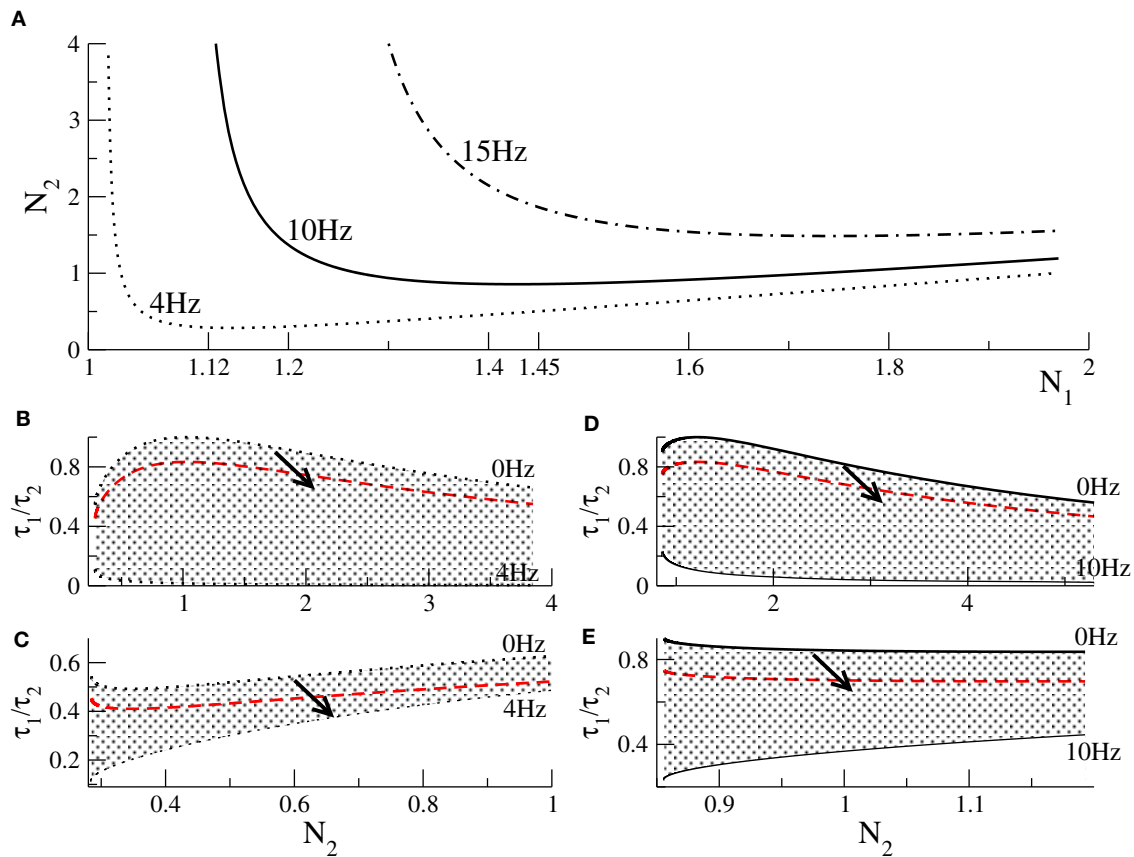


FIGURE 2 | Conditions for Hopf-instabilities and the effect of propofol action. (A) threshold of Hopf-instability at 4, 10, and 15 Hz dependent on N_1 , N_2 . Panels (B,C) show parameters for Hopf-instabilities at 4 Hz dependent on N_2 for which the corresponding values of N_1 obey $N_1 < N_{1,\min} = 1.12$ and (C) $N_1 \geq 1.12$, respectively. Panels (D,E) show parameters for Hopf-instabilities at 10 Hz for $N_1 < N_{1,\min} = 1.45$ and $N_1 \geq 1.45$, respectively. In each panel (B–E) the top line reflects no

oscillations, i.e., $f = 0$ Hz, and the bottom line denotes the corresponding stability threshold. Hence the shaded areas include parameters for which the system exhibits stable oscillations. The black solid and dotted lines reflect the situation of no anesthetic action, i.e., $p = 1.0$. The red-dashed line denotes the values of $\tau_1/\tau_2 \times p$ with $p = 1.3$. The arrows illustrate how the parameters change when anesthetic action is increased, i.e., when p is increased.

and $P(i\omega) = -\omega^2 - i\omega Tr + det$ using the definitions of Tr and det in Equations (12) and (13). In fact, $P(i\omega)$ in Equations (17) and (18) is the characteristic polynomial in Equation (11) for $\lambda = i\omega$.

In the following in a good approximation (Nunez and Srinivasan, 2006) we assume that the experimental encephalographic data that we want to model originates from the excitatory synapses, i.e., $x(t)$. This assumption is reasonable since the encephalographic activity is observed due to aligned apical dendritic branches and more excitatory synapses than inhibitory synapses are located on the apical branches of dendrites on cortical neurons. However, it is also possible to derive the power spectral density for the difference of excitatory and inhibitory potentials $x(t) - y(t)$ (Hutt and Longtin, 2009). Now the application of the residue theorem allows to compute the integrals in Equation (17) by

$$G_{11}(t) = \frac{1}{2\pi i} \sum_{n=1}^2 \text{Res}(z_n) \Theta(t)$$

with the residues

$$\begin{aligned} \text{Res}(z_1) &= \frac{1}{2\pi i} \frac{\lambda_1 + (N_2 + 1)/\tau_2}{\lambda_1 - \lambda_2} e^{\lambda_1 t}, \\ \text{Res}(z_2) &= \frac{1}{2\pi i} \frac{\lambda_2 + (N_2 + 1)/\tau_2}{\lambda_2 - \lambda_1} e^{\lambda_2 t} \end{aligned}$$

and the roots $\lambda_1 = \lambda_2^* = R + i\Omega$ of the characteristic polynomial $P(\lambda) = 0$ in Equation (11). This solution is valid if and only if $\text{Res}(\lambda_n) = R < 0$. These yields

$$G_{11}(t) = e^{Rt} \left(\frac{R+Z}{\Omega} \sin(\Omega t) + \cos(\Omega t) \right) \Theta(t)$$

with $Z = -(N_2 + 1)/\tau_2$, i.e., the Greens function of x , and hence $x(t)$ itself oscillates with frequency Ω and is damped with the factor $|R|$. In other words, the solution x oscillates with the imaginary part of the root of the characteristic equation, and this frequency is already determined in the stability analysis.

Finally, the power spectral density $S(\omega)$ of x is the Fourier transform of the auto-correlation function $\langle x(t)x(t-T) \rangle$:

$$S(\omega) = 2D\sqrt{2\pi}|\tilde{G}_{11}(\omega)|^2 = 2D\sqrt{2\pi} \frac{Z^2 + \omega^2}{(R^2 + \Omega^2 - \omega^2)^2 + 4R^2\omega^2} \quad (19)$$

where $\tilde{G}_{11}(\omega)$ is the Fourier transform of the Greens function $G_{11}(t)$.

3.5. THE FREQUENCY OF MAXIMUM POWER SPECTRAL DENSITY

According to the reversed-engineering approach motivated in the previous section, this section aims to derive further conditions on model constants for certain oscillations close to instabilities. This vicinity to the stability threshold guarantees a small damping factor R and thus the power peak is located close to Ω .

There is a maximum frequency ω_m of the oscillation frequency with respect to τ_2 given in Equation (15) and reached at

$$\frac{\tau_2}{\tau_1} = \frac{b}{c} \rightarrow \omega_m = \frac{1}{2}\sqrt{c/\tau_2^2 - a/\tau_1^2} \quad \text{if} \quad \frac{\tau_2}{\tau_1} < \sqrt{c/a}. \quad (20)$$

The last condition is identical to the stability condition (14), i.e., stable systems always have a non-vanishing maximum frequency $\omega_m > 0$. Together with Equation (16):

- If $b/c < \sqrt{c/a}$, then the frequency Ω may increase or decrease while increasing τ_2 with

$$\frac{d\Omega}{d\tau_2} > 0 \quad \text{for} \quad \frac{\tau_2^-}{\tau_1} \leq \frac{\tau_2}{\tau_1} < \frac{b}{c}, \quad (21)$$

$$\frac{d\Omega}{d\tau_2} \leq 0 \quad \text{for} \quad \frac{b}{c} \leq \frac{\tau_2}{\tau_1} \leq \sqrt{\frac{c}{a}}. \quad (22)$$

- If $b/c \geq \sqrt{c/a}$, then the frequency Ω increases only while increasing τ_2 with

$$\frac{d\Omega}{d\tau_2} > 0 \quad \text{for} \quad \frac{\tau_2^-}{\tau_1} \leq \frac{\tau_2}{\tau_1} \leq \sqrt{\frac{c}{a}}. \quad (23)$$

Figure 1 shows the case $b/c < \sqrt{c/a}$. There increasing τ_2 by increasing p from small frequencies on the left border for constant τ_1 increases the oscillation frequency until reaching the dotted line, i.e., $d\Omega/d\tau_2 > 0$. Then a further increase of p decreases the frequency of the oscillations again, $d\Omega/d\tau_2 < 0$. Although this reasoning assumes that $N_2(p)$ does not change with p , it gives a first insight into the dependence of the systems oscillation frequencies on p .

Now considering the power spectral density (Equation 19), for $R^4 + R^2(2\Omega^2 - 3Z^2) + \Omega^2(\Omega^2 + Z^2) > 0$ $S(\omega)$ has a global maximum at

$$\Omega_{\text{peak}} = \sqrt{\Omega^2 + R^2 \left(1 - 4 \frac{Z^2}{R^2 + \Omega^2 + Z^2} \right)}.$$

We learn that a vanishing real part of the characteristic root $|R|$ yields a global maximum of the power spectral density at the imaginary part of the characteristic root Ω . From a reversed-engineer point of view, one could say that a strong peak in the experimental power spectral density reflects a characteristic root in the underlying linear system with a small real part. This way to interpret the spectral results allows to find analytical conditions for physiological parameters, as will be seen below.

If $R^4 + R^2(2\Omega^2 - 3Z^2) + \Omega^2(\Omega^2 + Z^2) < 0$, then the system exhibits a global maximum at $\Omega_{\text{peak}} = 0$. Since R and Ω depend on the anesthetic concentration, i.e., the parameter p , it is interesting to examine how Ω_{peak} depends on R and Ω :

$$\frac{d\Omega_{\text{peak}}}{dR} > 0 \quad \text{for} \quad R^2 > Z^2(p) - \Omega^2 \quad \text{or} \quad R^2 < Z^2(p) - \Omega^2 \quad (24)$$

$$\frac{d\Omega_{\text{peak}}}{d\Omega} > 0 \quad \text{for} \quad R^2 > \sqrt{2}|Z(p)|\Omega - \Omega^2 + Z^2(p)$$

$$\text{or} \quad R^2 < -\sqrt{2}|Z(p)|\Omega - \Omega^2 + Z^2(p) \quad (25)$$

Figure 3 illustrates these conditions and shows that Ω_{peak} is increased or decreased subjected to values of R and Ω . Importantly, we observe in **Figure 3** that Ω_{peak} increases with R and Ω for large frequencies Ω .

It remains to examine how R^2 and Ω^2 depend on p to finally gain the full description how Ω_{peak} changes with p . To this end, we consider the specific assumption (Equation 9) on the anesthetic action yielding the specific dependence of Ω^2 and R^2 to p

$$\frac{d^2R}{dp} = -2 \frac{R}{\tau_2(1)} \frac{1}{p^2} > 0$$

$$\frac{d\Omega^2}{dp} = -\frac{2}{\tau_2(1)p^2} \left(A - \frac{1}{\tau_2(1)p} \right)$$

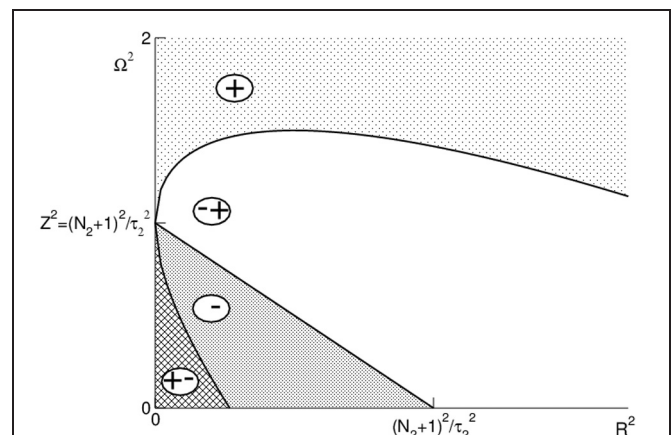


FIGURE 3 | Illustration of the conditions (24) and (25). The regions are marked as follows: (+): $d\Omega_{\text{peak}}/d\Omega > 0$, $d\Omega_{\text{peak}}/dR > 0$; (−): $d\Omega_{\text{peak}}/d\Omega < 0$, $d\Omega_{\text{peak}}/dR > 0$; (−): $d\Omega_{\text{peak}}/d\Omega < 0$, $d\Omega_{\text{peak}}/dR < 0$; (±): $d\Omega_{\text{peak}}/d\Omega > 0$, $d\Omega_{\text{peak}}/dR < 0$.

with $A = (1 - N_1)/\tau_1 - N_2(1)/\tau_2(1)$, $N_2(p) = N_2(1)p$, and $\tau_2 = \tau_2(1)p$. Defining $p_0 = (N_2(1) + \tau_2(1)(N_1 - 1)/\tau_1)^{-1}$ and re-calling $p \geq 1$, then there exist two distinct cases:

- If $p_0 < 1$, then $d\Omega^2/dp > 0$ for all p . Specifically, this stipulates

$$N_2(1) > 1 + \frac{\tau_2(1)}{\tau_1} - \frac{\tau_2(1)}{\tau_1} N_1.$$

- If $p_0 \geq 1$, then there is a small interval $1 \leq p \leq p_0$ for which $d\Omega^2/dp \leq 0$. For larger p , $d\Omega^2/dp > 0$. For instance, for $p_0 = 1.3$, i.e.,

$$N_2(1) = 1/1.3 + \frac{\tau_2(1)}{\tau_1} - \frac{\tau_2(1)}{\tau_1} N_1,$$

the frequency decreases with increasing p for clinically reasonable concentrations, i.e., $1 \leq p \leq 1.3$ (Hutt and Longtin, 2009).

Figure 4 illustrates the parameter space where $d\Omega^2/dp$ has different signs and we observe that the system may exhibit either increasing or decreasing frequencies while increasing p .

To elucidate the systems behavior at the stability threshold, we set $\tau_2(1)/\tau_1 = (N_2(1) + 1)/(N_1 - 1)$ according to Equation (14) and find $p_0 = 1/(2N_2(1) + 1) < 1$. Consequently, oscillations close to the stability threshold always increase their frequencies with increasing p . In contrast, for $1 < \tau_2(1)/\tau_1 \ll (N_2(1) + 1)/(N_1 - 1)$, i.e., systems far from the stability threshold, may exhibit values $p_0 > 1$.

Finally, **Figure 5** illustrates the temporal dynamics in two different frequency bands close to the corresponding

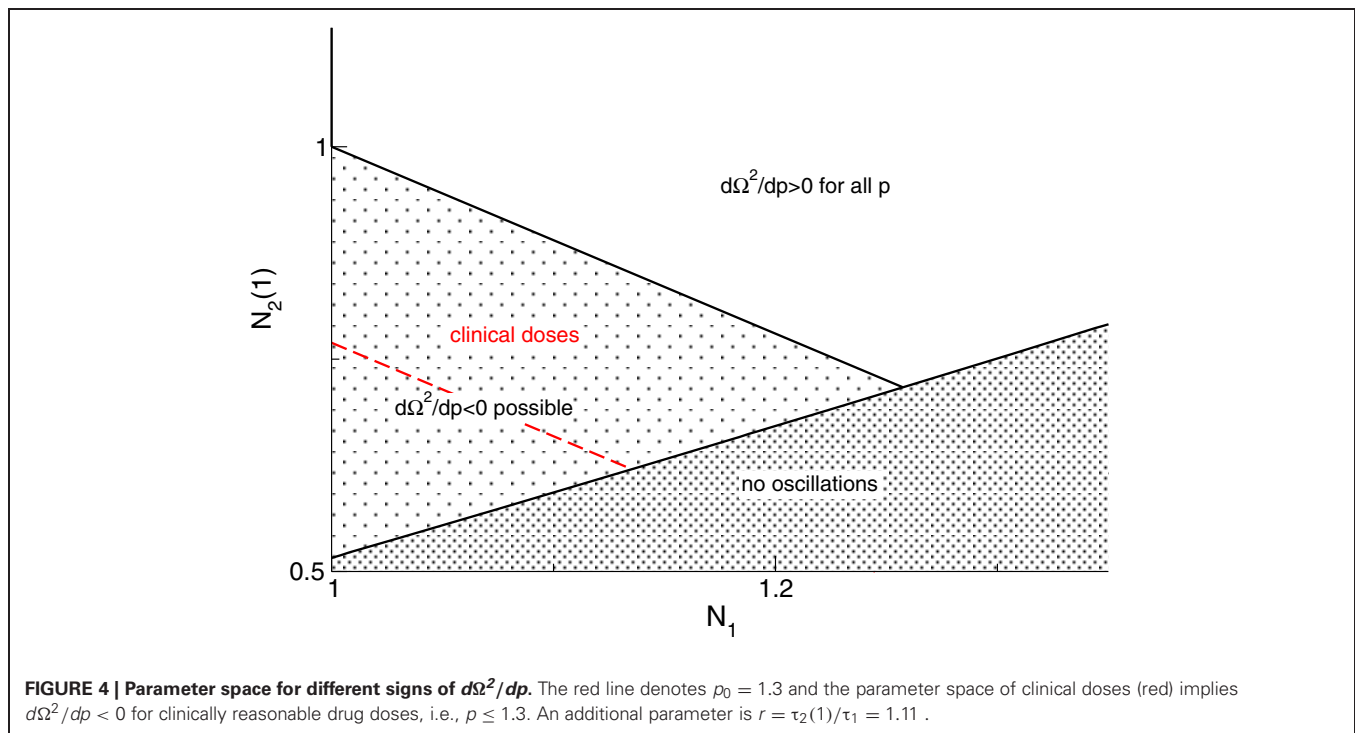
Hopf-instabilities and shows the effect of increased propofol concentration. The power increases and the peak of maximum power moves to larger frequencies as predicted by the theory and as observed in experiments (Hindriks and van Putten, 2012).

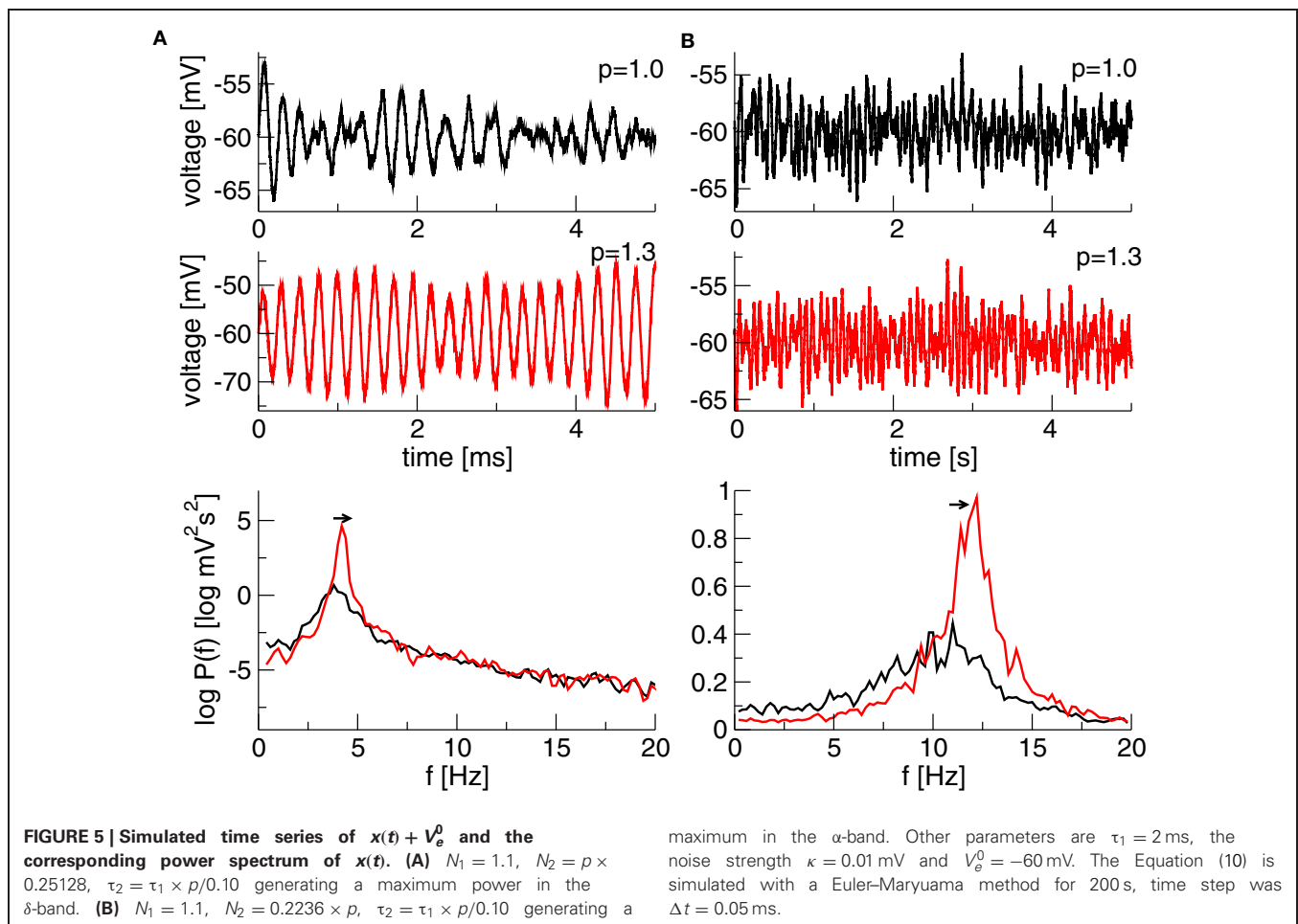
4. DISCUSSION

The introduced model considers first-order synaptic responses and take into account experimental findings on the propofol effect in synaptic GABA_A-receptors in cortical neurons. It describes the evolution of neural populations on a mesoscopic level involving major properties of underlying neurons and synapses on the microscopic description level.

The analytical study reveals that the frequency of maximum power may increase or decrease with increasing anesthetic concentration subjected to the physiological constants, cf. **Figure 3**. In detail, close to the oscillatory instability the frequency of maximum power always increases with increasing p as observed in EEG (Gugino et al., 2001; Feshchenko et al., 2004; Hindriks and van Putten, 2012), whereas far from the stability threshold the maximum power frequency may also decrease as observed recently in EEG (Ching et al., 2010; Cimenser et al., 2011), cf. **Figures 3, 4**. The analytical treatment shows clearly that these two findings depend strongly on the physiological parameters, which are derived analytically in section 3.5, i.e., the phenomena depend on the brain area in which they are generated.

Moreover, the analytically predicted increase of the power at higher frequencies explains the power enhancement in the α - and β -band in anesthesia (Gugino et al., 2001; McCarthy et al., 2008) by a dynamic oscillatory instability. In fact, the analytical treatment in the present work suggests that power enhancement always starts from oscillatory activity at lower frequencies





and are generated at slightly higher frequencies with increased concentration (**Figure 5**). This is in accordance to previous EEG-studies (Gugino et al., 2001; Hindriks and van Putten, 2012) showing power enhancement induced in the α - and β -band.

The analytical discussion in section 3.3 also predicts that decreasing the inhibitory time constant always moves the system toward an oscillatory instability, cf. **Figure 1**, and hence increases the spectral power, whereas increasing the charge transfer yields a stabilization of the system due to $dR/dP < 0$ and consequently a decrease of power. Hence the balance between decay prolongation and increased charge transfer decides on the change of the spectral power and the shift of the frequency peaks. Since many anesthetics share this balance in the major target GABA_A-receptor (Alkire et al., 2008) and exhibit similar EEG-change (Gugino et al., 2001; Kuizenga et al., 2001), the presented work suggests that this balance reflects one of the major underlying mechanisms during the sedation phase in GA.

Previous studies (Bojak and Liley, 2005; Hindriks and van Putten, 2012) already have explained the power enhancement in anesthesia by an oscillatory instability in high-dimensional neural models. As one of the first, the present work gives analytical conditions on physiological parameters for this effect, while Bojak and Liley (2005) and Hindriks and van Putten (2012) mainly performed numerical studies. The recent work of Hindriks and van

Putten (2012) resembles in some aspects the analytical approach of the present work by discussing in some detail the dynamics of superimposed oscillation modes subjected to the propofol concentration. However, no analytical conditions are given due to the higher model complexity.

It is important to point out that the current model is low-dimensional, physiologically reasonable and analytically treatable but still able to explain the neural phenomenon of the frequency shift to larger values. Bojak and Liley (2005) and Hindriks and van Putten (2012) have not performed a detailed analytical study of this phenomenon and have not derived analytical conditions under which it may occur. The current work shows that already a rather simple coupling of excitation and inhibition in cortical neural networks is sufficient to explain this phenomenon. This is concluded partially by Hindriks and van Putten (2012) based on a small numerical study, whereas the present work shows this explicitly. However, Hindriks and van Putten (2012) also argue that the cortico-thalamic feedback should be negative to gain this effect. The presented model does not need the thalamic feedback loop for the explanation.

In principle, the present work extends the work of Bojak and Liley (2005) studying just numerically a rather complicated model with tens of unknown parameters, while the current model allows to achieve some insights into the effect of few parameters.

For instance, the maximum of spectral power and the corresponding frequency are highly sensitive to modification of the relation between excitatory and inhibitory synaptic time scales. This confirms the general observation that many different anesthetics share effects on the activity spectrum.

The work both supports the hypothesis of cortical generation of α -activity and predicts the presence of oscillating neural circuits where each circuit generates a certain oscillation observed experimentally. This can be observed in **Figure 5** showing emerging δ - and α -activity for different parameters. Hence two neural circuits with different properties may explain the occurrence of both δ - and α -activity observed in experimental data. This interpretation of the results complements the findings of Hindriks and van Putten (2012) showing implicitly a linear decomposition into eigenmodes with corresponding eigenvalues and manifests the notion of interacting oscillation modes generated by interacting networks as observed experimentally (Fries, 2009; Spaak et al., 2012).

Of course the present model is limited since it cannot explain the increase of activity in the δ -band as observed in experiments and modeled by the previous studies. In the present work, we

observe clearly that this synchronous modeling of two rhythms is not possible since the model is too low-dimensional and just can describe a single rhythm such as the α - or the δ -rhythm. Consequently further neural elements should be considered to gain this additional rhythm such as the thalamic loop. This may be possible due to the linear superposition of oscillatory activity originating from different networks. This linear superposition proposes the interaction of different sub-networks each oscillating in a certain frequency band. Future work will consider such entangled neural networks on the basis of the presented neural field to explain the spontaneous emergence or diminution of spectral peaks in experimentally observed data such as the δ -rhythms or transient phenomena such as paradoxical excitation.

ACKNOWLEDGMENTS

The author thanks Meysam Hashemi for fruitful discussions and acknowledges funding from the European Research Council for support under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement No. 257253. The author thanks the referees for valuable comments.

REFERENCES

- Alkire, M., Hudetz, A., and Tononi, G. (2008). Consciousness and anesthesia. *Science* 322, 876–880.
- Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87.
- Angelucci, A., and Bressloff, P. (2006). The contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Prog. Brain Res.* 154, 93–121.
- Bojak, I., and Liley, D. (2005). Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* 71:041902. doi: 10.1103/PhysRevE.71.041902
- Boly, M., Moran, R., Murphy, M., Boveroux, P., Bruno, M. A., Noirhomme, Q., et al. (2012). Connectivity changes underlying spectral eeg changes during propofol-induced loss of consciousness. *J. Neurosci.* 32, 7082–7090.
- Ching, S., Cimenser, A., Purdon, P. L., Brown, E. N., and Kopell, N. J. (2010). Thalamocortical model for a propofol-induced-rhythm associated with loss of consciousness. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22665–22670.
- Cimenser, A., Purdon, P. L., Pierce, E. T., Walsh, J. L., Salazar-Gomez, A. F., Harrell, P. G., et al. (2011). Tracking brain states under general anesthesia by using global coherence analysis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8832–8837.
- Coombes, S., Venkov, N., Shiau, L., Bojak, I., Liley, D., and Laing, C. (2007). Modeling electrocortical activity through improved local approximations of integral neural field equations. *Phys. Rev. E* 76, 051901–051908.
- Ermentrout, G. (1998). Neural networks as spatio-temporal pattern-forming systems. *Rep. Progr. Phys.* 61, 353–430.
- Ermentrout, G., and Cowan, J. (1979). A mathematical theory of visual hallucination patterns. *Biol. Cybern.* 34, 137–150.
- Feshchenko, V. A., Veselis, R. A., and Reinsel, R. A. (2004). Propofol-induced alpha rhythm. *Neuropsychobiology* 50, 257–266.
- Franks, N., and Lieb, W. (1994). Molecular and cellular mechanisms of general anesthesia. *Nature* 367, 607–614.
- Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu. Rev. Neurosci.* 32, 209–224.
- Gugino, L. D., Chabot, R. J., Prichep, L. S., John, E. R., Formanek, V., and Aglio, L. S. (2001). Quantitative eeg changes associated with loss and return of consciousness in healthy adult volunteers anaesthetized with propofol or sevoflurane. *Br. J. Anaesth.* 87, 421–428.
- Hindriks, R., and van Putten, M. J. A. M. (2012). Meanfield modeling of propofol-induced changes in spontaneous eeg rhythms. *Neuroimage* 60, 2323–2344.
- Huang, X., Troy, W., Schiff, S., Yang, Q., Ma, H., Laing, C., et al. (2004). Spiral waves in disinhibited mammalian neocortex. *J. Neurosci.* 24, 9897–9902.
- Hutt, A. (2008). Additive noise may change the stability of nonlinear systems. *Europhys. Lett.* 84:34003. doi: 10.1209/0295-5075/84/34003
- Hutt, A. (ed.). (2011). *Sleep and Anesthesia: Neural Correlates in Theory and Experiment*. Springer Series in Computational Neuroscience, Vol. 15. New York, NY: Springer.
- Hutt, A. (2012). The population firing rate in the presence of gabaergic tonic inhibition in single neurons and application to general anaesthesia. *Cogn. Neurodyn.* 6, 227–237.
- Hutt, A., and Longtin, A. (2009). Effects of the anesthetic agent propofol on neural populations. *Cogn. Neurodyn.* 4, 37–59.
- Hutt, A., Longtin, A., and Schimansky-Geier, L. (2007). Additive global noise delays turing bifurcations. *Phys. Rev. Lett.* 98:230601. doi: 10.1103/PhysRevLett.98.230601
- Kitamura, A., Marszalec, W., Yeh, J., and Narahashi, T. (2002). Effects of halothane and propofol on excitatory and inhibitory synaptic transmission in rat cortical neurons. *J. Pharmacol.* 304, 162–171.
- Koch, C. (1999). *Biophysics of Computation*. Oxford: Oxford University Press.
- Kuizenga, K., Wierda, J., and Kalkman, C. (2001). Biphasic eeg changes in relation to loss of consciousness during induction with thiopental, propofol, etomidate, midazolam or sevoflurane. *Br. J. Anaesth.* 86, 354–360.
- Liley, D., and Bojak, I. (2005). Understanding the transition to seizure by modeling the epileptiform activity of general anaesthetic agents. *J. Clin. Neurophysiol.* 22, 300–313.
- McCarthy, M. M., Brown, E. N., and Kopell, N. (2008). Potential network mechanisms mediating electroencephalographic beta rhythm changes during propofol-induced paradoxical excitation. *J. Neurosci.* 28, 13488–13504.
- Murphy, M., Bruno, M.-A., Riedner, B. A., Boveroux, P., Noirhomme, Q., Landsness, E. C., et al. (2011). Propofol anesthesia and sleep: a high-density eeg study. *Sleep* 34, 283–291.
- Nunez, P., and Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford, NY: Oxford University Press.
- Orser, B. (2006). Extrasynaptic gabaa receptors are critical targets for sedative-hypnotic drugs. *J. Clin. Sleep Med.* 2, S12–S18.
- Schwabe, L., Obermayer, K., Angelucci, A., and Bressloff, P. (2006). The role of feedback in shaping the extra-classical receptive field of cortical neurons: a recurrent network model. *J. Neurosci.* 26, 9117–9126.
- Spaak, E., Bonnefond, M., Maier, A., Leopold, D. A., and Jensen, O. (2012). Layer-specific entrainment of gamma-band neural activity by the alpha rhythm in monkey visual cortex. *Curr. Biol.* 22, 2313–2318.

- Steyn-Ross, M., Steyn-Ross, D., and Sleigh, J. (2004). Modelling general anaesthesia as a first-order phase transition in the cortex. *Prog. Biophys. Molec. Biol.* 85, 369–385.
- Steyn-Ross, M., Steyn-Ross, D., Sleigh, J., and Wilcocks, L. (2001). Toward a theory of the general-anesthetic-induced phase transition of the cerebral cortex: I. a thermodynamic analogy. *Phys. Rev. E* 64:011917. doi: 10.1103/PhysRevE.64.011917
- Steyn-Ross, M., Steyn-Ross, D., and Sleigh, J. W. (2012). Gap junctions modulate seizures in a mean-field model of general anesthesia for the cortex. *Cogn. Neurodyn.* 6, 215–225.
- Steyn-Ross, M., Steyn-Ross, D., Sleigh, J. W., and Liley, D. T. J. (1999). Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: evidence for a general anesthetic-induced phase transition. *Phys. Rev. E* 60, 7299–7311.
- Wilson, H., and Cowan, J. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24.
- Ying, S., and Goldstein, P. (2005). Propofol suppresses synaptic responsiveness of somatosensory relay neurons to excitatory input by potentiating gabaa receptor chloride channels. *Mol. Pain* 1:2. doi: 10.1186/1744-8069-1-2
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 12 December 2012; accepted: 19 January 2013; published online: 05 February 2013.*
- Citation:* Hutt A (2013) The anesthetic propofol shifts the frequency of maximum spectral power in EEG during general anesthesia: analytical insights from a linear model. *Front. Comput. Neurosci.* 7:2. doi: 10.3389/fncom.2013.00002
- Copyright © 2013 Hutt. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



On the dynamics of cortical development: synchrony and synaptic self-organization

James Joseph Wright^{1,2*} and Paul David Bourke³

¹ Department of Psychological Medicine, Faculty of Medicine, The University of Auckland, Auckland, New Zealand

² Liggins Institute, The University of Auckland, Auckland, New Zealand

³ iVEC@UWA, University of Western Australia, Perth, WA, Australia

Edited by:

Peter Beim Graben,
Humboldt-Universität zu Berlin,
Germany

Reviewed by:

Maurizio Mattia, Istituto Superiore di
Sanità, Italy
Peter Beim Graben,
Humboldt-Universität zu Berlin,
Germany

*Correspondence:

James Joseph Wright, Department
of Psychological Medicine, Faculty
of Medical and Health Sciences,
School of Medicine, The University
of Auckland, PB 92019, Auckland,
New Zealand.
e-mail: james.wright@auckland.ac.nz

We describe a model for cortical development that resolves long-standing difficulties of earlier models. It is proposed that, during embryonic development, synchronous firing of neurons and their competition for limited metabolic resources leads to selection of an array of neurons with ultra-small-world characteristics. Consequently, in the visual cortex, macrocolumns linked by superficial patchy connections emerge in anatomically realistic patterns, with an ante-natal arrangement which projects signals from the surrounding cortex onto each macrocolumn in a form analogous to the projection of a Euclidean plane onto a Möbius strip. This configuration reproduces typical cortical response maps, and simulations of signal flow explain cortical responses to moving lines as functions of stimulus velocity, length, and orientation. With the introduction of direct visual inputs, under the operation of Hebbian learning, development of mature selective response “tuning” to stimuli of given orientation, spatial frequency, and temporal frequency would then take place, overwriting the earlier ante-natal configuration. The model is provisionally extended to hierarchical interactions of the visual cortex with higher centers, and a general principle for cortical processing of spatio-temporal images is sketched.

Keywords: synchronous oscillation, cortical development, synaptic organization, cortical response properties, cortical information flow

INTRODUCTION

During its embryological development the mammalian brain differentiates from a group of stem cells into an organized form ready to begin a life-long adaptive interaction with signals from the sensory environment. At the beginning of extra-uterine life, despite exposure to a limited milieu, it is somehow already organized to begin this engagement, as though a matrix of connections has formed in which signal flows are pre-adapted to learn specific recurring patterns of the experiential world. A large body of work, following the pioneering work of Hubel and Wiesel (1959), has addressed just this issue, taking as the main target for research the primary visual cortex (V1). The majority of this work has sought to understand the emerging connections in terms of stimulus “features”—that is, elementary properties of sensory stimuli—rather than as a process independent of sensation until the post-natal stage. Our approach depends on alternative assumptions. Here we summarize and extend our earlier work (Wright et al., 2006; Wright and Bourke, 2008, 2013; Wright, 2009, 2010) relating the basic dynamics of neuron firing and competition among developing neurons for the resources needed for their growth, to the emergent connections at birth.

Our model draws on two recent experimental observations. Firstly, neurons in neonatal cerebral cortical slices show increased apoptosis when their capacity to enter into synchronous firing is disrupted by pharmacological means (Heck et al., 2008). Secondly, embryonic neurons developing *in vitro* develop synchronous firing, and as their growth proceeds, show self-organization into “small world” networks (Downes et al., 2012).

We propose that the synchronous firing and protection from apoptosis are directly causally related, because during cortical embryogenesis there is competition among developing neurons and synapses, which, although mediated by trophic factors (Harris et al., 1997; van Ooyen and Willshaw, 1999; van Ooyen, 2001) is ultimately a competition for available metabolic energy and/or some other scarce resource needed to promote metabolism (Montague, 1996; Thomaïdou et al., 1997). We suppose that pre-synaptic pulse synchrony increases uptake of critical metabolic resources by some action not presently specified, and we argue that the assembly of cells that maximizes synchronous firing, and thus energy uptake, is also that which has the minimum metabolic cost per neuron in the length of axonal connections—the combination optimum for their survival.

Synchronous oscillation of pulses and local field potentials is a ubiquitous aspect of cortical activity (Eckhorn et al., 1988, 1990; Gray et al., 1989; Bressler et al., 1993; Singer, 1999) and has been proposed as a mechanism solving the “binding problem” of perceptual grouping and cognitive processing (Eckhorn et al., 1990; Singer, 1999; Crick and Koch, 2003). “Synchrony” refers to the broadband cross-correlation of neuron firing and field potentials at zero time-lag. The mechanism of origin of synchrony itself is controversial. In this paper we rely on an explanation that appears best applicable to the synchrony seen in neuron cultures, brain slices, or the early embryonic brain, and depends on a universal property of networks with summing junctions, including dendrites (Robinson et al., 1998; Wright et al., 2000; Chapman et al.,

2002). This type of synchrony appears in simulations that also accurately reproduce spectra, cross-correlations, and excitatory/inhibitory timings characteristic of activated cortex (Wright, 2009, 2010).

GEOMETRY OF RESPONSE ORGANIZATION IN THE DEVELOPED BRAIN

Since the discovery that individual cells in the primary visual cortex (V1) respond with an orientation preference (OP) to visual lines of differing orientation (Hubel and Wiesel, 1959), analysis of the response organization and its relationship to cortical function has remained both conceptually influential and controversial (von der Malsburg, 1973; Willshaw and von der Malsburg, 1976; Swindale, 1996). The surface organization of OP in V1 has recently been compared with appropriate random surrogates, and shown to exhibit significant hexagonal rotational periodicity, in which each roughly delineated macrocolumnar unit exhibits all values of OP arrayed around a pinwheel (Paik and Ringach, 2011; Muir et al., 2011). Varying chirality and orientation of the pinwheels achieves continuity of OP at the columnar margins, thus producing zones of irregular but continuously varying OP, known as linear zones and saddles.

Some species exhibit little or no sign of this hexagonal and continuous ordering, and because of the marked interspecies variation, serious doubt has been expressed that the pattern is of functional significance at all, since species showing little such organization have no apparent deficit in vision (Horton and Adams, 2005). Interspecies variation seems, in part, to be related to both variation in size of V1 between species, and a relative constancy of the size of macrocolumns, independently of species. Measurements of the average distance of separation of OP singularities (the singularity taken as demarcating the center of a macrocolumn) show this distance to be relatively constant over a 40-fold variation of body size, and related size of V1 (Kaschube et al., 2010; Keil et al., 2012). Models using symmetry arguments indicate that macrocolumns must undergo divisions during cortical development to maintain uniform surface density of singularities (Wolf and Geisel, 1998; Oster and Bressloff, 2006). Kaschube and colleagues conclude that self-organization has canalized the evolution of the underlying OP maps into a single common design—subject to the proviso that, from further symmetry arguments, this can only be the case where long-range interactions between developing macrocolumns, suppressing some possible connections, can take place. Thus, in animals with very small V1, this organization breaks down, creating a “pepper and salt” OP map pattern (Meng et al., 2012).

THE SUPERFICIAL PATCH SYSTEM

A related puzzle of V1 organization is posed by the superficial patch system. This system, composed of relatively long-range, largely excitatory (Hirsch and Gilbert, 1991; McGuire et al., 1991) patchy connections (Gilbert and Wiesel, 1979; Rockland and Lund, 1983) is ubiquitous in cortex (Muir and Douglas, 2011) and has a functional relationship to OP. Patchy connections develop before sensory afferents reach the cortex (Price, 1986; Callaway and Katz, 1990; Durack and Katz, 1996; Ruthazer

and Stryker, 1996) but do not arise or terminate in the vicinity of OP singularities. Instead, near singularities, connections are apparently diffuse and local (Sharma et al., 1995; Yousef et al., 2001; Mariño et al., 2005; Buzás et al., 2006; Muir and Douglas, 2011). Patchy connections link areas of common OP (“like-to-like”) over distances several times the diameter of a macrocolumn (Gilbert and Wiesel, 1989; Buzás et al., 2006; Muir et al., 2011), are periodic on roughly the same interval as OP, and are largely patch-reciprocal (Rockland and Lund, 1983; Angelucci et al., 2002). It has been shown that development of patchy connections must depend on the supply of organizing information from the neural field, and is not explicable from considerations of local neural growth *per se* (Muir and Douglas, 2011). Just as for maps of response properties, there is variation of patchy connection orderliness between species. Muir et al. (2011) have pointed out that those species with less orderliness have smaller visual cortices and/or less defined organization of “like-to-like” connections—an argument congruent with the findings on brain size, orderliness of response maps, and surface density of OP singularities cited above (viz. Kaschube et al., 2010; Keil et al., 2012, etc.).

PROBLEMS OF STANDARD MODELS OF FEATURE RESPONSES

Explanation of organization of OP has been undertaken in a group of now-classical theories, which we will refer to as “standard models,” following the comparative description of Swindale (Swindale, 1996). Descriptive dimension reduction methods (Kohonen, 1982; Durbin and Willshaw, 1987; Durbin and Mitchison, 1990) show that the response maps of OP, eye preference (OC), direction preference (DP), and spatial frequency preference (SF) are consequences of requiring continuity and completeness of representation of each response property, in a two-dimensional representation in which every type of response property occurs within any small area on the surface of V1 (Swindale, 1996; Carriera-Perpiñán et al., 2005). The same ordering is also explained as a consequence of competitive Hebbian learning among small neighborhood assemblies of excitatory neurons, driven by spatially filtered cortical noise. Separate spatial filters each distinguish a type of response, and total synaptic gain is conserved during the training (Grossberg and Olson, 1994).

Classical standard models depend on seeding with oriented lines, in one way or another (von der Malsburg, 1973; Swindale, 1982, 1992; Durbin and Mitchison, 1990; Obermayer et al., 1990, 1992; Tanaka, 1990; Miyashita and Tanaka, 1992; Grossberg and Olson, 1994) and recently, initial belief that primary response to static oriented lines in the visual field forms the basis of OP maps has been undermined in two ways:

Firstly, in large species particularly, maps of OP appear in the cortex prior to visual experience (Wiesel and Hubel, 1974; Blakemore and Van Sluyters, 1975; Sherk and Stryker, 1976). This problem has been addressed by arguments for the normal occurrence of line-like structure in ante-natal retinal input (Albert et al., 2008; Ringach, 2007; Paik and Ringach, 2011). In contrast to all the above models, Kang et al. (2003) have proposed a model which breaks with the traditional dependence on the primacy of lines, and depends instead on time-invariant correlations in cortical “Mexican Hat” inhibitory surrounds. This model accounts

successfully for the apparent isotropy of local intracortical connections and the observed uniformity of sharpness of definition of OP independent of proximity to singularities, and provides a mechanism which might plausibly operate before eye-opening. It requires instead, that LGM inputs to cortex become tuned according to orientation. A further model avoiding the problem of ascription of OP as a primary, stimulus dependent property, explains the conjoint development of OP and ocular dominance columns as a consequence of Hebbian connection formation driven by correlation of visual inputs as a declining function of retinotopic distance of separation at short distances, and reversed correlation of activity in ON and OFF V1 simple cells at greater distances (Erwin and Miller, 1998). All these models however, result in the emergence of OP as a property of line orientation alone, rather than as one attribute of some more complex mechanism of feature response.

Secondly, and more recently, Basole and colleagues, who tested OP using stimulus lines moving at different speeds, and oriented at differing angles to the line of movement of the stimulus, found OP to be a function of these variables to such a degree that for lines oriented non-orthogonally to the direction of movement, OP could vary progressively with increments of speed to an asymptotic limit of 90° (Basole et al., 2003, 2006). Longer lines showed less variation of OP with increasing speed. This finding challenged all models which depended on OP being a fixed “feature” of cortical response, whether or not direct visual stimuli was required to prime the process of self-organization. Basole and colleagues at first concluded that the primal stimulus characteristics are not isolated features such as orientation, direction and speed, but a single characteristic—the “spatio-temporal energy”—that is, the combined spatial and temporal Fourier components of the moving visual stimulus’ projection to V1. Subsequent workers explained these results by retaining OP as a primary characteristic, and adding separate consideration of the temporal and spatial frequencies associated with the moving stimuli (Baker and Issa, 2005; Mante and Carandini, 2005; Basole et al., 2006). This analysis was consistent with earlier single unit results, in which tuning of V1 neurons to spatial and temporal frequencies was demonstrated (DeAngelis et al., 1993). Issa and colleagues (Baker and Issa, 2005; Issa et al., 2008) reported that a total of six parameters are required to explain response maps—OP, SF preference, and temporal frequency preference, and the tuning bandwidths of all three. This account is referred to as the spatio-temporal filter model. How these response characteristics arise during cortical development and how neurons become tuned to just those features is the subject of continuing research (Rosenberg et al., 2010), and of this paper.

In common with the model of Erwin and Miller (1998) and that of Kang et al. (2003) the model reviewed here depends upon time-average correlations—that is, the common occurrence of synchronous oscillation in the cortex—although it does not share their other assumptions or conclusions. It seeks to avoid the ascription of “features” as primary characteristics, and to explain both the findings of Basole et al. (2003) and the empirical reduction to alternative feature attributes used in the spatio-temporal model, as well as explaining the emergence of the anatomical features described above.

DESCRIPTION OF MODEL

NEURAL FIELD EQUATIONS

As alternatives to neural network models, lumped neural models and neural field equations have been expressed in many forms (e.g., Wilson and Cowan, 1973; Freeman, 1975; Haken, 1996; Amari, 1977; Nunez, 1981; van Rotterdam et al., 1982; Jirsa and Haken, 1996; Robinson et al., 2001; Wright et al., 2003; beim Graben, 2008; Bressloff, 2012). These offer means of approximating the properties of ensembles of cells on a larger scale than neural networks *per se*. Here we have used a generic form of neural field equations to represent an idealized, isotropic, neural field, representing the developing cortex as if it were not subject to apoptosis—a potentiality from which connections are selected during development. The scale of the field is that of a cortical area such as V1, representing intracortical connections rather than cortico-cortical. Thus, the density of connection between neurons declines with increasing separation of their cell bodies (Braitenberg and Schüz, 1991). The high non-linearity of synapo-dendritic summations are linearized at the field level, and axonal conduction speed is considered single-valued. Subject to these strictures, the following equations include features relevant to the present context:

$$\varphi_p^{\mathbf{qr}'}(t) = f_p^{\mathbf{qr}'} \times Q_p \left(\mathbf{r}', t - \frac{|\mathbf{q} - \mathbf{r}'|}{v} \right) \quad (1)$$

$$\psi_p^{\mathbf{qr}'}(t) = M_p^{\mathbf{qr}'} * \varphi_p^{\mathbf{qr}'}(t) \quad (2)$$

$$\Psi_p(\mathbf{q}, t) = \int_{\mathbf{r}'} \psi_p^{\mathbf{qr}'}(t) d\mathbf{r}' \quad (3)$$

$$V_p(\mathbf{q}, t) = \sum_{p=e \wedge p=i} G_p * \Psi_p(\mathbf{q}, t) \quad (4)$$

$$Q_p(\mathbf{q}, t) = f_\Sigma(V_p(\mathbf{q}, t)) + E_p(\mathbf{q}, t) \quad (5)$$

Subscript $p = e, i$ refers to excitatory or inhibitory neurons; superscript \mathbf{qr}' refers to synaptic connection from \mathbf{r}' to \mathbf{q} where \mathbf{q}, \mathbf{r}' are cortical positions occupied by single neurons.

$\varphi_p^{\mathbf{qr}'}(t)$ is the flux of pulses reaching pre-synapses at the neuron at \mathbf{q} , from the neuron at \mathbf{r}' .

$\psi_p^{\mathbf{qr}'}(t)$ is the synaptic current generated by $\varphi_p^{\mathbf{qr}'}(t)$.

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

$V_p(\mathbf{q}, t)$ is the soma membrane potential (relative to the resting potential) generated at \mathbf{q} .

$Q_p(\mathbf{q}, t)$ is the pulse emission rate at \mathbf{q} .

$f_p^{\mathbf{qr}'}$ is the probability density of occurrence of pre-synapses generated by axons of the neuron at \mathbf{r}' terminating at \mathbf{q} .

v is axonal conduction speed.

$M_p^{\mathbf{qr}'}$ is the steady-state term in a convolution transforming pre-synaptic flux to synaptic current.

G_p is the steady-state term in a convolution transforming pre-synaptic flux into dendritic potentials.

$f_\Sigma(V_q(\mathbf{q}, t))$ is a sigmoid function describing the local conversion of dendritic potentials into the rate of generation of action potentials.

$E_p(\mathbf{q}, t)$ is a driving signal noise, arising from intrinsic random cell action potentials.

Restriction of the field to the scale of a cortical area carries several implications important for the model, all because the probability of connections between any two neurons declines with distance of separation. Firstly, descriptively we can consider “reciprocal couplings” as an idealization/representation of field coupling symmetry, and in many instances reciprocal couplings will in fact exist. Secondly because of more generally dense connections among near neighbors, smoothing at dendritic summation requires that $Q_p(\mathbf{q}, t)$ is spatially and temporally “brown”—i.e., has high correlation at short distances and times of separation. Thirdly, the average “degree” of separation—i.e., the average number of neighboring cells traversed by synaptic connections linking one cell to another—will also increase in proportion to physical distance of separation.

Experimental observations (Freeman, 1975, 1991; Hassenstaub et al., 2005) show intrinsic cortical oscillation arises from alternating excitatory cell and inhibitory cell firing at lags $\frac{1}{4}$ of the period of oscillation. Simulations of the oscillations (Wright, 2009, 2010) show that traveling waves are thus generated, the intersection of which produces broadband synchrony. In conditions of uniform cortical excitation without strong perturbation from external inputs the exchange of pulses between all cells reaches an equilibrium—that is, a steady-state of symmetrical exchange of signals between excitatory cells at any two positions on the cortex, so that over sufficient intervals, T ,

$$\frac{1}{T} \int_T (\varphi_{p1}(\mathbf{q}) - \bar{\varphi}_{p1}) dt = \frac{1}{T} \int_T (\varphi_{p2}(\mathbf{r}') - \bar{\varphi}_{p2}) dt \quad (6)$$

where $\bar{\varphi}_p$ is the time-average presynaptic flux, uniform throughout the cortical field. The equilibrium reached implies differences in timing between the firing of excitatory and inhibitory cells. The interaction of excitatory and inhibitory cells ($p1 \vee p2 = e$, and $p1 \vee p2 = i$) leads to closely correlated firing of both cells if they are very closely situated, as a consequence the similar local values of $E(\mathbf{q}, t)$ equation (5), while $\frac{1}{4}$ -cycle-out-of-phase oscillation develops between more separated excitatory and inhibitory cells. Inhibitory/inhibitory or excitatory/excitatory interactions ($p1 \wedge p2 = e$, or $p1 \wedge p2 = i$) between reciprocally connected neurons lead to zero-lag synchrony, and since conduction delays are short compared to the period of oscillation, the equality of equation (6) is generally approached even when T is smaller than the period of oscillation (Chapman et al., 2002). As there are equal time-lags in both directions of conduction excitatory pulse trains throughout the cortex have maximum correlation at zero lag—i.e., where \bar{Q}_e is the time-average firing rate—also uniform throughout the cortical field -

$$(Q_e - \bar{Q}_e)(\mathbf{r}', t) \approx (Q_e - \bar{Q}_e)(\mathbf{q}, t) \quad (7)$$

Figures 1 and 2 show these properties generated in a simulation of cortical dynamics with physiologically realistic parameters (Wright, 2009, 2010). In conditions of strong cortical excitation local oscillation is autonomous and corresponds to cortical gamma rhythm, while in conditions of lower cortical excitation,

damped gamma oscillation, and a predominance of background $1/f^2$ is seen.

MAGNITUDE OF PRE-SYNAPTIC PULSE SYNCHRONY

Zero-lag synchronous oscillation thus entails presynaptic pulse synchrony, with a magnitude of presynaptic flux variation which can be defined respectively for individual synapses, individual cells, and in aggregate, as

$$J^{\mathbf{q}\mathbf{r}'} = \left[\frac{1}{T} \int_T (\varphi_e^{\mathbf{q}\mathbf{r}'} - \bar{\varphi}_e)^2 dt \right]^{1/2} \quad (8)$$

$$J^{\mathbf{q}} = \left[\frac{1}{T} \int_T \int_{\mathbf{r}'} (\varphi_e^{\mathbf{q}\mathbf{r}'} - \bar{\varphi}_e)^2 dt d\mathbf{r}' \right]^{1/2} \quad (9)$$

$$J = \left[\frac{1}{T} \int_T \int_{\mathbf{r}'} \int_{\mathbf{q}} (\varphi_e^{\mathbf{q}\mathbf{r}'} - \bar{\varphi}_e)^2 dt d\mathbf{r}' d\mathbf{q} \right]^{1/2} \quad (10)$$

$J^{\mathbf{q}\mathbf{r}'}$ is RMS presynaptic flux variation between \mathbf{q} and \mathbf{r}' , $J^{\mathbf{q}}$ is the sum of $J^{\mathbf{q}\mathbf{r}'}$ at a single excitatory neuron, and J is the aggregate of $J^{\mathbf{q}}$ over the cortex.

SELECTION OF SCALE-FREE SMALL-WORLD CONFIGURATIONS OF NEURONS

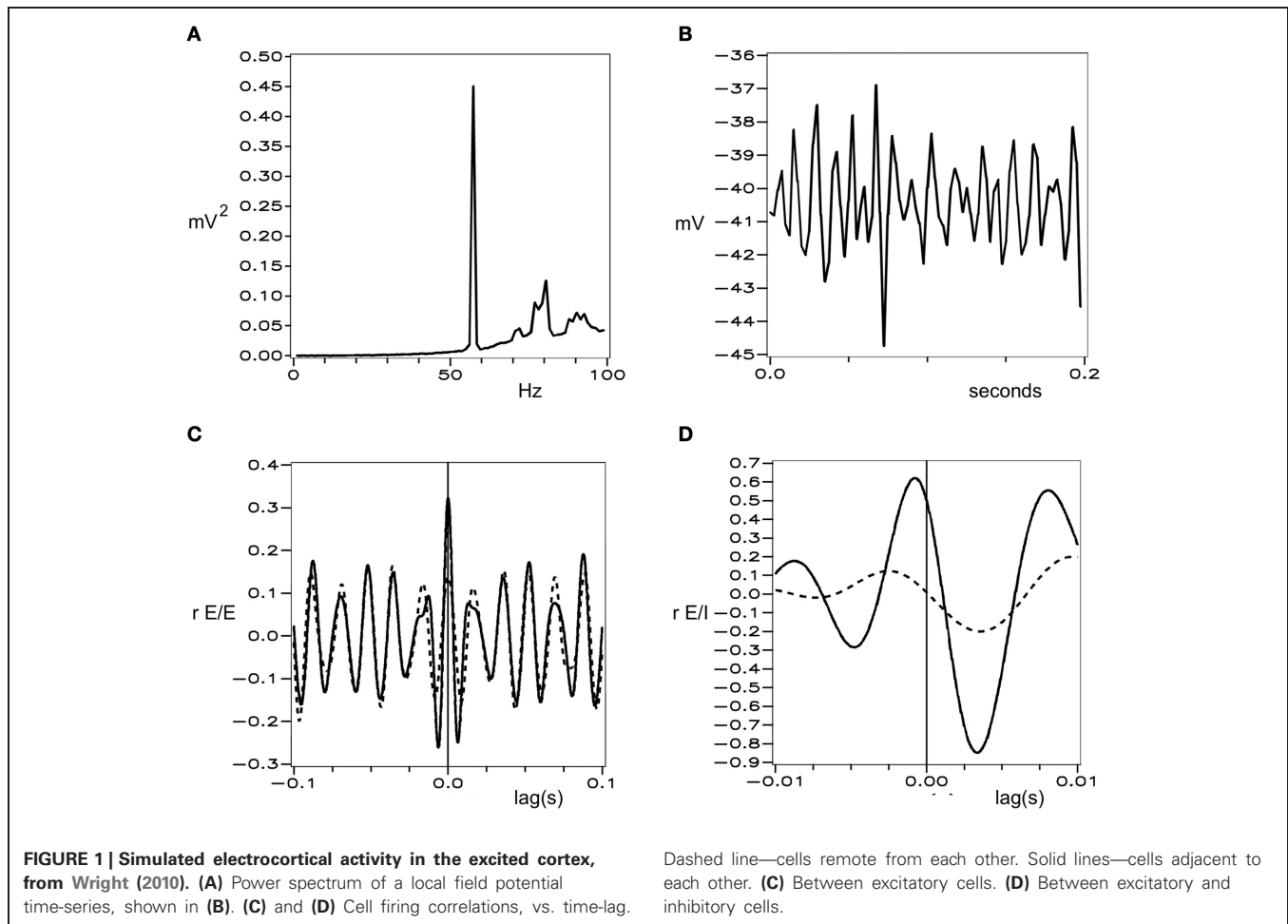
For any given level of cortical excitation, J is greatest for that ensemble of C connected neurons, in which excitatory pulses arrive at dendrites, from all sources at differing distances of separation, as closely in-phase as possible, so as to maximize their summation. Axonal delays, small compared to the period of gamma oscillation, contribute a phase difference between cell firing at \mathbf{r}' and the arrival of presynaptic pulses at \mathbf{q} , of

$$\Delta \Phi^{\mathbf{q}\mathbf{r}'} = 2\pi \frac{|\mathbf{q} - \mathbf{r}'|}{Pv} \quad (11)$$

where P is the period of oscillation. Therefore that ensemble selected by its capacity to maximize presynaptic synchrony must approach minimal total axonal length, $L = \int_{\mathbf{r}'} \int_{\mathbf{q}} |\mathbf{q} - \mathbf{r}'| d\mathbf{q} d\mathbf{r}'$, and minimization of this length minimizes the metabolic requirements of the axons.

It has been shown generally (Cohen and Havlin, 2003) for all systems of connected elements, the path length in a topological sense is at a minimum where degree distribution follows a power law. As was pointed out in conjunction with equations (1–5), in our idealized neural field, average degree of separation, in the topological sense, increases linearly as metric distance of separation of the cell bodies, so that if L , their total length of axonal connections, is minimal, then the path length in the topological sense is also minimal, and the degree distribution is that of a scale-free, or ultra-small world. Therefore, the connection density between cells vs. their metric distance of separation should also be approximated by a power-law distribution. Further, according to Cohen and Havlin

$$L \sim \log \log C \quad (12)$$



so the metabolic efficiency of the connection system is further enhanced if the surviving cells are linked into a continuum, as opposed to separate pools of neurons.

In accord with equation (1), the number of neighboring excitatory cells connected to a given excitatory neuron, as a function of distance of separation, is proportional to $2\pi \times f_e^{qr'}(|\mathbf{q} - \mathbf{r}'|)$ —so the ensemble of neurons selected by greatest synchrony must have a connection density function of the form:

$$f_e^{qr'} \sim (2\pi |\mathbf{q} - \mathbf{r}'|)^{-A} \quad A > 0 \quad (13)$$

Intracortical axonal trees have approximately exponential density/range relations (Scholl, 1956; Braitenberg and Schüz, 1991) and a power function is fitted exactly by an infinite sum of exponential functions—i.e.:

$$(2\pi |\mathbf{q} - \mathbf{r}'|)^{-A} = \frac{1}{\Gamma(A)} \int_0^\infty u^{A-1} \exp[-u2\pi |\mathbf{q} - \mathbf{r}'|] du \quad (14)$$

so an ultra-small-world connectivity can be achieved by sets of populations of cells with differing axonal characteristic lengths. During embryogenesis primal cells divide sequentially by layer (Rakic, 1988; Shi et al., 2012) with differences in growth pattern

and characteristic axonal length programmed in sequential cell divisions. For simplicity, we consider only two populations of excitatory cells, with cell bodies partially separated by layer, but with intermingled axonal and dendritic trees, and axonal tree connection probabilities described by:

$$f_\alpha^{qr'} = \frac{N_\alpha}{N} 2\pi \lambda_\alpha \exp[-\lambda_\alpha 2\pi |\mathbf{q} - \mathbf{R}|] \quad (15)$$

$$f_\beta^{qr'} = \frac{N_\beta}{N} 2\pi \lambda_\beta \exp[-\lambda_\beta 2\pi |\mathbf{q} - \mathbf{r}|] \quad (16)$$

$$f_e^{qr'} = f_\alpha^{qr'} + f_\beta^{qr'}$$

$f_\alpha^{qr'}$ refers to the axonal trees with longest axonal extensions, and $f_\beta^{qr'}$ refers to the axonal trees with short axonal extension, thus $\lambda_\alpha < \lambda_\beta$. $N = N_\alpha + N_\beta$ is the number of synapses received/generated by each cell. Distances from \mathbf{r}' to \mathbf{q} are substituted as \mathbf{r}, \mathbf{R} to indicate equal distances, $|\mathbf{q} - \mathbf{r}|$ and $|\mathbf{q} - \mathbf{R}|$, measured along the axonal trees of the respective populations.

The further defining characteristic of small-world connectivity—the occurrence of connection nodes—emerges as a consequence of the formation of the superficial patch system, as follows.

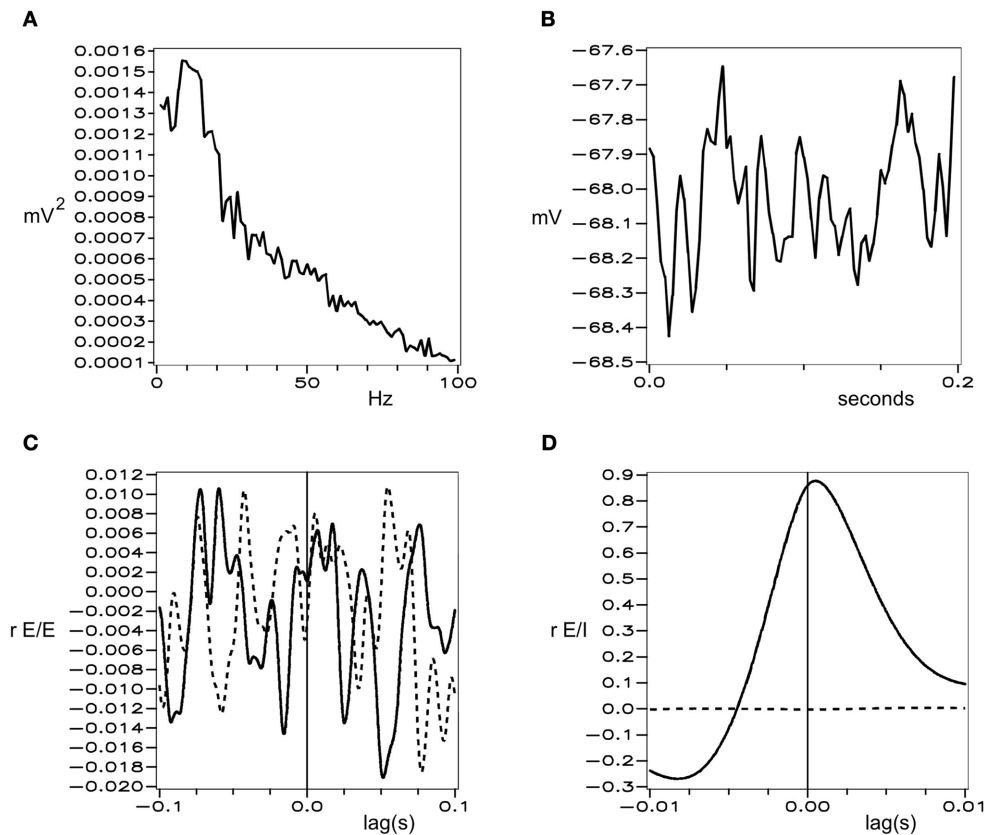


FIGURE 2 | Simulated background electrocortical activity, in conditions of low cortical excitation. Graphical format is the same as in **Figure 1**.

THE ORIGIN OF THE SUPERFICIAL PATCH SYSTEM

The two populations of cells described by equations (15) and (16), and the synapses they give rise to can be referred to as α -cells and synapses, and β -cells and synapses. We first make a provisional assumption (later justified on a species-specific basis) that $N_\beta \gg N_\alpha$, so that α -cells with long-range axons are embedded among much more numerous β -cells, all with sparse connectivity. Equation (10) can be written by separately summing contributions from α -cells at positions $\{\mathbf{q}\alpha\}$ and β -cells at positions $\{\mathbf{q}\beta\}$, to give:

$$J = \int_{\mathbf{q}\alpha} \int_{\mathbf{R}} J^{\mathbf{q}\mathbf{R}} d\mathbf{q}\alpha d\mathbf{R} + \int_{\mathbf{q}\beta} \int_{\mathbf{r}} J^{\mathbf{q}\mathbf{r}} d\mathbf{q}\beta d\mathbf{r} \quad (17)$$

so J is at a maximum if $\int_{\mathbf{q}\alpha} \int_{\mathbf{R}} J^{\mathbf{q}\mathbf{R}} d\mathbf{q}\alpha d\mathbf{R}$ and $\int_{\mathbf{q}\beta} \int_{\mathbf{r}} J^{\mathbf{q}\mathbf{r}} d\mathbf{q}\beta d\mathbf{r}$ are individually at maxima. Applying equations (15) and (16) via equation (1) to find values of $J^{\mathbf{q}\mathbf{r}}$ in equation (8) as functions of $|\mathbf{q} - \mathbf{r}, \mathbf{R}|$, shows that:

$$\begin{aligned} J^{\mathbf{q}\mathbf{r}} &= J^{\mathbf{q}\mathbf{R}} & \text{if } |\mathbf{q} - \mathbf{r}, \mathbf{R}| &= x \\ J^{\mathbf{q}\mathbf{r}} &> J^{\mathbf{q}\mathbf{R}} & \text{if } |\mathbf{q} - \mathbf{r}, \mathbf{R}| < x \\ J^{\mathbf{q}\mathbf{r}} &< J^{\mathbf{q}\mathbf{R}} & \text{if } |\mathbf{q} - \mathbf{r}, \mathbf{R}| > x \end{aligned} \quad (18)$$

where $x = -\ln\left(\frac{N_\alpha \lambda_\alpha}{N_\beta \lambda_\beta}\right) / 2\pi(\lambda_\beta - \lambda_\alpha)$

Consequently $\int_{\mathbf{q}\beta} \int_{\mathbf{r}} J^{\mathbf{q}\mathbf{r}} d\mathbf{q}\beta d\mathbf{r}$ is at a maximum if β -cells are clustered so they make reciprocal connections at minimum distance and maximum density (β -clusters). β -cells at the center of β -clusters, for which $J^{\mathbf{q}\mathbf{r}}$ attains the maximum possible value, must give and receive all their connections as β -connections to a radial distance of x .

Since β -cells are clustered, α -cells necessarily are also clustered (α -clusters), and since maximization of reciprocal β -connections excludes formation of short-range reciprocal α -connections, α -cells must form reciprocal synaptic connections at distances greater than x , to maximize $\int_{\mathbf{q}\alpha} \int_{\mathbf{R}} J^{\mathbf{q}\mathbf{R}} d\mathbf{q}\alpha d\mathbf{R}$. Similarly, reciprocal connections between α - and β -cells must occur at cluster margins, over distances approximate to x . Since we made the provisional assumption that $N_\beta \gg N_\alpha$, then fitting the sum of equations (15) and (16) to a power function requires $\lambda_\alpha \ll \lambda_\beta$. Consequently α -cells may form multiple patches of synaptic connections, skipping from α -cluster to α -cluster.

Since β -clusters have radius x and α -clusters are separated by distance x , α -clusters are necessarily placed at the vertices of hexagons tiling the cortical surface, with each hexagon embracing a β -cluster. Analogy to the superficial patch system in some species is apparent.

As noted earlier, hexagonal symmetry of OP and the superficial patch system is an idealization that is roughly approached in some species, while in others it is effectively absent (Horton

and Adams, 2005). Since approximation of a power law distribution by two populations of neurons requires $N_\alpha \ll N_\beta$ if $\lambda_\alpha \ll \lambda_\beta$, this case is more closely approached for larger cortical sizes, and the patchy connection system will have higher orderliness and hexagonal rotational symmetry. If $\lambda_\alpha < \lambda_\beta$ by only a small amount, as in animals with small cortical size, then N_β is not necessarily greater than N_α , and an ordered hexagonal structure need not be apparent. Such reduction of the apparent orderliness does not imply the absence of “small world” connectivity, nor imply impairment of function. The comparative invariance of distance between OP singularities across species reported by Kaschube et al. (2010) and Keil et al. (2012) implies that $x = -\ln\left(\frac{N_\alpha \lambda_\alpha}{N_\beta \lambda_\beta}\right) / 2\pi(\lambda_\beta - \lambda_\alpha)$ [equation (18)] is also relatively constant over species in the middle to large range of V1 size. Since the ratios N_β/N_α , and $\lambda_\alpha/\lambda_\beta$ must vary inversely in value in different species, according to cortical size, as required if the sum of the two synaptic distributions maintains a power law distribution, then comparative invariance of OP singularity density is to be expected.

LOCAL SYNAPTIC COMPETITION FOR METABOLIC SUPPLY

Turning from optimization of energy demand of axons, to that of dendrites, we can modify equation (2) to a form representing complex distinct processes of synaptic adaptation, impulse decay, and pre-synaptic synergy, including the limiting rate of metabolic energy supplied to excitatory synapses—viz:

$$\psi_e^{\mathbf{qr}'}(t) = \Gamma^{\mathbf{qr}'} M_e^{\mathbf{qr}'} * \varphi_e^{\mathbf{qr}'}(t) \quad (19)$$

$$M_e^{\mathbf{qr}'} = D \times S \quad (20)$$

$\Gamma^{\mathbf{qr}'}$ is the available fraction of the metabolic supply rate needed to attain maximum current flow. Since we have assumed increasing synaptic current in synchronously activated synapses increases the available metabolic supply, the value of $\Gamma^{\mathbf{qr}'}$ must follow that of $\psi_e^{\mathbf{qr}'}$.

$$D = \frac{1}{B} \exp[-Bt] \quad B > 0 \quad (21)$$

represents impulse decay following delivery of an afferent action potential, with time-integral of 1 (after Rennie et al., 2000).

$$S = 1/(1 + \exp[-g(J^{\mathbf{q}}(t))]) \quad (22)$$

is a sigmoid function with range 0–1, representing synaptic adaptation to the afferent pulse rate, and including the effect of pre-synaptic co-operation (Tsukada and Fukushima, 2010) upon individual synaptic current flow as $g(J^{\mathbf{q}})$ —a suitable ascending function in $J^{\mathbf{q}}$, such that if $J^{\mathbf{q}} = 0$, there is no current flow at the synapse.

As well as inter-cellular competition between assemblies of neurons, we assume competition takes place between adjacent individual synapses arising from the same neuron. Therefore those neurons that survive apoptosis must have found an efficient deployment of resource to the synapses best positioned to maximize the magnitude of synchrony. Any two adjacent synapses

arising from the same pre-synaptic neuron may terminate on the same, or different, post-synaptic neurons. If they terminate on the same neuron their conditions are essentially identical in terms of equations (19–22). If they terminate on different neurons, then the relevant values of $J^{\mathbf{q}}$ need not identical—and their competition for resources would lead, via the feedback between $\psi_e^{\mathbf{qr}'}$ and $\Gamma^{\mathbf{qr}'}$, to low synaptic current at one synapse, and high current at the other. Just what the physiological corollary of these opposite high and low-activity states is, and the critical metabolic component for which the synapses compete, we do not specify. A likely, but by no means unique contributing factor is the supply of extracellular calcium (Montague, 1996). Whatever the critical component(s), the important consequence is that, at synchronous equilibrium, closely situated neurons have either high, or low, pulse correlations with each other.

ORGANIZATION OF PRE-VISION RESPONSE PROPERTIES

We can now term those synapses that are transmitting impulses more strongly near equilibrium “saturated” synapses, and those which are more quiescent, but potentially able to be activated, “sensitive” synapses, and can consider what spatial patterns of saturated connections would best meet the requirement to maximize synchrony. Here a further property of the neural field commented on in relation to equations (1–5)—higher spatial cross-correlation of pulses and field potentials at shorter range—has a decisive impact on the equilibrium pattern of synaptic saturations. These emergent patterns, diagrammed in Figure 3, arise for the following reasons:

- (a) Maximum synchrony generation with highest cross-correlation among near-neighbors in each β -cluster requires saturated couplings link near-neighbor cells—but sensitive connections must also form between closely adjacent β -cells. Both requirements are met when saturated connections within each β -cluster form a re-entrant network analogous to a Möbius strip. A similar argument regarding connections formed within macrocolumns has been advanced earlier (Wright et al., 2006; Wright and Bourke, 2008).
- (b) The α -cluster system and each of the β -clusters must enter into maximum joint resonance. This requires the formation of a homeomorphic projection between scales. The projection must be homeomorphic, since spatial cross-correlation is constrained to decline with distance at both scales, and so if resonance is at a maximum, the projection map must be one preserving topological identity between scales. This is possible because a disk can be mapped to a Möbius strip. Thus saturated α -cell to β -cell synapses must systematically map limited angular ranges of the surrounding α -system onto limited angular ranges on the margins of each β -cluster, and receive reciprocal saturated β -cell to α -cell synapses. Such a mapping requires specification of an orientation and chirality for each β -cluster, and requires a reciprocal distribution of saturated and sensitive synapses from opposite sides of the α -system to neurons in a limited angular range within each β -cluster.
- (c) Maximum synchrony generation with high cross-correlation among near-neighbors in the α -system requires α -cells be

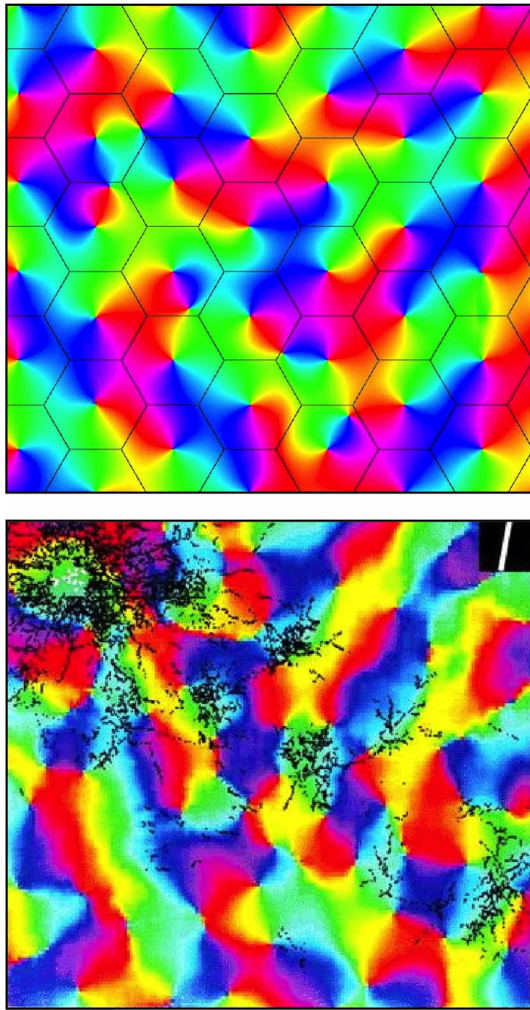


FIGURE 3 | Simulated and real maps of orientation preference in V1, from Wright et al. (2006). Top: Simulation. Colors of the spectrum, from red to violet, represent average OP of V1 neurons for slow-moving visual lines of orientation $0 - \pi$. Adjacent macrocolumns, of diameter $\sim 300 \mu\text{m}$ are set within a hexagonal frame (the patch system) with OP forming color wheels about OP singularities. Orientations and chiralities of the color wheels are arranged to approach a minimum total of angular disparity from mirror reflection of OP between each macrocolumn and its neighbors. **Bottom:** Real OP. Visualized in the tree shrew by Bosking et al. (1997). Superficial patchy connections are demarcated in black by a selective stain. Scale of macrocolumns is approximate to that of the simulation.

linked by saturated synapses. This requirement is concordant with deployment of the excess sensitive α -connections to neurons in β -clusters at positions outside the homeomorphic projection.

- (d) Saturated and sensitive $\beta - \beta$ connections between adjacent β -clusters must also be arranged to maximize resonance. Therefore β -clusters must project to each of their six neighbors as closely as possible to mirror symmetry, with both saturated and sensitive synapses linking homologous points—that is to say, points with similar OP as classically measured with low object speeds—within each cluster.

Perfect mirror symmetry is not possible between all adjacent clusters within a hexagonal array, so mirror-symmetry can be only approximate and irregular and the necessarily broken symmetry permits the particular pattern generated to be one of a large set of possible combinations.

Further analogy between the hypothetical α - and β -systems and real anatomical structures can now be drawn. As well as the α -system's congruence with the superficial patch system, the β -systems, each with a dense system of local connections that are centrally spared from patchy connections, are analogous to macrocolumns each centred about an OP singularity. The distribution of OP for lines of orientation $0 - \pi$ to angles $0 - 2\pi$ in pinwheels about a singularity finds analogy in the wrapping of a Euclidean plane onto a Möbius strip. It has also been earlier shown that arrangements of adjacent pinwheels in broken mirror symmetry match classical OP maps (Wright et al., 2006).

The structure of real patchy connections and classical OP response maps, contrasted with the results of simulating the arrangement of adjacent macrocolumnar structures in accord with the description above, are shown in **Figure 3**, while **Figure 4** shows diagrammatically the proposed arrangement of saturated and sensitive synapses, and foreshadows the effect of structured visual stimuli, once the post-natal phase of development begins—to be described in the next section.

Figure 5 shows a further impact upon response map organization—the emergence of OD columns.

Just as OP organization in some species is apparent before eye opening, so too is the organization into OD columns (Blakemore and Van Sluyters, 1975; Erwin and Miller, 1998). Explanation of this can be included in the present model by an argument similar to that of Erwin and Miller, who suppose the correlation of cell firing at short distances of separation of V1 cells to be greater than the correlation of visual inputs over a similar distance. This forces a columnar OD organization because of instability—in the present model's terms, the resulting disruption of the synchronous field at equilibrium produced by binocular inputs to the same cells—resolved by formation of columns in Turing patterns. A corollary of this effect is impact on the hexagonal arrangement, with broken mirror symmetry of OP organization, predicted in (d) above. The required alternation of OD columns would impose a frustration on the approach to hexagonal tiling of the cortical surface—forcing approach closer to a square tiling. The occurrence of mirror symmetry within a square tiling accounts for the way that lines of OP cross orthogonally between OP columns. (Obermayer and Blasdel, 1993). Following eyeopening inputs from the two eyes transmit images which are necessarily cross-correlated at a small spatial lag, because of angular disparity in their line of focus. Spatial lag correlation in their inputs at V1 level could then help maintain the columnar organization (Wright and Bourke, 2008).

WAVE TRANSMISSION OF VISUAL INFORMATION, FOLLOWING EYE-OPENING

We compactly express the emergent map by which the patchy connections over a part of V1 link to positions within each

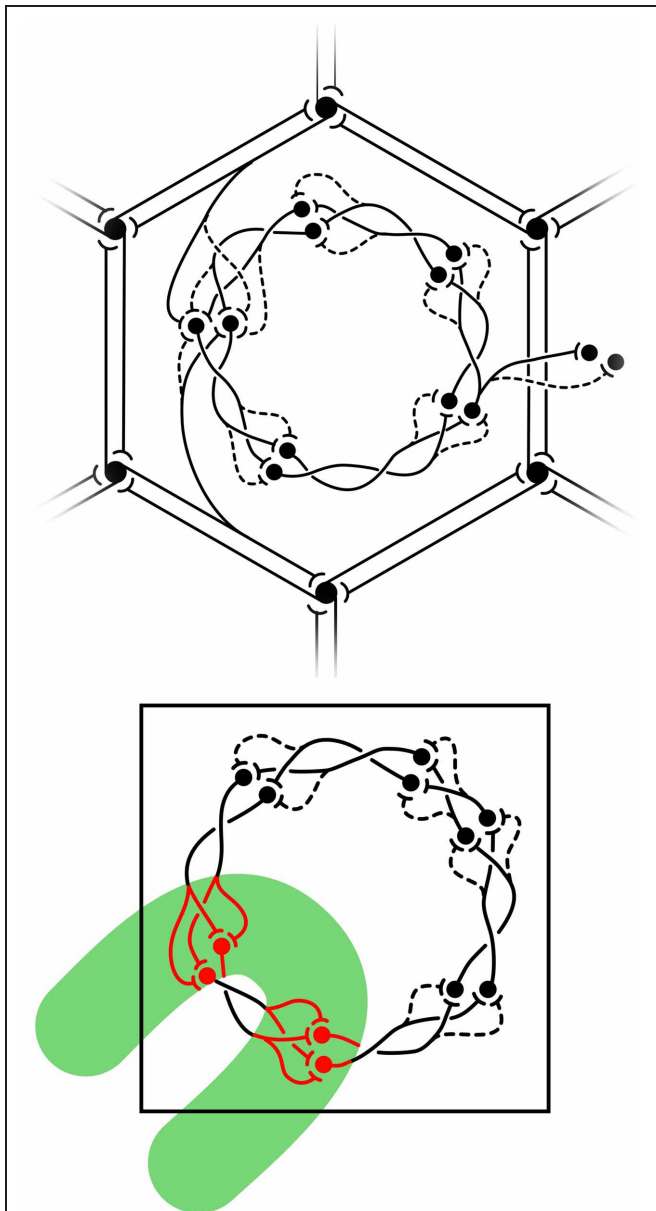


FIGURE 4 | Equilibrium distribution of synaptic activity, and the impact of visual inputs disrupting equilibrium from Wright and Bourke (2013). **Top:** equilibrium disposition of saturated and sensitive synapses. Black circles represent cell bodies and dendrites. Synapses are indicated as saturated (solid) or sensitive (dashed) terminations of axons. Reciprocal connections between α -patches (patchy connections) form the hexagonal array. (Other connections, although shown as unidirectional, are also reciprocal.) A representative pair of connections from α -cells to the β -patch is displayed in the upper-and lower-aspects of the figure. At the center of the figure, saturated and sensitive synapses show the network's analogy to a Möbius-strip within a β -patch (macrocolumn). To the right, representative links from the central macrocolumn to cells at homologous positions in neighboring macrocolumns. **Bottom:** exposed to strong transient signals conveyed over the superficial patch system, summing with direct visual inputs conveyed to the cRF, the equilibrium configuration breaks down. The green bar represents the field of excitation of cells by the contextual signals, within which cells also directly excited in the cRF, fire at high rates.

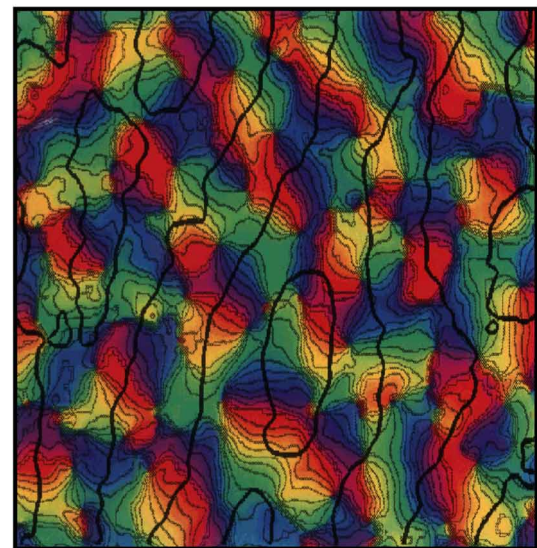
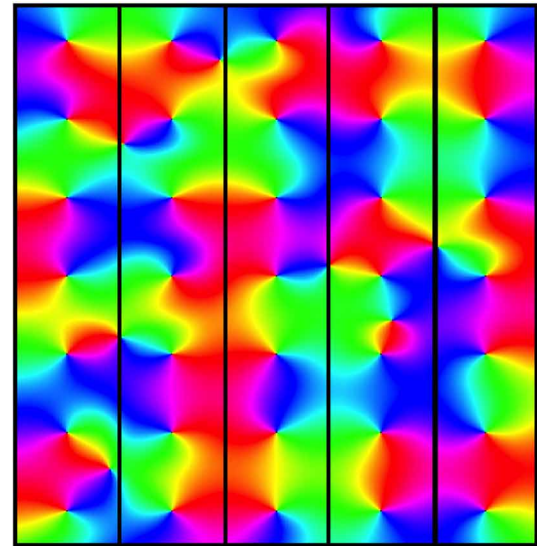


FIGURE 5 | Top: Simulation of OD columns in accord with Wright and Bourke (2008). **Bottom:** Real OD columns, visualized by Obermayer and Blasdel (1993). Color coding of OP and scale as for Figure 3. Black lines demarcate alternation of OD between columns. Fine black lines in the lower figure trace the way OP is aligned so it matches orthogonally across OD column boundary.

macrocolumn, as an homeomorphic projection from a disk on a Euclidean plane, \mathbf{P} , to a Möbius strip, $\mathbf{p}^{[2]}$ - the square brackets [2] indicating the map's resemblance, if viewed from a third dimension, to a 2:1 map formed by squaring a complex vector. Defined in polar co-ordinates,

$$\mathbf{P}(|\mathbf{R} - \mathbf{C}_j|, \vartheta) \rightarrow \mathbf{p}^{[2]}(|\mathbf{r} - \mathbf{C}_j|, \pm\vartheta + \varphi) \quad (23)$$

where \mathbf{C}_j is the origin of both \mathbf{P} and $\mathbf{p}^{[2]}$ for the j -th local map, and corresponds to the position of the OP singularity in that macrocolumn. ϑ is the polar angle of \mathbf{R} , chirality of the local

map is indicated by $\pm\vartheta$, and φ is the orientation of the local map relative to the global map. $\vartheta + \varphi$ is defined on the range $0 - 2\pi$ in both local and global maps, but is represented with apparent angle doubling in the local map. This describes a topology for “contextual” connections (Li et al., 2000; Angelucci and Bullier, 2003) to each macrocolumn.

Visual input after eye opening will cause departures from the equilibrium condition. Let $O(\mathbf{P}, t)$ be a visual image projected to V1 by the direct visual pathway. Laterally traveling waves of pulses and local field potentials relayed by the patchy connections can transmit that image to each local map with a point to point delay, $\frac{|\mathbf{R}-\mathbf{r}|}{v}$, where v now represents wave speed, so that

$$O(\mathbf{P}, t) \rightarrow O\left(\mathbf{p}^{[2]}, t + \frac{|\mathbf{R}-\mathbf{r}|}{v}\right) \quad (24)$$

Suppose $O(\mathbf{P}, t)$ is a segment of the image of a visual line, traveling with uniform velocity, \mathbf{V}_x , along an x -axis directed toward a macrocolumn with its singularity at \mathbf{C}_j . O has a component of its extension on the x -axis, O_x , and an orthogonal component of extension, on the y -axis, O_y . K_x is the dominant spatial frequency of O_x , and K_y is the dominant spatial frequency of O_y . Then the local map projection of O has a transformed spatial frequency in the x -axis but not in the y -axis—i.e.:

$$k_x \propto \frac{v}{v \pm \mathbf{V}_x} K_x \quad (25)$$

$$k_y \propto K_y \quad (26)$$

where k_x, k_y are the spatial frequencies in the local map projection of O , and the sign \pm in equation (25) depends on whether O is approaching or departing from \mathbf{C}_j . That is, O 's orientation in the global map is transformed to its projection to corresponding areas in the local map, by Doppler shift, with a difference in orientation, $\delta\vartheta$;

$$\delta\vartheta = |\tan^{-1}[K_y/K_x] - \tan^{-1}[k_y/k_x]| \quad (27)$$

INTERACTION OF CONTEXTUAL SIGNALS AND THE CLASSIC RECEPTIVE FIELD

Laterally transmitted contextual signals generally do not trigger cell firing, until the classic receptive field (cRF) is directly

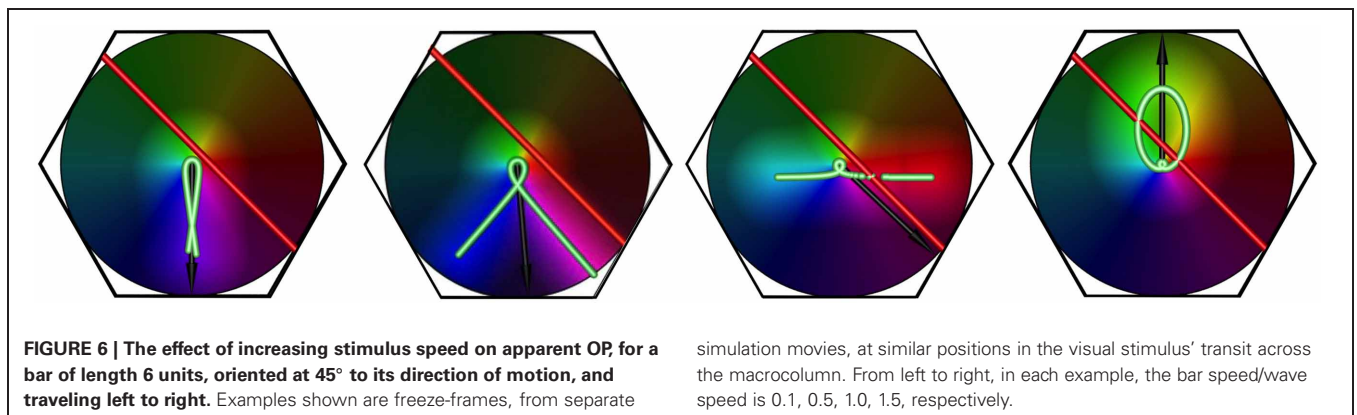
stimulated (Li et al., 2000; Angelucci and Bullier, 2003) via the visual pathway. Those cells that then fire within a macrocolumn are those that reflect the supra-threshold summations of sub-threshold signals conveyed over the contextual, patchy, connections, and the direct pathway. We next assume that the summation of contextual and direct cRF inputs acts as an impulse causing a transient breakdown of equilibrium, during which synapses that were in the sensitive state in equilibrium briefly generate substantial synaptic currents [See **Figure 4** (Bottom) and **Figure 6**]. Action potentials are triggered in surrounding cells, and subsequently there is a restoration toward the equilibrium state on withdrawal of the stimulus. During the breakdown the mapping of activity from the global to the local map becomes:

$$O(\mathbf{P}, t) \rightarrow O\left(\mathbf{p}^2, t + \frac{|\mathbf{R}-\mathbf{r}|}{v}\right) \quad (28)$$

The change from equation (24) made by removal of the square brackets from $\mathbf{p}^{[2]}$ represents the breakdown's form, as itself a map from global to local scale, resembling a 2:1 complex-multiplication map, as initially described by Alexander et al. (2004). The 2:1 map implies that single cells would show similar responses to a stimulus moving in either direction, but because firing is initiated over contextual connections in a 1:1 mapping, multi-cellular recordings would show that the spatial and temporal order of firing of neurons was unique for a given stimulus form and velocity.

POST-NATAL EFFECTS OF LEARNING, THE SPATIO-TEMPORAL FILTER MODEL, DIMENSION REDUCTION, AND “LIKE TO LIKE” CONNECTIONS

Equations (2,3,4, 19–22) contain state-variables required by mathematical expressions of physiological versions of the Hebb rule, and the spatio-temporal learning rule (Elliott and Shadbolt, 2002; O'Connor et al., 2005a,b; Enoki et al., 2009; Tsukada and Fukushima, 2010; Elliott, 2011). Following eye opening, stimuli with regularly repeated spatial and temporal structure reach V1, so we assume that exposure to a repeated stimulus leads to permanent synaptic consolidation of connections, overlaying those formed in the ante-natal, equilibrium condition. As remarked in the Introduction, Baker and Issa (2005) have shown that all V1 response features can be described in terms of six variables—optimal values of OP, spatial frequency preference, and temporal



frequency preference, each associated with a Gaussian bandwidth of tuning of the cortical response to these features. These define three hypothetical filter processes. However, stimulus variables in the present model have equivalents to those used in the spatio-temporal filter model. These are:

Spatio-temporal model	Present model
Object orientation	Orientation relative to the y-axis defined for equations (25, 26)
Object velocity	V_x
Object drift angle	$\tan^{-1}[K_y/K_x]$
Object spatial frequency	$K_x/ \cos(\tan^{-1}[K_y/K_x]) $
Object temporal frequency	$V_x K_x$

Repeated stimulation with a particular stimulus will therefore lead, under Hebbian learning, to maximization of the response to that stimulus, thus creating an apparent “tuning” of particular neurons to that particular combination of stimulus features. Thus, the spatio-temporal model can be regarded as a consequence of the present model. Optimization by learning of the parameters for each of the three filters must be competitive between adjacent cells, providing the necessary condition for fitting response maps with continuity and completeness, by dimension-reduction methods (Kohonen, 1982; Durbin and Willshaw, 1987; Durbin and Mitchison, 1990).

Finally, the consolidation of saturated long-range patchy connections by Hebbian learning would result in mature “like to like” connections.

SIMULATIONS—A CRITICAL TEST

A critical test of our model, then, is whether we can reproduce in simulation the results of Basole et al. (2003), without *a priori* feature-specific responses to orientation, spatial frequency, or temporal frequency. Our simulations assume the steady-state presence of the Mobius synaptic configuration and its perturbation by visual signals, intended to reflect the state of the visual cortex shortly after birth, when first exposed to visual stimuli.

Equation (28) was applied in simulations of an hexagonal array of seven adjacent macrocolumns. Results reported here are for the central macrocolumn of the array of 7. Examples are shown in **Figure 6**, which shows the orthogonal transformation of apparent OP from the lowest to the highest bar speed for a moving line stimulus oriented at 45° to its line of passage.

Diameter for each macrocolumn is 300 microns, and wave speed for transcortical polysynaptic propagation 0.1 m/s (Binguier et al., 1999). Units of length subsequently referred to, are multiples of the radius of a macrocolumn—150 microns. Simulation time-step was 0.1 ms.

A moving line in the visual field, relayed by the direct visual pathway to the cRF of each macrocolumn is represented as a red bar. In a single simulation the red bar traveled across the entire hexagonal array from left to right, with constant speed, direction and orientation. The orientation of the red bar to the line

of passage is measured as *bar angle* from 0°, where the bar is oriented orthogonally to the direction of travel, to ±90°, where the bar is oriented in the direction of travel.

The lag-transmitted image of the red bar, relayed as subthreshold activation to each macrocolumn via the superficial patch system, is shown in green, with illumination about the zone of subthreshold activation, to indicate that input to the cRF from the direct visual pathway and contextual signals caused triggering of action potentials. The average angle from the macrocolumn singularity to the centers of action potential generation (i.e., all points on the green line with illumination) was calculated at each time-step, and shown as a black arrow, thus indicating the part of the macrocolumn with a response preference (*apparent OP*) for the particular bar movement. (A change in the sector of the macrocolumn that is maximally stimulated is equivalent to an equal change in the angle of approach of the bar needed to maintain stimulation of the same sector). The black arrow angle was averaged over a window beginning after the red bar had passed the center of the macrocolumn by a distance equal to 10% of macrocolumn radius, and extending from the 10th percentile to the 20th percentile of that radius, thus obtaining an estimate of the apparent OP during the cRF activation time. The standard error (SE) of the black arrow angles was calculated from 11 equally spaced time steps through the averaging window.

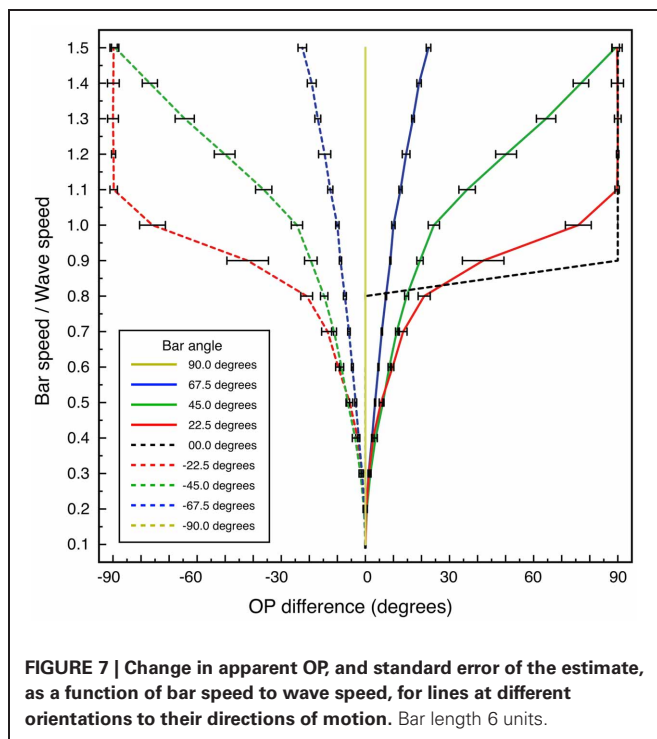
Combinations of bar-length, orientation of the bar to the direction of movement, and bar speed, were then systematically varied in separate simulations, results of which are supplied as supplementary animated movies. Their effects on OP, measured at the central local map of the hexagonal group, were obtained as *OP difference*, $\Delta\phi$ —a measure of the change in OP as a function of these variables—calculated as

$$\Delta\phi = \begin{cases} \phi_1 - \phi_0 - \pi & \text{when } \pi/2 < \phi_1 - \phi_0 \\ \phi_1 - \phi_0 & \text{when } -\pi/2 \leq \phi_1 - \phi_0 \leq \pi/2 \\ \phi_1 - \phi_0 + \pi & \text{when } \phi_1 - \phi_0 < -\pi/2 \end{cases} \quad (29)$$

The *reference OP*, $\phi_0 \in [0, \pi)$, was the OP found at the lowest bar speed applied (bar speed/wave speed = 0.1) and the *apparent OP*, $\phi_1 \in [0, \pi)$, was the OP found at higher speeds.

Systematic results are shown in **Figure 7**, which graphs OP difference vs. bar speed/wave speed, for bar angles 0 to ±90°, calculated for a bar length of 6 units. Variation of bar length showed progressive lessening of the effect of velocity on OP for greater bar lengths.

For the case of bar-angle zero degrees (a line oriented orthogonally to its direction of passage, as in classical measurements of OP) no OP difference is seen until, as bar speed approaches wave speed, a 90° change in apparent OP takes place at a single increment in speed. This corresponds to transition to a “motion streak,” as object movement blurs resolution in the direction of motion. Increasing OP difference with bar speed at other bar angles is a more gradual development of the same effect—that is, mixing of responses to object speed and to object orientation. The illuminated field of supra-threshold excitation generated is not that expected to accompany a Gaussian-shaped tuning



curve, but is roughly bimodal at medium speeds—e.g., bar-speed/wave-speed = 0.5. The form of the field of excitation is a combination of the classical preferred OP and the orthogonal orientation, expected as a consequence of Doppler shifts in the laterally-transmitted cortical signals generated by the moving visual input.

Variation of the window over which the apparent OP was estimated did not affect the qualitative results so long as averaging was conducted over a window beginning after the center point of the macrocolumn was crossed by the red bar. Variation of the estimate of wave speed was also without effect, so long as results were expressed in terms of bar speed/wave speed.

For comparative purposes similar simulations were performed in which contextual (green bar) responses were constrained to occur only with a limited angular response within a macrocolumn. That is, a restricted response to the line, according only to its orientation was imposed, in analogy to conventional models of OP, but with conduction delays of “like to like” fibers included. Then, systematic variation of OP with bar velocity did not occur.

These results match the findings of Basole et al. (2003) with respect to variation of OP peak responses as a function of line velocity and length. They do not reproduce the form of the experimentally observed Gaussian tuning curves, but as argued in the prior section, subsequent post-natal Hebbian learning progressively over-writing the Möbius configuration, and strengthening the peak response to the optimal visual signal, would concurrently strengthen responses to signals which are close to the optimum, resulting in Gaussian tuning curves in the more mature animals studied by Basole et al. (2003) and Issa et al. (2008).

INTER-AREAL INTERACTIONS OF V1 AND HIGHER VISUAL AREAS

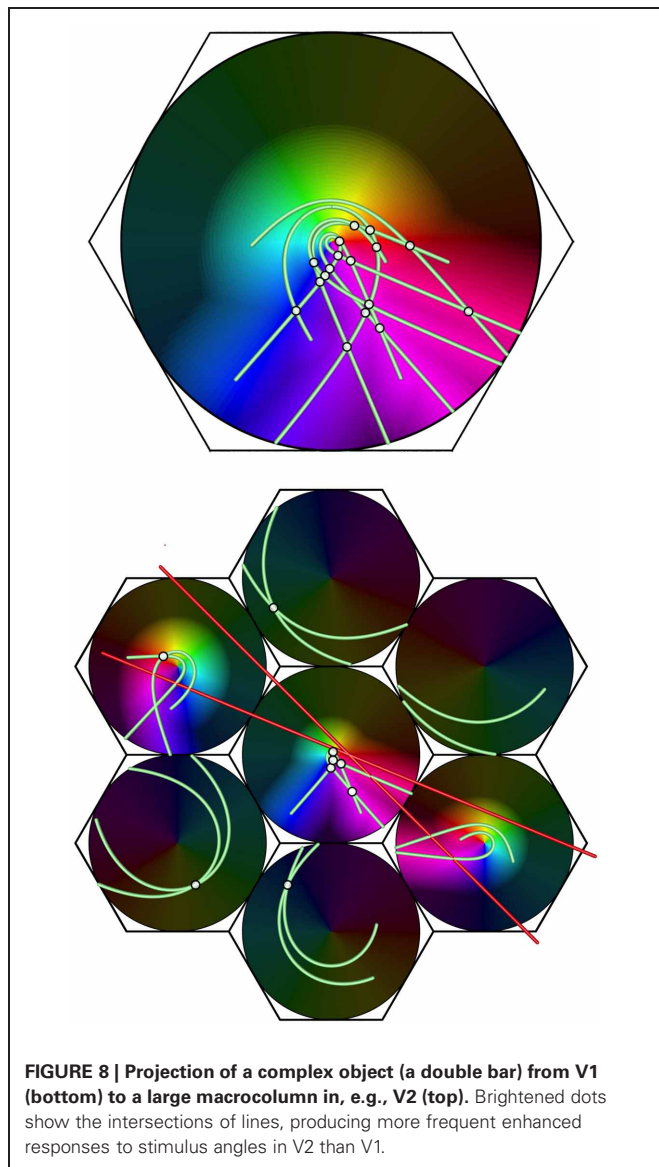
The principle underlying the development of connections between macrocolumns and the superficial patch system may be generalized to the emergence of inter-areal connections. To recap, taking V1 as an example we have argued above that, because co-variance of activity declines with metric distance at both the scale of the patchy connections and within a macrocolumn, a homeotypic mapping between scales can emerge. This requires that relative distances on the maps at each scale must be in the ratio of correlation lengths of synchronous oscillation at the two scales, and adjacent maps must themselves have a correlated structure over a distance approximate to the correlation length of the patch system. It then follows that superposition of adjacent local maps, with appropriate rotation and correction to a common chirality, would result in a further map with co-variance of activity declining with metric distance, over the correlation length of the patch system.

Inter-areal connections, made by cortico-cortical axonal projections, could permit maps of this type to arise during ante-natal development, with the composite map at the higher cortical-area level itself folded into the Möbius configuration. The selection of saturated connections, projecting between areas with normalization of rotations and chiralities, would be possible by selection from the larger set of possible connections made by branches of cortico-cortical axons, diverging from their cells of origin to their cells of termination, overlapping as they terminate, and generally reciprocal between areas (Braitenberg and Schüz, 1991; Boucsein et al., 2011). Thus, antenatally, sets of macrocolumns at both the lower, V1, level and higher levels, could resonate with, and form preferential connections with, superimposed and overlapping groups at the other level, in accord with the developmental selection requirement to maximize joint synchrony. With the occurrence of eye-opening, Hebbian learning would then begin to overwrite the equilibrium resonance configuration between areas, in analogy to the process at intra-areal level—with the added property of associating concurrent patterns of activity in the V1 macrocolumns.

Illustrating this effect, **Figure 8** shows, at the bottom, a system of seven macrocolumns at V1 level, driven via the direct visual pathway by a pair of intersecting lines in the visual field.

The top part of the figure shows a projection of activity in conjointly activated macrocolumns in V1, to a higher visual area, in which responses in the seven macrocolumns in V1 have been superimposed, with disparities in their orientation and chirality eliminated. Summations of points stimulated by both lines, shown by highlighted white points, occur frequently in the forward projection—much more so than at the level of V1 itself. These indicate response to angles of intersection of the lines at the lower level, and, commonly at the higher level, summing responses to time-lagged correlations between disparate positions of the moving visual stimulus. This effect is consistent with the preferential responses to angular and complex stimuli, characteristic of higher cortical levels (Merigan and Maunsell, 1993).

Conversely, since connections between higher and lower levels are generally reciprocal, a possible mechanism permitting control of attention (Rao and Ballard, 1997; Kveraga et al., 2007;



Swindale, 2008; Naci et al., 2012) is suggested, since the backward flow would continuously modify the forward flow of sensory information.

CONCLUSION

The model of cortical development we have outlined above is efficient from both energetic and information-processing perspectives, and has considerable anatomical and physiological explanatory power. It leads to an explanation of the spatial organization of signal flow in the cortex that differs from any other model. The proposed antenatal self-organization of cortical synapses leads to the creation of a *tabula rasa* on which homeomorphic maps, in a form disguised by the Möbius-strip-like folding of connections, occur in lateral connections at the millimetric scale, embedding the statistics of spatial organization of the sensory world to first approximation, before any detailed sensory inputs are received. The assumptions and

findings of the model overlap with, and although not necessarily contradictory to, are not identical to, those of other models (Erwin and Miller, 1998; Wolf and Geisel, 1998; Kang et al., 2003; Oster and Bressloff, 2006). Distinguishing features include the explanation of the relationship of superficial patch connections to macrocolumn centers, and their hexagonal rotational symmetry, and crucially, the findings of Basole et al. (2003), which cannot be explained by any model dependent on “like-to-like” connections between feature-specific neurons. Nor can any model with otherwise similar assumptions about the self-organizing effect of synchrony be formulated without introducing a Möbius configuration to the connections, since an equivalent model utilizing only Euclidian conformations would represent a given OP twice, rather than once, around a singularity.

In review, the assumptions and conclusions reached, were as follows. By assuming that cells surviving apoptosis are selected by competition for metabolic substrates, and that synchronous oscillation mediates the uptake of metabolic substrates, we showed the outcome was a neural system with ultra-small world axonal configuration. Further assuming the small world connections were necessarily constructed from neuron populations characterized by respective axonal length, we showed that long range patchy connections and regular macro-column-like areas with central sparing of patchy connections emerge, with some degree of hexagonal rotational symmetry, with species variation in orderliness according to cortical size, and were able to show that this result was consistent with anatomical observations of limited interspecies variation of singularity density. A crucial further assumption made, was that metabolic competition between synapses from the same neuron leads to particular configuration of synaptic current flows at equilibrium, in which active connection networks within each macrocolumn are arranged in a Möbius-strip-like conformation. Then, with the introduction of visual inputs, signals conveyed by contextual fibers transfer a visual image from the global map to each local map, determining the pattern of neuron firing induced by activation of the cRF, and synaptic consolidation on Hebbian principles begins—thus storing information based on visual experience—explaining how response maps for OP, SF, and TF become organized in accord with the spatio-temporal filter model (Baker and Issa, 2005; Issa et al., 2008), and how “like to like” anatomical connections emerge, as well as providing conditions for dimension-reduction description of response features. The model is also compatible with explanation of ocular columns and direction preference fractures, as proposed in our earlier work (Wright et al., 2006).

The resulting synaptic storage of learned information in local topological maps of Möbius configuration offers a further compression of format, adding to the efficiency of the “small world” arrangement, by minimizing the distance which need be spanned by connections between positions on the local map representing positions widely separated on the global map. The development of cross-links also offers large potential information storage, since the regular spatial organization of links in the Möbius configuration implies the synaptic connections have low joint entropy in their ante-natal state. With visual experience, and the storage of image information in cross-links, joint entropy could, in

principle, increase to a limit where all synaptic states are independent, and equally distributed about some mean connection strength, as implied by Montague's (1996) resource consumption principle. In effect, before eye-opening, the cortex has "learned" the underlying statistical structure of visual space—that of cross-correlation declining with metric distance—and subsequently stores information about departures from this "first component" of structure in the visual world.

The antenatal development of response maps (Wiesel and Hubel, 1974; Blakemore and Van Sluyters, 1975; Sherk and Stryker, 1976) presents no paradox in this model, since emergence of organized response properties within the Möbius configuration does not depend upon structured visual stimuli. EEG activity progressively matures toward alternating alert and sleeping states in the later antenatal period (Marks et al., 1995; Mirmiran, 1995) providing the widespread co-ordination of pre-synaptic activity required for initial synaptic self-organization. Conversely, overwriting by learning in the immediate and later post-natal periods explains why representation in adult response maps of stimuli to which the subject has not been exposed would not be present—as also seen experimentally (e.g., Blakemore and Van Sluyters, 1975).

No direct evidence yet exists of Möbius-like patterns of connections in cortex, yet this is scarcely surprising if transient dynamic couplings, present only in the equilibrium state, are overwritten by post-natal learning-related changes. However, relaxation toward the equilibrium condition is still to be expected in the mature state, so it is important whether or not some anatomical substrate exists in which the dynamic state of synapses may be capable of transient assembly into Möbius patterns. Markram and colleagues (Perin et al., 2011) found that pyramidal neuron networks cluster into multiple groups of a few dozen neurons each, with the neurons composing each group typically more than 100 μm apart, allowing for multiple groups to be interlaced in the same space. Connections within groups were largely reciprocal, and those between groups relatively sparse. Transient interlinkages between such interwoven linked groups could form Möbius-like networks. The temporal plasticity of synaptic connections near singularities (Dragoi et al., 2001) is also consistent with this interpretation. As well as plasticity of responses near OP singularities, Dragoi and colleagues found lack of plasticity in linear zones—the areas of strong patchy connection termination. This is to be expected if the patch system is composed of well consolidated connections suitable for consistent transmission with delay from fixed points in V1, while of the other hand, more complex, continually modified, information processing goes on in the areas around singularities. Consequently, an anatomical test of the model may be possible, in regard to the terminations of patchy connections in the periphery of the patch-free areas about singularities. As indicated in **Figure 3**, two populations of synaptic connections should be demonstrable in principle, by double injection/staining methods, near the singularity/patch edge. If some Hebbian consolidation occurs both antenatally and postnatally, then, in principle it should be possible to observe Möbius-like connections within macrocolumns antenatally, and the overwriting of these connections during post-natal learning.

If later testing supports this model, current conceptions of cortical information processing will require modification. Synchronous oscillation has been regarded as a mechanism for feature-binding—requiring that groups of cells in synchrony stand out in some way against a non-synchronized background. Instead, this model emphasizes synchrony as the organizer of a matrix of connections within which each macrocolumn gathers information from its surround, and organizes these connections systematically according to spatial position and time-lag, as functions of distance from each singularity. The topology of signal organization is markedly different to that of the association of "feature" neurons embedded in neural connections that are deployed on a Euclidean plane, as it implies that sensory images are not broken up into "features" which are subsequently associated in an abstract feature space, but retain, in modified form, an organization representing sensory space and time. Upon this more complicated matrix of connections, moment-by-moment states of autonomous local firing could interact with each other via traveling waves, generating internal images adding to those arising from sensation—all selectively strengthening preferred pathways by Hebbian learning, under the supervision of motivational systems. This gives a modified basis to Sherrington's "enchanted loom" (Sherrington, 1906, 1940), and a stage for the kind of neuro-dynamic events progressively observed and envisaged by Freeman for many years (Freeman, 1975; Freeman and Quiroga, 2013).

Hierarchical interaction of V1 with higher visual areas, by superposition of spatio-temporal images transmitted over convergent and divergent pathways might proceed to higher levels of abstraction, at higher cortical levels, and feedback interactions of ascending and descending signals in such a system might permit very complex image manipulation. Analogous processes may apply to other modalities throughout the cortex in general, since all sensory input systems are analogous to the visual system, in as much as they encode the sensory world by imposing a topological order to inputs as they arrive at the sensory cortices. Again, the ubiquitous distribution of patchy connections throughout the cortex, and the basic modular similarity of the paleo- and neo-cortex throughout, supports the notion that a single schema of information flow may be characteristic of all. The principle of organization might even extend to the motor cortex, with the efferent pyramidal motor neurons simply reversing the role of neurons in the direct visual pathway.

ACKNOWLEDGMENTS

A special debt of gratitude is owed to Adrienne Edith Wright.

This work was supported by the Frank P. Hixon Fund of the California Institute of Technology and the Oakley Foundation of New Zealand, and by iVEC, through the use of advanced computing resources located at the University of Western Australia.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Computational_Neuroscience/10.3389/fncom.2013.00004/abstract

REFERENCES

- Albert, M. V., Schnabel, A., and Field, D. J. (2008). Innate visual learning through spontaneous activity patterns. *PLoS Comput. Biol.* 4:e1000137. doi: 10.1371/journal.pcbi.1000137
- Alexander, D. M., Bourke, P. D., Sheridan, P., Konstantatos, O., and Wright, J. J. (2004). Intrinsic connections in tree shrew V1 imply a global to local mapping. *Vision Res.* 44, 857–876.
- Amari, S.-I. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87.
- Angelucci, A., and Bullier, J. (2003). Reaching beyond the classical receptive field of V1 neurons; horizontal or feedback axons? *J. Physiol.* 97, 141–154.
- Angelucci, A., Levitt, J. B., and Lund, J. S. (2002). Anatomical origins of the classic receptive field and modulatory surround field of single neurons in macaque visual cortical area V1. *Prog. Brain Res.* 136, 373–388.
- Baker, T. I., and Issa, N. P. (2005). Cortical maps of separable tuning properties predict population responses of complex visual stimuli. *J. Neurophysiol.* 94, 775–787.
- Basole, A., Kreft-Kerekes, V., White, L. E., and Fitzpatrick, D. (2006). Cortical cartography revisited: a frequency perspective on the functional architecture of visual cortex. *Prog. Brain Res.* 154, 121–134.
- Basole, A., White, L. E., and Fitzpatrick, D. (2003). Mapping of multiple features in the population response of visual cortex. *Nature* 423, 986–990.
- beim Graben, P. (2008). “Foundations of neurophysics,” in *Lectures in Supercomputational Neuroscience: Dynamics in Complex Brain Networks*, eds P. Graben, C. Zhou, M. Thiel, and J. Kurths (Berlin: Springer), 3–48.
- Blakemore, C., and Van Sluyters, R. C. (1975). Innate and environmental factors in the development of the kitten's visual cortex. *J. Physiol.* 248, 663–716.
- Bosking, W. H., Zhang, Y., Schofield, B., and Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* 17, 2112–2127.
- Boucsein, C., Nawrot, M., Schnepel, P., and Aertsen, A. (2011). Beyond the cortical column: abundance and physiology of horizontal connections imply a strong role for inputs from the surround. *Front. Neurosci.* 5:32. doi: 10.3389/fnins.2011.00032
- Braitenberg, V., and Schüz, A. (1991). *Anatomy of the Cortex: Statistics and Geometry*. Berlin, New York: Springer.
- Bressler, S. L., Coppola, R., and Nakamura, R. (1993). Episodic multiregional cortical coherence at multiple frequencies during visual task performance. *Nature* 366, 153–156.
- Bressloff, P. (2012). Spatiotemporal dynamics of continuum neural fields. *J. Phys. A Math. Theor.* 45:033001. doi: 10.1088/1751-8113/45/3/033001
- Bringuier, V., Chavane, F., Glaeser, L., and Frégnac, Y. (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science* 283, 695–699.
- Buzás, P., Kovács, K., Ferecskó, A. S., Budd, J. M. L., Eysel, U. T., and Kisvárdy, Z. F. (2006). Model-based analysis of excitatory lateral connections in the visual cortex. *J. Comp. Neurol.* 499, 861–881.
- Callaway, E. M., and Katz, L. C. (1990). Emergence and refinement of clustered horizontal connections in cat striate cortex. *J. Neurosci.* 10, 1134–1153.
- Carriera-Perpiñán, M. Á., Lister, R. J., and Goodhill, G. J. (2005). A computational model for development of multiple maps in primary visual cortex. *Cereb. Cortex* 15, 1222–1233.
- Chapman, C. L., Bourke, P. D., and Wright, J. J. (2002). Spatial eigenmodes and synchronous oscillation: coincidence detection in simulated cerebral cortex. *J. Math. Biol.* 45, 57–78.
- Cohen, R., and Havlin, S. (2003). Scale-free networks are ultra-small. *Phys. Rev. Lett.* 90:058701. doi: 10.1103/PhysRevLett.90.058701
- Crick, F. C., and Koch, C. (2003). A framework for consciousness. *Nat. Neurosci.* 6, 119–126.
- DeAngelis, G. C., Ohzawa, I., and Freeman, R. D. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J. Neurophysiol.* 69, 1091–1117.
- Downes, J. H., Hammond, M. W., Xydas, D., Spencer, M., Becerra, V. M., Warwick, K., et al. (2012). Emergence of a small-world functional network in cultured neurons. *PLoS Comput. Biol.* 8:e1002522. doi: 10.1371/journal.pcbi.1002522
- Dragoi, V., Rivadulla, C., and Sur, M. (2001). Foci of orientation plasticity in visual cortex. *Nature* 411, 80–86.
- Durack, J. C., and Katz, L. C. (1996). Development of horizontal projections in layer 2/3 of ferret visual cortex. *Cereb. Cortex* 6, 178–183.
- Durbin, R., and Mitchison, G. (1990). A dimension reduction framework for understanding cortical maps. *Nature* 343, 644–647.
- Durbin, R., and Willshaw, D. J. (1987). An analogue approach to the travelling salesman problem using an elastic net method. *Nature* 326, 689–691.
- Eckhorn, R., Bauer, R., Jordon, W., Brosch, M., Kruse, W., Monk, M., et al. (1988). Coherent oscillations: a mechanism of feature linking in the visual cortex? *Biol. Cybern.* 60, 121–130.
- Eckhorn, R., Reitboeck, H. J., Arndt, M., and Dicke, P. (1990). Feature linking via synchronization among distributed assemblies: simulations of results from cat visual cortex. *Neural Comput.* 2, 293–307.
- Elliott, T. (2011). Stability against fluctuations: scaling, bifurcations, and spontaneous symmetry breaking in stochastic models of synaptic plasticity. *Neural Comput.* 23, 674–734.
- Elliott, T., and Shadbolt, N. R. (2002). Multiplicative synaptic normalization and a nonlinear Hebb rule underlie a neurotrophic model of competitive synaptic plasticity. *Neural Comput.* 14, 1311–1322.
- Enoki, R., Hu, Y.-L., Hamilton, D., and Fine, A. (2009). Expression of long-term plasticity at individual synapses in hippocampus is graded, bi-directional, and mainly pre-synaptic: optic quantal analysis. *Neuron* 62, 242–253.
- Erwin, E., and Miller, K. D. (1998). Correlation-based development of ocular matched orientation and ocular dominance maps: determination of required input activities. *J. Neurosci.* 18, 9870–9895.
- Freeman, W. J. (1975). *Mass Action in the Nervous System*. New York, NY: Academic Press.
- Freeman, W. J. (1991). “Predictions on neocortical dynamics derived from studies in paleocortex,” in *Induced Rhythms of the Brain*, eds E. Basar and T. H. Bullock (Boston, MA: Birkhäuser), 183–195.
- Freeman, W. J., and Quiroga, R. Q. (2013). *Imaging Brain Function*. New York, Heidelberg, Dordrecht, London: Springer.
- Gilbert, C. D., and Wiesel, T. N. (1979). Morphology and intracortical projections of functionally characteristic neurons in cat visual cortex. *Nature* 280, 120–125.
- Gilbert, C. D., and Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J. Neurosci.* 9, 2432–2442.
- Gray, C. M., König, P., Engel, A. K., and Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties. *Nature* 388, 334–337.
- Grossberg, S., and Olson, S. J. (1994). Rules for the cortical map of ocular dominance and orientation columns. *Neural Netw.* 7, 883–894.
- Haken, H. (1996). *Principles of Brain Functioning*. Berlin: Springer.
- Harris, A. E., Ermentrout, G. B., and Small, S. L. (1997). A model of ocular column development by competition for trophic factor. *Proc. Natl. Acad. Sci. U.S.A.* 94, 9944–9949.
- Hassenstaub, A., Shu, Y., Haider, B., Krauschaar, U., Duque, A., and McCormick, D. A. (2005). Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron* 47, 423–435.
- Heck, N., Golbs, A., Riedemann, T., Sun, J.-J., Lessmann, V., and Luhmann, H. J. (2008). Activity dependent regulation of neuronal apoptosis in neonatal mouse cerebral cortex. *Cereb. Cortex* 18, 1335–1349.
- Hirsch, J. A., and Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *J. Neurosci.* 11, 1800–1809.
- Horton, C. H., and Adams, D. L. (2005). The cortical column: a structure without a function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 837–862.
- Hubel, D. H., and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* 148, 574–591.
- Issa, P., Rosenberg, A., and Husson, T. R. (2008). Models and measurements of functional maps in V1. *J. Neurophysiol.* 99, 2745–2754.
- Jirsa, V. K., and Haken, H. (1996). Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* 77, 960–963.
- Kang, K., Shelley, M., and Sompolsky, H. (2003). Mexican hats and pinwheels in visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 100, 2848–2853.
- Kaschube, M., Schnabel, M., Löwel, S., Coppola, D. M., White, L. E., and Wolf, F. (2010). Universality in the evolution of orientation columns in the visual cortex. *Science* 330, 1113.
- Keil, W., Kaschube, M., Schnabel, M., Kisvárdy, Z. F., Löwel, S., Coppola,

- D. M., et al. (2012). Response to comment on "Universality in the evolution of orientation columns in the visual cortex". *Science* 336, 413.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol. Cybern.* 43, 59–69.
- Kveraga, K., Ghuman, A. S., and Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain Cogn.* 65, 145–168.
- Li, W., Their, P., and Wehrhahn, C. (2000). Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *J. Neurophysiol.* 83, 941–954.
- Mariño, J., Schummers, J., Lyon, D. C., Schwabe, L., Beck, O., Wiesel, P., et al. (2005). Invariant computations in local cortical networks with balanced excitation and inhibition. *Nat. Neurosci.* 8, 194–201.
- Marks, G. A., Shaffery, J. P., Oksenberg, A., Speciale, S. G., and Roffwarg, H. P. (1995). A functional role for REM sleep in brain maturation. *Behav. Brain Res.* 69, 1–11.
- McGuire, B. A., Gilbert, C. D., Rivlin, P. K., and Wiesel, T. N. (1991). Targets of horizontal connections in macaque primary visual cortex. *J. Comp. Neurol.* 305, 370–392.
- Mante, V., and Carandini, M. (2005). Mapping of stimulus energy in primary visual cortex. *J. Neurophysiol.* 94, 788–798.
- Meng, Y., Tanaka, S., and Poon, C.-S. (2012). Comment on "Universality in the evolution of orientation columns in the visual cortex". *Science* 336, 413. Available online at: www.sciencemag.org/cgi/content/full/336/6080/413-d
- Merigan, W. H., and Maunsell, J. H. R. (1993). How parallel are the primate visual pathways? *Annu. Rev. Neurosci.* 16, 369–402.
- Mirmiran, M. (1995). The function of fetal/neonatal rapid eye movement sleep. *Behav. Brain Res.* 69, 13–22.
- Miyashita, M., and Tanaka, S. (1992). A mathematical model for the self-organization of orientation columns in visual cortex. *Neuroreport* 3, 69–72.
- Montague, P. R. (1996). The resource consumption principle: attention and memory in volumes of neural tissue. *Proc. Natl. Acad. Sci. U.S.A.* 93, 3619–3623.
- Muir, D. R., Da Costa, N. M. A., Girardin, C. C., Naaman, S., Omer, D. B., Ruesch, E., et al. (2011). Embedding of cortical representations by the superficial patch system. *Cereb. Cortex* 21, 2244–2260.
- Muir, D. R., and Douglas, R. J. (2011). From neural arbours to daisies. *Cereb. Cortex* 21, 1118–1133.
- Naci, L., Taylor, K. I., Cusack, R., and Tyler, L. K. (2012). Are the senses enough for sense? Early high-level feedback shapes our comprehension of multisensory objects. *Front. Integr. Neurosci.* 6:82. doi: 10.3389/fnint.2012.00082
- Nunez, P. L. (1981). *Electric Fields of the Brain*. London: Oxford University Press.
- Obermayer, K., and Blasdel, G. G. (1993). Geometry of orientation and ocular dominance columns in monkey striate cortex. *J. Neurosci.* 13, 4114–4129.
- Obermayer, K., Ritter, H., and Schulten, K. (1990). A principle for the formation of the spatial structure of cortical feature maps. *Proc. Natl. Acad. Sci. U.S.A.* 87, 8345–8349.
- Obermayer, K., Ritter, H., and Schulten, K. (1992). A model for the development of the spatial structure of retinotopic maps and orientation columns. *IEICE Trans. Fundamentals* E75A, 537–545.
- O'Connor, D. H., Wittenberg, G. M., and Wang, S. S.-H. (2005a). Dissection of bidirectional synaptic plasticity into saturable unidirectional processes. *J. Neurophysiol.* 94, 1565–1573.
- O'Connor, D. H., Wittenberg, G. M., and Wang, S. S.-H. (2005b). Graded bidirectional synaptic plasticity is composed of switch-like unitary events. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9679–9684.
- Oster, A. M., and Bressloff, P. C. (2006). A developmental model of ocular dominance column formation on a growing cortex. *Bull. Math. Biol.* 68, 73–98.
- Paik, S.-B., and Ringach, D. L. (2011). Retinal origin of orientation maps in visual cortex. *Nat. Neurosci.* 14, 919–925.
- Perin, R., Berger, T. K., and Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. *Proc. Natl. Acad. Sci. U.S.A.* 108, 5419–5424.
- Price, D. J. (1986). The postnatal development of clustered intrinsic connections in area 18 of the visual cortex in kittens. *Dev. Brain Res.* 24, 31–38.
- Rakic, P. (1988). Specification of cerebral cortical areas. *Science* 241, 170–176.
- Rao, R. P. N., and Ballard, D. H. (1997). Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Comput.* 9, 721–763.
- Rennie, C. J., Wright, J. J., and Robinson, P. A. (2000). Mechanisms of cortical electrical activity and the emergence of gamma rhythm. *J. Theor. Biol.* 205, 17–35.
- Ringach, D. L. (2007). On the origin of the functional architecture of the cortex. *PLoS ONE* 2:e251. doi: 10.1371/journal.pone.0000251
- Robinson, P. A., Rennie, C. J., and Wright, J. J. (1998). Synchronous oscillations in the cerebral cortex. *Phys. Rev. E* 57, 4578–4588.
- Robinson, P. A., Rennie, C. J., Wright, J. J., Bahramali, H., Gordon, E., and Rowe, D. L. (2001). Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E* 63, 701–702.
- Rockland, K. S., and Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *J. Comp. Neurol.* 216, 303–318.
- Rosenberg, A., Husson, T. R., and Issa, N. P. (2010). Subcortical representation of non-Fourier image features. *J. Neurosci.* 30, 1985–1993.
- Ruthazer, E. S., and Stryker, M. P. (1996). The role of activity in the development of long-range horizontal connections in area 17 of the ferret. *J. Neurosci.* 16, 7253–7269.
- Scholl, D. A. (1956). *The Organization of the Cerebral Cortex*. New York, NY: Wiley.
- Sharma, J., Angelucci, A., Rao, S. C., and Sur, M. (1995). *Relationship of Intrinsic Connections to Orientation Maps in Ferret Primary Visual Cortex: Iso-Orientation Domains and Singularities*. San Diego, CA: Presented at Society for Neuroscience.
- Sherk, H., and Stryker, M. P. (1976). Quantitative study of orientation selectivity in visually inexperienced kittens. *J. Neurophysiol.* 39, 63–70.
- Sherrington, C. S. (1906). *The Integrative Action of the Nervous System*. New Haven, CT: Yale University Press.
- Sherrington, C. S. (1940). *Man on His Nature. Gifford Lecture, Edinburgh, 1937–1938*. Cambridge, UK: Cambridge University Press.
- Shi, Y., Kirwan, P., Smith, J., Robinson, H. P. C., and Livesey, F. J. (2012). Human cerebral cortex development from pluripotent stem cells to functional cortical synapses. *Nat. Neurosci.* 15, 477–486.
- Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24, 49–65.
- Swindale, N. V. (1982). A model for the formation of orientation columns. *Proc. R. Soc. Lond. B Biol. Sci.* 215, 211–230.
- Swindale, N. V. (1992). A model for the coordinated development of columnar systems in primate striate cortex. *Biol. Cybern.* 66, 217–230.
- Swindale, N. V. (1996). The development of topography in the visual cortex: a review of models. *Network* 7, 161–247.
- Swindale, N. V. (2008). Feedback decoding of spatially structured population activity in cortical maps. *Neural Comput.* 20, 176–204.
- Tanaka, S. (1990). Theory of self-organization of cortical maps: mathematical framework. *Neural Netw.* 3, 625–640.
- Thomaidou, D., Mione, M. C., Cavanagh, J. F. R., and Parnavelas, J. G. (1997). Apoptosis and its relation to the cell cycle in the developing cerebral cortex. *J. Neurosci.* 17, 1075–1085.
- Tsukada, M., and Fukushima, Y. (2010). A context dependent mechanism in hippocampal CA1 networks. *Bull. Math. Biol.* 73, 417–435.
- van Ooyen, A. (2001). Competition in the development of nerve connections: a review of models. *Network* 12, R1–R47.
- van Ooyen, A., and Willshaw, D. J. (1999). Competition for neurotrophic factor in the development of nerve connections. *Proc. Biol. Sci.* 266, 883–892.
- van Rotterdam, A., Lopes da Silva, F. H., van den Ende, J., Viergever, M. A., and Hermans, A. J. (1982). A model of the spatio-temporal characteristics of the alpha rhythm. *Bull. Math. Biol.* 44, 283–305.
- von der Malsburg, C. (1973). Self organization of orientation sensitive cells in the striate cortex. *Kybernetik* 14, 85–100.
- Wiesel, T. N., and Hubel, D. H. (1974). Ordered arrangement of orientation columns in monkeys lacking visual experience. *J. Comp. Neurol.* 158, 307–318.
- Willshaw, D. J., and von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization. *Proc. R. Soc. Lond. B Biol. Sci.* 194, 431–435.
- Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80.
- Wolf, F., and Geisel, T. (1998). Spontaneous pinwheel annihilation during visual development. *Nature* 395, 73–78.
- Wright, J. J. (2009). Generation and control of cortical gamma: findings from simulation at two scales. *Neural Netw.* 22, 373–384.

- Wright, J. J. (2010). Attractor dynamics and thermodynamic analogies in the cerebral cortex: synchronous oscillation, the background EEG, and the regulation of attention. *Bull. Math. Biol.* 73, 436–457.
- Wright, J. J., Alexander, D. M., and Bourke, P. D. (2006). Contribution of lateral interactions in V1 to organization of response properties. *Vision Res.* 46, 2703–2720.
- Wright, J. J., and Bourke, P. D. (2008). An outline of functional self-organization in V1: synchrony, STLR and Hebb rules. *Cogn. Neurodyn.* 2, 147–157.
- Wright, J. J., and Bourke, P. D. (2013). A model for embryogenesis of cortical macrocolumns and superficial patchy connections: consequent neuronal responses at maturity. *Cereb. Cortex* (in press).
- Wright, J. J., Bourke, P. D., and Chapman, C. L. (2000). Synchronous oscillation in the cerebral cortex and object coherence: simulation of basic electrophysiological findings. *Biol. Cybern.* 83, 341–353.
- Wright, J. J., Rennie, C. J., Lees, G. J., Robinson, P. A., Bourke, P. D., Chapman, C. L., et al. (2003). Simulated electrocortical activity at microscopic, mesoscopic and global scales. *Neuropsychopharmacology* 28, S80–S93.
- Yousef, T., Tóth, É., Rausch, M., Eysel, U. T., and Kisvárdy, Z. F. (2001). Topography of orientation centre connections in the primary visual cortex of the cat. *Neuroreport* 12, 1693–1699.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 12 October 2012; accepted: 24 January 2013; published online: 15 February 2013.
- Citation: Wright JJ and Bourke PD (2013) On the dynamics of cortical development: synchrony and synaptic self-organization. *Front. Comput. Neurosci.* 7:4. doi: 10.3389/fncom.2013.00004
- Copyright © 2013 Wright and Bourke. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



How adaptation shapes spike rate oscillations in recurrent neuronal networks

Moritz Augustin^{1,2*†}, Josef Ladenbauer^{1,2*†} and Klaus Obermayer^{1,2}

¹ Department of Software Engineering and Theoretical Computer Science, Technische Universität Berlin, Berlin, Germany

² Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany

Edited by:

Peter Robinson, The University of Sydney, Australia

Reviewed by:

Alessandro Treves, Scuola Internazionale Superiore di Studi Avanzati, Italy

Peter Robinson, The University of Sydney, Australia

*Correspondence:

Moritz Augustin and Josef Ladenbauer, Department of Software Engineering and Theoretical Computer Science, Technische Universität Berlin, Neural Information Processing Group, Marchstr. 23, MAR 5-6, 10587 Berlin, Germany.
e-mail: augustin@ni.tu-berlin.de; jl@ni.tu-berlin.de

[†] These authors have contributed equally to this work.

Neural mass signals from *in-vivo* recordings often show oscillations with frequencies ranging from <1 to 100 Hz. Fast rhythmic activity in the beta and gamma range can be generated by network-based mechanisms such as recurrent synaptic excitation-inhibition loops. Slower oscillations might instead depend on neuronal adaptation currents whose timescales range from tens of milliseconds to seconds. Here we investigate how the dynamics of such adaptation currents contribute to spike rate oscillations and resonance properties in recurrent networks of excitatory and inhibitory neurons. Based on a network of sparsely coupled spiking model neurons with two types of adaptation current and conductance-based synapses with heterogeneous strengths and delays we use a mean-field approach to analyze oscillatory network activity. For constant external input, we find that spike-triggered adaptation currents provide a mechanism to generate slow oscillations over a wide range of adaptation timescales as long as recurrent synaptic excitation is sufficiently strong. Faster rhythms occur when recurrent inhibition is slower than excitation and oscillation frequency increases with the strength of inhibition. Adaptation facilitates such network-based oscillations for fast synaptic inhibition and leads to decreased frequencies. For oscillatory external input, adaptation currents amplify a narrow band of frequencies and cause phase advances for low frequencies in addition to phase delays at higher frequencies. Our results therefore identify the different key roles of neuronal adaptation dynamics for rhythmogenesis and selective signal propagation in recurrent networks.

Keywords: spike frequency adaptation, adaptation, oscillations, rate models, network dynamics, Fokker-Planck, mean-field, recurrent network

INTRODUCTION

A prominent characteristic of cortical activity is its rhythmicity as shown by electroencephalography or the local field potential. Dominant oscillation frequencies in these signals range from <1 to 100 Hz and reflect synchronous activity of populations of neurons. Such oscillations are linked to behavioral states (Wang, 2010) and involved in a variety of cognitive functions (Engel et al., 2001; Fries, 2001; Melloni et al., 2007; Ghazanfar et al., 2008; Wang, 2010) as well as pathological conditions (Hammond et al., 2007; Zijlmans et al., 2009; Uhlhaas and Singer, 2010). It is therefore important to understand the mechanisms of oscillations in neuronal networks, how they are initiated and terminated, and how their frequency is determined.

Fast rhythmic activity in the beta and gamma band (>20 Hz) can be generated by network-based mechanisms, such as synaptic excitation-inhibition loops or by feedback inhibition alone (Isaacson and Scanziani, 2011). In these scenarios the oscillation frequency is largely determined by the inhibitory decay time constant (Brunel and Wang, 2003; Tiesinga and Sejnowski, 2009). Low-frequency oscillations, on the other hand, could depend on slow transmembrane outward currents (Compte et al., 2003; Gigante et al., 2007b; Destexhe, 2009), which are mediated by low-threshold voltage-dependent muscarinic (M) and high-threshold calcium-gated afterhyperpolarization (AHP) K⁺

channels, respectively (Brown and Adams, 1980; Connors et al., 1982; Stocker, 2004). These currents cause spike frequency adaptation and are typically more pronounced in cortical regular spiking pyramidal (excitatory) neurons compared to fast spiking (inhibitory) interneurons (La Camera et al., 2006). Both, the M and AHP type K⁺ currents, are susceptible to cholinergic modulation (McCormick, 1992). Their kinetic time constants range from milliseconds to seconds (Abel et al., 2004; Manuel et al., 2005) and can be pharmacologically manipulated (Pedarzani et al., 2001).

Here we study the interplay of the dynamics of such adaptation currents with synaptic excitation and inhibition in recurrent networks of excitatory and inhibitory neurons. Specifically, we ask (1) how adaptation can generate slow oscillations, (2) how it modulates faster rhythms based on synaptic interaction, and (3) how adaptation affects resonance properties of the network.

In-vivo recordings from behaving animals have revealed that even when the population activity oscillates, the spike trains of the constituent neurons are rather irregular and display Poisson-like characteristics (Fries, 2001; Wang, 2010). This stochasticity in neuronal responses allows us to derive a mean-field model from a recurrent network of adaptive spiking model neurons coupled through conductance-based synapses with heterogeneous strengths and delays. Our approach is based on the

Fokker–Planck (FP) formalism (Brunel, 2000; Deco et al., 2008) and efficiently describes the activity of large networks where the features of the spiking neurons (i.e., the model parameters) are retained. Using this method we analyze network responses to constant as well as rhythmic external input. In particular we describe asynchronous irregular states with constant steady-state activity as well as oscillatory states and their properties. We validate our mean-field results qualitatively by large-scale network simulations.

METHODS

We first describe our network model containing two populations (excitatory and inhibitory) of adaptive spiking neurons with delayed conductance-based synaptic coupling. Based on that model we then derive mean-field model equations and solve them numerically to obtain distributions of the membrane potentials and instantaneous spike rates.

NETWORK MODEL

We consider a network of $N = N_{\mathcal{E}} + N_{\mathcal{I}}$ adaptive exponential integrate-and-fire neurons (aEIF) proposed by Brette and Gerstner (2005), where $N_{\mathcal{E}}$ and $N_{\mathcal{I}}$ are the numbers of excitatory and inhibitory neurons, respectively. The dynamics of the i -th neuron of population $\alpha \in \{\mathcal{E}, \mathcal{I}\}$ is described by

$$C \frac{dV_i^\alpha}{dt} = I_{\text{ion}}(V_i^\alpha) - w_i^\alpha + I_{\text{syn},i}^\alpha(V_i^\alpha, t) \quad (1)$$

$$\tau_w \frac{dw_i^\alpha}{dt} = a(V_i^\alpha - E_L) - w_i^\alpha \quad (2)$$

with reset condition

$$\text{if } V_i^\alpha > V_{\text{cut}} \text{ then } \begin{cases} V_i^\alpha := V_r \\ w_i^\alpha := w_i^\alpha + b. \end{cases} \quad (3)$$

The first Equation (1) is for the membrane potential V_i^α , where the capacitive current through the membrane with capacitance C equals the sum of ionic currents I_{ion} , the adaptation current w_i^α and the synaptic current $I_{\text{syn},i}^\alpha$. The ionic currents are given by

$$I_{\text{ion}}(V) := g_L(E_L - V) + g_L \Delta_T e^{\frac{V - V_T}{\Delta_T}}, \quad (4)$$

where the first term on the right-hand side describes an Ohmic leak current with conductance g_L and reversal potential E_L . The exponential term with threshold slope factor Δ_T and threshold potential V_T approximates the Na^+ -current which is responsible for the generation of spikes, assuming that the activation of Na^+ -channels is instantaneous and neglecting their inactivation (Fourcaud-Trocme et al., 2003). Equation (2) governs the dynamics of the adaptation current w_i^α , where τ_w denotes the adaptation time constant and a quantifies a conductance that mediates sub-threshold adaptation. A spike is said to occur at the time when V_i^α diverges to infinity, but in practice a finite “cutoff” value V_{cut} is chosen. When V_i^α crosses V_{cut} from below, V_i^α is set to the reset potential V_r and w_i^α is incremented by b , cf. condition (3). In this way spike-triggered adaptation is included in the

model. Immediately after the reset, V_i^α and w_i^α are clamped for a refractory period T_{ref} .

The aEIF model has been shown to reproduce a broad range of subthreshold dynamics (Touboul and Brette, 2008) and spike patterns of cortical neurons (Naud et al., 2008) and can well predict their spike times (Jolivet et al., 2008) and post-stimulus time histograms (Pospischil et al., 2011). Importantly, the subthreshold and spike-triggered adaptation components of this model have been shown to capture the effects of the M and AHP currents in a detailed biophysical neuron model, respectively (Ladenbauer et al., 2012).

Neuron i of population α receives total synaptic current

$$I_{\text{syn},i}^\alpha(V_i^\alpha, t) := \sum_j I_{ij}^{\alpha,\text{ext}} + \sum_j I_{ij}^{\alpha,\mathcal{E}} + \sum_j I_{ij}^{\alpha,\mathcal{I}}, \quad (5)$$

which is the superposition of synaptic inputs $I_{ij}^{\alpha,\text{ext}}$ from K_{ext} external excitatory neurons, $I_{ij}^{\alpha,\mathcal{E}}$ from $K_{\mathcal{E}}$ excitatory neurons of the network and $I_{ij}^{\alpha,\mathcal{I}}$ from $K_{\mathcal{I}}$ inhibitory neurons of the network. j is the index of the respective presynaptic neuron. The synaptic current $I_{ij}^{\alpha,\gamma}$ caused by neuron j of population $\gamma \in \{\text{ext}, \mathcal{E}, \mathcal{I}\}$ is modeled using delta functions,

$$I_{ij}^{\alpha,\text{ext}}(V_i^\alpha, t) := C J_{ij}^{\alpha,\text{ext}} \sum_k \delta(t - t_j^k) (E_{\mathcal{E}} - V_i^\alpha) \quad (6)$$

$$I_{ij}^{\alpha,\beta}(V_i^\alpha, t) := C J_{ij}^{\alpha,\beta} \sum_k \delta(t - t_j^k - d_{ij}^{\alpha,\beta}) (E_\beta - V_i^\alpha), \quad (7)$$

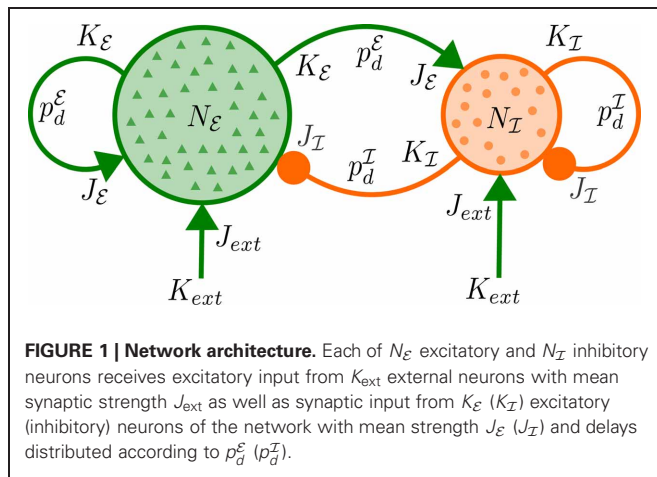
where $\beta \in \{\mathcal{E}, \mathcal{I}\}$ denotes the presynaptic population. $J_{ij}^{\alpha,\gamma}$ are dimensionless synaptic efficacies drawn from a Gaussian distribution with mean $J_{\alpha,\gamma}$ and standard deviation $\Delta J_{\alpha,\gamma}$. Here we consider that $J_{\alpha,\gamma} \equiv J_\gamma$ and $\Delta J_{\alpha,\gamma} \equiv \Delta J_\gamma$ depend only on the presynaptic population γ . t_j^k is the k -th spike time of neuron j from the respective population. $E_{\mathcal{E}}$ and $E_{\mathcal{I}}$ denote the excitatory and inhibitory reversal potentials, respectively. $d_{ij}^{\alpha,\beta}$ is the synaptic delay, sampled using a bi-exponential probability density

$$p_d^{\alpha,\beta}(d) := \frac{1}{\tau_d - \tau_r} \left(e^{-\frac{d-d_0}{\tau_d}} - e^{-\frac{d-d_0}{\tau_r}} \right) \quad (8)$$

for positive delays d , where d_0 is the minimal delay and τ_r , τ_d are the rise and decay time constants, for each pair of populations. In the model we use two different delay distributions $p_d^{\mathcal{E}}$ and $p_d^{\mathcal{I}}$ which do not depend on the postsynaptic population as for the synaptic weights. For a schematic diagram of the network, see **Figure 1**.

We assume the neurons from the external population generate spike times according to Poisson processes with rates $r_{\text{ext}}^\alpha(t)$. The spike rate of each population $\alpha \in \{\mathcal{E}, \mathcal{I}\}$ at time t is given by the average number of spikes of neurons from the corresponding population in the interval $[t, t + \Delta t]$,

$$r_\alpha^{\Delta t}(t) := \frac{1}{N_\alpha \Delta t} \sum_{j=1}^{N_\alpha} \int_t^{t+\Delta t} \sum_k \delta(s - t_j^k) ds. \quad (9)$$



In the mean-field limit $N \rightarrow \infty$, $\Delta t \rightarrow 0$ we obtain a continuous population spike rate $r_\alpha(t)$ (see below).

We selected the following parameters for the neuron model: $C = 200$ pF, $g_L = 10$ nS, $E_L = -70$ mV, $\Delta_T = 1$ mV, $V_T = -50$ mV, $V_r = -70$ mV, $V_{cut} = -40$ mV, and $T_{ref} = 1.4$ ms (Badel et al., 2008; Destexhe, 2009). For excitatory neurons the adaptation parameters were varied within reasonable ranges: $\tau_w \in [5, 1000]$ ms, $a \in [0, 10]$ nS, $b \in [0, 50]$ pA. For inhibitory neurons adaptation was neglected ($a = b = 0$) since it was found to be weak in fast spiking interneurons compared to pyramidal cells (La Camera et al., 2006).

The network parameter values were $N_E = 40,000$, $N_I = 10,000$, $K_{ext} = 1600$, $K_E = 1600$, $K_I = 400$, $E_E = 0$ mV, $E_I = -80$ mV, $J_{ext} = 0.003$, $J_E = 0.003$, and $\Delta J_\gamma = 0.1 J_\gamma$ with $\gamma \in \{ext, E, I\}$ (Brunel and Wang, 2003). To adjust the balance of recurrent synaptic excitation and inhibition we introduce the parameter

$$g := \frac{J_I |E_L - E_I|}{J_E |E_L - E_E|}, \quad (10)$$

which is the ratio of total charges induced at rest (Kumar et al., 2008). g determines J_I and thus ΔJ_I for fixed J_E and was varied in $[0.8, 4]$ which yields a physiological range of inhibitory postsynaptic potential amplitudes (Tamas et al., 1997). Note that the value of g that corresponds to balanced mean recurrent excitatory and inhibitory synaptic currents depends on the mean membrane potential for each population. The effect of a spike of presynaptic neuron j on neuron i is mediated by a delayed instantaneous increment or decrement of the postsynaptic membrane potential, cf. Equations (1), (5), and (7). This implies that $d_{ij}^{\alpha,\beta}$ reflects the conduction delay as well as delays in the synaptic kinetics. We therefore chose the parameter values of p_d^E and p_d^I such that conduction delays as well as typical time courses of excitatory AMPA and inhibitory GABA_A synaptic receptors are taken into account. The values we selected were $d_0 = 1$ ms, $\tau_r^E \in [1.25, 1.5]$ ms, $\tau_d^E \in [1.5, 2]$ ms, $\tau_r^I \in [0.55, 1.25]$ ms, and $\tau_d^I \in [1.5, 5]$ ms. The input rate of the excitatory population r_{ext}^E was varied in $[1, 12.5]$ Hz. r_{ext}^I was chosen

such that $r_E = r_I$ in case of uncoupled populations of neurons, i.e., $J_E = J_I = 0$.

MEAN-FIELD MODEL

We reduce the two-population network of aEIF neurons to the mean-field model in three steps. First, we replace the synaptic current fluctuations by a Gaussian white noise process via the diffusion approximation. Next, we take a mean-field limit to formulate the stochastic network model in terms of two coupled deterministic scalar partial differential equations (PDE). Finally, to allow for efficient numerical computation we reduce the number of variables in these equations using an adiabatic approximation.

Diffusion approximation

We approximate the total synaptic current $I_{syn,i}^\alpha$ of Equation (5) by its mean plus a fluctuating Gaussian part, which is justified by the following physiologically plausible assumptions: (1) The number of synaptic inputs to a neuron is large, i.e., $K_{ext}, K_E, K_I \gg 1$ (Destexhe et al., 2003) and (2) the postsynaptic potential amplitudes elicited by individual presynaptic spikes are small, i.e., $J_{ext}|E_E - V|$, $J_E|E_E - V|$, $J_I|E_I - V| \ll V_{cut} - V_r$ (Williams and Stuart, 2002). We further assume that (3) the network connectivity is random and sparse, i.e., $K_E, K_I \ll N$, and that (4) presynaptic spike times are represented by Poisson processes which are homogeneous in each small time interval. The total synaptic current can then be written as (Brunel, 2000; Nykamp and Tranchina, 2000; Renart et al., 2004; Richardson, 2004; Gigante et al., 2007b)

$$I_{syn,i}^\alpha \approx \mu_{\alpha,i}(V_i^\alpha, t) + \sigma_{\alpha,i}(V_i^\alpha, t)\eta_i(t), \quad (11)$$

where $\mu_{\alpha,i}$ and $\sigma_{\alpha,i}$ are the infinitesimal mean and standard deviation of $I_{syn,i}^\alpha$, respectively, and η_i is a Gaussian white noise process with δ -autocorrelation. The infinitesimal mean is given by

$$\begin{aligned} \mu_{\alpha,i} &:= \lim_{\Delta t \rightarrow 0} \frac{\left\langle \int_t^{t+\Delta t} I_{syn,i}^\alpha(s) ds \right\rangle}{\Delta t} \\ &= \mu_{\alpha,i}^{ext} + \mu_{\alpha,i}^E + \mu_{\alpha,i}^I \end{aligned} \quad (12)$$

with

$$\mu_{\alpha,i}^{ext} = C(E_E - V_i^\alpha)J_{ext}K_{ext}r_{ext}^\alpha(t) \quad (13)$$

$$\mu_{\alpha,i}^\beta = C(E_\beta - V_i^\alpha)J_\beta K_\beta(r_\beta * p_\beta)(t), \quad (14)$$

where $\langle \cdot \rangle$ denotes the expectation operator. The infinitesimal variance is

$$\begin{aligned} \sigma_{\alpha,i}^2 &:= \lim_{\Delta t \rightarrow 0} \frac{\left\langle \left(\int_t^{t+\Delta t} I_{syn,i}^\alpha(s) ds \right)^2 \right\rangle}{\Delta t} + O(\Delta t^2) \\ &= (\sigma_{\alpha,i}^{ext})^2 + (\sigma_{\alpha,i}^E)^2 + (\sigma_{\alpha,i}^I)^2 \end{aligned} \quad (15)$$

with

$$\sigma_{\alpha,i}^{\text{ext}} = C(E_{\mathcal{E}} - V_i^{\alpha}) \sqrt{(J_{\text{ext}}^2 + \Delta J_{\text{ext}}^2)} K_{\text{ext}} r_{\text{ext}}(t) \quad (16)$$

$$\sigma_{\alpha,i}^{\beta} = C(E_{\beta} - V_i^{\alpha}) \sqrt{(J_{\beta}^2 + \Delta J_{\beta}^2)} K_{\beta} (r_{\beta} * p_{\beta})(t), \quad (17)$$

where $\beta \in \{\mathcal{E}, \mathcal{I}\}$ and $*$ denotes convolution. In Equations (13), (14), (16), and (17) we have used that the presynaptic Poisson processes, the synaptic weights and delays are independent.

Mean-field limit

We analyze networks of sparsely coupled neurons, i.e., the probability for a connection between any pair of neurons is low, cf. assumption (3) above. For large N correlations between the fluctuations of synaptic currents of different neurons become negligible, i.e., $\langle \eta_i(t) \eta_j(t) \rangle = 0$ for $i \neq j$. In the mean-field limit $N \rightarrow \infty$ the network model Equations (1)–(4), Equations (11)–(17) can be described by two FP equations—one for each population α —which are delay-coupled by the population spike rates $r_{\mathcal{E}}$ and $r_{\mathcal{I}}$,

$$\frac{\partial p_{\alpha}}{\partial t} + \frac{\partial S_{\alpha}^V}{\partial V} + \frac{\partial S_{\alpha}^w}{\partial w} = 0 \quad (18)$$

with

$$S_{\alpha}^V := \left(\frac{I_{\text{ion}}(V) - w + \mu_{\alpha}}{C} - \frac{\sigma_{\alpha}}{2C^2} \frac{\partial \sigma_{\alpha}}{\partial V} \right) p_{\alpha} \quad (19)$$

$$S_{\alpha}^w := \frac{a(V - E_L) - w}{\tau_w} p_{\alpha}. \quad (20)$$

$p_{\alpha}(V, w, t)$ is the probability density to find a neuron of population α in the state (V, w) at time t . $S_{\alpha}^V(V, w, t)$ and $S_{\alpha}^w(V, w, t)$ are the probability fluxes in positive V and w direction, respectively. Note that we used the Stratonovich interpretation of the underlying stochastic equations (Risken, 1996; Richardson, 2004). To account for the reset condition (3) the flux through the cutoff voltage V_{cut} at w is re-injected after the refractory period T_{ref} at $V_r, w + b$, i.e.,

$$\lim_{V \downarrow V_r} S_{\alpha}^V(V, w + b, t) - \lim_{V \uparrow V_r} S_{\alpha}^V(V, w + b, t) = S_{\alpha}^V(V_{\text{cut}}, w, t - T_{\text{ref}}) \quad \forall w \in \mathbb{R}. \quad (21)$$

This implies that in general p_{α} is not differentiable at the line $V = V_r$. The boundary conditions are reflecting for $w \rightarrow \pm\infty$, $V \rightarrow -\infty$ and absorbing for $V = V_{\text{cut}}$,

$$\lim_{w \rightarrow \pm\infty} S_{\alpha}^w(V, w) = 0 \quad \forall V \in (-\infty, V_{\text{cut}}] \quad (22)$$

$$\lim_{V \rightarrow -\infty} S_{\alpha}^V(V, w) = 0 \quad \forall w \in \mathbb{R} \quad (23)$$

$$p_{\alpha}(V_{\text{cut}}, w) = 0 \quad \forall w \in \mathbb{R} \quad (24)$$

The spike rate of population α is given by the integral of the cutoff fluxes,

$$r_{\alpha}(t) = \int_{\mathbb{R}} S_{\alpha}^V(V_{\text{cut}}, w, t) dw. \quad (25)$$

At any timepoint t the histogram of the membrane potentials of neurons in population α can be seen as a sample drawn from the probability density $p_{\alpha}(V, t)$ which is governed by the FP equation.

Adiabatic approximation

Solving the $2 + 1$ dimensional PDE (Equations 18–20) with corresponding reset and boundary conditions (21)–(24) numerically is possible but computationally demanding. We therefore reduce the dimensionality of the FP system Equations (18)–(20) assuming the timescales of membrane voltage and adaptation current dynamics are separable. This is justified by the observation that the dynamics of neuronal adaptation is significantly slower than the other in the model system such as membrane time constant and average inter-spike interval (Womble and Moises, 1992; Stocker, 2004). Under this assumption, the adaptation current of each neuron can be seen as an efficient integrator that filters the fluctuations in the neuronal activity. We approximate $w_i^{\alpha}(t)$ in Equation (2) by its population average $w_{\alpha}(t)$, which evolves according to

$$\tau_w \frac{dw_{\alpha}}{dt} = a(\langle V \rangle_{p_{\alpha}(V, t)} - E_L) - w_{\alpha} + \tau_w b r_{\alpha}(t), \quad (26)$$

where $\langle \cdot \rangle_p$ denotes the average over the density p (Brunel et al., 2003; Gigante et al., 2007b). The probability density $p_{\alpha}(V, t)$ then satisfies the $1 + 1$ dimensional FP equation

$$\frac{\partial p_{\alpha}}{\partial t} + \frac{\partial S_{\alpha}^V}{\partial V} = 0, \quad (27)$$

where again S_{α}^V is the probability flux defined in Equation (19) and $w := w_{\alpha}(t)$ appears as a system parameter. The reset condition is

$$\lim_{V \downarrow V_r} S_{\alpha}^V(V, t) - \lim_{V \uparrow V_r} S_{\alpha}^V(V, t) = S_{\alpha}^V(V_{\text{cut}}, t - T_{\text{ref}}). \quad (28)$$

and the boundary conditions (23)–(24) become

$$\lim_{V \rightarrow -\infty} S_{\alpha}^V(V) = 0, \quad (29)$$

$$p_{\alpha}(V_{\text{cut}}) = 0. \quad (30)$$

The population spike rates are given by the corresponding fluxes through the cutoff voltage,

$$r_{\alpha}(t) = S_{\alpha}^V(V_{\text{cut}}, t). \quad (31)$$

Note that the adiabatic approximation described above could be applied repeatedly for additional slow variables.

NUMERICAL SOLUTION

We solved the reduced FP Equation (27) subject to conditions (28)–(30) and mean adaptation current dynamics (Equation 26) forward in time until either steady states $r_{\mathcal{E}}^{\infty}, r_{\mathcal{I}}^{\infty}$ with $r_{\alpha}^{\infty} := \lim_{t \rightarrow \infty} r_{\alpha}(t)$ or stable oscillatory states were reached. The probability densities $p_{\mathcal{E}}, p_{\mathcal{I}}$ were initialized using normalized Gaussians with mean $0.5 \cdot (V_r + V_T)$ and standard deviation $0.2 \cdot (V_T - V_r)$. We applied a first-order finite volume method on a finite and non-uniform grid $V_0 < V_1 < \dots < V_{N_V}$ using upwind-fluxes to stabilize the numerical solution (LeVeque, 2002). Time was discretized using the implicit Euler method on an equidistant grid, i.e., $t_{n+1} - t_n \equiv \Delta t$. The resulting linear equation systems were solved with a preconditioned Krylov subspace method in each time step. Specifically, BiCGSTAB (van der Vorst, 1992) was used in combination with an incomplete LU decomposition preconditioner (Saad, 2003) that strongly improved the convergence speed.

$w_{\mathcal{E}}$ was initialized with values $w_{\mathcal{E}}(0) \in [0, 500]$ pA (and $w_{\mathcal{I}} \equiv 0$). The other parameters were $\Delta t = 50 \mu\text{s}$, $\min_m \Delta V_m = 1 \mu\text{V}$ with $\Delta V_m := V_{m+1} - V_m$, $V_0 := -100 \mu\text{V}$, $V_{N_V} = V_{\text{cut}}$ and $N_V = 256$.

We complemented the mean-field results with numerical simulations of the network model Equations (1)–(4) using a Runge-Kutta second order method implemented in Brian 1.4 (Goodman and Brette, 2009) with a time step of $50 \mu\text{s}$.

In case of stable periodic population spike rates the oscillation frequency was determined by the dominant frequency of the Fourier spectrum of $r_{\mathcal{E}}$ over the last 2 s of runtime.

RESULTS

ADAPTATION MEDIATES OSCILLATIONS

To examine how the interplay of adaptation and recurrent synaptic input shapes network dynamics we vary the type, strength and

timescale (parameters a , b , and τ_w) of adaptation for excitatory neurons as well as the strength of synaptic inhibition (parameter g) across networks. Adaptation currents are disregarded for inhibitory neurons, which is supported by experimental observations, see the section Methods. We consider constant rates $r_{\text{ext}}^{\mathcal{E}}, r_{\text{ext}}^{\mathcal{I}}$ for the external Poisson-inputs and identical delay distributions $p_d^{\mathcal{E}} \equiv p_d^{\mathcal{I}}$. First, we examine steady-state spike rates, oscillation amplitudes and frequencies for networks with different values of spike-triggered adaptation b and inhibition strength g , see **Figure 2A**. All networks without adaptation ($a = b = 0$) settle into asynchronous states with constant population rates that decrease with increasing g . For networks with increased b slow oscillatory states become stable if recurrent excitation is sufficiently strong. The larger b is, the less recurrent excitation is necessary for sustained oscillations. Amplitude and period of the oscillatory rate decrease with an increase of b and g , respectively. Thus, in networks where recurrent synaptic excitation dominates inhibition at least slightly, spike-triggered adaptation b generates spike rate oscillations. The dynamics of an example network is shown in **Figure 2B**. The evolution of the population spike rates $r_{\mathcal{E}}, r_{\mathcal{I}}$, membrane potential probability densities $p_{\mathcal{E}}, p_{\mathcal{I}}$ and adaptation current $w_{\mathcal{E}}$ display periodic bursts of population activity. As a validation of the findings above using the mean-field model the activity of simulated large networks of spiking neurons is shown in **Figure 2C**. The raster plots reveal population bursts when b is increased and g is small. An asynchronous state with low population activity occurs if g is increased. If in addition adaptation is removed ($a = b = 0$) the network settles into an asynchronous state with increased spike rates.

The mechanism that generates these oscillations is a loop of recurrent excitation, build up and decay of adaptation current as indicated in **Figure 2B**. A low level of population activity is initiated by the external input $r_{\text{ext}}^{\mathcal{E}}$ and recurrent synaptic excitation

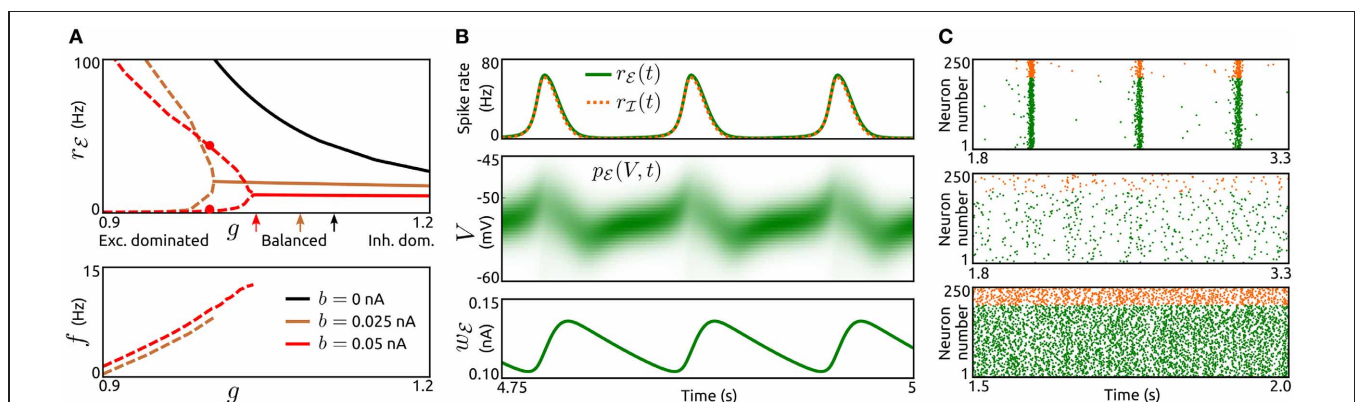


FIGURE 2 | Population bursts caused by spike-triggered adaptation.

(A) Top: Spike rate $r_{\mathcal{E}}$ of the excitatory population as a function of the strength of inhibition g for networks without spike-triggered adaptation ($b = 0$, black) and with increased levels of b (0.025 nA, brown and 0.05 nA, red). In case of stable oscillatory states the maxima and minima of the periodic $r_{\mathcal{E}}$ are shown by dashed lines. Solid lines represent asynchronous states. Arrows indicate balance of recurrent excitation and inhibition for both populations. Bottom: Corresponding oscillation frequencies f . $\tau_w = 200$ ms, $a = 0$, and $r_{\text{ext}}^{\mathcal{E}} = 6.25$ Hz. The parameter values for both delay distributions $p_d^{\mathcal{E}}, p_d^{\mathcal{I}}$ were $\tau_r = 1.5$ ms and $\tau_d = 2$ ms. For other model

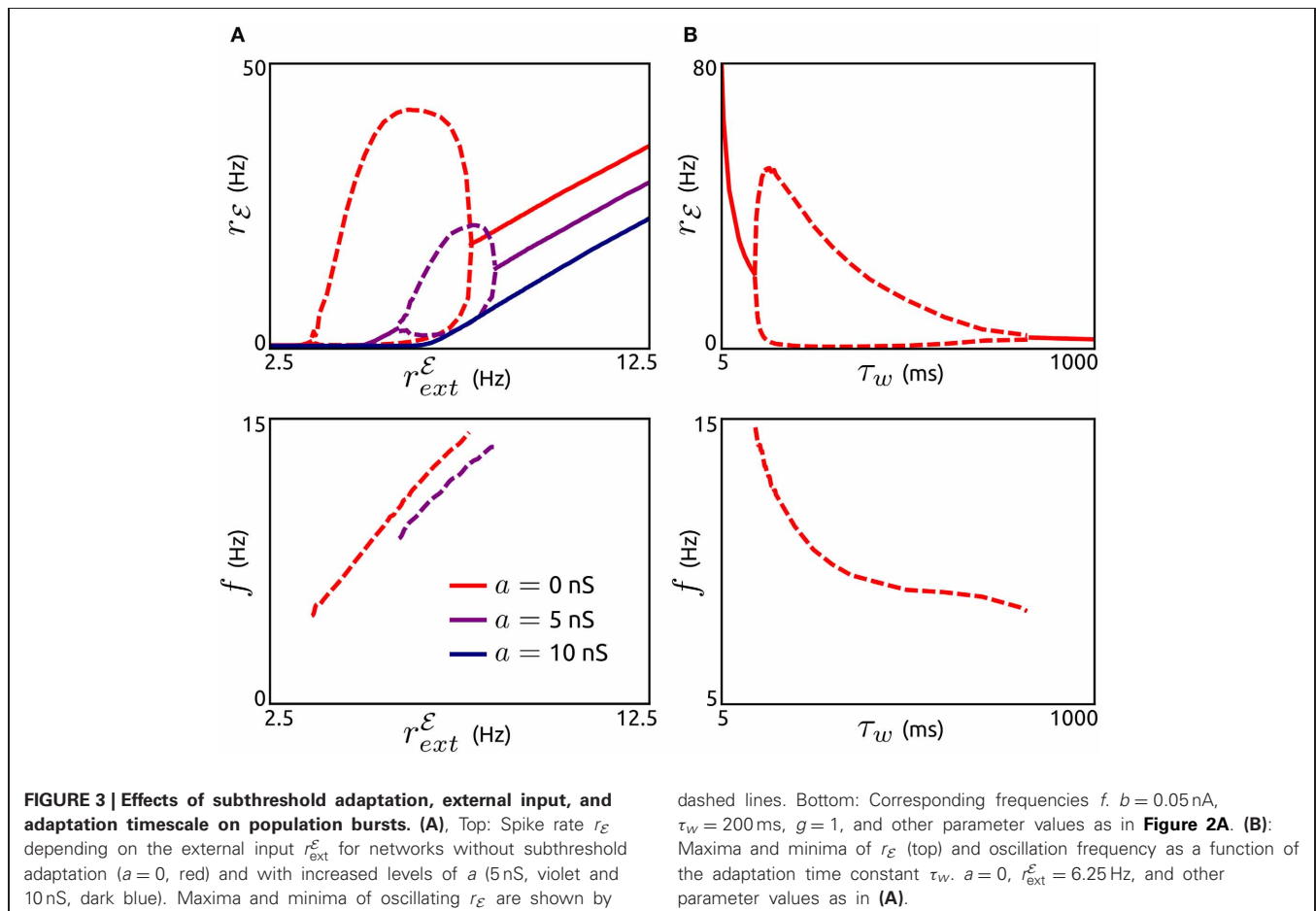
parameters see the section Methods. **(B)**, Top: Time-dependent spike rates $r_{\mathcal{E}}(t)$ (green) and $r_{\mathcal{I}}(t)$ (orange, dashed) for the parameter values $b = 0.05$ nA and $g = 1$, as indicated in **(A)** by red dots. Center: Corresponding membrane potential density $p_{\mathcal{E}}(V, t)$. Bottom: Corresponding mean adaptation current $w_{\mathcal{E}}(t)$. **(C)**: Raster plots of simulated networks of $N = 50,000$ aEIF neurons for $b = 0.05$ nA, $g = 0.85$ (top), $b = 0.05$ nA, $g = 1.05$ (center) and $b = 0$, $g = 1$ (bottom). The spike times of 200 excitatory neurons and 50 inhibitory neurons, all randomly selected, are shown by green and orange dots, respectively. $\tau_w = 200$ ms, $a = 0$, and $r_{\text{ext}}^{\mathcal{E}} = 3.75$ Hz. Other parameter values as in **(A)**.

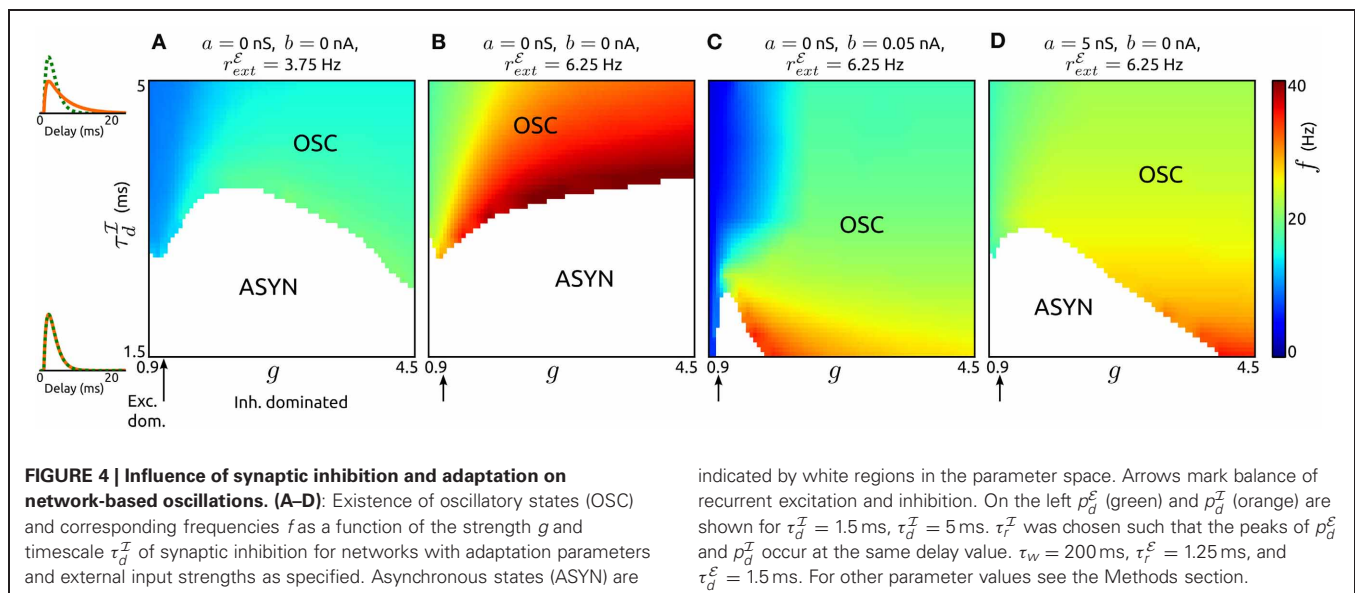
boosts the activity, thereby increasing the adaptation current $w_{\mathcal{E}}$ through b in a spike rate dependent way. The adaptation current in turn acts as a negative feedback which eventually outweighs the recurrent excitation. The population activity drops rapidly and the adaptation current decays slowly. Upon recovery from the adaptation current the cycle starts again.

Next, we investigate how these oscillations are affected by the external input $r_{\text{ext}}^{\mathcal{E}}$, the subthreshold adaptation conductance a and the adaptation timescale τ_w , see **Figure 3**. The existence of adaptation-induced oscillations is quite sensitive to the level of $r_{\text{ext}}^{\mathcal{E}}$ (**Figure 3A**). Periodic activity is stable for small values of $r_{\text{ext}}^{\mathcal{E}}$ (above threshold). While oscillation frequencies increase monotonically with increasing $r_{\text{ext}}^{\mathcal{E}}$, oscillation amplitudes increase initially for a small interval of $r_{\text{ext}}^{\mathcal{E}}$ values and decrease over the following interval. For larger values of $r_{\text{ext}}^{\mathcal{E}}$ oscillatory activity is destabilized and asynchronous states occur. Interestingly, an increase in a does not lead to oscillations. On the contrary, periodic population bursts are destabilized by a . The dependence of oscillation amplitude and frequency on τ_w is shown in **Figure 3B**. Stable oscillations exist for a large range of values of τ_w , where the frequencies decrease with increasing τ_w . Oscillations are unstable for small adaptation timescales in the range of the membrane time constant and for very large values of τ_w .

ADAPTATION MODULATES FREQUENCIES OF NETWORK-BASED OSCILLATIONS

Here we study the influence of adaptation on oscillations generated by recurrent synaptic excitation-inhibition (\mathcal{E} - \mathcal{I}) loops. The pace of such oscillations is believed to be largely determined by the decay of inhibition. To describe their dependence on the timescale of inhibition for various recurrent network regimes (from excitation dominated to inhibition dominated) we first consider networks of neurons without an adaptation current ($a = b = 0$), see **Figures 4A,B**. By varying the decay $\tau_d^{\mathcal{I}}$ of inhibition and its strength (by parameter g) across networks we find that stable oscillatory states occur if inhibition is sufficiently slow in comparison to excitation. The oscillation frequencies increase with increasing external input spike rate $r_{\text{ext}}^{\mathcal{E}}$, increasing g and decreasing $\tau_d^{\mathcal{I}}$, respectively. A low value of $r_{\text{ext}}^{\mathcal{E}}$ leads to frequencies in the low beta band (**Figure 4A**), for a higher value of $r_{\text{ext}}^{\mathcal{E}}$ the frequencies span the beta and low gamma bands (**Figure 4B**). Note that the network parameters can be adjusted to obtain higher oscillation frequencies. The generating mechanism underlying the oscillations is a loop of recurrent synaptic excitation and inhibition, initiated by the excitatory external input. We verified this by removing the recurrent excitatory input to the inhibitory population, which lead to a destabilization of the oscillations. For larger values of g as the ones used in **Figure 4**, the





\mathcal{E} - \mathcal{I} -loop mechanism is replaced by an \mathcal{I} - \mathcal{I} -loop that does not depend on recurrent excitation (not shown). Since adaptation is only exhibited by excitatory neurons, we disregard the parameter space where \mathcal{I} - \mathcal{I} -loop-based rhythmic activity occurs and focus on \mathcal{E} - \mathcal{I} -loop-based oscillations instead.

An increase of spike-triggered adaptation or subthreshold current stabilizes oscillatory states also for faster recurrent inhibition, see **Figures 4C,D**. This change in single neuron dynamics causes oscillations in large parts of explored (g , τ_d^I)-space. In particular, for spike-triggered adaptation asynchronous states only occur in a small region of the parameter space. Interestingly, the oscillation frequencies are significantly reduced by either type of adaptation.

Next, we investigate how the timescale of adaptation τ_w affects oscillations mediated by an \mathcal{E} - \mathcal{I} -loop. In **Figure 5A** we show the dependence of amplitude and frequency of such oscillations on τ_w for networks with both adaptation components increased ($a = 5$ nS, $b = 0.05$ nA) and either dominant recurrent excitation ($g = 1.05$) or inhibition ($g = 1.5$). In both cases, stable oscillatory states exist for a large range of time constants. As τ_w increases the oscillation frequencies decrease while the amplitudes first increase abruptly and then decrease. The networks settle into asynchronous states for small τ_w (in the order of the membrane time constant) or large τ_w (several hundreds of milliseconds). Note that these effects of τ_w are similar if either a or b is increased individually (not shown). We validated these effects by simulations of aEIF neuron networks, see **Figure 5B**. The raster plots show that an increase in τ_w leads to a decrease in oscillation frequency and amplitude.

ADAPTATION PROMOTES PERIODIC SIGNAL PROPAGATION

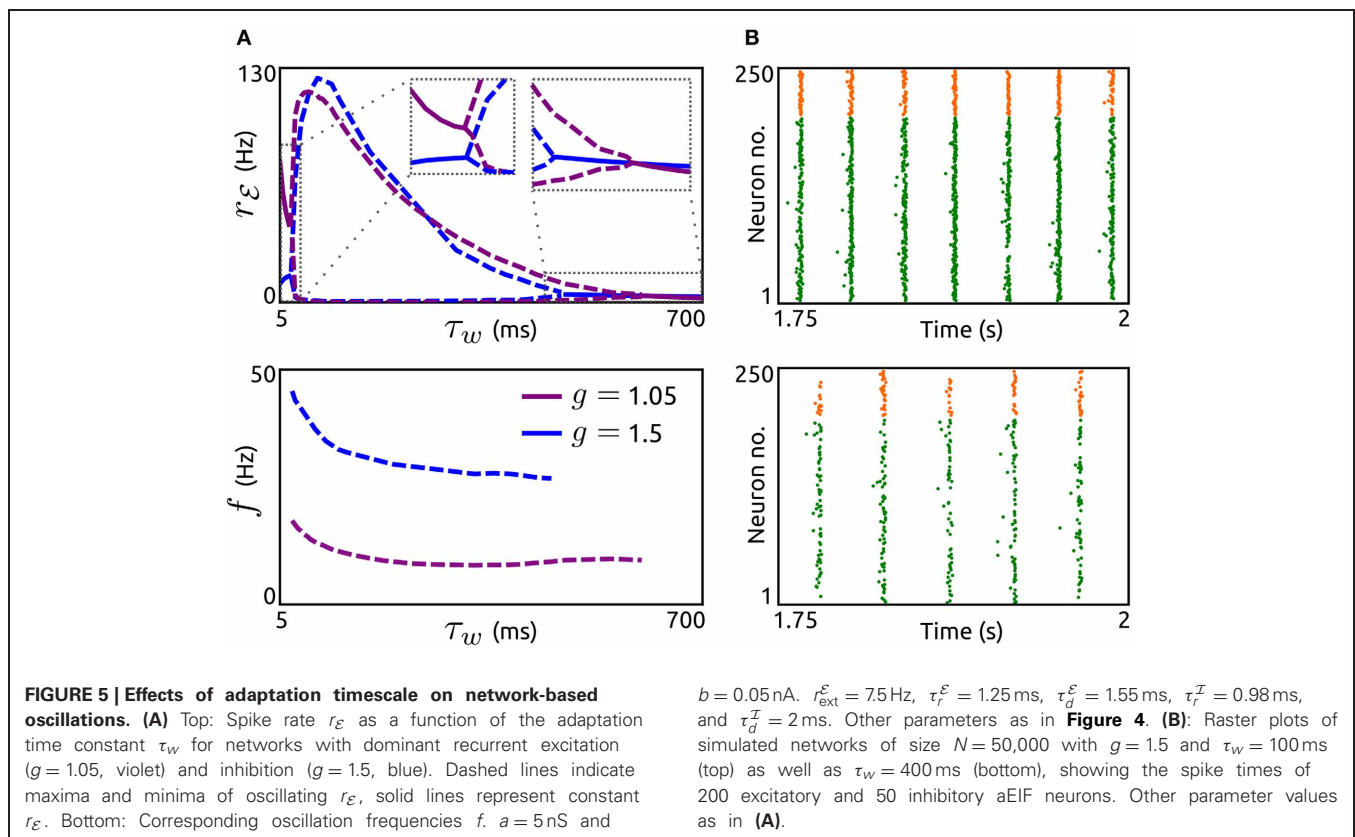
To analyze how the resonance properties of recurrent networks in asynchronous states are influenced by adaptation currents, we here consider external Poisson-inputs with oscillatory rates with frequency f . Gain of input spike rate and phase difference between network and input spike rates as a function of input frequency for networks without ($a = b = 0$) and with adaptation ($a = 5$ nS,

$b = 0.05$ nA) considering two adaptation time constants are presented in **Figures 6A,B**. Excitation dominated networks without adaptation do not exhibit resonance at any frequency and show only phase delays. The presence of an adaptation current leads to a significant amplification of oscillations in the input which is particularly strong at lower frequencies (of the beta band). This effect is pronounced for an increased adaptation timescale. In addition, adaptation causes a phase advance for low oscillation frequencies.

In networks where recurrent inhibition dominates excitation on the other hand even in the absence of adaptation currents resonance is shown for a high frequency band and phase advances for lower frequencies. Adaptation greatly enhances resonance and shifts the preferred frequency band to the high gamma range. The resonance effect is even stronger if the adaptation current is slower, i.e., τ_w increased. Although these effects of adaptation on resonance properties of recurrent networks are similar when either the subthreshold (a) or spike-triggered adaptation component (b) is increased individually, the dominant contribution to the frequency amplifications comes from b (not shown). We additionally examined the response of single neurons to oscillatory noisy inputs using our mean-field model and found that adaptation mediates resonance even in the absence of recurrent input (not shown). These results emphasize the importance of adaptation for the amplification and thus propagation of oscillatory signals in neuronal networks.

DISCUSSION

In this work we have investigated the role of neuronal adaptation currents in shaping spike rate oscillations in large recurrent networks of excitatory and inhibitory neurons. Based on a network of aEIF model neurons sparsely coupled through conductance-based synapses with heterogeneous delays and strengths driven by noisy external input, we used a mean-field method taking advantage of the FP equation. We simplified the problem by applying an adiabatic approximation and solved the resulting equations numerically. Using this method we obtain membrane potential



distributions and population averages of spike rates and adaptation currents. At the same time, the dynamical properties of single neurons, i.e., the neuron model parameters, are retained in the derived mean-field network model.

Alternative mean-field methods have been developed for conductance-based model neurons (Robinson et al., 2008) and recurrent networks thereof in asynchronous states (Shriki et al., 2003), where spike rates are obtained without having to solve a PDE. Our approach based on the FP equation on the other hand treats noise in the synaptic inputs in more detail and allows for the calculation of membrane potential distributions in addition to spike rates.

We chose the aEIF model because it provides a rich yet low-dimensional description of neuronal dynamics and includes a proper phenomenological description of the M and AHP adaptation currents. The effects of subthreshold (a) and spike-triggered adaptation (b) on response properties of aEIF neurons (measured by spike rate-input current relationships and phase response curves) match those of M and AHP adaptation currents in a Hodgkin-Huxley type neuron model, respectively (Ladenbauer et al., 2012). Furthermore, fitting the aEIF model parameters to a detailed biophysical model using standard electro-physiological paradigms revealed a clear relationship between parameter a and the conductance for the M current as well as between parameter b and the AHP current (not shown).

Our method is based on several assumptions which allow to derive the mean-field equations. The Poisson approximation of spike train statistics is justified by experimental findings (Tolhurst

et al., 1983; McAdams and Maunsell, 1999) although spiking seems to be more regular in some cortical areas (Maimon and Assad, 2009). The sparse random connectivity implies vanishing noise correlations between neurons in the large network limit and an experimental study in primary visual cortex of awake monkeys has reported almost zero noise correlations (Ecker et al., 2010). However, there is an ongoing debate about the strength of correlations in experimental data (Cohen and Kohn, 2011). We have used an adiabatic approximation, which relies on separable time scales of adaptation current and membrane voltage. Although this assumption is violated for small values of τ_w , numerically solving the unreduced FP system, Equations (18)–(24), showed that our results are robust regarding the violation of this assumption. The results we obtained by simulations of aEIF networks and the mean-field results show quantitative differences. However, the presented effects described using the mean-field model are validated qualitatively by the network simulations.

We have shown that spike-triggered adaptation provides a mechanism to generate spike rate oscillations in a low frequency range (alpha band and lower) if recurrent excitation is sufficiently strong. Increased subthreshold adaptation on the other hand does not contribute to this mechanism but rather dampens such oscillations. The type of adaptation current therefore strongly determines rhythmic activity in excitation dominated networks. The importance of activity-driven adaptation for slow oscillations is consistent with results from simulations of detailed (thalamo-)cortical spiking neuron network models (Bazhenov et al., 2002; Compte et al., 2003; Destexhe, 2009), mean-field

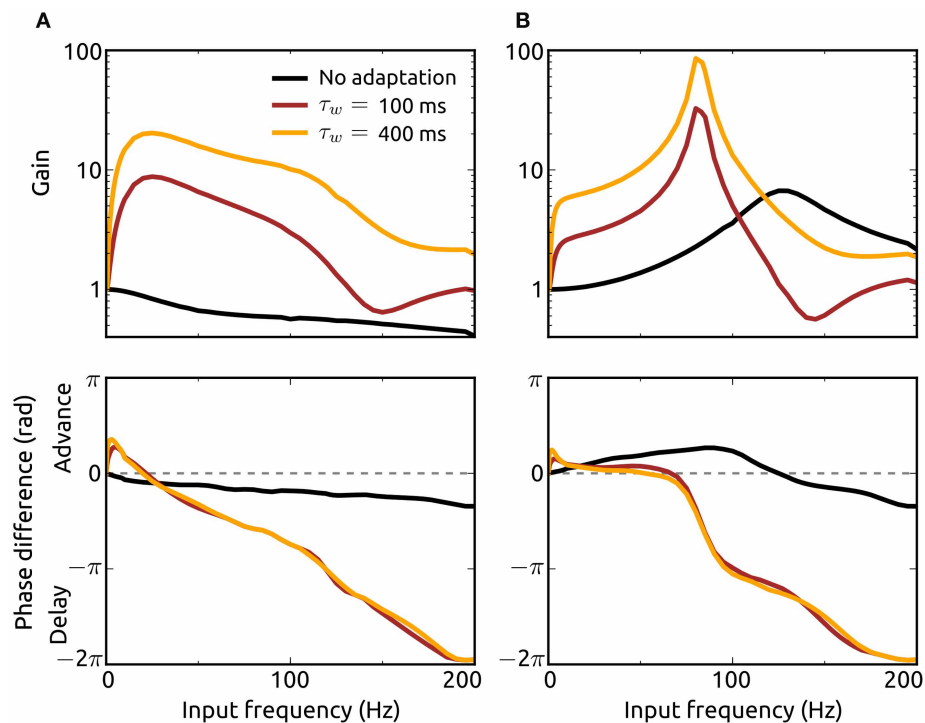


FIGURE 6 | Effects of adaptation on resonance properties of recurrent networks. Gain (top) and phase shift (bottom) of the spike rate r_E for networks with dominant recurrent excitation [$g = 1.05$, (A)] and inhibition [$g = 1.5$, (B)] as a function of the input frequency f . The gain is defined as the quotient of the oscillation amplitude in r_E for the input with frequency f and the amplitude for the lowest frequency ($f_{\min} = 0.5$ Hz). Adaptation parameter values are $a = b = 0$ (black), $a = 5$ nS, $b = 0.05$ nA, $\tau_w = 100$ ms (dark red),

and $a = 5$ nS, $b = 0.05$ nA, $\tau_w = 400$ ms (orange). Delay distributions are identically parameterized ($p_d^E \equiv p_d^I$) with $\tau_r = 1.25$ ms and $\tau_d = 1.5$ ms. The Poisson-input rates r_{ext}^E , r_{ext}^I each consist of a baseline rate plus a sinusoidal component of small amplitude (1/1000th of the baseline) with frequency f . The baseline of r_{ext}^E is chosen to yield a steady-state spike rate r_{ext}^E of 50 Hz with constant input rate. The baseline rate of r_{ext}^I is chosen as explained in the Methods section.

studies based on networks of excitatory neurons under the assumption sparse (Gigante et al., 2007b) and all-to-all connectivity (Nesse et al., 2008), as well as phenomenological rate models (Latham et al., 2000). We have further shown that reducing inhibitory synaptic strength leads to a reduction on oscillation frequency, which is in agreement with similar experimental findings (Sanchez-Vives et al., 2010).

The M and AHP K^+ currents, which mediate spike frequency adaptation in pyramidal neurons, are known to be deactivated by acetylcholine (McCormick, 1992), with the AHP current showing higher sensitivity. Since the adaptation parameter b is strongly related to AHP type adaptation, our results support the hypothesis that the cholinergically induced activating transition from slow-wave oscillations to asynchronous irregular states (Lee and Dan, 2012) is mediated (at least in part) by a reduction of spike-triggered adaptation (Destexhe, 2009).

We have demonstrated that an increase of either type of adaptation current leads to a reduction in the frequency of oscillations generated by a loop of recurrent excitation and inhibition. This shows that the dynamical properties of neurons in addition to coupling characteristics strongly affect the network frequency. Also the passive (integrative) membrane properties significantly influence such networks oscillations as has been described previously (Geisler et al., 2005). Our additional finding of decreased

frequencies for increased adaptation time constants is consistent with the results from a computational study on clustering effects of spike-triggered adaptation in gamma oscillations (Kilpatrick and Ermentrout, 2011).

Low input frequencies have been shown to be suppressed in the output of single excitatory neurons with increased spike-triggered (Gigante et al., 2007a) or subthreshold adaptation (Richardson et al., 2003; Prescott and Sejnowski, 2008), which we confirmed using our aEIF-based mean-field model. Such a high pass property of single neurons has also been found using a more general model of adaptation (Benda and Herz, 2003). We have demonstrated that both adaptation currents cause spike rate resonance in excitation dominated recurrent networks. Inhibition dominated networks, on the other hand, exhibit resonance without adaptation and we have shown that increased adaptation of excitatory neurons strongly amplifies this resonance. A similar effect has been described for purely inhibitory networks (Richardson, 2009). In addition, our results show that adaptation shifts the resonance frequency to lower values.

In excitation dominated networks, adaptation further leads to phase advances for low input frequencies in addition to phase delays for higher frequencies as observed in previous studies on single excitatory neurons (Fuhrmann et al., 2002; Gigante et al., 2007a). These adaptation-induced phase advances enable

synchronization of periodic activity between distant neurons (and populations of neurons) in different areas of the brain if the strength of adaptation is controlled appropriately, e.g., through cholinergic neuromodulation.

Here we have considered one adaptation current for each neuron of the excitatory population. To account for the multimodal distribution of adaptation timescales found experimentally (La

Camera et al., 2006) our approach can be easily extended to include multiple adaptation currents.

ACKNOWLEDGMENTS

We thank Timm Lochmann for helpful comments on the manuscript. This work was supported by the DFG Collaborative Research Center SFB910.

REFERENCES

- Abel, H. J., Lee, J. C. F., Callaway, J. C., and Foehring, R. C. (2004). Relationships between intracellular calcium and afterhyperpolarizations in neocortical pyramidal neurons. *J. Neurophysiol.* 91, 324–335.
- Badel, L., Lefort, S., Brette, R., Petersen, C. C. H., Gerstner, W., and Richardson, M. J. E. (2008). Dynamic I-V curves are reliable predictors of naturalistic pyramidal-neuron voltage traces. *J. Neurophysiol.* 99, 656–666.
- Bazhenov, M., Timofeev, I., Steriade, M., and Sejnowski, T. J. (2002). Model of thalamocortical slow-wave sleep oscillations and transitions to activated states. *J. Neurosci.* 22, 8691–8704.
- Benda, J., and Herz, A. V. M. (2003). A universal model for spike-frequency adaptation. *Neural Comput.* 15, 2523–2564.
- Brette, R., and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.* 94, 3637–3642.
- Brown, D. A., and Adams, P. R. (1980). Muscarinic suppression of a novel voltage-sensitive K⁺ current in a vertebrate neurone. *Nature* 283, 673–676.
- Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.* 8, 183–208.
- Brunel, N., Hakim, V., and Richardson, M. (2003). Firing-rate resonance in a generalized integrate-and-fire neuron with subthreshold resonance. *Phys. Rev. E* 67, 051916.
- Brunel, N., and Wang, X.-J. (2003). What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *J. Neurophysiol.* 90, 415–430.
- Cohen, M. R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat. Neurosci.* 14, 811–819.
- Compte, A., Sanchez-Vives, M. V., McCormick, D. A., and Wang, X.-J. (2003). Cellular and network mechanisms of slow oscillatory activity (<1 Hz) and wave propagations in a cortical network model. *J. Neurophysiol.* 89, 2707–2725.
- Connors, B., Gutnick, M., and Prince, D. (1982). Electrophysiological properties of neocortical neurons *in vitro*. *J. Neurophysiol.* 48, 1302–1320.
- Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., and Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* 4:e1000092. doi: 10.1371/journal.pcbi.1000092
- Destexhe, A. (2009). Self-sustained asynchronous irregular states and up-down states in thalamic, cortical and thalamocortical networks of nonlinear integrate-and-fire neurons. *J. Comput. Neurosci.* 27, 493–506.
- Destexhe, A., Rudolph, M., and Paré, D. (2003). The high-conductance state of neocortical neurons *in vivo*. *Nat. Rev. Neurosci.* 4, 739–751.
- Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., Logothetis, N. K., and Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science* 327, 584–587.
- Engel, A., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* 2, 704–716.
- Fourcaud-Trocme, N., Hansel, D., van Vreeswijk, C., and Brunel, N. (2003). How spike generation mechanisms determine the neuronal response to fluctuating inputs. *J. Neurosci.* 23, 11628–11640.
- Fries, P. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291, 1560–1563.
- Fuhrmann, G., Markram, H., and Tsodyks, M. (2002). Spike frequency adaptation and neocortical rhythms. *J. Neurophysiol.* 88, 761–770.
- Geisler, C., Brunel, N., and Wang, X.-J. (2005). Contributions of intrinsic membrane dynamics to fast network oscillations with irregular neuronal discharges. *J. Neurophysiol.* 94, 4344–4361.
- Ghazanfar, A., Chandrasekaran, C., and Logothetis, N. (2008). Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *J. Neurosci.* 28, 4457–4469.
- Gigante, G., Del Giudice, P., and Mattia, M. (2007a). Frequency-dependent response properties of adapting spiking neurons. *Math. Biosci.* 207, 336–351.
- Gigante, G., Mattia, M., and Del Giudice, P. (2007b). Diverse population-bursting modes of adapting spiking neurons. *Phys. Rev. Lett.* 98, 1–4.
- Goodman, D. F. M., and Brette, R. (2009). The brain simulator. *Front. Neurosci.* 3, 192–197. doi: 10.3389/neuro.01.026.2009
- Hammond, C., Bergman, H., and Brown, P. (2007). Pathological synchronization in Parkinson's disease: networks, models and treatments. *Trends Neurosci.* 30, 357–364.
- Isaacson, J. S., and Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron* 72, 231–243.
- Jolivet, R., Schürmann, F., Berger, T. K., Naud, R., Gerstner, W., and Roth, A. (2008). The quantitative single-neuron modeling competition. *Biol. Cybern.* 99, 417–426.
- Kilpatrick, Z. P., and Ermentrout, B. (2011). Sparse gamma rhythms arising through clustering in adapting neuronal networks. *PLoS Comput. Biol.* 7:e1002281. doi: 10.1371/journal.pcbi.1002281
- Kumar, A., Schrader, S., Aertsen, A., and Rotter, S. (2008). The high-conductance state of cortical networks. *Neural Comput.* 20, 1–43.
- La Camera, G., Rauch, A., Thurbon, D., Lüscher, H.-R., Senn, W., and Fusi, S. (2006). Multiple time scales of temporal response in pyramidal and fast spiking cortical neurons. *J. Neurophysiol.* 96, 3448–3464.
- Ladenbauer, J., Augustin, M., Shiau, L., and Obermayer, K. (2012). Impact of adaptation currents on synchronization of coupled exponential integrate-and-fire neurons. *PLoS Comput. Biol.* 8:e1002478. doi: 10.1371/journal.pcbi.1002478
- Latham, P. E., Richmond, B. J., Nelson, P. G., and Nirenberg, S. (2000). Intrinsic dynamics in neuronal networks. I. Theory. *J. Neurophysiol.* 83, 808–827.
- Lee, S.-H., and Dan, Y. (2012). Neuromodulation of brain states. *Neuron* 76, 209–222.
- LeVeque, R. (2002). *Finite Volume Methods for Hyperbolic Problems*. Cambridge: Cambridge University Press.
- Maimon, G., and Assad, J. A. (2009). Beyond Poisson: increased spike-time regularity across primate parietal cortex. *Neuron* 62, 426–440.
- Manuel, M., Meunier, C., Donnet, M., and Zytnicki, D. (2005). How much afterhyperpolarization conductance is recruited by an action potential? A dynamic-clamp study in cat lumbar motoneurons. *J. Neurosci.* 25, 8917–8923.
- McAdams, C. J., and Maunsell, J. H. (1999). Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron* 23, 765–773.
- McCormick, D. A. (1992). Neurotransmitter actions in the thalamus and cerebral cortex and their role in thalamocortical activity. *Progr. Neurobiol.* 39, 337–388.
- Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., and Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *J. Neurosci.* 27, 2858–2865.
- Naud, R., Marcille, N., Clopath, C., and Gerstner, W. (2008). Firing patterns in the adaptive exponential integrate-and-fire model. *Biol. Cybern.* 99, 335–347.
- Nesse, W. H., Borisyuk, A., and Bressloff, P. C. (2008). Fluctuation-driven rhythmogenesis in an excitatory neuronal network with slow adaptation. *J. Comput. Neurosci.* 25, 317–333.
- Nykamp, D. Q., and Tranchina, D. (2000). A population density

- approach that facilitates large-scale modeling of neural networks: analysis and an application to orientation tuning. *J. Comput. Neurosci.* 8, 19–50.
- Pedarzani, P., Mosbacher, J., Rivard, A., Cingolani, L. A., Oliver, D., Stocker, M., et al. (2001). Control of electrical activity in central neurons by modulating the gating of small conductance $\text{Ca}(2+)$ -activated K^+ channels. *J. Biol. Chem.* 276, 9762–9769.
- Pospischil, M., Piwkowska, Z., Bal, T., and Destexhe, A. (2011). Comparison of different neuron models to conductance-based post-stimulus time histograms obtained in cortical pyramidal cells using dynamic-clamp *in vitro*. *Biol. Cybern.* 105, 167–180.
- Prescott, S. A., and Sejnowski, T. J. (2008). Spike-rate coding and spike-time coding are affected oppositely by different adaptation mechanisms. *J. Neurosci.* 28, 13649–13661.
- Renart, A., Brunel, N., and Wang, X.-J. (2004). “Mean-field theory of irregularly spiking neuronal populations and working memory in recurrent cortical networks,” in *Computational Neuroscience: A Comprehensive Approach*, ed J. Feng (Boca Raton, FL: CRC Press), 431–490.
- Richardson, M. (2004). Effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons. *Phys. Rev. E* 69, 1–8.
- Richardson, M. (2009). Dynamics of populations and networks of neurons with voltage-activated and calcium-activated currents. *Phys. Rev. E* 80, 1–16.
- Richardson, M. J. E., Brunel, N., and Hakim, V. (2003). From subthreshold to firing-rate resonance. *J. Neurophysiol.* 89, 2538–2554.
- Risken, H. (1996). *The Fokker-Planck Equation: Methods of Solutions and Applications*. Berlin: Springer.
- Robinson, P. A., Wu, H., and Kim, J. W. (2008). Neural rate equations for bursting dynamics derived from conductance-based equations. *J. Theor. Biol.* 250, 663–672.
- Saad, Y. (2003). *Iterative Methods for Sparse Linear Systems*. Philadelphia, PA: SIAM.
- Sanchez-Vives, M. V., Mattia, M., Compte, A., Perez-Zabalza, M., Winograd, M., Descalzo, V. F., et al. (2010). Inhibitory modulation of cortical up states. *J. Neurophysiol.* 104, 1314–1324.
- Shriki, O., Hansel, D., and Sompolinsky, H. (2003). Rate models for conductance-based cortical neuronal networks. *Neural Comput.* 15, 1809–1841.
- Stocker, M. (2004). $\text{Ca}(2+)$ -activated K^+ channels: molecular determinants and function of the SK family. *Nat. Rev. Neurosci.* 5, 758–770.
- Tamas, G., Buhl, E. H., and Somogyi, P. (1997). Fast IPSPs elicited via multiple synaptic release sites by different types of GABAergic neurone in the cat visual cortex. *J. Physiol.* 500, 715–738.
- Tiesinga, P., and Sejnowski, T. J. (2009). Cortical enlightenment: are attentional gamma oscillations driven by ING or PING? *Neuron* 63, 727–732.
- Tolhurst, D. J., Movshon, J. A., and Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* 23, 775–785.
- Touboul, J., and Brette, R. (2008). Dynamics and bifurcations of the adaptive exponential integrate-and-fire model. *Biol. Cybern.* 99, 319–334.
- Uhlhaas, P. J., and Singer, W. (2010). Abnormal neural oscillations and synchrony in schizophrenia. *Nat. Rev. Neurosci.* 11, 100–113.
- van der Vorst, H. A. (1992). Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.* 13, 631–644.
- Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol. Rev.* 90, 1195–1268.
- Williams, S. R., and Stuart, G. J. (2002). Dependence of EPSP efficacy on synapse location in neocortical pyramidal neurons. *Science* 295, 1907–1910.
- Womble, M. D., and Moises, H. C. (1992). Muscarinic inhibition of M-current and a potassium leak conductance in neurones of the rat basolateral amygdala. *J. Physiol.* 457, 93–114.
- Zijlmans, M., Jacobs, J., Zelmann, R., Dubeau, F., and Grotman, J. (2009). High-frequency oscillations mirror disease activity in patients with epilepsy. *Neurology* 72, 979–986.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 December 2012; paper pending published: 30 December 2012; accepted: 08 February 2013; published online: 27 February 2013.

Citation: Augustin M, Ladenbauer J and Obermayer K (2013) How adaptation shapes spike rate oscillations in recurrent neuronal networks. *Front. Comput. Neurosci.* 7:9. doi: 10.3389/fncom.2013.00009

Copyright © 2013 Augustin, Ladenbauer and Obermayer. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Inferring network properties of cortical neurons with synaptic coupling and parameter dispersion

Dipanjan Roy^{1,2,3*} and Viktor Jirsa¹

¹ Theoretical Neuroscience Group, Faculté de Médecine, Institut de Neurosciences des Systèmes, Inserm UMR1106, Aix-Marseille Université, Marseille, France

² Bernstein Center for Computational Neuroscience, Berlin, Germany

³ Department of Software Engineering and Theoretical Computer Science, Technische Universität Berlin, Berlin, Germany

Edited by:

Dimitris Pinotsis, University College
London, UK

Reviewed by:

Mukesh Dhamala, Georgia State
University, USA
Roxana Stefanescu, University of
Michigan, USA

*Correspondence:

Dipanjan Roy, Department of
Software Engineering and
Theoretical Computer Science,
Technische Universität Berlin, Neural
Information Processing Group,
Marchstrasse 23, MAR 5-6, 10587
Berlin, Germany.
e-mail: dipanjan@ni.tu-berlin.de

Computational models at different space-time scales allow us to understand the fundamental mechanisms that govern neural processes and relate uniquely these processes to neuroscience data. In this work, we propose a novel neurocomputational unit (a mesoscopic model which tell us about the interaction between local cortical nodes in a large scale neural mass model) of bursters that qualitatively captures the complex dynamics exhibited by a full network of parabolic bursting neurons. We observe that the temporal dynamics and fluctuation of mean synaptic action term exhibits a high degree of correlation with the spike/burst activity of our population. With heterogeneity in the applied drive and mean synaptic coupling derived from fast excitatory synapse approximations we observe long term behavior in our population dynamics such as partial oscillations, incoherence, and synchrony. In order to understand the origin of multistability at the population level as a function of mean synaptic coupling and heterogeneity in the firing rate threshold we employ a simple generative model for parabolic bursting recently proposed by Ghosh et al. (2009). Further, we use here a mean coupling formulated for fast spiking neurons for our analysis of generic model. Stability analysis of this mean field network allow us to identify all the relevant network states found in the detailed biophysical model. We derive here analytically several boundary solutions, a result which holds for any number of spikes per burst. These findings illustrate the role of oscillations occurring at slow time scales (bursts) on the global behavior of the network.

Keywords: multispikes, self-organization, transients, firing rate, parabolic burst, network synchrony, generative model, oscillations

1. INTRODUCTION

The neuronal spike-burst activity is characterized by recurrent transitions between rest state and firing state where bursts are temporal groupings of multiple spikes. Certain cells in the mammalian brain, for example, neurons in the thalamus during periods of drowsiness, attentiveness, and sleep are known to exhibit this type of spike-burst behavior (Sherman and Koch, 1986; Steriade and Llinás, 1988; McCormick and Feese, 1990; Steriade et al., 1993; Amzica and Steriade, 1998). Autonomously bursting neurons are found in a variety of neural systems, from the mammalian cortex (Morris and Lecar, 1981; Dhamala et al., 2004a,b) to brainstem (Hindmarsh and Rose, 1984; Wang, 1994; Izhikevich, 2007; Jirsa and McIntosh, 2007; Jirsa, 2008). When neurons are coupled with each other, they produce different modes of behavior, including synchrony and phase-locking, which have been implicated in memory, cognition, sensory processing, motor planning, and execution (McCormick and Feese, 1990; Wang, 1994; Jirsa and McIntosh, 2007). Many neurological diseases, on the other hand, including Parkinson, schizophrenia, and epilepsy, are the result of abnormal synchronization (Uhlhaas and Singer, 2006; Jensen et al., 2007), which suggests that a better understanding of the basic mechanisms producing synchrony and phase locking will be a stepping stone toward the repair of brain function. Modeling attempts using large scale networks to understand

emergence of cognitive states rely heavily on the approximation of the dynamics as a neural ensemble. The concept of a neural mass like abstraction (Hebb, 1949; Beurle, 1956) designates a group of Co-activated neurons capable of acting like a closed system when performing a certain function. A small scale network of this kind is sometimes referred to as a “neurocomputational unit.” In large scale brain networks, these mesoscopic units of operation serve as the network nodes (see for instance, Deco et al., 2008, 2011; Ghosh et al., 2008). On intermediate spatial scales of few cm, neural activations along the spatially continuous cortical sheets are described by neural fields, for which the connectivity is assumed to be translationally invariant (see, Wilson and Cowan, 1972; Nunez, 1974; Amari, 1977; Jirsa and Haken, 1997; Feng et al., 2006; Jirsa, 2009; Robinson, 2011). To define such small neurocomputational units, simplified neuron models, known as phase models, offer an attractive tool for the study of network modes, since they allow for detailed mathematical analysis of network dynamics (Breakspear et al., 2010). As an example, Carbal et al. have explored the role of local network oscillations in resting-state functional connectivity by using such phase oscillators in the respective nodes of the simulated network. They have shown when these oscillatory units are integrated in the network, they behave as weakly coupled oscillators. Moreover, for a set of network parameters they found subsets of nodes tend to

synchronize although the network is not globally synchronized (Cabral et al., 2011). For the present work we use a recently proposed phenomenological model that admits parabolic bursting in one dimension, which is a type of bursting observed in the R-15 neuron in abdominal ganglion of aquatic mollusc *Aplysia Californica* (Ermentrout and Kopell, 1986; Izhikevich, 2000; Ghosh et al., 2009). This type of bursting can arise even without bistability in the generation of spikes. The investigation carried out in this work with a detailed neuron model capable of displaying spiking and bursting behavior and a minimal model that not only reproduces the mean field amplitude of the original networks but also capture the most important temporal features of its dynamics. The detailed model used here is extensively discussed in Rinzel and Ermentrout (1989). On the other hand, our phase model is a minimal model that captures the generality of the mechanism of bursting present in the detailed model. As we vary network parameters including mean field coupling strength and dispersion, both networks display various temporal dynamics. In order to understand these states in mathematically tractable terms we take advantage of the mean field coupled network of phase model. Our goal is to identify to what degree this mean field model serves as a reliable neurocomputational unit and captures the qualitative features of temporal dynamics of the full network as a function of the investigated network parameters. Mean field analysis for singleton burst reveals solutions such as incoherence and partial oscillation which can be completely described analytically. However, as we are interested in a multispike system where analytical calculation is rather non-trivial and therefore, we combine semi-analytical approach with numerics to derive the stability diagram. Mean field phase network allow us to identify the mechanism of transitions between various network states that appear as solutions of the full network. Stability diagram is independent of number of spikes per burst and qualitatively commensurates well with the findings in our full network. The paper is structured as follows. In the next section, we introduce the Rinzel–Ermentrout model (Rinzel and Ermentrout, 1989) for parabolic bursting and describe the model in details. In the following section, we couple individual neurons via global coupling and present our analysis of this network model. In the subsequent section, we set up a generic network of bursters coupled to their mean field and derive semianalytically all the network states and corresponding phase transition boundaries. In the next section, we derive numerically a stability diagram using global phase coherence measure. In the final section, we summarize the results obtained from mean field descriptions and link them systematically with the network states obtained from biophysical model network.

2. MATERIALS AND METHODS

2.1. SINGLE NEURON MODEL

A dynamical system with multiple time scales (for example, a neuron with spiking-bursting behaviors) can be written in a singularly perturbed form: $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{y})$, $\dot{\mathbf{y}} = r\mathbf{g}(\mathbf{x}, \mathbf{y})$, where \mathbf{x} is the vector of fast variables, \mathbf{y} the vector of slow variables that modulate the fast activity, and $r \ll 1$ is a ratio of fast/slow time scales. A system which has been proposed to describe parabolic bursting behavior is known as Rinzel model (1989). Single neuron

model parameters used here are exactly as described in Rinzel and Ermentrout (1989).

$$\begin{aligned}\dot{V} &= (I - I_{Ca} - (g_K w + g_{KCa} z)(V - V_K) - g_l(V - V_l))/c \\ \dot{w} &= \phi(w_\infty - w)/\tau_w \\ \dot{Ca} &= \epsilon(-\mu I_{Ca} - Ca) \\ \dot{n} &= \epsilon(n_\infty(V) - n)/\tau_n\end{aligned}\quad (1)$$

where $I_{Ca} = (g_{Ca} m_\infty(V) + g_{sCa} n)(V - V_{Ca})$, $z = \frac{Ca}{Ca + Ca_0}$ and gating functions are

$$\begin{aligned}m_\infty(V) &= 0.5(1 + \tanh((V - v_1)/v_2)) \\ w_\infty(V) &= 0.5(1 + \tanh((V - v_3)/v_4)) \\ n_\infty(V) &= 0.5(1 + \tanh((V - v_5)/v_6)) \\ \tau_w(V) &= 0.5(1 + \tanh((V - v_3)/2v_2))\end{aligned}\quad (2)$$

where V is the membrane potential, w is associated with the fast current, Na^+ or K^+ , Ca and n are the two slow currents, Model parameters which are held fixed throughout our simulations are, $V_K = -84$, $V_l = 60$, $V_{Ca} = 120$, $g_K = 8$, $g_l = 2$, $c = 20$, $v_1 = 1.2$, $v_2 = 18$, $v_6 = 24$, $v_5 = 12$, $v_3 = 12$, $v_4 = 17.4$, $\tau_n = 0.05$, $\phi = 0.06666666$, $g_{Ca} = 4.0$, $\mu = 0.025$, $Ca_0 = 1$, $\epsilon = 0.0005$, and $g_{KCa} = 1$, $g_{sCa} = 1$.

I is the applied input current. The ionic currents are given by an ohmic leak current, determined by the leak conductance g_l and leak reversal potential V_l , and a Na^+ current which is responsible for the generation of spikes. The dynamics of this model which is relevant to our study is outlined as follows. When the input current I exceeds a critical value I_c a single neuron described by Equation (1) undergoes a Saddle-node bifurcation on an invariant circle (SNIC). This same system for two different parameterization of I and in the presence of the slow currents can exhibit both spiking as well as parabolic bursting behavior. Spiking behaviors are elicited for a slightly higher value of the external drive. For example, to observe a typical burst-like pattern in this system we held the input current to the values $I = 68$ and for spikes $I \geq 70$. **Figure 1** displays the relationship between the applied input current and a parabolic bursting pattern that is observed in the single neuron dynamics.

2.2. PHASE MODEL

The generality of the underlying mechanism for parabolic bursting is investigated in details by numerous authors (Ermentrout and Kopell, 1986; Baer et al., 1995; Izhikevich, 2000). In many such formulations, parabolic bursting neurons are typically in their canonical form described as:

$$\begin{aligned}\dot{\theta} &= [1 - \cos(\theta) + f(x, y)] \\ \dot{x} &= \mu_x[x_\eta(\theta) - x] \\ \dot{y} &= \mu_y[y_\eta(\theta) - y]\end{aligned}\quad (3)$$

where function $f(x, y)$ in the above equation couples to spike generative mechanism depending on the slow variables x, y dynamics,

respectively. The function $f(x, y)$ is a smoothly varying periodic function alternating signs such that the system undergoes a SNIC to generate parabolic burst at the single neuron level. Recently Ghosh et al. (2009) has also proposed a simpler model that in principle captures the underlying mechanism of parabolic bursting involving only a circular phase variable θ and moreover, involve only one slow term to allow the fast dynamics to enter or get out of repetitive firing. Motivation for using such a model is primarily mathematical tractability. Parameter space of this model cannot be directly linked to the biophysical parameters, however, qualitatively it may account for the transient and longterm behavior of more detailed biophysical models. In this model a single neuron is described by the following equation,

$$\dot{\theta} = I - \cos \theta - \cos \frac{\theta}{n} \quad (4)$$

In Equation (4) a slow variable activation term is represented by a modulation term $\cos(\frac{\theta}{n})$ which mimics the entire slow subsystem instead of describing it as a separate dynamical system, I is the applied input current and n is an integer, which determines the number of spikes per burst. In our simulation with this model all the results are for $n = 5$ spikes per burst unless otherwise specified.

2.3. FULL NETWORK MODEL

Golomb and Rinzel (1993) considered a heterogeneous network of all-to-all coupled inhibitory bursting neurons and found regimes of synchronous, anti-synchronous and asynchronous behavior when the width of the heterogeneity was changed (Golomb and Rinzel, 1993; Stefanescu and Jirsa, 2008, 2011; Smeal et al., 2010; Jirsa and Stefanescu, 2011). We describe our network equations via a fast instantaneous coupling. N synaptically coupled (all-to-all) parabolic bursting neurons are described by a similar set of non-linear differential equations with fast chemical synapse. To this end we formally describe:

$$\begin{aligned} \dot{V}_i &= (bI_i - I_{Ca} - (g_K w + g_{KCa} z)(V_i - V_K) \\ &\quad - g_l(V_i - V_l) + KS(V_i - V_{th}))/c \\ \dot{w}_i &= \phi(w_\infty - w_i)/\tau_w \\ \dot{Ca}_i &= \epsilon(-\mu I_{Ca} - Ca_i) \\ \dot{n}_i &= \epsilon(n_\infty(V) - n_i)/\tau_n \\ \dot{s}_i &= a_s(V_i)(1 - s_i) - \frac{s_i}{\beta} \end{aligned} \quad (5)$$

where all the parameters and the gating variables inherit from the single neuron model Equation (1, 2) and b is a rescaling factor to applied drive to cross the threshold and elicit spike/burst in the uncoupled system. Stimulus that all the neurons see $I_i > 0$ are drawn from a uniform distribution assumed to be symmetrically distributed over the interval $I_i \in [2.1 - \Delta I, 2.1 + \Delta I]$. Where ΔI is the spread of the applied stimulus parameter. ΔI introduces a heterogeneity in the spike threshold. The synaptic coupling appears as an ensemble average given by $S = \frac{1}{N} \sum_{i=1}^N s_i$, where $a_{s_i}(V_i) = \frac{1}{(1 + \exp(-V_i/2))}$ is a sigmoidal activation function.

The synaptic strength K is the same for all the neurons. For the entire simulation, we fixed the reversal potential of potassium ions to $v_{th} \approx 0.0$ (for purely excitatory connectivity).

Analysis is carried out for a fast synapse (AMPA-type glutamate receptors), such as those found in the auditory system, the rise time is instantaneous, and post-synaptic responses commence almost instantaneously after the start of presynaptic action potential (Nunez, 1974; Morris and Lecar, 1981). This brisk communication is a consequence of rapid calcium-channel kinetics, which allows significant calcium entry during the upstroke of the presynaptic action potential (Sabatini and Regehr, 1996). Under the fast synapse approximation the variable s_i relaxes much more rapidly than V_i , in which case we may apply a quasi-static approximation to (Equation 5) (e), $\dot{s}_i \approx 0$, allowing us to adiabatically eliminate the synaptic variable via $s_i = \frac{\beta}{(1 + \beta + \exp(-V_i/2))}$. The time course of the postsynaptic conductivity caused by an activation of AMPA receptors can be captured by a rise time $\beta_{rise} = 0.09$ ms and decay time $\beta_{decay} = 1.5$ ms (Gabbiani et al., 1994; Parnas and Parnas, 1994). Numerical results in **Figure 3** provides a good approximation for β in the range between [0.01 ms, 0.5 ms]. Although, we have provided here the details about the fast excitatory synaptic connectivity, our approach can be readily extended to inhibitory connectivity as well. In the continuum limit, a mean field formulation with inhibitory synaptic coupling is provided in details in Appendix.

2.4. MEAN FIELD COUPLED PHASE MODEL

Each generic neuron is coupled to this mean field and typically their response to the mean field expressed as $R(\theta)$ explicitly dependent on θ , and implicitly on time. In absence of any coupling, their vector field flow on a real line is governed by $F(\theta) = \omega - \cos(\theta) - \cos(\theta)/n$. In the absence of the term $\cos(\theta)/n$ expression reduces to a mathematical description used in Roy et al. (2011). Together, we can write for N (still finite) such neurons:

$$\dot{\theta}_i = F(\theta_i) - \Gamma R(\theta_i), \quad (6)$$

Recently, we have proposed a formulation for mean synaptic activation term under fairly general setting and taking advantage of instantaneous activation, deactivation between pre and postsynaptic events. It allows one to describe synaptic activation variable $s_i = \frac{\beta}{1 + \beta + \exp(-\frac{V_i}{2})}$ as a non-linear transfer function of membrane voltage (Roy et al., 2011). Moreover, we have described how the mean field coupled spiking neurons can be described mathematically with this synaptic coupling. Details of this formulation is described elsewhere, (Roy et al., 2011). Collective activity of synapses is described by a mean field. For a given population of neurons is expressed more formally as,

$$\Gamma = \frac{K}{N} \sum_{l=1}^N \frac{\beta}{(1 + \beta + \exp(-\frac{\cos \theta_l}{2}))}, \quad i \neq l. \quad (7)$$

where Γ is the mean field influence function. Coupling K is the same for all the neurons. In our previous work, response to

such mean field coupling explicitly described as $R(\theta_i) = \sin \theta_i (\cos \theta_i - \nu_{th})$,

$$\dot{\theta}_i = F(\theta_i) - \Gamma \sin \theta_i (\cos \theta_i - \nu_{th}) + O(\epsilon), \quad (8)$$

where $O(\epsilon)$ contains non-circular deviations of the order ϵ that results due to perturbations. $\nu_{th} \approx 0.0$ for all simulations and analytical calculations unless mentioned otherwise. It is important to note that the couplings in the phase descriptions retain their mathematical expression in the full model plus some linearly added correction terms, which scale with the degree of order of deviation from the circle (Roy et al., 2011). Hence, in application it is rather suitable when phase perturbations are close to the circular orbit. The above equation further can be written combining the terms containing a single Fourier harmonic in the coupling plus the higher order Fourier terms.

$$\begin{aligned} \dot{\theta}_i &= \omega_i - \sin \theta_i - \sin(\theta_i/n) + P(\theta_l) \sin \theta_i \nu_{th} \\ &\quad + O(2\theta_i) + O(\epsilon), \quad (9) \\ P(\theta_l) &= \frac{K}{N} \sum_{l=1}^N \frac{\beta}{\left(1 + \beta + \exp\left(-\frac{\cos \theta_l}{2}\right)\right)}, \quad i \neq l. \quad (10) \end{aligned}$$

See for details (Roy et al., 2011). Where, in Equation (6) the frequencies $I_i \geq 0$ are assumed to be symmetrically distributed over the interval $I_i \in [I - \Delta I, I + \Delta I]$ according to a uniform probability distributions.

2.5. CHARACTERIZATION OF SPIKE/BURST COHERENCE IN BIOPHYSICAL NETWORK MODEL WITH MEAN FIELD COUPLING

The bursting coherence and incoherence is quantitatively characterized in terms of a statistical-mechanical spike-based measure. We consider an excitatory population of neurons coupled to a common mean field drive and heterogeneity in their threshold for spikes/bursts. By varying the strength of the coupling K and the stimulus spread ΔI we investigate the emergence of spike/burst coherence. Emergence of collective spiking/bursting coherence may be well described by the (population-averaged) global potential,

$$V_{\text{mean}}(t) = \frac{1}{N} \sum_{i=1}^N V_i \quad (11)$$

In the thermodynamic limit ($N \rightarrow \infty$), a collective state becomes coherent if $\delta V_{\text{mean}}(t) \equiv [V_{\text{mean}}(t) - \overline{V_{\text{mean}}(t)}]$ is non-stationary (i.e., an oscillating global potential V_{mean} appears for a coherent case), where the overbar represents the time average, and also, the correlated mean field $\Gamma(t)$ activity appears oscillatory. Otherwise (i.e., when V_{mean} is time independent or stationary), it either becomes incoherent (IN) or partial oscillatory (PO). In $N \rightarrow \infty$ limit both these states converges to a stationary solution. Thus, the mean square deviation of the global potential is a global marker for mean burst coherence for the entire population described here. More formally one can write it as (i.e., time-averaged fluctuations of V_{mean}),

$$R(t) = \overline{(V_{\text{mean}}(t) - \overline{V_{\text{mean}}(t)})^2} \quad (12)$$

plays the role of an order parameter used for describing the coherence-incoherence transition (Manrubia et al., 2004). For the coherent (IN) state, the order parameter $R(t)$ approaches a non-zero (zero) limit value as N goes to the infinity. We compute $R(t)$ in Equation (12) as a function of mean field coupling strength K and dispersion parameter ΔI for the full system. We vary both K , ΔI from 0 to 1 in a step size of 0.01. Subsequently, computed values of $R(t)$ is plotted in grid size of 100×100 . Contour plot is colorcoded from low values at zero (blue) to high values at 1 (red). Nearly (in phase or anti phase) synchronized population spike/burst activity is lumped into a regime with labeled as SR and IN population spike/burst activity is lumped into a regime called IN activity. In the IN regime as described above $R(t)$ values stays close to zero with substantial subthreshold fluctuations. Partial bursty regime is labeled as PO observed for $R(t)$ values stationary and close to values other than zero. This regime displays dynamical behaviors far from synchrony, such as multi-clustering (some of the neurons are firing incoherently while others are not firing at all) in the phase for instance. Depending on the heterogeneity in stimulus spread we get random distribution of phases such that individual members can exhibit cluster hopping. Multiclustering in our model can reliably be captured using an ensemble average quantity rotation number ρ_i given by Equation (14).

2.6. CHARACTERIZATION OF SPIKE/BURST COHERENCE IN PHASE NETWORK MODEL WITH MEAN FIELD COUPLING

The bursting coherence and incoherence is quantitatively characterized in terms of statistical mechanical order parameter coherence measure. As an alternative to storing and plotting many time series data $\theta_i(t)$, $i = 1, \dots, N$ for all $N = 1000$ variables, we define an order parameter

$$R_\theta(t) = \frac{1}{N} \sum_{i=1}^N \cos \theta_i \quad (13)$$

Equation (13) measures the population dynamics. The advantage of using such a formulation becomes apparent immediately. Let's say our model system has periodic orbit then $\theta_i(t) \theta_i(t+T)$, where T periodic pacing spikes or bursts (latency). Then in order parameter space one can detect this state in a straight forward manner as a solution $R_\theta(t) R_\theta(t+T)$. This result holds for all i , t . In this case, R_θ dynamics is dominated mostly by the x co-ordinate dynamics. Absolute values of mean order parameter $\text{mod } R_\theta \leq 1$. There is a mathematical relationship of macroscopic global phase measure with macroscopic $V_{\text{mean}}(t)$ in Equation (11). The interval between each microscopic spike/burst in an arbitrary i th stripe of spike/burst can be determined in a statistical-mechanical way by taking into consideration its contribution to the macroscopic global membrane potential $V_{\text{mean}}(t)$. In this interpretation, the time series of the global potential $V_{\text{mean}}(t)$ has a local maxima and minima, respectively and strictly bounded between $[0,1]$. The global cycle in the suprathreshold regime starting from the minimum of $V_{\text{mean}}(t)$ which appears first after the transient time is regarded as the first global cycle, which is denoted by G_1 . The 2nd global cycle G_2 begins from the next following right minimum of G_1 , and so on. Then, we can introduce an instantaneous global phase measure

$\theta(t)$ of $V_{\text{mean}}(t)$ via a linear interpolation in the two successive subregions forming a complete global cycle (Lim and Kim, 2011). A microscopic spike makes the most constructive (inphase) contribution to V_{mean} when the corresponding global phase θ_k for k th cycle of spikes/burst is $2n\pi$ ($n = 0, 1, 2, \dots$), while it makes the most destructive (anti-phase) contribution to $V_{\text{mean}}(t)$ when θ_i for an arbitrary i th cycle of burst is $2(n+1/2)\pi$. By averaging the contributions of all microscopic spikes within a burst in the i th burst stripe to V_{mean} , we can obtain the following degree of ordering of spikes/bursts. Hence, the contribution of k th microscopic burst occurring at the time t_k is ordered by $R_\theta(t_k)$. If the degree of synchrony is high between the bursts/spikes then $R_\theta(t_k) \rightarrow 1$. We quantify the average firing frequency to compare the long-term behavior of individual neurons in the population model. We compute the average frequency (also known as the rotation number) of population of neurons using

$$\rho_i = \lim_{t \rightarrow \infty} \frac{\theta_i}{t}, \quad i = 1, \dots, N. \quad (14)$$

Averaging is carried out over about 1000 neurons starting from random initial conditions after the transient have died out. Collective states of ensemble of $N = 1000$ neurons with spikes per burst $n = 5$ as indicated by their rotation numbers with uniform distribution of frequency I in the interval $[2.1 - \Delta I, 2.1 + \Delta I]$. Different branch of rotation index indicate different dynamical states of the network as a function of mean field coupling strength K , ΔI . We carry out a grid search in the 2D parameter space K , ΔI . Our goal is to obtain a phase transition diagram to understand long-term collective behavior of Equation (8) for large N , as a function of the coupling strength $K \geq 0$ and the stimulus spread $\Delta I \in [0, 1)$. Global order parameter $R_\theta(t)$ is computed for different parameterization of K , ΔI and embedded on a contour plot. Color spectrum is the same as the one used for displaying phase diagram in the full network. The values which are high and close to 1 are indicated by red and the values which are close to zero are indicated by blue.

2.7. CLUSTERING ANALYSIS IN N COUPLED FULL AND PHASE NETWORK MODEL

We describe firing patterns in large networks (finite N) with excitatory mean field coupling in terms of array diagrams. Array diagrams are obtained by simulating a coupled system consisting of mean field coupled biophysical neurons ($N = 100$) governed by the Equation (5). All the coupling coefficients are the same K where $i = 1, \dots, N$. In the arrays the intensity of the voltage variables V_1, \dots, V_i have been encoded in color spectrum. Two different color spectrums are used for the biophysical network (see Figure 4). In the first color spectrum blue part of the array values implies the quiescent activity of the spikes where the voltage variables have relatively lower values. All the other colors in the spectrum indicates the higher values for the voltage variables, consequently these pixels in the array imply the spike activity. The horizontal line of the array shows the time with increasing epochs of activity. The second color spectrum used here shows burst depiction in the nearly coherent parameter regime. Green colors in the array indicate completely silent neurons. Purple pixels on the green background shows burst activity. On the vertical

axis neuron index are aligned and again, on the horizontal axis gives the direction of time. These diagrams were obtained from a phase network by monitoring phases of individual neurons $i = 1, \dots, N$ and aligning them on the vertical axis. The choice of the color spectrum used for phases is given by a colorbar with uniformly distributed phase values. In Figure 9 red color index in the spectrum corresponds to higher phase values of θ (close to π) and orange color index are for lower phase values (close to $-\pi$). First initial conditions $\theta_i(0)$ is generated randomly and then they are sorted according to their neuron index and subsequently distributed uniformly about $[-\pi, \pi]$. The parameters K , ΔI , for both realizations are chosen from SR, IN regime of the respective phase diagrams.

2.8. NUMERICAL PROCEDURES AND VISUALIZATION OF THE SYSTEM DYNAMICS

Two network models were implemented in Matlab, numerically integrated using second order Runge Kutta routine and Euler-Maruyama (EM) method (Higham, 2001). The simulations were performed with a fixed time step of $dt = 0.05$. The first 200 time points of the simulation are disregarded to set the network to a steady state. Thus, the results within this time were ignored. The membrane potential $V(t)$, standard deviation of membrane potential $\text{std } V$, mean field $\Gamma(t)$, order parameters $R(t)$, $R_\theta(t)$ are captured for the entire population. For full network, simulation is carried out for $N = 100$ neurons and for the phase network for $N = 1000$ neurons. Numerical Phase diagrams are obtained using parallel for loops implemented in Matlab. Coupled mean field Phase model represented in Equation (8) can be visualized as a collection of N points rotating around the unit circle, where the estimated phase for each neuron $\theta_i(t)$ denotes their position on a ring or a circle at time t . This alternate representation of the dynamical system (as N points moving along a circular reference frame, instead of a single point tracing out a trajectory in an N -dimensional phase space) is possible because the system's state space, the N -torus, is equivalent to N copies of the unit circle. It is worth noting that for most other N -dimensional state spaces such a reduced representation is not feasible. In order to distinguish between oscillators with different natural frequencies, we color the dots according to the standard color spectrum: the neurons correspond to the low end of the spectrum (close to $-\pi$) (red), neurons at the high end (close to π) (blue), while those in between occupy the middle part of the spectrum (orange/yellow/green). To show how the system evolves from one instant to the next, we plot a series of snapshots of the system at different times (see Figures 11B–D, for example). This allows us to observe the behavior of individual neurons at the same time as we witness the collective evolution of the system toward an attractive state.

3. RESULTS

3.1. SINGLE NEURON BURST DYNAMICS

We first examine the behavior of single neuron model Equation (2.1) as the applied input current I is brought close to the threshold for generating spikes or bursts. For the given parameters in Equation (2.1) a neuron is excitable. Figure 1 depicts the relationship between applied input and parabolic bursting pattern. We are only interested in the behavior of this

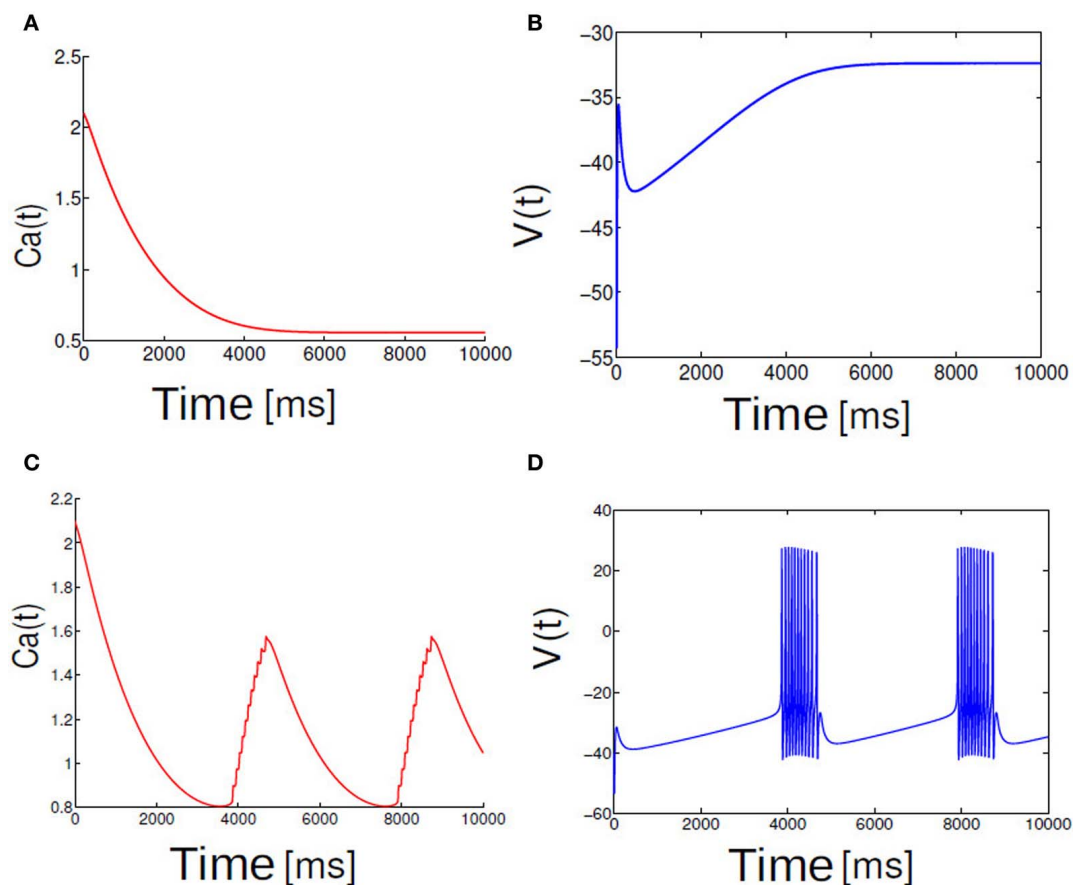


FIGURE 1 | Shown here trace of membrane potential and calcium dynamics. Fast spikes rides on a slow modulation of calcium. Slow subsystem moves Ca back and forth across SNIC bifurcation points

(A,C). In (B,D) membrane potential dynamics is shown for two different cases (A) $I < I_c$ and (B) $I > I_c$ in the single neuron model.

system for low current values where the resting state of membrane voltage is sufficiently depolarized below -40 mV. When the applied input current I is below a critical value membrane potential $V(t)$ maintains their steady state value and for values greater than the threshold exhibits bursting behavior. For the parameterization used here we find that at $I \geq 60$ steady state destabilizes exhibiting multispikes. When the applied input current is further increased a neuron make transition from bursting to spiking behavior. In order to observe a typical spiking behavior we set $I \geq 70$. To get an intuitive understanding about the relationship between slow and fast subsystems, Rinzel and Lee analyzed this model by varying ca (a variable in the slow subsystem) as a bifurcation parameter to report that parabolic bursting is obtained from an oscillation in the slow subsystem that periodically moves the ca variable back and forth across the SNIC bifurcation, to link the steady state solution of this system to (quiescence state in **Figure 1A**) the branch of periodic solutions (**Figure 1B**) and vice versa. Time series of fast variable shows that the interspike interval is relatively longer at the beginning and end of each burst. As has been shown by numerous authors oscillation for the fast dynamics is obtained when the slow variables are held fixed; it is where the saddle-node-loop

bifurcation occurs. There is a clear threshold below which there is a unique stable fixed point. Parabolic bursting can occur without having any bistability in the spike generating process. One way to achieve parabolic bursting behavior without requiring any bistability in the generating process and moreover, mathematically tractable would require a generic description like the one shown in Equation (2.2) (see section 2). From numerical results we find that as the applied input current $I \rightarrow 2$, time period $T \rightarrow \infty$. Applied input current can be tuned such that it is possible to obtain parabolic bursts of desired interburst gap. The time evolution of a single neuron activity is shown in **Figure 2**, where a membrane potential like variable $V(t) = -\cos[\theta(t)]$ is plotted by numerically integrating Equation (2.2). Temporal dynamics shows regular parabolic bursting behavior. For $I < I_c$ a neuron fires few spikes before it settles into a steady state. For $I > I_c$ ($I = 2.01$, $n = 5$) neurons exhibits parabolic bursting behavior. Based on the qualitative similarity in the burst pattern with parabolic bursting neurons (At the start and the end of the active phase the spike frequency is smaller compared to the middle of the active phase as can be seen in **Figure 2** detailed model is substituted to investigate the network effects.

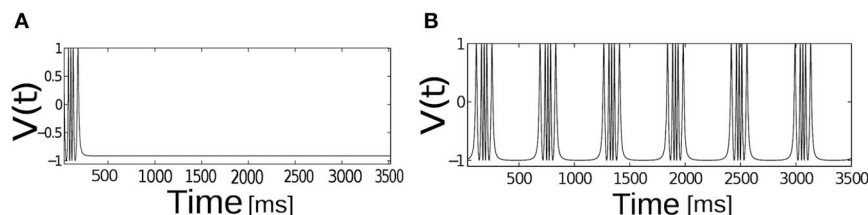


FIGURE 2 | The temporal dynamics of the phase model of spike-burst neuron. $V(t) = -\cos[\theta]$ is plotted as a function of time, $l_c = 2.01$, $n = 5$, for (A) $l < l_c$ and for (B) $l > l_c$.

3.2. NETWORK DYNAMICS OF PARABOLIC BURSTING NEURONS WITH HETEROGENEITY

To understand the influence of heterogeneity in the applied input current and the coupling strength in a network of single neurons exhibiting parabolic bursting we use Equation (5) and parameters as described in section 2.4. We use fast excitatory synapses to couple these units. When the synaptic coupling is sufficiently fast, the coupling tends to push the neurons toward anti-synchrony (Wang and Rinzel, 1992; Friesen, 1994; Van Vreeswijk et al., 1994). Moreover, several studies have observed emergence of multistable solutions in their mean field network with parameter heterogeneity (Assisi et al., 2005; Jirsa and Stefanescu, 2011). Our motivation is to go toward this particular direction to capture the relevant network dynamics at the population level. In particular to understand the combined effect of heterogeneity in the firing rate threshold (biophysical model) with the fast time scale of activation-deactivation of synapses in the coupling; the interplay between these two critical factors in spike/burst timing at the population level is largely unknown. In our formalism their individual and combined influence on the network dynamics become clearly visible. Typical time course of such responses of synaptic variable in our model simulation are shown in **Figure 2**. Fast synapse approximation holds as long as the variable s_i relaxes much more rapidly than V_i , in which case we may apply a quasi-static approximation to reduce s_i further in Equation (5), $s_i \simeq 0$ allowing us to adiabatically eliminate β , and set the synaptic variable via an approximation as $s_i = \frac{\beta}{1 + \beta + \exp\left(\frac{-V_i}{2}\right)}$. The mean

synaptic action can be formulated as $\Gamma = \frac{1}{N} \sum_{i=1}^N s_i$, where $a_s(V_i) = \frac{1}{(1 + \exp(-V_i/2))}$. The synaptic constant K is the same for all the neurons. **Figures 3A–D** shows kinetics of excitatory synaptic variable s_i (plotted with black solid lines) for different β values. Mean synaptic variable (plotted with dotted lines) for the same set of values of time constant β shows dissimilar temporal response compared to s_i for higher time constant values. For smaller time constant values simulation provides relatively better agreement as can be seen from **Figure 3**. We numerically integrate the above network to investigate how the mean population burst changes with time as a function of spread of applied stimulus ΔI and mean field coupling strength K . Firing patterns in this network are shown with array diagrams in **Figures 4A,B**. For small spread in the applied stimulus and sufficient coupling strength $\Delta I = 0.001$, $K = 0.7$ nearly burst synchronization takes place. Moreover, in the array diagram we detect clusters of synchronous

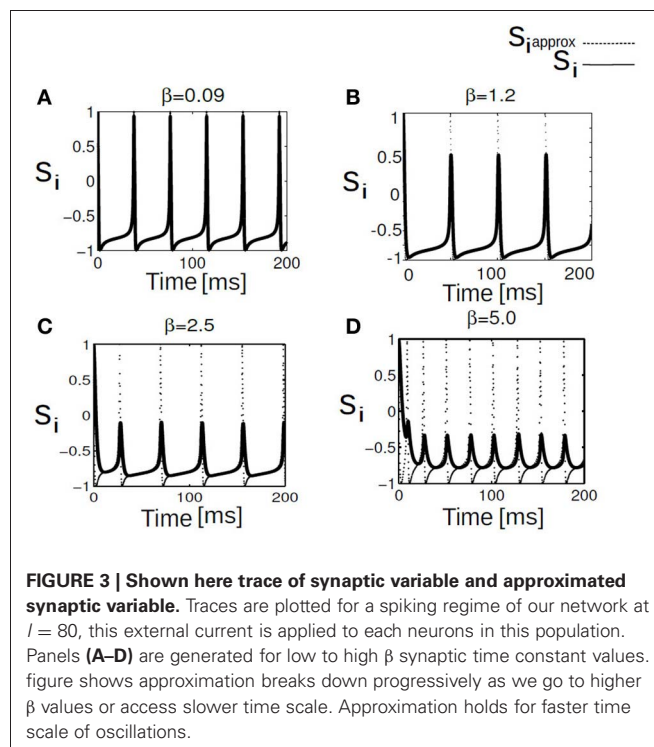
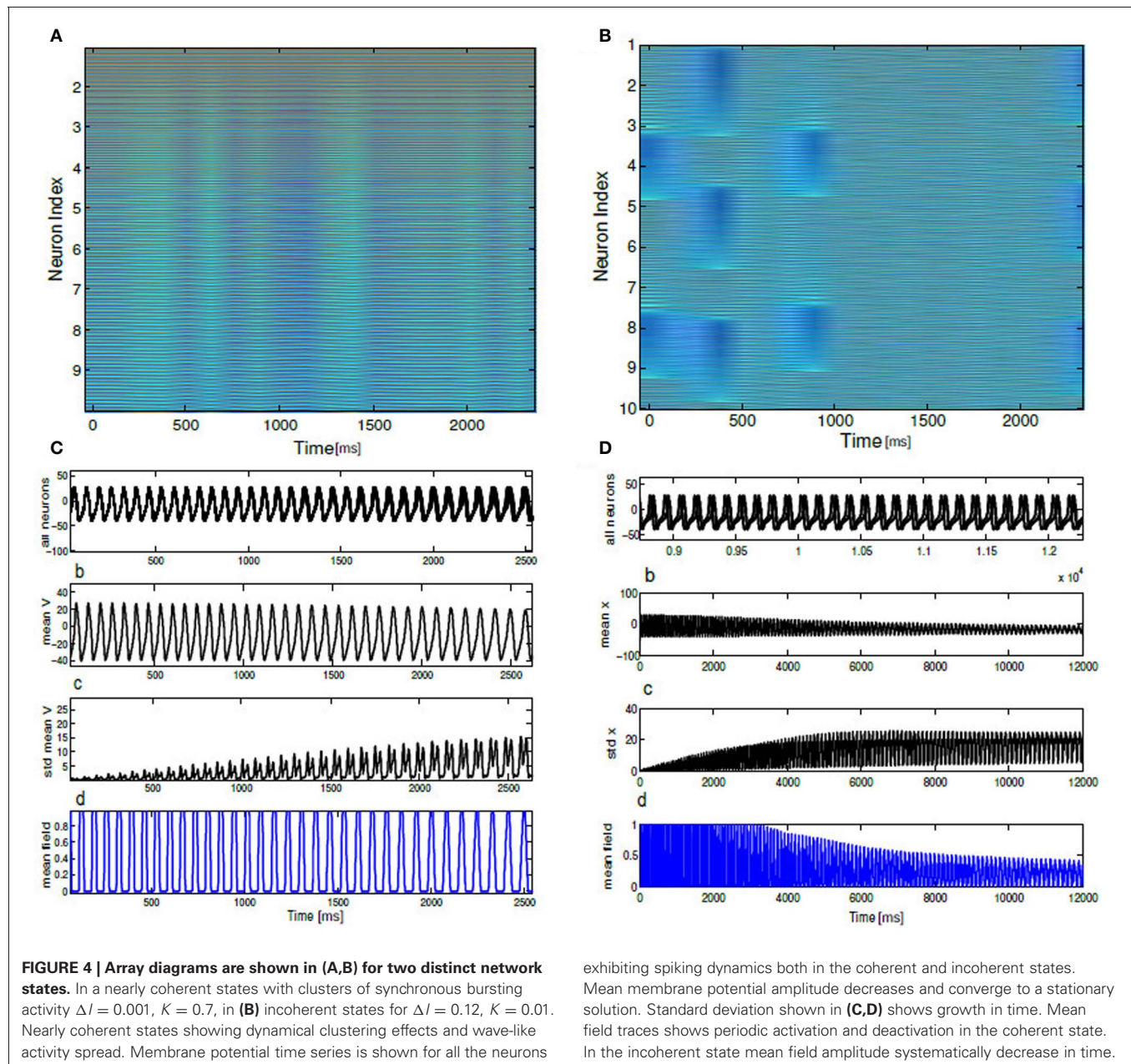


FIGURE 3 | Shown here trace of synaptic variable and approximated synaptic variable. Traces are plotted for a spiking regime of our network at $l = 80$, this external current is applied to each neurons in this population. Panels (A–D) are generated for low to high β synaptic time constant values. figure shows approximation breaks down progressively as we go to higher β values or access slower time scale. Approximation holds for faster time scale of oscillations.

states which fires in a wave-like pattern. Corresponding time series of mean quantities such as the membrane potential $V_{\text{mean}}(t)$ in Equation (11), mean field $\Gamma(t)$ shows periodic activity in **Figure 4C**. Membrane potential spiking activity is nearly synchronized across population of neurons in **Figure 4**. On the other hand, for the IN state mean membrane potential fast decays to zero and shows subthreshold fluctuations about mean zero. Response of mean membrane potential is more suppressed compared to their mean field oscillations between $[0,1]$. Amplitude of mean field $\Gamma(t)$ changes in time systematically but fluctuates about the mean value of 0.5 instead of approaching zero values as can be seen in **Figure 4D**. Population burst synchrony is observed for many different parameterization, for one such choice of parameter $\Delta I = 0.002$, $K = 0.8$, an array diagram is computed and plotted in **Figure 5A**. As can be seen in the figure a wave-like spread of activity. In **Figure 5B** various time series plots of population burst synchrony is shown across 10 neurons. In order to identify different network states for all possible combination



of two parameters K , ΔI we carry out a grid search and compute the values of R in Equation (12). Global order parameter measure identifies three distinct network states in the parameter space as shown in **Figure 6A**. For low coupling values K , order parameter shows fluctuations about mean zero. In this regime each neurons activity is mainly driven by their firing rate threshold and displays largely incoherence. For medium values of both coupling strength K and stimulus spread ΔI network exhibits a hybrid state (some neurons are firing and some of them are silent). For very small values of stimulus spread and medium to high K values nearly burst synchrony appears. Temporal dynamics of membrane potential activity $V(t)$ for four neurons are plotted in **Figures 6B–D** for three arbitrary parameterization of our network model. In **Figures 6B,D** PO state is shown where

one neuron is spiking or bursting and three neurons are silent. In **Figure 6C** all neurons are showing nearly synchronized parabolic bursting behavior.

3.3. NETWORK DYNAMICS OF PARABOLIC BURSTING PHASE MODEL WITH HETEROGENEITY

In this section, we use a phase network with mean field coupling to get some insights about the novel network states observed in $(K, \Delta I)$ the parameter space of the full network model. Coupling between each phase neuron via a mean field is formulated in section 2.4. Numerically we integrate Equation (6) to compute time averaged membrane potential, mean field Γ as in Equation (10) (see section 2), global measure of coherence R_θ as a function of K , ΔI a parameter combination which is used in

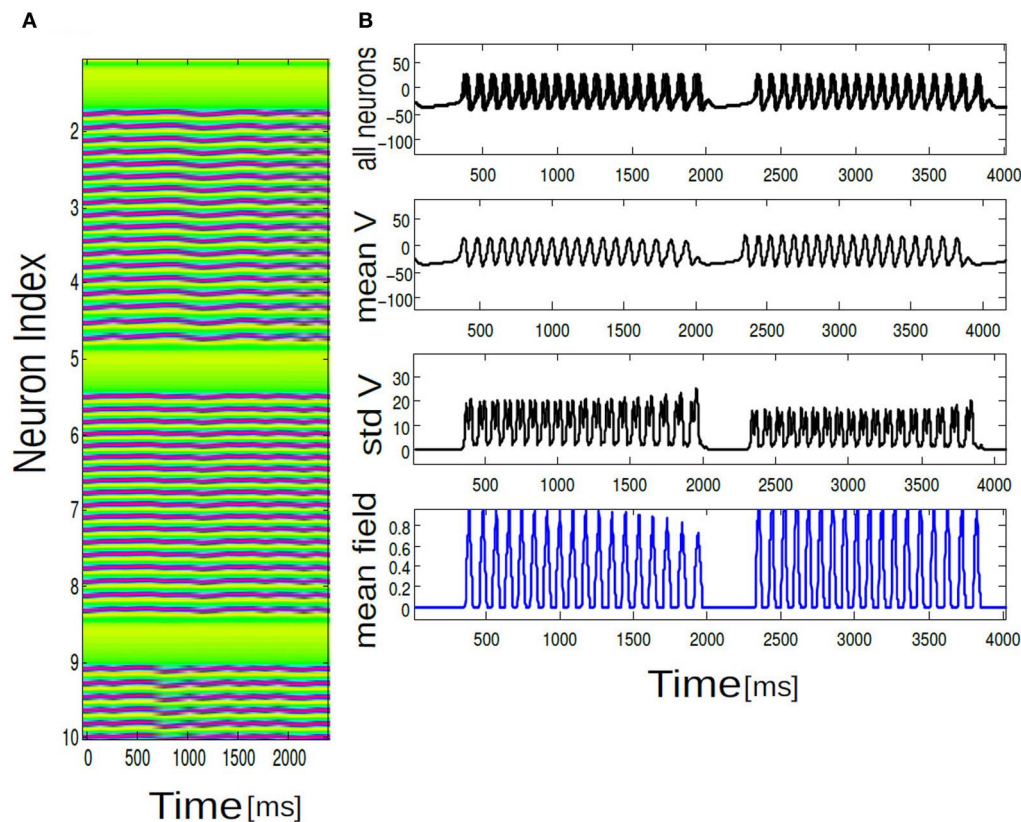


FIGURE 5 | In (A) array diagram showing firing pattern in a population of 100 neurons. Only 10 neuron index are shown for clarity. Horizontal axis is always time and vertical axis is labeled as neuron index. Green color corresponds to no firing activity or quiescence. Purple pixels corresponds to

parabolic bursting activity of each individual neurons which are locked in time. In **(B)** time series data for membrane potential of $V(t)$, $V_{\text{mean}}(t)$, $\Gamma(t)$, and $\text{std } V_{\text{mean}}(t)$ are plotted for 10 neurons. Mean population burst synchronizes in time.

the detailed network model. Time evolution of the above quantiles are shown in **Figure 7** for a parameterization $K = 0.8$, $\Delta I = 0.001$. The parameter choice is the same as the full network investigations. With this combination of parameters all the neurons synchronously spikes. Mean membrane potential-like quantity $V_{\text{mean}}(t)$ oscillates in phase with synchronized spike activity as plotted in **Figure 7B**. Here, n a quantity which determines the number of spikes per burst is kept at $n = 1$. Mean field Γ also shows up and down states (Locked in time) and act as an oscillating drive to each individual neurons. The time series of the global order parameter $R_0(t)$ for synchronized spiking is periodic in **Figure 7E**. Next, we show in **Figure 8** temporal evolution of the mean quantities for the choice of $K = 0.8$, $\Delta I = 0.5$. For medium values of mean field coupling strength and stimulus spread network shows PO behavior, where some of the neurons are firing incoherently and others are completely silent. This means for some parameterization network has two stable branches of solutions. It is important to note PO state of the network was observed in the full network for a comparable parameterization (see **Figure 6**). Time series for 10 neurons and their order parameter evolution in time is plotted in **Figure 8**. Three neurons are completely silent while other seven neurons are bursting with variable inter-burst intervals. As there is no noise in this system

and coupling magnitude is set at high values as in the case of sync, this variability must be introduced by the heterogeneity in their individual firing rate threshold via stimulus spread.

Figure 6 shows the parameter space diagram for the full and phase models presented in Equations (5) and (6–9). Phase boundaries are calculated by computing the mean field for both full and the phase model for different combination of (K, I) values on a two-dimensional grid. In the following subsection we would lay out the details for obtaining the phase transition boundaries semi analytically. Interestingly, Over a wide range of $(K - \Delta I)$ values the collective dynamics of the two networks primarily show three distinct regions of interest which are close to each other in the parameter space. For sufficiently large K values holding ΔI fixed to a narrow range of values near zero, the system converges to a state of partial oscillations in which the some of the neurons are not firing altogether, while the others display IN oscillations. Very large K values result in damping of oscillation activity and all the neurons stops firing altogether. The stability state of locking is much more difficult to achieve and in fact we found distinct branches in their rotation number, these states should all be regarded as variants of 1:1 locking, and therefore we lump them together in the locked region of the stability diagram. With further increase in ΔI , parameter heterogeneity,

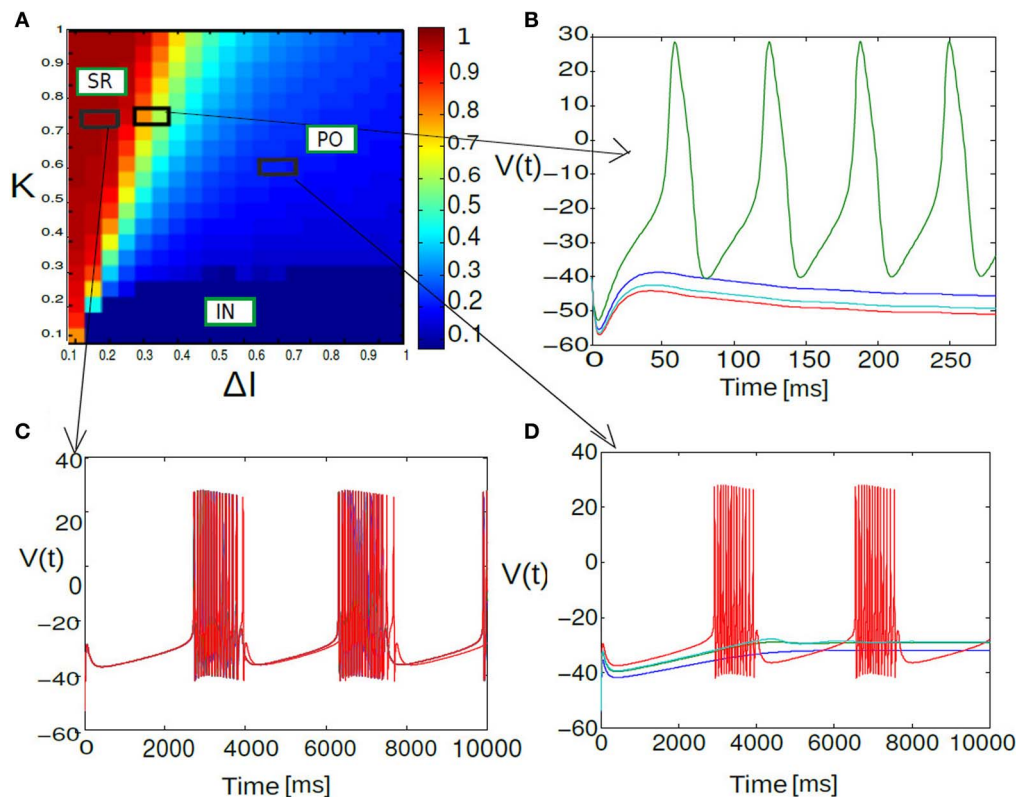


FIGURE 6 | Phase diagram of mean synaptic action variable is shown as a function of 2D parameter space of stimulus spread ΔI and excitatory coupling strength K . In the partial burst regime labeled as PO in (A), a subset of neurons are not firing at all as their respective drives are below their individual firing thresholds. Heterogeneous dynamics between synchronized population spiking activity and oscillation frequency death response for PO state is displayed in (B). Nearly

synchronized population of spike/burst activity lumped in a regime labeled as SR [corresponding time series is displayed in (C)] and incoherent population spike/burst activity is lumped into a regime called IN. In the incoherent regime mean field values stay close to zero with substantial subthreshold fluctuations. In (D) multi stability of PO state is displayed again; now between population burst and fixed point dynamics for an entirely different parameterization.

successively more neurons peel away until eventually the entire population is IN.

4. PHASE DIAGRAM USING SEMI ANALYTICAL METHODS FOR MEAN FIELD PHASE MODEL

Mean field coupled neurons in phase model is described in Equation (6). Let's rewrite the mean field equation explicitly.

$$\dot{\theta}_i = (F(\theta_i) - \Gamma \sin \theta_i (\cos \theta_i - V_{th})) \quad (15)$$

$$\text{where } \Gamma = \frac{K}{N} \sum_{j=1}^N \frac{\beta}{\left(1 + \beta + \exp\left(-\frac{\cos \theta_j}{2}\right)\right)}.$$

In a semianalytical approach we would like to understand the phase transitions between three distinct network states discovered in two networks. For the IN states where the average firing frequency increases monotonically plotted in **Figure 10**, the θ_i are all distributed across the closed orbit in a unit circle. This leads to the following phase evolution equation SR state may undergo instability either through parameter changes of K or ΔI and make phase transition to either IN or PO state. Mean field Γ approaches a stationary density as the number of neurons are increased in both PO and IN state (see **Figure 4D**). Hence,

Γ approaches some positive real number for these two states. When varying K , we consider small perturbations μ to the SR solution $\theta = \theta_i = 0$. With $\theta = \theta_i = 0 + \mu$ Equation (15) becomes $\dot{\theta}_i = \dot{\mu} = F(\mu) + \frac{\Gamma}{2} \sin(2\mu)$ and linearization yields $\dot{\theta}_i = (F'(0) + \Gamma)\mu$. Moreover, SR state may get phase locked at $\theta = \pi$ (subpopulation clusters). Hence, $\theta = \pi$ may get destabilized as we changed the width of heterogeneity by changing ΔI or the coupling strength K . Similarly, we consider small perturbations μ about solution $\theta = \pi$. Hence, we can write $\theta_i = \pi + \mu$, $\dot{\theta}_i = \dot{\mu} = F(\pi + \mu) + \frac{\Gamma}{2} \sin(2\pi + 2\mu)$ and linearization yields $\dot{\theta}_i = (F'(\pi) + \Gamma)\mu$. With $F(\theta) \approx I - \cos \theta - \cos \frac{\theta}{n}$ for the SR state, we find that $F'(\theta) \approx -\sin(\theta) - \frac{1}{n} \sin \frac{\theta}{n}$ will be generally small for $\theta = 0, \pi$. SR state solution hence becomes unstable when $F'(0, \pi) + \Gamma = 0$, which suggest almost a vertical critical line between SR and PO, IN state. The bifurcation route from PO (multistable state) to IN solutions as the parameter ΔI increases is less conclusive in the framework of the circular approximation, since in the previous stability analysis the only I -dependent term is $F'(0, \pi)$, which is very small, hence higher orders of the approximation must be considered. We use the following ansatz: If r is the radius of a unit circle, any smooth deformation from a unit circle can be approximated as, $r(\theta_i) = 1 + \epsilon h(\theta_i)$. Hence we can

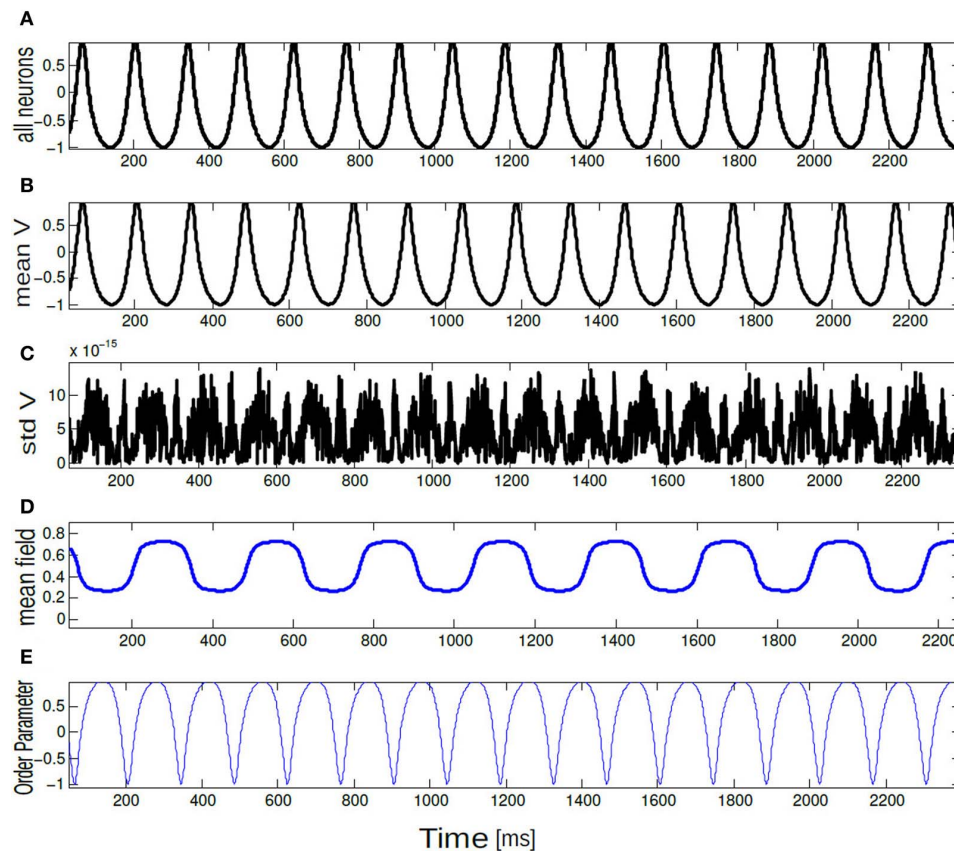


FIGURE 7 | Time series of (A) $V(t) = -\cos \theta_i(t)$, (B) $V_{\text{mean}}(t)$, (C) $\text{std } V_{\text{mean}}(t)$, (D) mean field Γ , and (E) order parameter $R_0(t)$ are plotted for 10 neurons. All neurons are spiking in synchrony and time locked. The parameter values are $K = 0.8$, $\Delta I = 0.001$.

compute the non-linear flow contribution with the above first order correction term as $F(\theta_i) + \epsilon H(\theta_i)$. It is possible to explicitly determine $H(\theta_i)$ for a certain choice of $h(\theta_i)$ and moreover, $H(\theta_i)$ has a periodicity of π , that is $H(\theta + \pi) = H(\theta)$. Thus the linear stability analysis about the fixed point $\theta_i^* = 0 + \mu$ gives

$$\dot{\mu} = (F'(0) + \epsilon H'(0) + \Gamma)\mu = (F'(\pi) + \epsilon H'(\pi) + \Gamma)\mu \quad (16)$$

From the above equation with $F'(0) \approx 0$ and the π -periodicity of $H(\theta_i)$, we find that the two fixed points at $0, \pi$ lose stability at the same time for increasing ΔI and as a result leads directly to the IN state. Since $H'(\theta_i) \sim I$, scales linearly for fixed μ , we can also estimate the critical line of transition in the parameter space in **Figure 11** which separates PO state from IN state. For the critical line: $H'(\theta_i) = m(I - I_c)^p$ where m is the slope of this line and $m > 0$ allows for destabilization. Hence the critical condition is $\epsilon H'(0) + \Gamma = 0$. By substituting the dependence of $H'(\theta_i)$ on (I, I_c) and in turn dependence on ΔI one can write $\epsilon m(I - I_c)^p + \Gamma = 0$. This implies coupling strength $K = -\epsilon(\Delta I + \Delta I_c)^p$ for $(m > 0)$ and p is some exponent representing a scaling relationship near saddle-node bifurcation. Thus the critical condition is $|K| = \epsilon m(\Delta I - \Delta I_c)$, which serves as a convenient guide to numerically compute the stability line separating PO region from IN. Next we try to obtain analytically

the stability boundary between INC and PO oscillation states in the infinite- N limit. it turns out that the IN and partial oscillation states can be made steady in our system. The possibility of doing so was suggested by the numerical results. In numerics we observed that as the number of neurons N is increased, the order parameter R_i approaches a constant for both these states **Figure 8** and the oscillators tend to arrange themselves in a stationary distribution around the circle **Figure 11**. The way to approach these two states analytically, therefore, is to first write down the appropriate infinite- N analog of our model.

$$\frac{\partial}{\partial t} f + \frac{\partial}{\partial \theta} \left\{ \left[F(\theta) - \left(\left(\int_0^{2\pi} \int_{I-\Delta I}^{I+\Delta I} \Gamma p(\theta', t, I') g(I') dI' d\theta' \right) \times \sin(2\theta) \right) p \right] \right\} = 0 \quad (17)$$

The above equation is the infinite- N analog of continuity equation introduced earlier. It is a non-linear partial integro-differential equation for the number density $f(\theta, t, \omega)$. In addition we demand f to be non-negative, 2π periodic in θ , and we impose the normalization

$$\int_0^{2\pi} f(\theta, t, I) d\theta = 1, \quad (18)$$

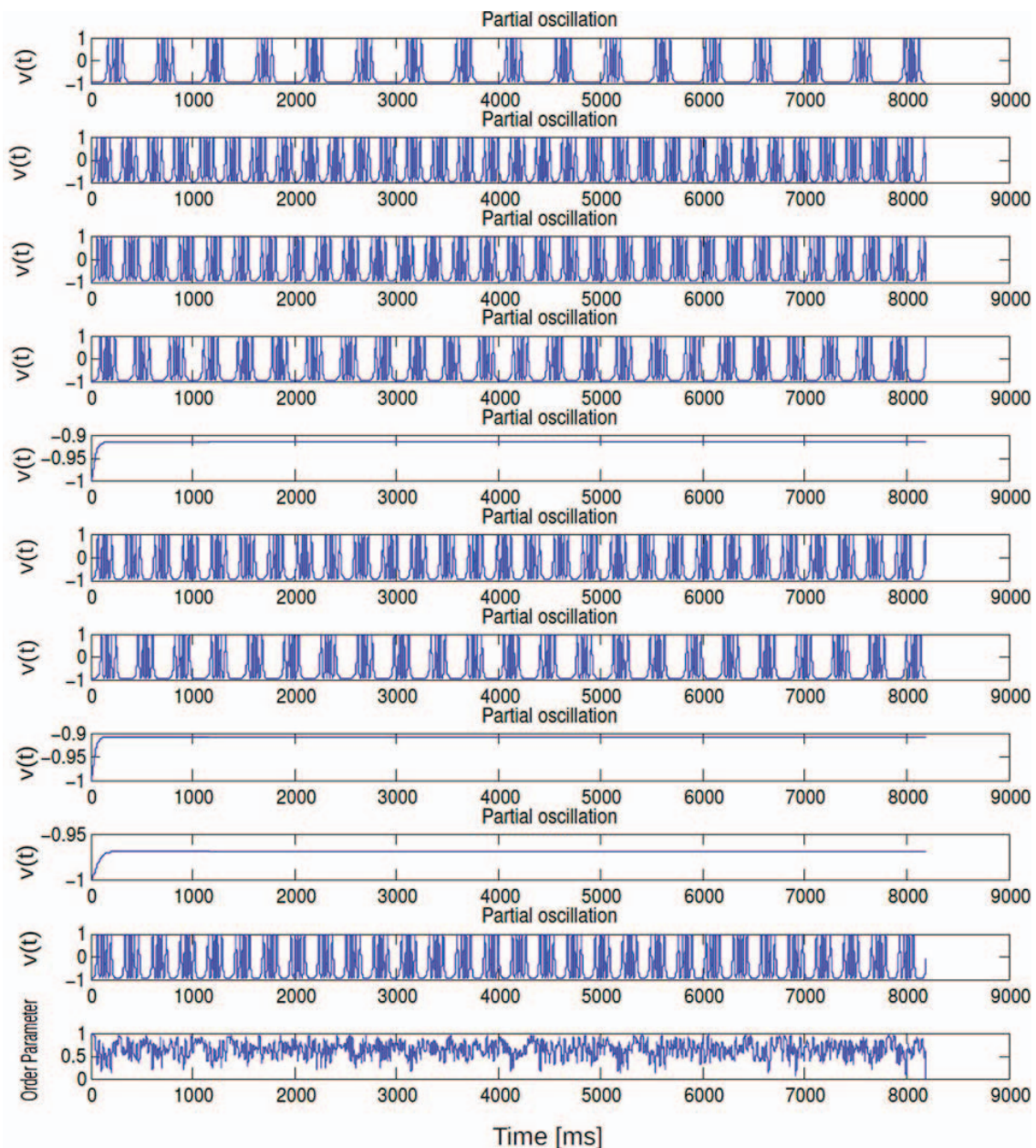


FIGURE 8 | Temporal evolution $V_i(t)$ for $N = 10$ neurons are shown here for an arbitrary parameterization $K = 0.8$, $\Delta I = 0.5$. K value is unchanged from previous figure. Stimulus spread ΔI is changed. Time series of order

parameter $R_0(t)$ undergoes statistical fluctuations of magnitude $O(\frac{1}{\sqrt{N}})$ about some positive constant value. After 8000 time points, dynamics is truncated assuming network dynamics settles into a steady state.

For incoherence and partial oscillation the above system tends toward a stationary distribution of phases in time. The above two states are the fixed points of the stationary density in the continuum limit. To solve for the fixed points we set $\frac{\partial}{\partial t}f = 0$ in Equation (10). let's assume that $f_0(\theta, \omega)$ be the stationary phase density and $v_0 = [F(\theta) - ((\int_0^{2\pi} \int_{I-\Delta I}^{I+\Delta I} \Gamma f(\theta', t, I') g(I') dI' d\theta')) \sin(2\theta)]$ be the velocity field. Then one can write

$$\frac{\partial}{\partial \theta}(f_0 v_0) = 0 \Rightarrow f_0 v_0 = L(I) \quad (19)$$

where $L(I)$ is a constant which is determined exactly by using normalization condition. Depending on it's applied drive I , neuron's steady state behavior falls in the following two categories:

Case (i) When $I \ll \Gamma$ implies

$$v_0(\theta, I) = F(\theta) - \Gamma \sin(2\theta) = 0 \quad (20)$$

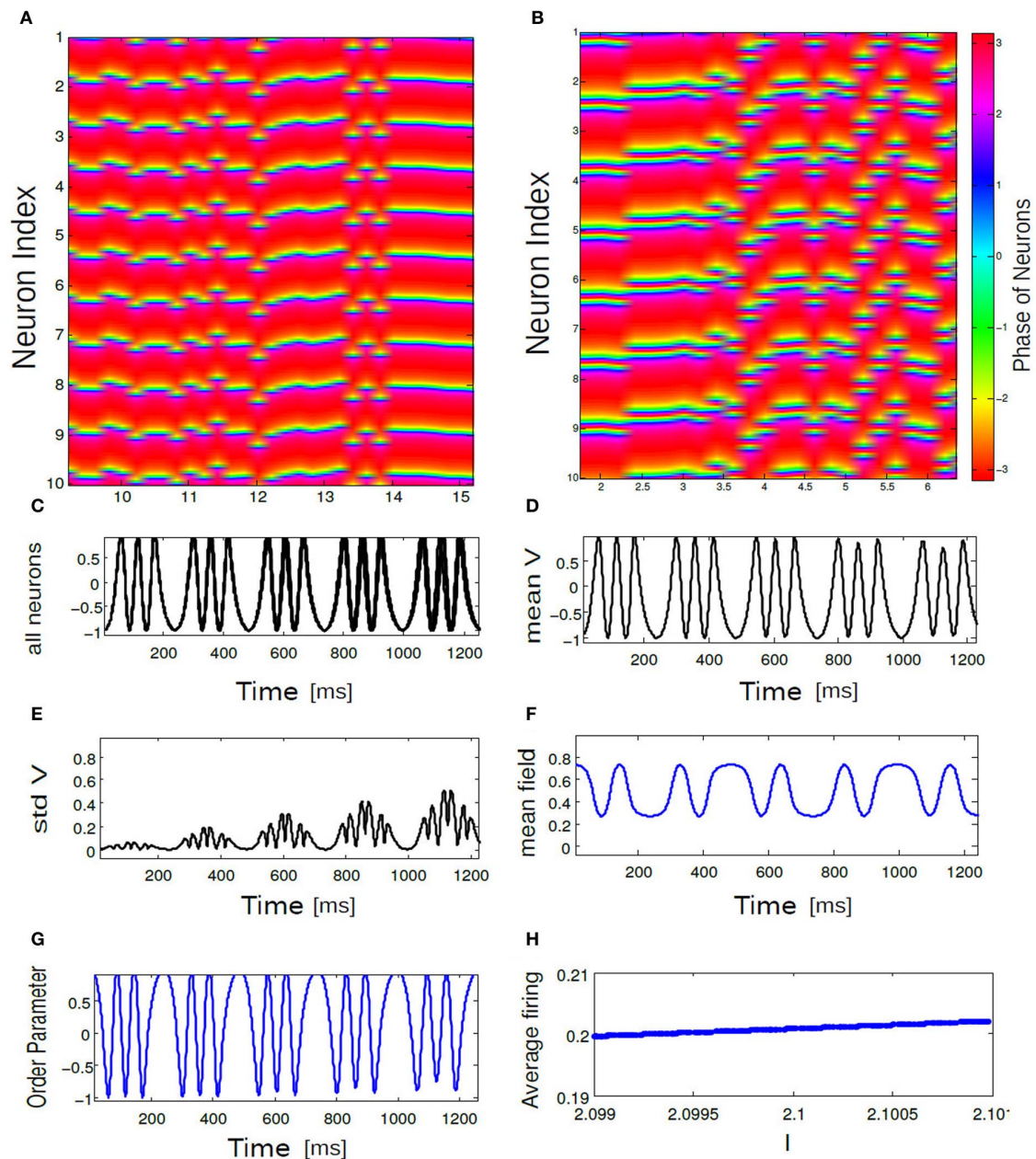


FIGURE 9 | In (A,B) an array diagram is shown for phase network model for a parameter combination. In (A) $K = 0.61$, $\Delta I = 0.001$, $n = 1$ spikes only and in (B) $K = 0.001$, $\Delta I = 0.15$, $n = 3$ bursts only. Almost always, near synchronous burst states are observed for high K and low ΔI values. In (C) corresponding time evolution of $V_i(t)$ is shown for all 10 bursting neurons. In (D) temporal response of V_{mean} is shown for the coherent state of our

network. In (E) standard deviations of V_{mean} is plotted as a function of time. In (F) mean field Γ vs. time for the coherent state is shown. (G) displays temporal dynamics of order parameter. Average firing frequency as described in Equation (14) is plotted in (H) for the parameter combination of $K = 0.001$, $\Delta I = 0.15$. Panel (H) further demonstrates phase locking behavior among all the neurons.

Case (ii) When $I \gg \Gamma$ neuron fires incoherently and typically individual phases follows an uniform distribution about the unit circle. In this case the velocity field turns out to be,

$$v_0(\theta, I) = F(\theta) - \Gamma \sin(2\theta) \quad (21)$$

Fixed point solution demands that the density must be inversely proportional to the velocity:

$$f_0(\theta, I) = \frac{L(I)}{F(\theta) - \Gamma \sin(2\theta)} \quad (22)$$

In the IN state, neurons driven by different external drives are firing at different phases, however, their collective state is close to being stationary. Every neuron belong to Case (ii) as described above. Further, it is possible to derive nearly an exact relationship between K , ΔI that gives the transition from case (i) to

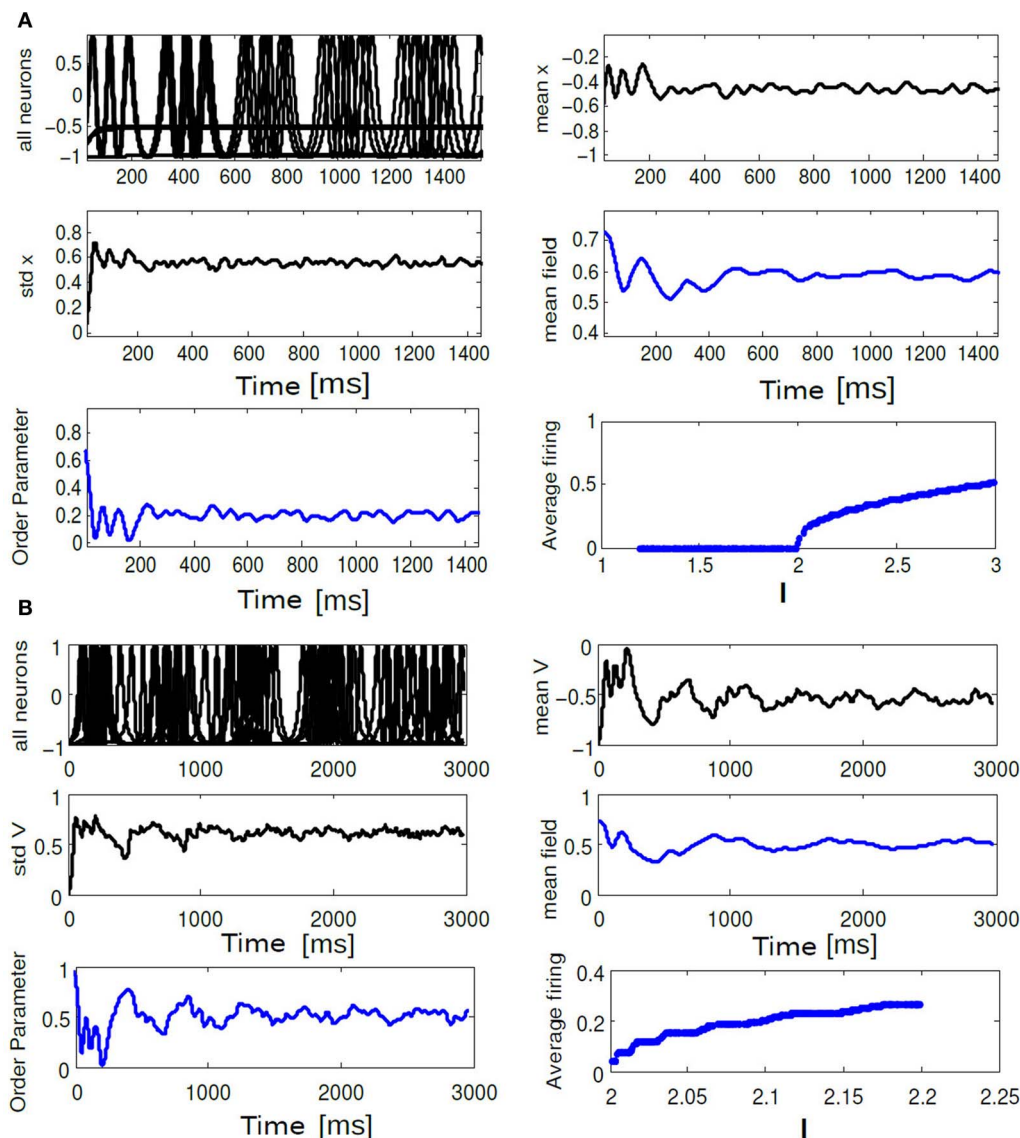


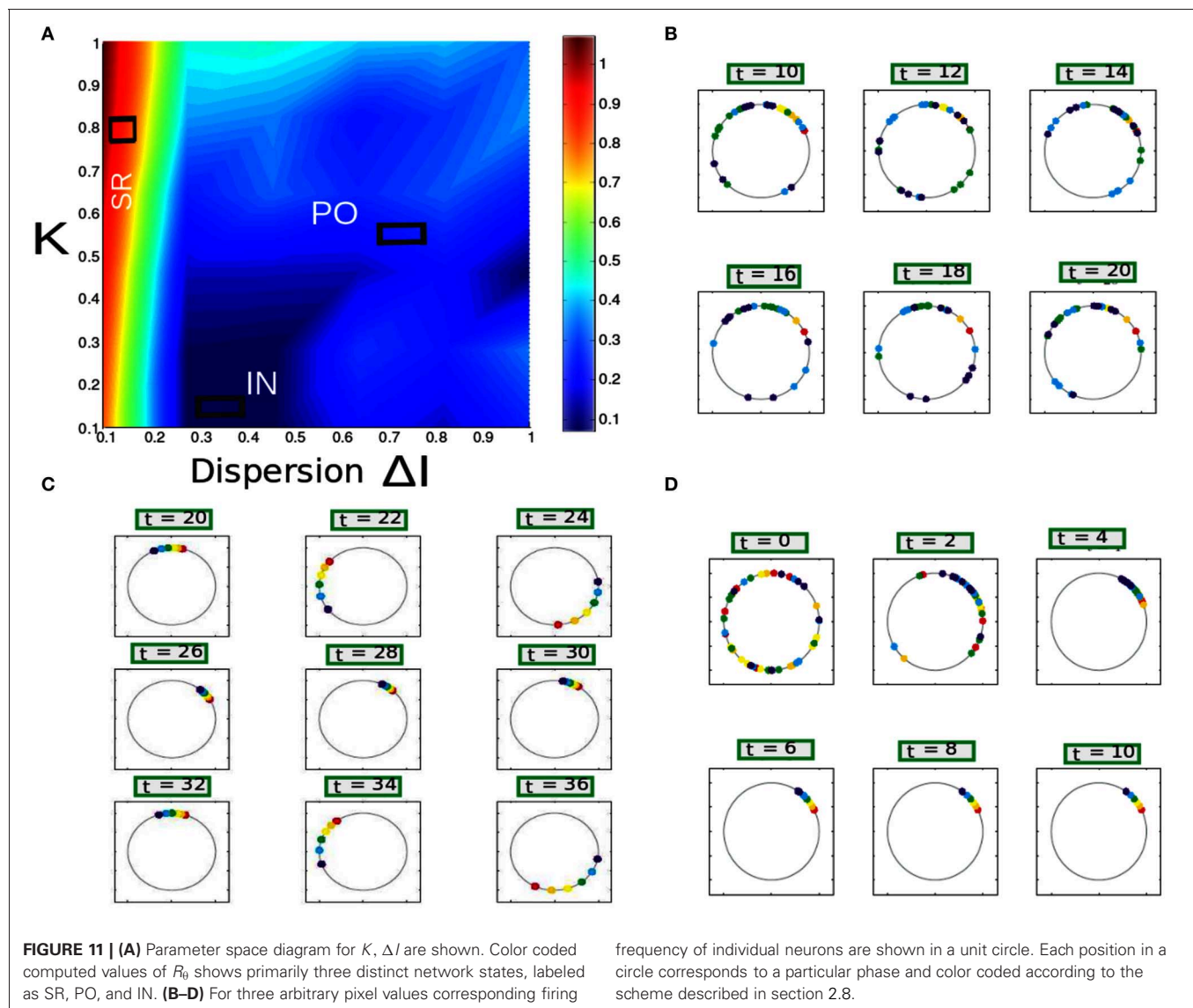
FIGURE 10 | In (A) time evolution of $V_i(t)$, $\text{std } V_{\text{mean}}$, order parameter $R_0(t)$, V_{mean} , mean field Γ , Average frequency are shown for the choice of $K = 0.6$, $\Delta I = 0.3$. In (B) time evolution of the same quantities are shown for

$K = 0.01$, $\Delta I = 0.1$. In (A), average firing frequency plot shows clusters of neurons firing incoherently while another cluster of neurons are completely silent. In (B) same subfigure shows a monotonic increase in average firing frequency.

case (ii) as described above. As shown before, in case of a finite size network such a relationship in the first order perturbation $\epsilon m(I - I_c)^p + \Gamma = 0$ does exist. In this scenario those neurons with a minimum bound on their applied drive I_{\min} reach cessation of firing as we find from numerical simulations. They then fall into the Case (i) above where mean field Γ exerts much bigger influence on the dynamics and overall effect is damping of firing activity. The first neurons to stop firing are the ones which do not cross the threshold for firing which in this case $I > 2$. Then the boundary that separates IN from PO in the phase diagram is almost a straight line given by,

$$|K| = \epsilon m(\Delta I - \Delta I_c) \quad (23)$$

Hence, both finite and infinite analog of our network identifies the putative transition boundary between IN and PO states. Now from numerical simulations we find Andronov–Hopf (AH) bifurcations leads to the transition from INC to SR solutions in the **Figure 6** near K, I values close to zero. It is equivalent to look at the imaginary eigensolutions that arise due to the instability of the IN state. This instability requires calculation of higher order perturbation terms of the stationary density obtained at the IN state of our network. This is out of the scope of our paper, however, we show a numerical fitting result which gives an empirical relationship between K and ΔI to quantify the transition boundary between IN and SR states. Assuming ϵ is the perturbation to the IN



solution we can express a relationship between K and ΔI as follows,

$$|K| = a_0\epsilon + a_1\epsilon^2 + a_2\epsilon^3 + O(\epsilon^4) \quad (24)$$

Equation (24) gives us an empirical relationship between parameters upto fourth order perturbations for the bifurcation of a limit cycle. Optimization of the above equation gives coefficients $a_0 = \frac{8}{\pi}$, $a_1 = 0$, $a_2 = \frac{128}{\pi^3}$, respectively. Next, we substitute the amount of dispersion ΔI into the perturbative term ϵ to obtain the boundary between IN and SR state. Taken together we can write,

$$|K| = \frac{8}{\pi}\Delta I + \frac{128}{\pi^3}\Delta I^3 + O(\Delta I^4) \quad (25)$$

Results are shown in **Figure 12** in the $(K, \Delta I)$ plane using Equations (25) and 23. Critical lines obtained semi analytically qualitatively agrees well with the numerical results that captures various network states in both these models with purely

excitatory coupling. In Appendix, we show a stability calculation for an inhibitory coupled mean field network in the infinite analog limit. From numerical simulations we find that the results are independent of the number of spikes n per burst.

5. DISCUSSION

One of the most frequent assumption employed in simulations of large neural networks is that the whole network can be lumped into small aggregates of collective unit (sometimes called a “neurocomputational unit”) exhibit a sufficiently similar dynamical behavior. Consequently, the network that instantiates this ensemble, consisting of thousands of excitatory and inhibitory neurons, it is considered to display a synchronized behavior with no other significant temporal features for the dynamics of the large scale network. The main reason for this assumption, is the impractical large computational time arising from too many details considered in the large network properties. In this paper, we

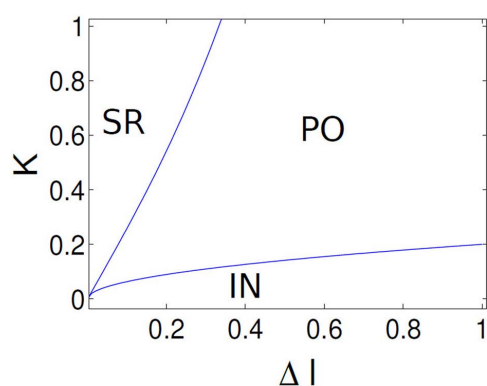


FIGURE 12 | Parameter boundaries are shown in $(K, \Delta I)$ parameter space using analytic results. Critical lines separates three distinct network states, labeled as SR, PO, and IN. Critical line separating PO, IN states corresponds to a saddle-node bifurcation and the line that separates IN, SR states corresponds to a Hopf bifurcation.

have analyzed the behavior of a neural network that serves as a good example of such a unit, namely a mean field coupled bursting ensemble. First, we have investigated a Hodgkin–Huxley type detailed biophysical model widely employed in theoretical and computational neuroscience with global coupling. We found that the dynamical features of the network are far more complex than the ones corresponding to synchronized or rest state behavior. The network dynamics depends critical on the balance between firing rate threshold dispersion and mean field synaptic coupling strength; in fact, the synchronized state can be found only for a specific range of parameters typically involving a large or medium values for the coupling strength and low values of dispersion. On the other hand, for large dispersion and weak coupling strength values both networks display purely IN behavior. In the IN state, individual neurons are driven by different external drives results in firing at different phases, however, their collective state is close to being stationary. This stationarity in the density distribution led us to formulate scaling relationship between coupling strength and dispersion parameter. One interesting finding is that, when mean field exerts a greater influence than parameter dispersion; it causes shutting down of the neural activity in some neurons. In this parameter range, we find interesting dynamical behavior such as partial activity. In order to address the problem of the high computational cost of such an implementation, we have further developed a self-consistent mathematically tractable mean field coupled phase model following (Assisi et al., 2005; Ghosh et al., 2009; Jirsa and Stefanescu, 2011), but incorporating a higher degree of realism. Rather than finding the most appropriate type and number of dimensions that could minimize certain error functions or capture statistical variance in the full network, we have focused our attention on understanding a phenomenological burst generation model system which captures the most important network dynamics of bursting units at the population level. Collective activity of synapses is described by a mean field which relies on instantaneous rise and decay time (Roy et al., 2011). This mean field is then employed in the coupling to individual neurons

to describe phase network. Together, we investigate this population of neurons coupled to a common mean field drive and heterogeneity in their threshold for spikes/bursts. Our detailed analysis demonstrated that the reduced representation manages to recreate correctly the topology of the mean field amplitudes of the original system for various parameter scenarios. In the full network, in the thermodynamic limit ($N \rightarrow \infty$), a collective state becomes coherent if $\delta V_{\text{mean}}(t) \equiv [V_{\text{mean}}(t) - \overline{V_{\text{mean}}}(t)]$ is non-stationary (i.e., an oscillating global potential V_{mean} appears for a coherent case) and also, the correlated mean field $\Gamma(t)$ activity appears oscillatory. In the phase network, global order parameter is computed by averaging the contributions of all microscopic spikes within a burst in order to obtain a similar degree of ordering of spikes/bursts as in the full model for identical parameterization. Hence, for a dynamical behavior such as coherence-incoherence transition macroscopic order parameter gives us a crude approximation of burst timing. From a more general perspective, despite its limitations because of the consideration of purely excitatory or inhibitory network, it can be readily extended to study networks with mixed coupling. Moreover, the analytical approach to estimate the scaling relationship and transition boundaries between the IN-PO-SR states is not restricted to small scale network only. With global coupling, as the size of the network grows the boundaries may shift leading to a different parameterization than the one displayed here; however, underlying bifurcations remain the same. We have demonstrated this in our work by analytically deriving a low dimensional mean field amplitude reduction for an inhibitory coupled mean field network in the continuum limit. In this case, all the relevant dynamics of an infinite dimensional network in Equations (29) and (30) is captured by a two dimensional representation of the reduced mean field population given by Equation (40). Thus, using this approach, we derive analytically a low dimensional representation of the network dynamics and we show that the main features of the neural population's collective behavior can be captured well by the dynamics of a few cortical nodes exhibiting spiking as well as bursting behavior. While it is true that strong reductionist assumptions are common (sacrificing many of the biological realism of a network node's dynamics) in large-scale network modeling, these assumptions are usually made *ad-hoc* on the network node's dynamics and limit the network dynamics to a small range. We emphasize here that because of the “near to synchrony” assumption, neural mass models cannot capture complex dynamical features such as multi-clustering, oscillator death or multi-time scale synchronization. Evidently a reduced small scale network model is desirable to serve as a node in a large scale network simulation whereby displaying a sufficiently rich dynamic repertoire. Here it is of less importance to find a quantitatively precise reduced description of a neural population; rather more importantly, we seek a computationally inexpensive population model (this means typically low-dimensional) which is able to display the major qualitative dynamic behaviors (synchronization, rest state, multi-clustering, etc.) for realistic parameter ranges as observed in the total population of neurons. Our approach may offer a viable alternative to the neural mass models currently used in the literature. By comparison, our model offers the possibility to account for such

features (temporal details of their spiking activity considered irrelevant for the dynamics of the large network) at a very low computational cost. Therefore, the type of reduced representation discussed in this paper qualifies as a good candidate for a “neural unit” in computational simulations of large scale neural networks.

REFERENCES

- Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87.
- Amzica, F., and Steriade, M. (1998). Electrophysiological correlates of sleep delta waves. *Electroencephalogr. Clin. Neurophysiol.* 107, 69–83.
- Antonsen, T. M., Faghih, R. T., Girvan, M., Ott, E., and Plati, J. (2008). External periodic driving of large systems of globally coupled phase oscillators. *Chaos* 18:037112. doi: 10.1063/1.2952447
- Assisi, C. G., Jirsa, V. K., and Kelso, J. A. S. (2005). Synchrony and clustering in heterogeneous networks with global coupling and parameter dispersion. *Phys. Rev. Lett.* 94:018106. doi: 10.1103/PhysRevLett.94.018106
- Baer, S. M., Rinzel, J., and Carrillo, H. (1995). Analysis of an autonomous phase model for neuronal parabolic bursting. *J. Math. Biol.* 33, 309–333.
- Beurle, R. L. (1956). Properties of a mass of cells capable of regenerating pulses. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 240, 55–94.
- Breakspear, M., Heitmann, S., and Daffertshofer, A. (2010). Generative models of cortical oscillations: neurobiological implications of the Kuramoto model. *Front. Hum. Neurosci.* 4:190. doi: 10.3389/fnhum.2010.00190
- Cabral, J., Hugues, E., Sporns, O., and Deco, G. (2011). Role of local network oscillations in resting-state functional connectivity. *Neuroimage* 57, 130–139.
- Deco, G., Jirsa, V. K., and McIntosh, A. R. (2011). Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* 12, 43–56.
- Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., and Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* 4:e1000092. doi: 10.1371/journal.pcbi.1000092
- Dhamala, M., Jirsa, V., and Ding, M. (2004a). Enhancement of neural synchrony by time delay. *Phys. Rev. Lett.* 92, 6–9.
- Dhamala, M., Jirsa, V. K., and Ding, M. (2004b). Transitions to synchrony in coupled bursting neurons. *Phys. Rev. Lett.* 92, 2–5.
- Ermentrout, G. B., and Kopell, N. (1986). Parabolic bursting in an excitable system coupled with a slow oscillation. *SIAM J. Appl. Math.* 46, 233–253.
- Feng, J., Jirsa, V. K., and Ding, M. (2006). Synchronization in networks with random interactions: theory and applications. *Chaos* 16:015109. doi: 10.1063/1.2180690
- Friesen, W. O. (1994). Reciprocal inhibition: a mechanism underlying oscillatory animal movements. *Neurosci. Biobehav. Rev.* 18, 547–553.
- Gabbiani, F., Midtgard, J., and Knöpfel, T. (1994). Synaptic integration in a model of cerebellar granule cells. *J. Neurophysiol.* 72, 999–1009.
- Ghosh, A., Roy, D., and Jirsa, V. K. (2009). Simple model for bursting dynamics of neurons. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 80(4 Pt 1):041930. doi: 10.1103/PhysRevE.80.041930
- Ghosh, A., Rho, Y., McIntosh, A. R., Kötter, R., and Jirsa, V. K. (2008). Cortical network dynamics with time delays reveals functional connectivity in the resting brain. *Cogn. Neurodyn.* 2, 115–120.
- Golomb, D., and Rinzel, J. (1993). Dynamics of globally coupled inhibitory neurons with heterogeneity. *Phys. Rev. E* 48, 4810–4814.
- Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. Vol 44. New York, NY: Wiley.
- Higham, D. J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.* 43, 525–546.
- Hindmarsh, J. L., and Rose, R. M. (1984). A model of neuronal bursting using three coupled first order differential equations. *Proc. R Soc. Lond B* 221, 87–102.
- Izhikevich, E. M. (2000). Neural excitability, spiking and bursting. *Int. J. Bifurcat. Chaos* 10, 1171–1266.
- Izhikevich, E. M. (2007). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. Vol 25. Cambridge: MIT press.
- Jensen, O., Kaiser, J., and Lachaux, J. P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci.* 30, 317–324.
- Jirsa, V. K. (2008). Dispersion and time delay effects in synchronized spike-burst networks. *Cogn. Neurodyn.* 2, 29–38.
- Jirsa, V. K. (2009). Neural field dynamics with local and global connectivity and time delay. *Philos. Trans. R Soc. A Math. Phys. Eng. Sci.* 367, 1131–1143.
- Jirsa, V. K., and Haken, H. (1997). A derivation of a macroscopic field theory of the brain from the quasi-microscopic neural dynamics. *Physica D* 99, 503–526.
- Jirsa, V. K., and McIntosh, A. R. (2007). *Handbook of Brain Connectivity*. Berlin: Springer.
- Jirsa, V. K., and Stefanescu, R. A. (2011). Neural population modes capture biologically realistic large scale network dynamics. *Bull. Math. Biol.* 73, 325–343.
- Lim, W., and Kim, S.-Y. (2011). Statistical-mechanical measure of stochastic spiking coherence in a population of inhibitory subthreshold neurons. *J. Comput. Neurosci.* 31, 667–677.
- Manrubia, S. C., Mikhailov, A. S., and Zanette, D. H. (2004). *Emergence of Dynamical Order. Synchronization Phenomena in Complex Systems*, Singapore: World Scientific Publishing Co.
- McCormick, D. A., and Feeser, H. R. (1990). Functional implications of burst firing and single spike activity in lateral geniculate relay neurons. *Neuroscience* 39, 103–113.
- Morris, C., and Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* 35, 193–213.
- Nunez, P. L. (1974). The brain wave equation: a model for the eeg. *Math. Biosci.* 21, 279–297.
- Ott, E., and Antonsen, T. M. (2009). Long time evolution of phase oscillator systems. *Chaos* 19:023117. doi: 10.1063/1.3136851
- Parnas, H., and Parnas, I. (1994). Neurotransmitter release at fast synapses. *J. Mem. Biol.* 142, 267–279.
- Rinzel, J., and Ermentrout, G. (1989). *Analysis of Neural Excitability and Oscillations*. Cambridge, MA: MIT Press.
- Robinson, P. A. (2011). Neural field theory of synaptic plasticity. *J. Theor. Biol.* 285, 156–163.
- Roy, D., Ghosh, A., and Jirsa, V. K. (2011). Phase description of neural oscillators with global electric and synaptic coupling. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 83, 1–10.
- Sabatini, B. L., and Regehr, W. G. (1996). Timing of neurotransmission at fast synapses in the mammalian brain. *Nature* 384, 170–172.
- Sherman, S. M., and Koch, C. (1986). The control of retinogeniculate transmission in the mammalian lateral geniculate nucleus. *Exp. Brain Res.* 63, 1–20.
- Smeal, R. M., Ermentrout, G. B., and White, J. A. (2010). Phase-response curves and synchronized neural networks. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 365, 2407–2422.
- Stefanescu, R. A., and Jirsa, V. K. (2008). A low dimensional description of globally coupled heterogeneous neural networks of excitatory and inhibitory neurons. *PLoS Comput. Biol.* 4:e1000219. doi: 10.1371/journal.pcbi.1000219
- Stefanescu, R. A., and Jirsa, V. K. (2011). Reduced representations of heterogeneous mixed neural networks with synaptic coupling. *Phys. Rev. E* 83, 1–12.
- Steriade, M., and Llinás, R. R. (1988). The functional states of the thalamus and the associated neuronal interplay. *Physiol. Rev.* 68, 649–742.
- Steriade, M., McCormick, D., and Sejnowski, T. J. (1993). Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262, 679–685.
- Uhlhaas, P. J., and Singer, W. (2006). Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron* 52, 155–168.
- Van Vreeswijk, C., Abbott, L. F., and Ermentrout, G. B. (1994).

ACKNOWLEDGMENTS

We would like to thank Anandamohan Ghosh and Mohit Adhikari for the helpful corrections on this manuscript. The research reported herein was supported by the Brain Network Recovery Group through the James S. McDonnell Foundation and the FP7-ICT BrainScales.

- When inhibition not excitation synchronizes neural firing. *J. Comput. Neurosci.* 1, 313–321.
- Wang, X. J. (1994). Multiple dynamical modes of thalamic relay neurons: rhythmic bursting and intermittent phase-locking. *Neuroscience* 59, 21–31.
- Wang, X. J., and Rinzel, J. (1992). Alternating and synchronous rhythms in reciprocally inhibitory model neurons. *Neural Comput.* 4, 84–97.
- Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 28 January 2013; accepted: 05 March 2013; published online: 26 March 2013.
- Citation: Roy D and Jirsa V (2013) Inferring network properties of cortical neurons with synaptic coupling and parameter dispersion. *Front. Comput. Neurosci.* 7:20. doi: 10.3389/fncom.2013.00020
- Copyright © 2013 Roy and Jirsa. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

MEAN FIELD REDUCTION FOR INHIBITORY SYNAPTIC COUPLING

Here, we extend our network in the continuum limit in the presence of inhibitory coupling. V_{th} is held negative. ($N \rightarrow \infty$), where the state of the coupled system can be described by a density function $f(\theta, I, t)$, where f is defined such that the fraction of neurons with phases lying between θ and $d\theta$ and applied drive between I and dI is given by $f(\theta, I, t)d\theta dI$ (Antonsen et al., 2008; Ott and Antonsen, 2009). The applied stimulus are drawn from a distribution $g(I)$ such that

$$\int_{-\infty}^{\infty} \int_0^{2\pi} f(\theta, I, t) d\theta dI = 1 \quad (26)$$

$$\int_0^{2\pi} f(\theta, I, t) d\theta = g(I) \quad (27)$$

For the conservation of currents I the continuity equation is written as

$$\frac{\partial f}{\partial t} + \frac{\partial(fv)}{\partial \theta} = 0. \quad (28)$$

In order to make the coupling amenable to analytical study we use a pulse-like function for the mean field $\Gamma = a1 + b1(1 + \cos \theta)$. Response to the mean field by individual neuron's $R(\theta) = V_{th} \sin(\theta)$, containing only single Fourier component, a choice motivated primarily due to the tractability of the resulting model. Further, $V_{th} = -1$ for the convenience of calculations without losing any generality of our results. The velocity $v(\theta, I, t)$ in Equation (28) is now written as

$$v(\theta, I, t) = a + \epsilon \int_{-\infty}^{\infty} \int_0^{2\pi} (1 + \cos(\hat{\theta})) (-\sin \theta) f(\hat{\theta}, \hat{I}, t) d\hat{\theta} d\hat{I} - F \sin(\theta) - F \sin(\theta/n) \quad (29)$$

where without loss of any generality we are using sin functions instead of cos functions in Equation (6).

In the continuum limit the order parameter z can be defined as

$$z(t) = \int_{-\infty}^{\infty} \int_0^{2\pi} \frac{(e^{i\theta} + e^{-i\theta})}{2} f(\theta, I, t) d\theta dI \quad (30)$$

It's a linear sum of two complex order parameters and one could in principle unfold the entire dynamics of the network in any one of the manifold given above. Here,

$$z1(t) = \int_{-\infty}^{\infty} \int_0^{2\pi} e^{i\theta} f(\theta, I, t) d\theta dI \quad (31)$$

$$z2(t) = \int_{-\infty}^{\infty} \int_0^{2\pi} e^{-i\theta} f(\theta, I, t) d\theta dI \quad (32)$$

Using the above it is easy to see that the expression for velocity becomes

$$v(\theta, I, t) = I + \frac{1}{2i} \left[\left(1 + \frac{\epsilon}{2} (z2 + z1^*) + F \right) e^{-i\theta} - \left(1 + \frac{\epsilon}{2} (z1 + z2^*) + F \right) e^{i\theta} \right] - F \left(\frac{e^{i\theta/n}}{2i} - \frac{e^{-i\theta/n}}{2i} \right) \quad (33)$$

* indicates the complex conjugate. The distribution function can be expressed as a Fourier series

$$f(\theta, I, t) = \frac{g(I)}{2\pi} \left[1 + \sum_{k=1}^{\infty} f_k(I, t) e^{ik\theta} + c.c. \right] \quad (34)$$

The above infinite dimensional system is difficult to analyze. However, the “amazing” ansatz of Ott and Antonsen (2009) has been shown to be successful in obtaining the low-dimensional description of the globally coupled phase oscillators. The ansatz impose a restriction on the fourier coefficients:

$$f_k(I, t) = (\psi(I, t))^k \quad (35)$$

for $k \geq 1$ and has been shown to be a reasonable guess under different scenarios (Ott and Antonsen, 2009). This restricted class of functions readily reduces our continuity equation to an θ -independent form

$$\frac{d\psi}{dt} = \frac{1}{2} \left(1 + \frac{\epsilon}{2} z1 + F \right)^* - iI\psi - \frac{1}{2} \left(1 + \frac{\epsilon}{2} z1 + F \right) \psi^2 - F \left(\frac{\psi^{1+1/n}}{2} - \frac{\psi^{1-1/n}}{2} \right) \quad (36)$$

with $z1$ satisfying

$$z1(t) = \int_{-\infty}^{\infty} \psi^*(I, t) g(I) dI. \quad (37)$$

If we assume that $g(I)$ is a Lorentzian distribution function

$$g(I) = \frac{1}{\pi[(I - I_0)^2 + 1]}. \quad (38)$$

$z(t)$ can be evaluated by contour integration with poles at $I = I_0 - i$ and we obtain the exact evolution equation of order parameter z

$$\frac{dz1}{dt} = iI_0 z1 - z1 + \frac{1 + \frac{\epsilon}{2} z1 + F}{2} - \frac{1 + \frac{\epsilon}{2} z1^* + F}{2} - F \left(\frac{z1^{1+1/n}}{2} - \frac{z1^{1-1/n}}{2} \right) \quad (39)$$

The above equation can be expressed in polar coordinates if we substitute $z_1 = \rho_1 \exp(i\phi_1)$ giving evolution equations for ρ_1 and ϕ_1

$$\begin{aligned} \frac{d\rho_1}{dt} = & \frac{\epsilon}{2} \rho_1 (1 - \rho_1^2) - \rho_1 + \frac{F}{2} (1 - 2\rho_1^2) \cos \phi_1 + \frac{1}{2} \cos(\phi_1) \\ & + \frac{F\rho_1}{2} (\rho_1^{-1/n} - \rho_1^{1/n}) \cos(\phi_1/n) \end{aligned} \quad (40)$$

$$\begin{aligned} \frac{d\phi_1}{dt} = & I_0 - \frac{F}{2} \left(\rho_1 + \frac{1}{\rho_1} \right) \sin \phi_1 - \frac{\rho_1}{2} \sin(\phi_1) \\ & - \frac{F}{2} \left(\rho_1^{1/n} + \frac{1}{\rho_1^{1/n}} \right) \sin(\phi_1/n). \end{aligned} \quad (41)$$

For the Lorentzian distribution function the above equation is exact. However, we do not find any deviation of the above results for any other unimodal distributions of our firing threshold (such as uniform distribution) such as the one considered in the numerical simulations with excitatory coupling. The above two dimensional system can be solved numerically to identify the full network states and the corresponding transition boundaries.



Ketamine, propofol, and the EEG: a neural field analysis of HCN1-mediated interactions

Ingo Bojak^{1,2*}, Harry C. Day¹ and David T. J. Liley^{3,4}

¹ Centre for Computational Neuroscience and Cognitive Robotics, School of Psychology, University of Birmingham, Birmingham, UK

² Donders Centre for Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

³ Brain and Psychological Sciences Research Centre, Faculty of Life and Social Sciences, Swinburne University of Technology, Hawthorn, VIC, Australia

⁴ Cortical Dynamics Ltd., Hawthorn, VIC, Australia

Edited by:

Dimitris Pinotsis, University College London, UK

Reviewed by:

Julien Modolo, Lawson Health Research Institute and Western University, Canada
Sacha J. Van Albada, Research Center Jülich, Germany
Jamie Sleight, University of Auckland, New Zealand

*Correspondence:

Ingo Bojak, Centre for Computational Neuroscience and Cognitive Robotics, School of Psychology, University of Birmingham, Birmingham West Midlands B15 2TT, UK.
e-mail: i.bojak@bham.ac.uk

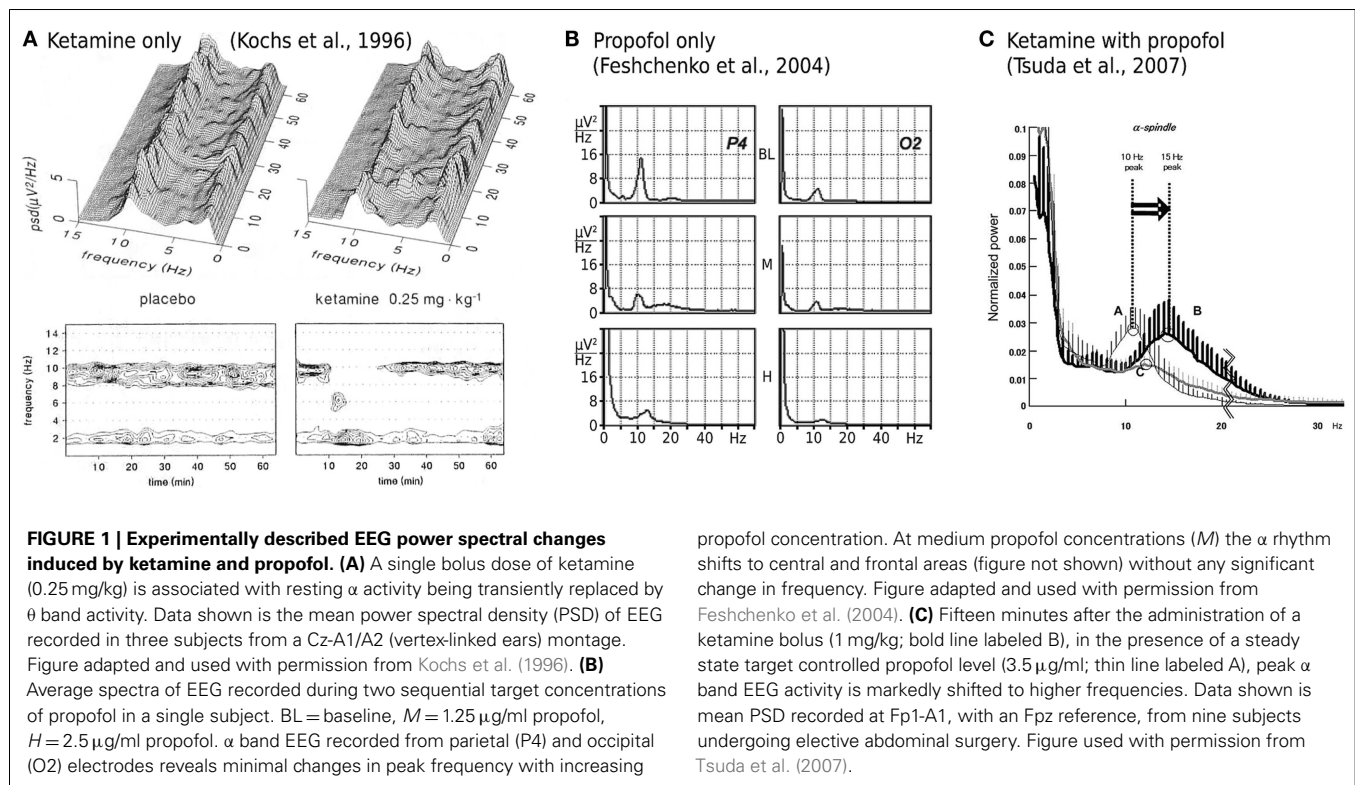
Ketamine and propofol are two well-known, powerful anesthetic agents, yet at first sight this appears to be their only commonality. Ketamine is a dissociative anesthetic agent, whose main mechanism of action is considered to be *N*-methyl-D-aspartate (NMDA) antagonism; whereas propofol is a general anesthetic agent, which is assumed to primarily potentiate currents gated by γ -aminobutyric acid type A (GABA_A) receptors. However, several experimental observations suggest a closer relationship. First, the effect of ketamine on the electroencephalogram (EEG) is markedly changed in the presence of propofol: on its own ketamine increases θ (4–8 Hz) and decreases α (8–13 Hz) oscillations, whereas ketamine induces a significant shift to beta band frequencies (13–30 Hz) in the presence of propofol. Second, both ketamine and propofol cause inhibition of the inward pacemaker current I_h , by binding to the corresponding hyperpolarization-activated cyclic nucleotide-gated potassium channel 1 (HCN1) subunit. The resulting effect is a hyperpolarization of the neuron's resting membrane potential. Third, the ability of both ketamine and propofol to induce hypnosis is reduced in HCN1-knockout mice. Here we show that one can theoretically understand the observed spectral changes of the EEG based on HCN1-mediated hyperpolarizations alone, without involving the supposed main mechanisms of action of these drugs through NMDA and GABA_A, respectively. On the basis of our successful EEG model we conclude that ketamine and propofol should be antagonistic to each other in their interaction at HCN1 subunits. Such a prediction is in accord with the results of clinical experiment in which it is found that ketamine and propofol interact in an infra-additive manner with respect to the endpoints of hypnosis and immobility.

Keywords: ketamine, propofol, EEG, HCN1, neural field theory, drug interaction, anesthesia, infra-additivity

INTRODUCTION

Ketamine, a phenylcyclohexylpiperidine (PCP) derivative, is a powerful psychoactive drug that is predominantly used as a sedative and general anesthetic agent in humans and animals (Sinner and Graf, 2008). Ketamine occurs as two stereoisomers, R(–) and S(+), in which the latter is found to be some three to four times more potent (White et al., 1985), but despite such differences in potency the drug is generally made available clinically as a racemate (racemic mixture) that contains both stereoisomers in equal proportion. Ketamine is classified as a dissociative agent due to its ability to induce hallucinations and perceptual/environmental detachment (Wolff and Winstock, 2006). Because of these properties it has become popular recreationally. At sufficiently high doses it has been reported to induce a state of dissociation comparable to that of schizophrenia, and as a consequence has found use as a pharmacological model for psychosis (Bubenikova-Valesova et al., 2008; Corlett et al., 2011). More recently its therapeutic use has been re-evaluated in light of evidence suggesting that sub-anesthetic doses may aid in the treatment of bipolar affective disorder and major depression (Mathew et al., 2012; Murrrough, 2012; Murrrough et al., 2012).

While ketamine is widely believed to act principally through the non-competitive antagonism of *N*-methyl-D-aspartate (NMDA) receptor mediated glutamatergic activity (Irifune et al., 1992; Oye et al., 1992), two significant pieces of empirical evidence have emerged that challenge such a unitary view. Firstly, dizocilpine (also known as MK801), an even more potent non-competitive NMDA antagonist, produces no significant hypnotic effect (Kelland et al., 1993; Irifune et al., 2007). Secondly, ketamine's effect on spontaneous electroencephalogram (EEG) activity is qualitatively altered when administered in the presence of propofol, a widely used intravenous general anesthetic agent that, at clinically meaningful concentrations, has little or no effect on NMDA mediated currents. Ketamine alone has been shown to reduce spectral edge frequencies, an effect that is driven predominately by increases in absolute θ band (4–8 Hz) power at the expense of α band (8–13 Hz) power (Schuttler et al., 1987; Kochs et al., 1996), see **Figure 1A**. In contrast, ketamine administered in the presence of steady state propofol levels is associated with a definite acceleration of α band activity; increasing its peak frequency by up to 4.7 Hz (Hayashi et al., 2007; Tsuda et al., 2007), see **Figure 1C**. Propofol on its own roughly maintains the α peak frequency with an



anteriorization of power (decrease occipital, increase frontal), see **Figure 1B**; though an additional broadband “beta buzz” just above α frequencies, “biphasic” response dynamics and smooth transitions to lower frequencies can confound the picture (Schwender et al., 1996; Kuizenga et al., 1998, 2001; Feshchenko et al., 2004; Breshears et al., 2010; Cimenser et al., 2011). We assume here from previous theoretical studies (Liley et al., 2003; Hutt and Schimansky-Geier, 2008; Hutt and Longtin, 2010; Hindriks and van Putten, 2012) that these complications can be accounted for by mechanisms not considered in this work, in particular the prominent γ -aminobutyric acid type A (GABA_A) agonism of propofol that affects dominantly inhibitory postsynaptic currents (Kitamura et al., 2003). Furthermore, the acceleration due to ketamine observed by Hayashi et al. (2007) and Tsuda et al. (2007) that we wish to describe occurred on top of a clear α rhythm at steady propofol concentration, see **Figure 1C**. Thus we assume in the following that the action of propofol is largely neutral concerning the α peak frequency (while unspecified concerning total α band spectral power).

Recently a number of alternative, behaviorally relevant, molecular targets for ketamine action have been identified (Schnoebel et al., 2005; Hevers et al., 2008; Chen et al., 2009). Of particular significance is the identification of hyperpolarization-activated cyclic nucleotide-gated (HCN) potassium channel subunits as a target for ketamine action (Chen et al., 2009). HCN subunits, of which there are four isoforms (HCN1–4), assemble to form a tetrameric ion channel that mediates an inward (i.e., depolarizing) hyperpolarization-activated pacemaker current I_h implicated in neuronal rhythmogenesis (Biel, 2009; Biel et al., 2009). In

particular the HCN1 isoform has been identified as a molecular substrate for the actions of ketamine (Chen et al., 2009): ketamine causes inhibition of HCN1-mediated I_h currents, and hence membrane hyperpolarization, in pyramidal neurons from wild-type but not HCN1-knockout mice. The potency of ketamine to provoke a loss of the righting reflex (the ability to regain footing from a back position), which is a behavioral correlate of hypnosis, is also strongly reduced in HCN1-knockout mice. Hence a causal relationship between ketamine-induced membrane hyperpolarization and its clinical effects can be made. The existence of such a causal relationship is made more likely by evidence indicating that the hypnotic potency of propofol is also reduced in HCN1-knockout mice, in approximate proportion to its ability to inhibit HCN1-mediated membrane depolarization (Chen et al., 2009).

It should be noted though that the hypnotic response was not abolished entirely in HCN1-knockout mice by either ketamine or propofol (Chen et al., 2009), thus other effects like the mentioned GABA_A agonism will be required to fully understand the hypnotic action of these agents. However, etomidate, which has no effect on HCN1 channels, showed no loss of hypnotic effect in HCN1-knockout mice (Chen et al., 2009), suggesting the specific involvement of HCN1-mediated I_h currents for ketamine and propofol. The identification of a shared molecular target for ketamine and propofol action could offer a new possibility to account for the qualitatively disparate electroencephalographic effects of ketamine alone and in the presence of propofol. On this basis we hypothesized that by modeling the differential effects of ketamine and propofol on neuronal membrane hyperpolarization, in the

context of an established theory of resting EEG (Liley et al., 2002, 2010, 2011; Bojak and Liley, 2005), we would be able to describe the observed effects on the EEG at least qualitatively. Because current depth of anesthesia monitoring approaches are either insensitive (Faraoni et al., 2009; Nonaka et al., 2012), or respond anomalously (Hans et al., 2005; Sengupta et al., 2011), to the hypnotic effects of ketamine, understanding the mechanism by which ketamine and propofol interact electroencephalographically will ultimately assist in the development of improved approaches to clinically monitor the hypnotic effects of combinations of these drugs. The combination of propofol and ketamine (often referred to as *ketofol*) is becoming increasingly important in the procedural sedation setting where rapid and effective sedation and analgesia, with minimal cardiorespiratory/hemodynamic compromise, is required (Hui et al., 1995; Frizelle et al., 1997; Sakai et al., 1999; Phillips et al., 2010).

MATERIALS AND METHODS

MODELING DRUG RESPONSE AND INTERACTIONS

The simplest pharmacodynamic model of drug effect involving two or more agonists is that of competitive ligand-receptor binding. It is easily shown for two full agonists competing for the same receptor binding site, that the fractional receptor occupancy θ , as a function of the respective drug concentrations (D_1 , D_2) is (Shafer et al., 2008)

$$\theta = \frac{k_2 D_1 + k_1 D_2}{k_2 D_1 + k_1 D_2 + k_1 k_2}, \text{ or} \quad (1)$$

$$\frac{\theta}{1 - \theta} = \frac{D_1}{k_1} + \frac{D_2}{k_2}. \quad (2)$$

For $D_1 \rightarrow \infty$ and/or $D_2 \rightarrow \infty$, one then finds $\theta \rightarrow 1$, i.e., full receptor occupancy. k_1 , $k_2 > 0$ are the respective drug-receptor dissociation constants, which are equivalent to single drug concentrations that produce 50% receptor occupancy, i.e., $\theta = 1/2$. In general a pharmacodynamic effect E is assumed to be some monotonic function of θ , i.e., $E = f(\theta) \equiv g(D_1, D_2)$. For a fixed effect E the locus of points (D_1 , D_2) defines a *response isobole* and $E = g(D_1, D_2)$ a *response surface*. Now consider the case of competitive binding and drug interaction (Greco et al., 1995)

$$\frac{\theta}{1 - \theta} = \frac{D_1}{k_1} + \frac{D_2}{k_2} + \frac{\eta D_1 D_2}{k_1 k_2}, \quad (3)$$

where η defines an interaction term. It can be easily demonstrated that

$$\begin{aligned} \frac{\theta-1}{\theta} < \eta < 0 &\Rightarrow \text{infra-additivity / antagonism,} \\ \eta = 0 &\Rightarrow \text{additivity,} \\ \eta > 0 &\Rightarrow \text{synergy.} \end{aligned} \quad (4)$$

Inspired by these considerations, we chose here to describe the general pharmacodynamic effect of our two ligands by the following bilinear form

$$E = c_1 D_1 + c_2 D_2 + c_{12} D_1 D_2. \quad (5)$$

This ansatz represents the simplest extension beyond the purely additive; and the sign of c_{12} then has the same interpretation as the sign of η in Eq. 4. We will use this bilinear form below to parameterize the dependence of the induced hyperpolarizations on normalized concentrations of propofol and ketamine, respectively.

One can however relate Eqs 3 and 5 more directly. Assume first that the pharmacodynamic effect is directly proportional to receptor occupancy, i.e., $E \propto \theta$. Then k_1 and k_2 become the respective “half maximum effective concentrations” (EC50s) at which 50% of the maximum response is observed for each drug applied alone. Furthermore, assume that the receptor occupancy remains relatively small $\theta \approx \theta/(1 - \theta) = D_1/k_1 + D_2/k_2 + \eta D_1 D_2/(k_1 k_2) < 1/2$, so that the effect $E < E_{\max}/2$. The half-maximal inhibition of HCN1 subunit-mediated ionic currents by racemic ketamine occurs at a concentration of approximately 16 μM (Chen et al., 2009), which is significantly greater than the estimated minimum free plasma concentrations of 2.9 μM required to produce surgical anesthesia in humans (Grant et al., 1983). Data for the half-maximal inhibition of HCN1-mediated ionic currents by propofol is to our knowledge not available. However, because HCN1-knockout mice are significantly less sensitive to the effects of propofol than wild-type ones, we can speculate that the ED50 (the “half maximum effective dose”) for unresponsiveness with propofol in wild-type mice corresponds roughly to the half maximum of the neuronal changes (EC50). Chen et al. (2009) found this to be approximately 7 mg/kg. Using the volume of distribution of 1.38 l/kg (Cox et al., 1998) in the rat (murine values not available), EC50 is then about 5.1 mg/l or 29 μM , which is significantly greater than the minimum free plasma concentration $\sim 8.5 \mu\text{M}$ for surgical anesthesia. Thus $D_{\text{ketamine}}/k_{\text{ketamine}} + D_{\text{propofol}}/k_{\text{propofol}} + \eta D_{\text{ketamine}} D_{\text{propofol}}/(k_{\text{propofol}} k_{\text{ketamine}}) < 1/2$ is approximately satisfied as long as $\eta < \frac{1}{2}$. While the E_{\max} for ketamine-induced membrane hyperpolarization in murine pyramidal neurons is of the order of -4 mV , the actual value of E_{\max} will depend on the species and the recording conditions/preparation. In the absence of any information to the contrary one can assume that $E < E_{\max}/2$. Thus our ansatz Eq. 5 can be considered as following from Eq. 3 under a range of reasonable assumptions.

LILEY MODEL AND EIGENSPECTRUM CALCULATION

We base our investigation in this paper on the Liley et al. (2002) model, which is a typical neural field model (Deco et al., 2008; Coombes, 2010; Bressloff, 2012; Liley et al., 2012). In Bojak and Liley (2005) 73,454 different parameter sets, which produce biologically plausible resting state activity, were found for this model. We use here also the “eigenspectrum” approach introduced in Bojak and Liley (2005) to directly predict EEG power spectral densities (PSDs) from a given parameter set. In the following we will briefly review a few key features of the Liley et al. (2002) model and of the eigenspectrum approach that will play a role for the analysis in this paper, and refer the reader to the original reference for more detail. The Liley et al. (2002) model can be written concisely as follows:

$$\tau_k \frac{\partial}{\partial t} h_k(\mathbf{x}, t) = h_k^r - h_k(\mathbf{x}, t) + \sum_{l=e,i} \frac{h_{lk}^{eq} - h_k(\mathbf{x}, t)}{|h_{lk}^{eq} - h_k^r|} I_{lk}(\mathbf{x}, t), \quad (6)$$

$$\left(\frac{1}{\gamma_{lk}} \frac{\partial}{\partial t} + 1 \right)^2 I_{lk}(\mathbf{x}, t) = \frac{\Gamma_{lk} e}{\gamma_{lk}} \left\{ \frac{N_{lk}^{\beta} S_l^{\max}}{1 + e^{\sqrt{2}[h_l(\mathbf{x}, t) - \mu_l]/\sigma_l}} + \Phi_{lk}(\mathbf{x}, t) + p_{lk}(\mathbf{x}, t) \right\}, \quad (7)$$

$$\left[\left(\frac{1}{v_{lk} \Lambda_{lk}} \frac{\partial}{\partial t} + 1 \right)^2 - \frac{3}{2} \frac{1}{\Lambda_{lk}^2} \nabla^2 \right] \Phi_{lk}(\mathbf{x}, t) = \frac{N_{lk}^{\alpha} S_l^{\max}}{1 + e^{\sqrt{2}[h_l(\mathbf{x}, t) - \mu_l]/\sigma_l}}. \quad (8)$$

In all these equations $l, k = e, i$ serve as indices for excitatory and inhibitory neural populations, respectively, and \mathbf{x} gives their position on a two-dimensional cortical sheet. The mean excitatory soma membrane potential $h_e(\mathbf{x}, t)$ of Eq. 6 is taken to predict the EEG. In the absence of postsynaptic inputs these potentials $h_k(\mathbf{x}, t)$ decay to their resting values h_k^r . The inputs $I_{lk}(\mathbf{x}, t)$ correspond to postsynaptic potentials and are weighted by ionic driving forces $h_{lk}^{eq} - h_k(\mathbf{x}, t)$, where the h_{lk}^{eq} are the respective Nernst potentials. These weights are normalized at rest to +1 (excitatory) and -1 (inhibitory), respectively. A postsynaptic input in Eq. 7 uses double indices to indicate source and target (for example, $I_{ei}(\mathbf{x}, t)$ is excitatory input to an inhibitory neural population). Γ_{lk} is the mean peak amplitude induced by a single presynaptic pulse $\delta(t - t_p)$, and $1/\gamma_{lk}$ the corresponding rise time to this peak of a postsynaptic “ α form” response $I(\mathbf{x}, t) \propto \gamma^2 t e^{-\gamma t} \Theta(t - t_p)$, where Θ is the Heaviside step function and δ the Dirac delta function. Extra-cortical input is given by $p_{lk}(\mathbf{x}, t)$, and is here assumed to be shaped noise (p_{ee}), static (p_{ei}), or absent (p_{ik}). The noise represents the average of uncorrelated input to the many neurons in the neural mass. For simplicity it is imposed only on the excitatory extra-cortical input to excitatory neurons, which is sufficient to generate the full dynamical range of the model. Finally, activity is propagated cortico-cortically via Eq. 8 with a standard damped wave equation (Jirsa and Haken, 1996; Robinson et al., 1997). The activity propagation through $\Phi_{lk}(\mathbf{x}, t)$ represents a synaptic footprint which falls off exponentially with characteristic distance scale Λ_{lk} , and fibers having conduction velocity v_{lk} . Since there are no long-range inhibitory fibers, we can set $\Phi_{ik} \equiv 0$ in the following. Short range connectivity is both excitatory and inhibitory, and is represented by the first term in the curly brackets of Eq. 7. Note that Eq. 8 can be improved upon (Bojak and Liley, 2010), but its main role is in this case to include a larger variety of EEG wavelengths as will become apparent. Our main conclusions are not affected even for the radical choice of an entirely homogeneous cortex, i.e., upon removing all spatial dependence.

The eigenspectrum approach (Bojak and Liley, 2005) assumes that Eqs 6–8 have a “fixed point” solution for a homogeneous cortex with static p_{ee} . All variables are then linearly expanded around this solution, and auxiliary variables $\tilde{I}_{lk} = \partial I_{lk} / \partial t$ and $\tilde{\Phi}_{ek} = \partial \Phi_{ek} / \partial t$ are used to turn Eqs 6–8 into 14 first order ODEs.

One can then Fourier-transform in space and time, and obtains an equation for the 14-dimensional state vector \mathbf{s} in the form

$$i\omega \mathbf{s}(\omega, \mathbf{k}) = \mathbf{J}(\mathbf{k}^2) \cdot \mathbf{s}(\omega, \mathbf{k}) + \mathbf{P}(\omega, \mathbf{k}), \quad (9)$$

where \mathbf{J} is the Jacobian matrix and $\mathbf{P}(\omega, \mathbf{k})$ contains the remainder of the extra-cortical input, i.e., the variation of p_{ee} with subtracted mean. Note that the only spatial derivative here is the Laplacian in Eq. 8, hence the Fourier-transformed Jacobian is a function of the square of wavenumber \mathbf{k} . One can then show (Bojak and Liley, 2005) that

$$|h_e(\omega, k)|^2 = \left| \sum_{n=1}^{14} \frac{c_n(k)}{i\omega - \lambda_n(k)} \right|^2, \quad (10)$$

where both the coefficients c_n and the eigenvalues λ_n can be obtained from a decomposition of the Jacobian in both left and right eigenmatrices, and $k \equiv |\mathbf{k}|$.

Furthermore, if one makes the simplifying assumption that an EEG electrode aggregates the contributions of a disk-shaped part of the cortical sheet with radius R , then one can compute a prediction of the PSD as follows (Bojak and Liley, 2005)

$$\text{PSD}(f) = 2\pi R^2 \int_0^\infty \frac{dk}{k} J_1^2(kR) |h_e(\omega \equiv 2\pi f, k)|^2, \quad (11)$$

where J_1 is a Bessel function of the first kind. In practice we evaluate the integral Eq. 11 numerically using a 64 point $0 < k_i < 14.14/\text{cm}$ Gauss–Legendre quadrature, and hence need to evaluate Eq. 10 for all these $k = k_i$. PSDs calculated in this manner from 10 parameter sets selected out of the 73,454 in Bojak and Liley (2005) are shown in **Figure 2**. We call a solution *stable* if for all $64 k = k_i$, as well as for homogeneous cortex $k = 0$ cm, the eigenvalues are such that $\forall n: \Re \lambda_n(k) < 0$. Only for stable cases do all the approximations leading up to Eq. 11 make sense. The largest contributions to Eq. 10 arise when $\omega = \Im \lambda_n(k)$; and if one disregards the $c_n(k)$, then the “least stable” eigenvalue with largest $\Re \lambda_n(k) < 0$ will contribute most.

Consider now only those λ_m that have non-zero frequencies $f_m \equiv \Im \lambda_m(k = 0/\text{cm})/(2\pi) \neq 0$: due to the selection process (Bojak and Liley, 2005), the “least stable” λ_{\max} of these λ_m will have $8 \text{ Hz} \leq f_{\max} \leq 13 \text{ Hz}$, i.e., a frequency in the α rhythm range. If we change from parameter set $\{\mathbf{P}_1\}$ to $\{\mathbf{P}_2\}$, we can compute the resulting frequency shift $\Delta f \equiv f_{\max}^{(\mathbf{P}_2)} - f_{\max}^{(\mathbf{P}_1)}$ of this eigenvalue α frequency. We find that for the parameter changes considered below, this “theoretical” α frequency shift Δf estimated directly from the eigenvalues provides a reasonable approximation for a more “experiment-like” calculation of the α peak shift. In an experiment one would typically seek the maxima of the measured PSDs in the 8–13 Hz range, and then compute their difference in frequency in order to determine an α frequency shift, cf. **Figure 1C** (Hayashi et al., 2007; Tsuda et al., 2007). We can do something similar here by evaluating the full PSDs with Eq. 11 for $\{\mathbf{P}_1\}$ and $\{\mathbf{P}_2\}$, respectively, and then compute the difference of the maxima of these theoretical predictions. However, we use the “theoretical” Δf

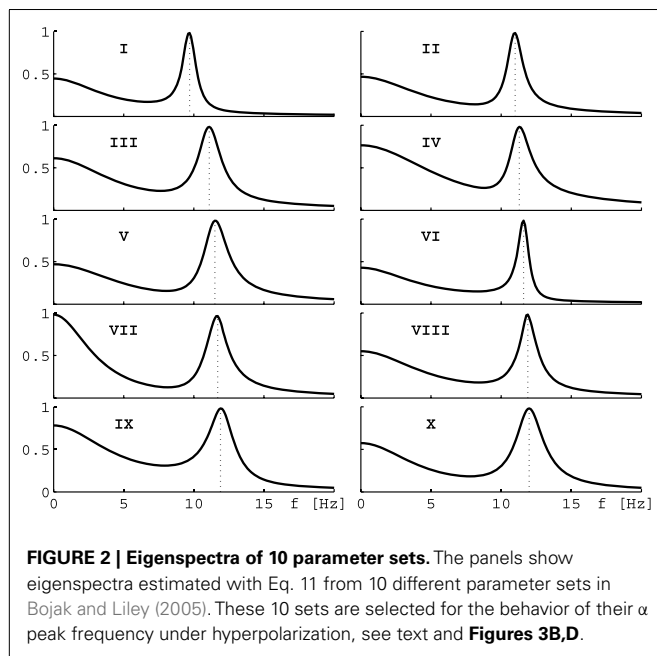


FIGURE 2 | Eigenspectra of 10 parameter sets. The panels show eigenspectra estimated with Eq. 11 from 10 different parameter sets in Bojak and Liley (2005). These 10 sets are selected for the behavior of their α peak frequency under hyperpolarization, see text and **Figures 3B,D**.

Table 1 | α peak frequency shifts predicted from the leading eigenvalue (Δf) and the full PSD, respectively, for the parameter sets of Figure 2.

	Propofol only $P = 1.2, K = 0$		Ketamine only $P = 0, K = 1.4$		Both $P = 1.2, K = 1.4$	
	Δf (Hz)	PSD (Hz)	Δf (Hz)	PSD (Hz)	Δf (Hz)	PSD (Hz)
I	0.25	0.07	-1.42	-1.70	1.84	1.85
II	0.40	0.29	-1.60	-1.69	2.70	2.61
III	0.24	0.13	-1.37	-1.69	1.91	2.27
IV	0.45	0.37	-1.10	-0.96	2.55	2.54
V	-0.20	-0.21	-1.39	-1.30	1.51	1.40
VI	0.12	-0.18	-1.00	-1.44	1.47	1.44
VII	0.10	0.16	-2.12	-2.11	2.72	2.81
VIII	0.02	-0.15	-1.33	-1.50	1.45	1.59
IX	0.20	-1.46	-1.89	-3.02	2.09	2.05
X	0.15	0.15	-1.38	-1.66	1.86	2.14

P and K are normalized propofol and ketamine concentrations, respectively.

in the following. It is much easier to compute, since it involves only one eigendecomposition for $k = 0/\text{cm}$ compared to 64 for $k = k_i$ needed in the numerical PSD integration Eq. 11. Furthermore, the “theoretical” Δf separates the change of the α peak frequency from other changes to the spectrum. “Experiment-like” calculations of shifts directly from local maxima in the full spectrum can be confounded easily by other spectral changes, and a prior subtraction of the spectral “background” around these maxima would closely match our “theoretical” procedure. For the parameter sets shown in **Figure 2**, a comparison between “theoretical” and “experiment-like” frequency shifts is provided by **Table 1**. How these shifts are generated will be discussed in the following, but note for now that most results are quite similar. The big discrepancies for parameter

set IX are caused precisely by a rise of the spectral “background,” as discussed.

DRUG EFFECT PARAMETERIZATION AND SELECTION OF SETS

The effect of the action of both ketamine and propofol on HCN1 channels is to hyperpolarize the resting membrane potentials of pyramidal (excitatory) cells (Chen et al., 2009). Consider Eq. 6 in the absence of synaptic inputs $I_{jk}(\mathbf{x}, t) \equiv 0$, then $\lim_{t \rightarrow \infty} h_k(\mathbf{x}, t) = h_k^r$. Thus h_e^r and h_i^r parameterize the excitatory and inhibitory resting membrane potentials, respectively. In the spirit of Eq. 5 we hence use the following ansatz:

$$\begin{aligned} \Delta h_e^r &\equiv h_e^r|_{P,K} - h_e^r|_{P=K=0} = -(a_1 P + a_2 K + a_{12} PK) \\ &= -\Delta h \cos \theta, \end{aligned} \quad (12)$$

$$\begin{aligned} \Delta h_i^r &\equiv h_i^r|_{P,K} - h_i^r|_{P=K=0} = -(b_1 P + b_2 K + b_{12} PK) \\ &= -\Delta h \sin \theta, \end{aligned} \quad (13)$$

where P, K are normalized (dimensionless) concentrations of propofol and ketamine, respectively; and $\Delta h_e^r, \Delta h_i^r$ are changes of the excitatory and inhibitory resting membrane potentials, respectively, due to these drugs. For convenience we have factored out the sign corresponding to hyperpolarization, and we have assumed that inhibitory neurons would react qualitatively like the pyramidal cells, i.e., $\Delta h_e^r, \Delta h_i^r \leq 0$ mV in the considered ranges $0 \leq P \leq P_{\max}$ and $0 \leq K \leq K_{\max}$, while quantitative differences are expressed by potentially different coefficients. Since the drugs applied individually lead to hyperpolarization, we must have coefficients $a_1, a_2, b_1, b_2 > 0$, whereas the sign of the interaction coefficients a_{12}, b_{12} carries the same meaning as that of η in Eq. 4. In the following it often will be useful to express the “Cartesian” $\Delta h_e^r, \Delta h_i^r \leq 0$ mV in the corresponding “polar coordinate” form as $\Delta h \equiv \sqrt{(\Delta h_e^r)^2 + (\Delta h_i^r)^2} \geq 0$ mV and $\theta \equiv \arctan \frac{\Delta h_i^r}{\Delta h_e^r} \in [0^\circ, 90^\circ]$.

As a first step, we have investigated which of the 73,454 human α rhythm sets from Bojak and Liley (2005) can be extended viably via Eq. 12 and Eq. 13. Chen et al. (2009) found for rat pyramidal neurons that $\Delta h_e^r = -4.0$ mV for ketamine at $20 \mu\text{M}$ concentration and $\Delta h_e^r = -3.7$ mV for propofol at $5 \mu\text{M}$ concentration. Assuming that in humans (and in inhibitory neurons) hyperpolarizations of similar sizes occur, we varied both h_e^r and h_i^r away from their original values in steps of -0.05 mV up to a hyperpolarization of -6 mV, while the remaining parameters were left unchanged. This leads to a grid of 121×121 hyperpolarization combinations $(\Delta h_e^r, \Delta h_i^r)$, for which we tested whether the changed parameter sets remain stable, i.e., we computed eigendecompositions for $64 + 1$ values of k and made sure that all eigenvalues had negative real parts. We also calculated the resulting shift in the α peak frequency as compared to the original parameter set in the “theoretical” manner discussed above: $\Delta f(\Delta h_e^r, \Delta h_i^r) \equiv f_{\max}^{\{P(\Delta h_e^r, \Delta h_i^r)\}} - f_{\max}^{\{P(0,0)\}}$.

We find that of the 73,454 parameter sets only 1,627 remain stable for all 121×121 combinations of hyperpolarizations up to -6 mV. This does not mean that the other parameter sets are thereby rejected on biological or physiological grounds;

rather their PSDs cannot be calculated with the eigenspectrum approximation used here, but would have to be estimated from explicit simulations with the fully non-linear Eqs 6–8. This ordinary numerical procedure is several orders of magnitude slower and hence not employed here. **Figure 3A** displays the average $\langle \Delta f(\Delta h_e^r, \Delta h_i^r) \rangle$ over the 1,627 stable sets. The color bar indicates the corresponding frequency values. We can see that in this average there is little effect of Δh_i^r , whereas decreasing Δh_e^r (increasing the hyperpolarization of the pyramidal neurons) leads to an increasingly negative $\langle \Delta f \rangle$. The lowest average value for the 1,627 sets in **Figure 3A** is $\langle \Delta f(-6 \text{ mV}, -6 \text{ mV}) \rangle = -2.03 \text{ Hz}$, whereas the highest is $\langle \Delta f(0 \text{ mV}, -6 \text{ mV}) \rangle = 0.336 \text{ Hz}$.

Since more substantial increases in frequency are expected for the interaction of ketamine and propofol (Hayashi et al., 2007; Tsuda et al., 2007), we introduce the following cut: a set will be kept only if for at least one of the 121×121 hyperpolarization combinations $(\Delta h_e^r, \Delta h_i^r)$ we find $\Delta f > 1.6 \text{ Hz}$. Similarly, since ketamine on its own should introduce a decrease in Δf (Schuttler et al., 1987; Kochs et al., 1996), we require that for at least one other hyperpolarization combination $\Delta f < -0.8 \text{ Hz}$. Finally, propofol on its own is assumed here to not change the α frequency significantly $\Delta f \approx 0 \text{ Hz}$ (Schwender et al., 1996; Kuizenga et al., 1998, 2001; Feshchenko et al., 2004; Breshears et al., 2010; Cimenser et al., 2011), at least not by a HCN1-mediated mechanism, as was discussed in the Introduction. It is more difficult to introduce a simple cut for this property, since for small hyperpolarizations by definition one finds $\Delta f \approx 0 \text{ Hz}$. We orient ourselves here to $\Delta h_e^r = -3.7 \text{ mV}$ for propofol from Chen et al. (2009), and require that at least for one combination with $\Delta h_i^r \leq -4.3 \text{ mV}$ one has $|\Delta f(-3.7 \text{ mV}, \leq 4.3 \text{ mV})| < 0.4 \text{ Hz}$. Considered individually, the low frequency cut for ketamine eliminates only 80 parameter sets, whereas the high frequency cut for the interaction of ketamine and propofol leaves only 66 parameter sets. Combining these two cuts then leaves 64 parameter sets in total. Individually, the cut for propofol limiting the frequency shift leaves 149 parameter sets. Combined with the other two cuts, we arrive at 10 parameter sets. Their original PSDs are the ones that were displayed previously in **Figure 2**, and we display their parameter values in **Table A1** in the Appendix. We show the resulting $\langle \Delta f \rangle$, now averaging over only the 10 selected sets, in **Figure 3B**. It is immediately apparent that there are now three zones: for small Δh_e^r but large (negative) Δh_i^r one sees large increases in frequency, for large (negative) Δh_e^r but small Δh_i^r large decreases in frequency, and in between there is a corridor with little change in frequency. This same basic structure is found in all 10 selected sets individually.

Now we can use this structure to determine the coefficients in Eq. 12 and Eq. 13. Starting with the case of giving ketamine only, we can write

$$P = 0 : \tan \theta_K = \frac{b_2}{a_2}. \quad (14)$$

Thus the effect of increasing ketamine concentration in the plane of hyperpolarizations is to move out along a line through the origin with angle θ_K . Ketamine on its own is supposed to deliver shifts to low frequencies, for which we have set a cut $\Delta f < -0.8 \text{ Hz}$ above. We now determine for every

hyperpolarization combination how many of the 10 selected parameter sets have $\Delta f(\Delta h_e^r, \Delta h_i^r) < -0.8 \text{ Hz}$. This leads to a 121×121 grid of values between 0 and 10. In **Figure 3C** this is shown by blue contour lines for 4, 7, and 10 sets fulfilling this cut. We choose the mean of all $(\Delta h_e^r, \Delta h_i^r)$ in the “maximal fulfillment” (10 sets) region (tip of blue arrow) to determine $\tan \theta_K = 0.315$. Given that we do not know the dependence of hyperpolarizations on ketamine concentrations in humans, we choose $\Delta h_e^r \equiv -4 \text{ mV}$ at $K = 1$ and thus consequently $a_2 \equiv 4 \text{ mV}$. This implies an unknown normalization $K \equiv c_K/c_K^*$, so that at a ketamine concentration c_K^* one finds $\Delta h_e^r = -4 \text{ mV}$ in humans. Given this choice, we have $b_2 \equiv 1.26 \text{ mV}$ from the ketamine angle $\tan \theta_K = 0.315$. Since $a_2 > b_2$, we can now also find $K_{\max} = (6 \text{ mV})/a_2 = 1.5$ as the largest value for the normalized ketamine concentration for which both hyperpolarizations remain below -6 mV .

In a similar manner we can deal with the case of propofol as the sole drug. Then we find the angular dependence:

$$K = 0 : \tan \theta_P = \frac{b_1}{a_1}. \quad (15)$$

Figure 3C shows contour lines for 4, 7, and 10 parameter sets fulfilling the cut for an α frequency shift $|\Delta f(\Delta h_e^r, \Delta h_i^r)| < 0.4 \text{ Hz}$, this time in green color. Since the cut was evaluated for $\Delta h_e^r = -3.7 \text{ mV}$ only to find these sets, we determine the mean of combinations $(\Delta h_e^r \equiv -3.7 \text{ mV}, \Delta h_i^r)$ that have “maximal fulfillment” (10 sets) in order to obtain $\tan \theta_P = 1.297$, indicated by the tip of the green arrow. We choose $\Delta h_e^r \equiv -3.7 \text{ mV}$ at $P \equiv c_P/c_P^* = 1$, so that $a_1 \equiv 3.7 \text{ mV}$ and at an unknown propofol concentration c_P^* one finds $\Delta h_e^r = -3.7 \text{ mV}$ in humans. Then $b_1 \equiv 4.8 \text{ mV}$, and since $b_1 > a_1$ it follows that $P_{\max} = (6 \text{ mV})/b_1 = 1.25$.

Finally, **Figure 3C** shows red contour lines for 4, 7, and 10 parameter sets fulfilling $\Delta f(\Delta h_e^r, \Delta h_i^r) > 1.6 \text{ Hz}$. Again we find the mean of “maximal fulfillment” (10 parameter sets), as indicated by the tip of the red arrow. These mean values are $\Delta h_e^r \equiv -0.177 \text{ mV}$ and $\Delta h_i^r \equiv -5.903 \text{ mV}$ in this case. We now extend to the -6 mV hyperpolarization limit by setting $\Delta h_i^{\max} \equiv -6 \text{ mV}$ and $\Delta h_e^{\max} \equiv (-6 \text{ mV})(-0.177 \text{ mV})/(-5.903 \text{ mV}) = -0.180 \text{ mV}$. We can now solve the following two equations

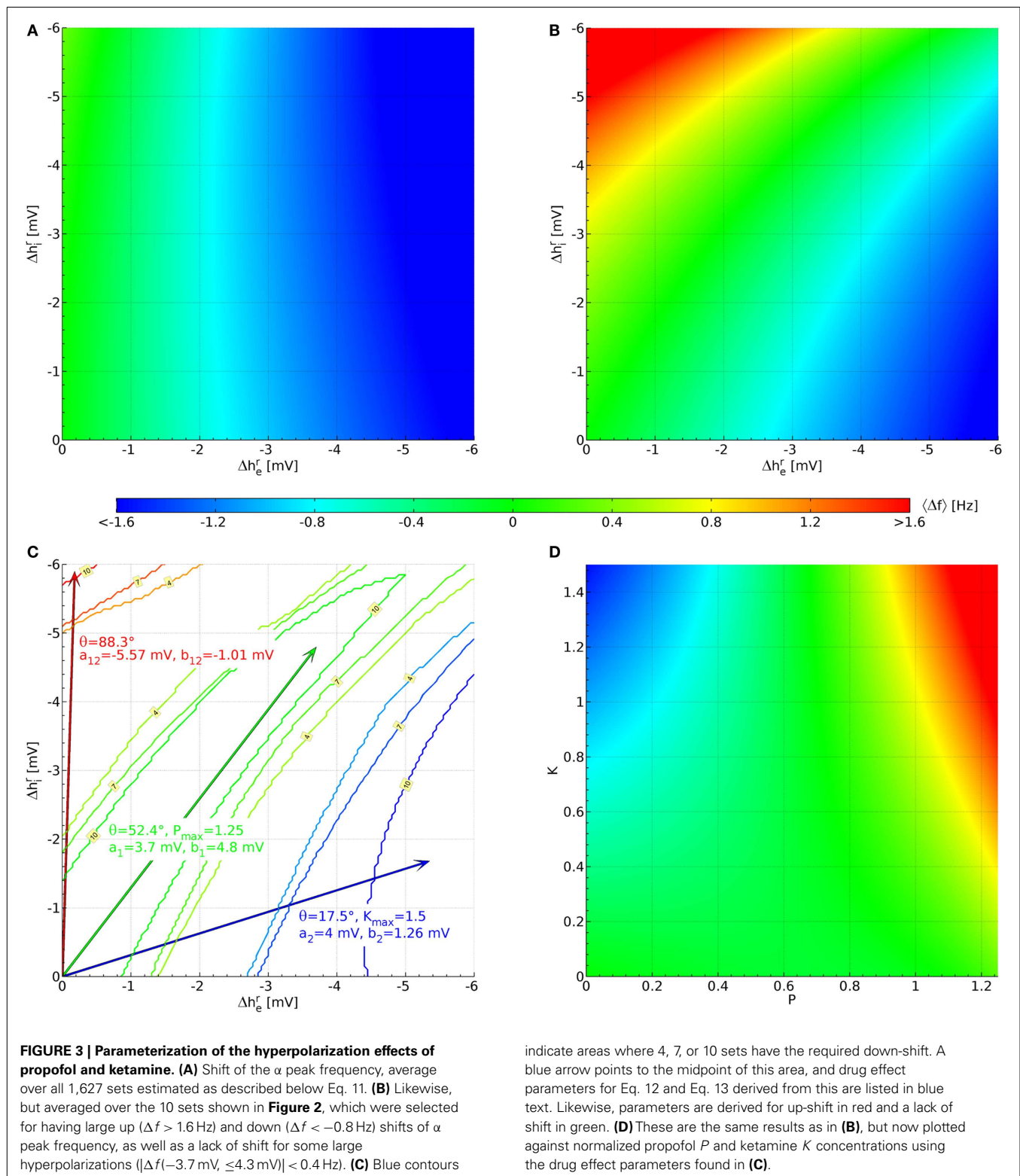
$$\Delta h_e^{\max} = -(a_1 P_{\max} + a_2 K_{\max} + a_{12} P_{\max} K_{\max}), \quad (16)$$

$$\Delta h_i^{\max} = -(b_1 P_{\max} + b_2 K_{\max} + b_{12} P_{\max} K_{\max}). \quad (17)$$

This will then mean that our entire hyperpolarization grid $-6 \text{ mV} \leq \Delta h_e^r, \Delta h_i^r \leq 0 \text{ mV}$ will be projected onto a rectangular area bounded by $0 \leq P \leq P_{\max}$ and $0 \leq K \leq K_{\max}$, respectively. Solving Eq. 16 and Eq. 17 with our previous results yields $a_{12} = -5.57 \text{ mV}$ and $b_{12} = -1.01 \text{ mV}$. **Figure 3D** shows the projected $\langle \Delta f(P, K) \rangle$. Clearly the intended α frequency shifts are now achieved: negative ones for only ketamine, none for only propofol, and positives ones for propofol and ketamine together.

RESULTS

We have parameterized the HCN1-mediated hyperpolarizations of neuron membrane potentials in order to reproduce



the observed EEG effects of ketamine and propofol, and in particular of their interaction when concurrent. The coefficients that we have obtained for Eqs 12–13 afford the following interpretation: pyramidal neurons react similarly to

ketamine and propofol ($a_1 = 0.925 \times a_2$), whereas inhibitory neurons react much more strongly to propofol than to ketamine ($b_1 = 3.81 \times b_2$). Furthermore, and perhaps most interestingly, there is an antagonism of ketamine and propofol (a_{12} ,

$b_{12} < 0$), which leads to infra-additivity in the investigated effect of HCN1-mediated hyperpolarization, cf. Eq. 4. This antagonism is stronger in pyramidal neurons $a_{12}/(a_1 P_{\max} + a_2 K_{\max}) = 4.10$ $b_{12}/(b_1 P_{\max} + b_2 K_{\max})$, though the precise proportion depends on the given concentrations of the drugs. Intuitively it makes sense however that in inhibitory neurons, where one drug is much more effective than the other, the antagonism between the drugs is less pronounced.

To illustrate these results we look again at the “theoretical” estimates of the α peak frequency in **Figure 4**, where we compare now the effects of changing propofol and ketamine concentration on the 10 selected sets (red) with those computed for all the valid 1,627 sets (gray). Note that the 1,627 sets include the 10 selected ones. Quantile bands are computed to summarize the results for the individual parameter sets, as indicated by the legend. Starting from a baseline without drugs, four phases are being considered: first, propofol concentration is increased linearly; then propofol is maintained at maximum concentration and ketamine

concentration is increased linearly; next propofol concentration is decreased linearly while ketamine is maintained at maximum concentration, and finally ketamine concentration is decreased linearly for a return to the baseline. It should be noted that no attempt at modeling the pharmacodynamics/pharmacokinetics of ketamine and propofol drug action beyond drug interaction has been made here. Furthermore, the eigenspectrum approach assumes that the system has reached equilibrium for the given parameters. Thus every single $f_{\max}(P, K)$ predicted here, and consequently every single quantile band value, represents a “steady state” result for that particular drug concentration combination. Hence one can for example view **Figure 4** from right to left, beginning with an increase in ketamine concentration, followed by an increase in ketamine at maximum propofol concentration, and so forth.

Comparing now the red with the gray quantile bands, we see that our cuts selected sets that react particularly dramatically to the concurrence of propofol and ketamine (phases 2 and 3), while

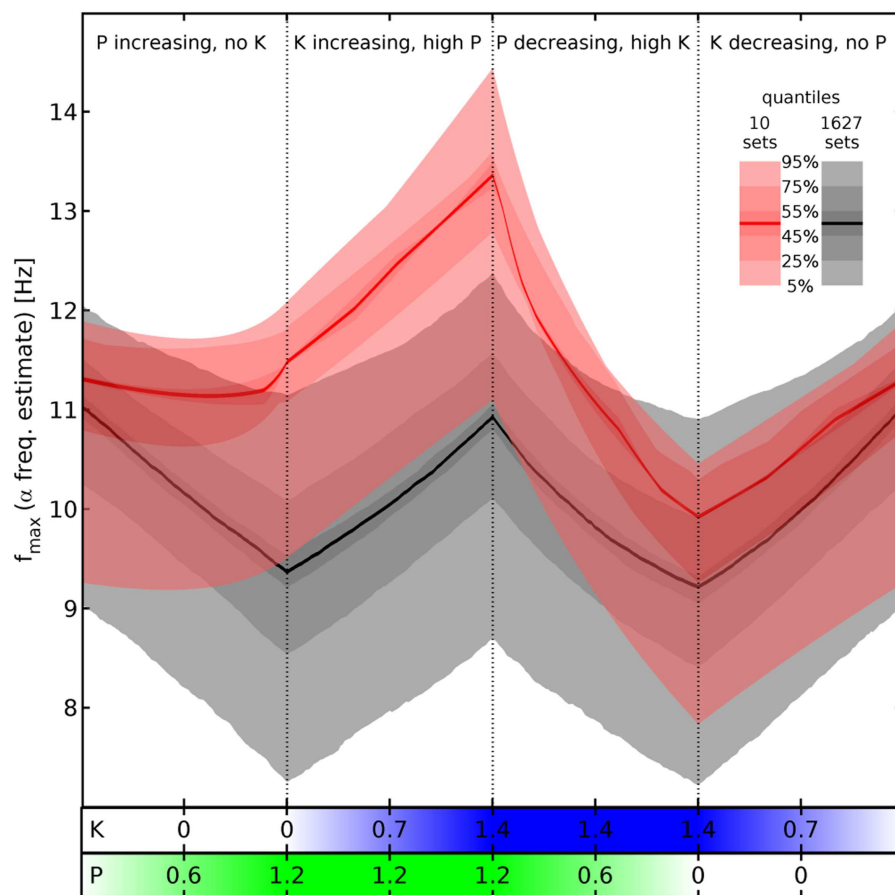


FIGURE 4 | Estimated α peak frequency shifts. Shifts of the α peak frequency for normalized propofol P and ketamine K concentrations estimated as described below Eq. 11, using the hyperpolarizations in Eq. 12 and Eq. 13. Either all 1,627 (gray) or the 10 selected sets (red) are used to compute quantile bands, as indicated by the legend. The median value is shown by a thick black or red line, respectively. There are four phases of drug variation, as indicated by the titles and dotted

lines, quantified by bars below the main panel: first, $P = 0 \rightarrow 1.2$ linearly, while $K = 0$. Then $K = 0 \rightarrow 1.4$ linearly, while $P = 1.2$. Next $P = 1.2 \rightarrow 0$, while $K = 1.4$. Finally, $K = 1.4 \rightarrow 0$, while $P = 0$. No pharmacodynamics has been modeled here, so every (P, K) combination yields an independent “steady state” result. Hence for example an increase of P at high K is shown by the third phase viewed from right to left.

being unresponsive to propofol alone (phase 1). Nevertheless, it is not the case that the results for the 1,627 sets show a totally divergent response pattern. In fact, the median rise of estimated α peak frequency in phase 2 (upon introducing ketamine at maximum propofol concentration) is comparable: 1.88 Hz for the selected sets (from 11.48 to 13.36 Hz) vs. 1.56 Hz for all sets (from 9.37 to 10.93 Hz). Thus the predicted boost of α peak frequencies due to the interaction between ketamine and propofol is a robust result for all sets given our drug effect parameterization, which is infra-additive concerning HCN1-mediated hyperpolarization. The main difference appears to be rather that the α peak frequencies of the selected sets do not react significantly to propofol, whereas they are similar to all other sets in the reaction to ketamine and the interaction between these drugs.

Turning to results for full PSDs from Eq. 11, we will consider the 10 selected sets only due to the higher computational demands. **Figure 5** shows results for one individual set (Set III of **Figure 2**) under three variations of drug concentration. In **Figure 5A** we see that as desired and estimated, the α peak frequency stays roughly the same during propofol anesthesia (Schwender et al., 1996; Kuizenga et al., 1998, 2001; Feshchenko et al., 2004; Breshears et al., 2010; Cimenser et al., 2011). The damping seen here would be more characteristic of occipital than frontal regions, though other processes in particular related to the GABA_A agonism could

modify these results. The “theoretical” $\Delta f = 0.24$ Hz at $P = 1.2$ is larger than the “experiment-like” shift of 0.13 Hz. In **Figure 5B** we can see the reaction to increasing ketamine concentration. As expected, the α peak gets shifted to lower frequencies. The “theoretical” $\Delta f = -1.37$ Hz estimate at $K = 1.4$ is somewhat lower than the “experiment-like” shift of the local maxima of the PSDs of -1.69 Hz. While the α oscillations get dampened, they contribute to a net increase in the θ frequency range (Schuttler et al., 1987; Kochs et al., 1996) thanks to their downward frequency shift. But one sees also a general rise in power at lower frequencies. **Figure 5C** shows that adding ketamine in the presence of high doses of propofol leads to a shift of the α peak into the beta band, as observed by Hayashi et al. (2007) and Tsuda et al. (2007). The “theoretical” $\Delta f = 1.91$ Hz at $P = 1.2$ and $K = 1.4$ is smaller than the “experiment-like” shift of 2.27 Hz. We predict here an increase in power, unlike the experiment, which observed a small but significant reduction, and our frequency shift of 2.27 Hz is less than half the observed 4.7 Hz. But this could be explained easily by the missing GABA_A and NMDA mechanisms, or the precise circumstances of the experiment. A bolus of ketamine was given by intravenous injection in the experiment, whereas here we calculate “steady state” results. Furthermore, while the frequency shifts from our “theoretical” Δf method are largely in agreement with those obtained from the local α maxima of the PSDs, they do not

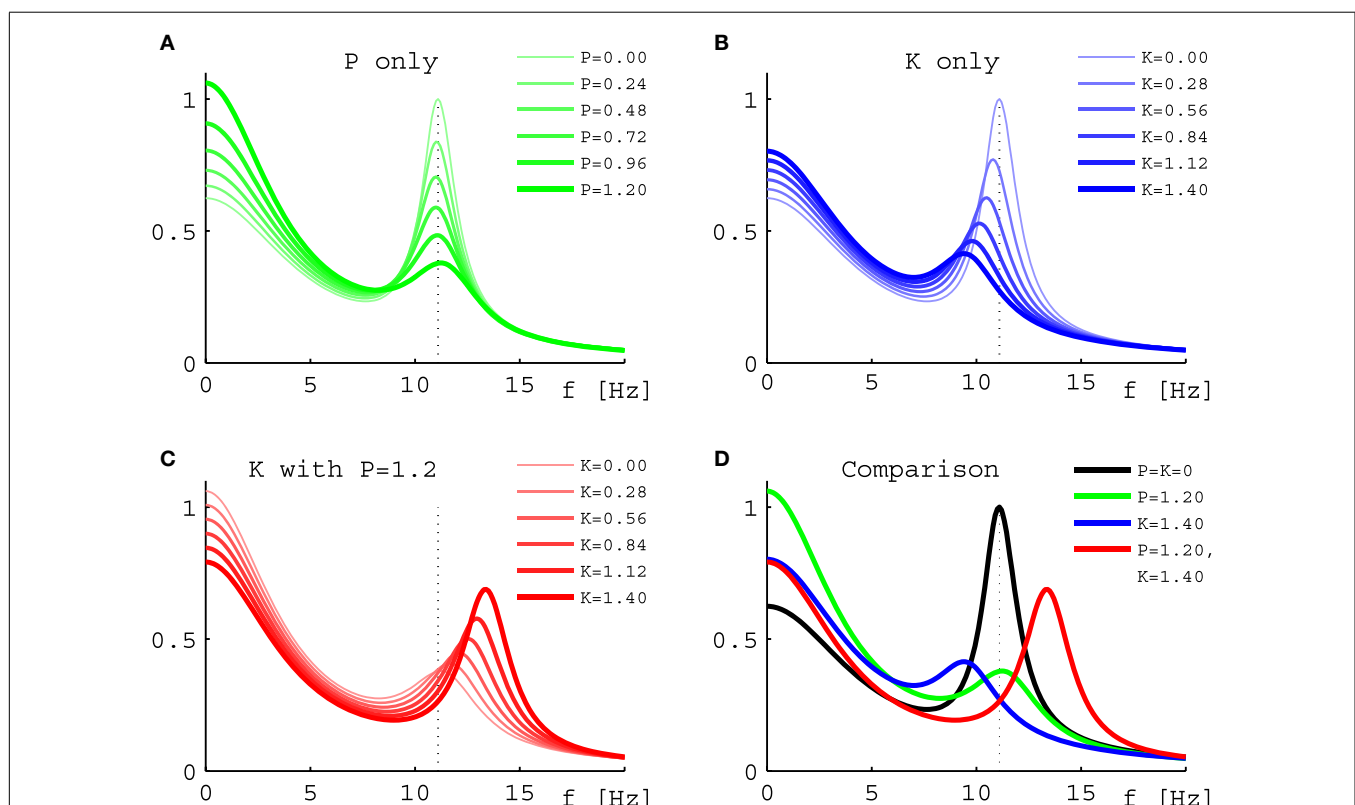


FIGURE 5 | Power spectral densities for Set III under drug variations.

(A) PSDs for increasing normalized propofol concentration from none (thinnest green line) to 1.2 (thickest green line). (B) PSDs for increasing normalized ketamine concentration from none (thinnest blue line) to 1.4 (thickest blue line). (C) PSDs for increasing normalized ketamine concentration

from none (thinnest red line) to 1.4 (thickest red line), while normalized propofol concentration is held constant at 1.2. (D) Comparison of the PSDs representing the highest normalized concentrations from (A) in green, (B) in blue, and (C) in red. The black curve is the PSD without drugs. In all four panels the dotted line represents the position of the α peak of this curve.

agree perfectly – in spite of both being derived using the same eigenspectrum technique. The difference is that by looking at the maxima the results are influenced by changes to the overall spectrum, which provide the “background” on which the α resonance sits. Table 1 gives Δf and “experiment-like” α peak frequency shifts for the 10 sets at the three highlighted drug concentrations.

Finally, in Figure 6 we show similar results for all the 10 selected sets. We follow here the same scheme of changing drug concentrations as in Figure 4. We see that the α peaks of the full PSDs (here shown in decibels by color) of the individual sets indeed follow the “zigzag” shape we saw in the quantile bands of Figure 4. Panel III in Figure 6 can be directly compared to Figure 5, which we have just discussed. For example, Figure 5A corresponds to the first phase in panel III here. Overall we see that while the sets clearly change in a similar way, they all have individual features that set them apart from the others. For example, parameter set IV shows particularly strong changes in the low frequency range, whereas parameter set VII reacts with a particularly strong lowering of the α peak frequency in the presence of ketamine. These variations can be considered as representing the variations that one can also observe in humans.

DISCUSSION

We have shown that observed changes of the EEG α peak frequency induced by the presence of the anesthetic agents propofol

and ketamine, but in particular also by their interaction when given concurrently, can be explained based on the modeling of HCN1-mediated hyperpolarizations alone, at least qualitatively. This is surprising, since the main mechanism of action of these drugs is supposed to be through NMDA antagonism (ketamine) and GABA_A agonism (propofol), respectively. However, since HCN1-knockout mice are indeed less sensitive to the hypnotic effects of both drugs (Chen et al., 2009), this would indicate that the EEG remains useful as an indicator of anesthetic action. It is perhaps interesting to note that while ketamine is famous for its hallucinatory action (Wolff and Winstock, 2006), and hence is considered a dissociative anesthetic agent, propofol as a classic general anesthetic agent is also capable of inducing a range of hallucinatory phenomena (Balasubramaniam and Park, 2003). Hence it is possible that psychotropic HCN1-mediated effects are simply masked behaviorally more by propofol’s GABA_A agonism than by ketamine’s NMDA antagonism, but that the EEG is particularly sensitive to these underlying changes.

Only a fraction of all considered parameter sets (1,627 of the 73,454 parameter sets from Bojak and Liley (2005) proved “stable” under the HCN1-mediated hyperpolarization changes up to -6 mV on both excitatory and inhibitory neurons. However, this is at least partly due to the computational methods used here: the eigenspectrum method (Bojak and Liley, 2005) can only be used for “fixed point” dynamics. Rejected sets could follow the

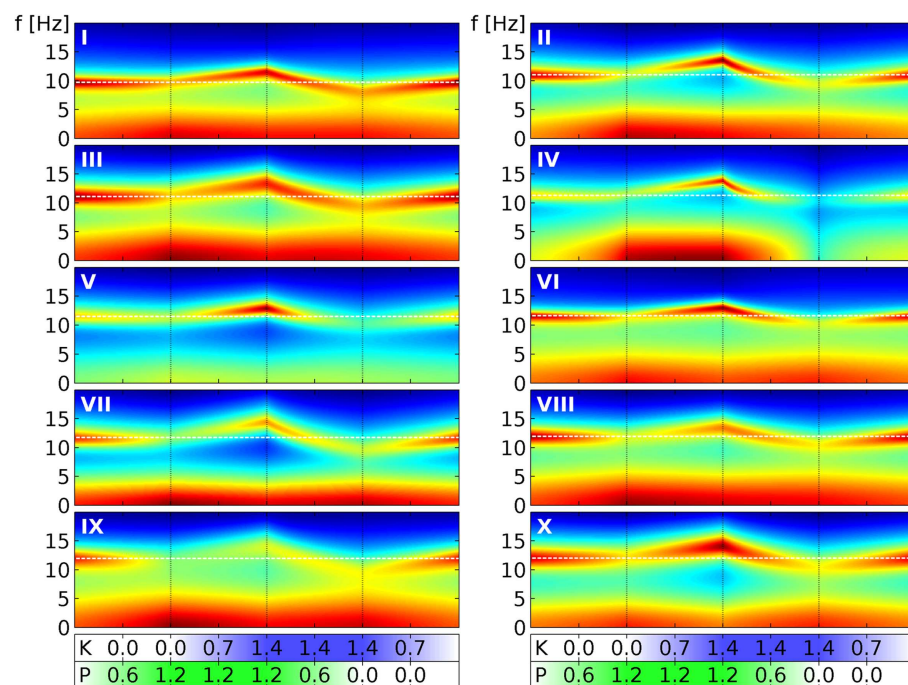


FIGURE 6 | Power spectral densities for all 10 selected parameter sets under drug variation. We use here the same four phases of drug variation as in Figure 4, as indicated by the dotted lines and quantified by bars below the main panels: first, $P = 0 \rightarrow 1.2$ linearly, while $K = 0$. Then $K = 0 \rightarrow 1.4$ linearly, while $P = 1.2$. Next $P = 1.2 \rightarrow 0$, while $K = 1.4$. Finally, $K = 1.4 \rightarrow 0$, while $P = 0$. Every panel corresponds to 1 of the 10 selected parameter sets, as indicated by a white roman numeral. The PSD for one

specific (P , K) combination is indicated in the panel by a colored vertical line corresponding to frequencies from 0 to 20 Hz. Colors here indicate decibels of the PSD, with dark red corresponding to large, green to medium and dark blue to small values. (The “jet” colormap of Matlab has been mapped for each panel individually, to the full range of PSD decibel values shown in the panel.) A white dashed line indicates the α peak frequency in the absence of drugs.

same kind of drug-induced changes, but be inaccessible with our method chosen for its computational speed. In principle it would be desirable to carry out fully non-linear calculations instead, but under extra-cortical noise input it takes about a minute to estimate a sufficiently detailed and accurate power spectrum on a regular PC. For the 121×121 hyperpolarization changes of **Figure 3** that would lead to about 10 days of calculations even for a single parameter set – but we investigate here 73,454 different ones. The eigenspectrum method is several orders of magnitude faster even for a single neural mass. Yet we include here also effects due to the spatial distribution of neural masses, see Eq. 11. The computational load scales roughly linearly with the number of integration points for the eigenspectrum method, but roughly as a square for an equivalent fully non-linear simulation based on spatial grids. This increases the difference in computation speed even further. Thus for the investigations carried out here only the eigenspectrum method proves practicable. In addition, limit cycle or chaotic dynamics are often more representative of seizures or other pathological brain states, which in clinical practice would lead to the termination of the pharmacological intervention that we intend to describe here. Other than by numerical simulation in every single case, we do not know how to determine the characteristics of the dynamics past the point of instability.

Furthermore, our current investigation does not include the NMDA and GABA_A actions commonly assumed to be dominant in these drugs. We speculate that a more complete simulation could allow the use of a larger fraction of the 73,454 parameter sets, since these omitted actions can affect the required stability. A prolongation of the inhibitory postsynaptic potentials due to GABA_A, for example, could suppress excessive excitation and thus stabilize a parameter set. These neglected stabilizing effects would increase also in due proportion to the agent concentration, just as the potentially destabilizing hyperpolarizations we have modeled here do. In order to obtain spectral changes that demonstrate clearly the expected frequency shifts, we introduced three further selection cuts, leaving us with only 10 parameter sets out of the 1,627. Again we speculate that NMDA and GABA_A actions may ameliorate this reduction. If this is not the case, then this may point to underlying correlations between neural parameters or functional properties that were not considered in Bojak and Liley (2005), but which now prove crucial for a realistic description. Note that in terms of the hyperpolarizations, cf. **Figures 3B,C**, the non-reactivity to propofol means that for the 10 selected parameter sets an increase in the hyperpolarization of excitatory neurons can be compensated by an increase in the hyperpolarization of inhibitory neurons. This could suggest a particular intrinsic balance of excitation and inhibition maintaining functional stability against extrinsic disturbances.

As is apparent from **Figure 2**, most of our 10 selected sets have relatively high α peak frequencies. However, this simply reflects an underlying bias in the original 73,454 parameter sets, cf. **Figure 8** in Bojak and Liley (2005); and we were able to lower base α frequencies through the adjustment of a few parameters (e.g., cortico-cortical connectivity N_{lk}^B) within physiological limits, without thereby qualitatively changing the relative frequency shifts due to propofol and ketamine. Other spectral features that could be selected for are also unlikely to affect the α frequency

shifts here qualitatively. The constraints used by Bojak and Liley (2005) leave plenty of room for such adjustments: effectively only four out of 14 system eigenvalues were used to establish the “ $1/f$ ” background and the α resonance. Adding a weak beta frequency resonance to enhance realism, for example, would only require the adjustment of two further system eigenvalues. Thus we expect that our results here would hold true if one were to redo the parameter space search of Bojak and Liley (2005) first with additional constraints on the base power spectra. Finally, for concurrent application of propofol and ketamine we predict increases of the α peak frequency of around 2 Hz, falling short of the 4.7 Hz seen experimentally (Hayashi et al., 2007; Tsuda et al., 2007). Yet our calculations were for “steady state” concentrations, and for rapid increases in dosage, as in the experimental injection of a single ketamine bolus here, one often finds a more complex response due the pharmacodynamics and transient neural responses. The so-called biphasic responses to anesthetic agents (Kuizenga et al., 1998, 2001) has received theoretical attention from several groups (Bojak and Liley, 2005; Wilson et al., 2006; Molaee-Ardekani et al., 2007; Hutt and Longtin, 2010; Steyn-Ross and Steyn-Ross, 2010), see also the review in Foster et al. (2008). Hints of such a biphasic response could be visible in **Figure 5** of Tsuda et al. (2007), which shows a significant drop of the α peak shift with time from the initial 4.7 Hz to values around 2 Hz. However, this would have to be disentangled from the decrease in ketamine concentration due to natural clearance after the bolus. We intend to investigate all the mentioned issues in future work.

We found that we could account for the heterogeneous effects of ketamine on the EEG if we assumed that propofol and ketamine interacted in an *infra-additive* or *antagonistic* manner in their inhibition of HCN1-mediated neuronal membrane hyperpolarization. While most anesthetic and sedative agents are reported to interact synergistically ketamine is well-known to be a major exception (Hendrickx et al., 2008). The interaction between ketamine and GABA_A agonists (most sedative/anesthetic agents) to produce hypnosis is reported to range from additivity to infra-additivity/antagonism (Hendrickx et al., 2008). On the basis of the limited clinical data available it appears that the GABA_A agonist propofol can interact infra-additively to produce a given hypnotic endpoint. For instance Hui et al. (1995), in a study involving 180 female patients presenting for minor gynecological surgery, calculated quantal dose-response curves for propofol and ketamine administered alone and in combination. On the basis of logarithmic regression of a response surface model, it was first suggested that the dose-response for the combination was best explained by additivity, but on reanalysis (Hendrickx et al., 2008), by infra-additivity (significantly so for immobility, as trend for hypnosis). In apparent contradiction, Sakai et al. (1999) claim a significantly additive interaction between propofol and ketamine for the endpoints of unresponsiveness to vocal command and the loss of eyelash reflex. However, the former study determined the probability of a given hypnotic endpoint in response to *single doses* of fixed proportions of ketamine and propofol, whereas the latter gave *continuous infusions* of ketamine and propofol until a given hypnotic endpoint was reached. Another study by Frizelle et al. (1997) used smaller doses of ketamine and targeted a lower level of sedation (rousable to verbal stimuli) with a bolus of propofol

and ketamine followed by concomitant infusion. They found no statistical evidence that the addition of ketamine reduced the required amount of propofol to reach their intended sedation level, suggesting once more infra-additivity. Clearly, additional work regarding the pharmacodynamic interactions of ketamine and propofol are required, as it is difficult to reconcile the results of these studies.

The use of neural field/mass approaches to modeling drug action on the EEG is emerging as a powerful explanatory framework (Liley and Bojak, 2005; Foster et al., 2008; Hutt, 2011), which is able to retain meaningful connections to the brain's physiology despite its mesoscopic scale of description. Because neural field/mass models generally have much smaller parameter and state spaces than biophysically plausible neural network

models, they are not only much easier to parameterize and simulate, but offer a simpler framework from which to make predictions and derive hypotheses that can be empirically tested. Our ability to account qualitatively for the effects that propofol and ketamine have on the EEG adds to a growing list of phenomena that are amenable to neural field/mass description (Deco et al., 2008; Coombes, 2010; Bressloff, 2012; Liley et al., 2012). Ultimately it is hoped that by accounting for large scale phenomena using neural field/mass models, genuine and enduring insights into brain function and physiology will emerge.

ACKNOWLEDGMENTS

Harry C. Day is supported by a scholarship from the College of Life and Environmental Sciences at the University of Birmingham.

REFERENCES

- Balasubramaniam, B., and Park, G. R. (2003). Sexual hallucinations during and after sedation and anaesthesia. *Anaesthesia* 58, 549–553.
- Biel, M. (2009). Cyclic nucleotide-regulated cation channels. *J. Biol. Chem.* 284, 9017–9021.
- Biel, M., Wahl-Schott, C., Michalakakis, S., and Zong, X. (2009). Hyperpolarization-activated cation channels: from genes to function. *Physiol. Rev.* 89, 847–885.
- Bojak, I., and Liley, D. T. J. (2005). Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 71, 041902.
- Bojak, I., and Liley, D. T. J. (2010). Axonal velocity distributions in neural field equations. *PLoS Comput. Biol.* 6:e1000653. doi:10.1371/journal.pcbi.1000653
- Breshears, J. D., Roland, J. L., Sharma, M., Gaona, C. M., Freudenburg, Z. V., Tempelhoff, R., et al. (2010). Stable and dynamic cortical electrophysiology of induction and emergence with propofol anesthesia. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21170–21175.
- Bressloff, P. C. (2012). Spatiotemporal dynamics of continuum neural fields. *J. Phys. A* 45, 033001.
- Bubenikova-Valesova, V., Horacek, J., Vrajova, M., and Hoschl, C. (2008). Models of schizophrenia in humans and animals based on inhibition of NMDA receptors. *Neurosci. Biobehav. Rev.* 32, 1014–1023.
- Chen, X., Shu, S., and Bayliss, D. A. (2009). HCN1 channel subunits are a molecular substrate for hypnotic actions of ketamine. *J. Neurosci.* 29, 600–609.
- Cimenser, A., Purdon, P. L., Pierce, E. T., Walsh, J. L., Salazar-Gomez, A. F., Harrell, P. G., et al. (2011). Tracking brain states under general anesthesia by using global coherence analysis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8832–8837.
- Coombes, S. (2010). Large-scale neural dynamics: simple and complex. *Neuroimage* 52, 731–739.
- Corlett, P. R., Honey, G. D., Krystal, J. H., and Fletcher, P. C. (2011). Glutamatergic model psychoses: prediction error, learning, and inference. *Neuropsychopharmacology* 36, 294–315.
- Cox, E. H., Knibbe, C. A., Koster, V. S., Langemeijer, M. W., Tukker, E. E., Lange, R., et al. (1998). Influence of different fat emulsion-based intravenous formulations on the pharmacokinetics and pharmacodynamics of propofol. *Pharm. Res.* 15, 442–448.
- Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., and Friston, K. J. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* 4:e1000092. doi:10.1371/journal.pcbi.1000092
- Faraoni, D., Salengros, J. C., Engelman, E., Ickx, B., and Barvais, L. (2009). Ketamine has no effect on bispectral index during stable propofol-remifentanyl anaesthesia. *Br. J. Anaesth.* 102, 336–339.
- Feshchenko, V. A., Veselis, R. A., and Reinsel, R. A. (2004). Propofol-induced alpha rhythm. *Neuropsychobiology* 50, 257–266.
- Foster, B. L., Bojak, I., and Liley, D. T. J. (2008). Population based models of cortical drug response: insights from anaesthesia. *Cogn. Neurodyn.* 2, 283–296.
- Frizelle, H. P., Duranteau, J., and Samii, K. (1997). A comparison of propofol with a propofol-ketamine combination for sedation during spinal anesthesia. *Anesth. Analg.* 84, 1318–1322.
- Grant, I. S., Nimmo, W. S., McNicol, L. R., and Clements, J. A. (1983). Ketamine disposition in children and adults. *Br. J. Anaesth.* 55, 1107–1111.
- Greco, W. R., Bravo, G., and Parsons, J. C. (1995). The search for synergy: a critical review from a response surface perspective. *Pharmacol. Rev.* 47, 331–385.
- Hans, P., Dewandre, P. Y., Brichant, J. F., and Bonhomme, V. (2005). Comparative effects of ketamine on bispectral index and spectral entropy of the electroencephalogram under sevoflurane anaesthesia. *Br. J. Anaesth.* 94, 336–340.
- Hayashi, K., Tsuda, N., Sawa, T., and Hagiwara, S. (2007). Ketamine increases the frequency of electroencephalographic bicoherence peak on the alpha spindle area induced with propofol. *Br. J. Anaesth.* 99, 389–395.
- Hendrickx, J. F., Eger, E. I. II, Sonner, J. M., and Shafer, S. L. (2008). Is synergy the rule? *Anesth. Analg.* 107, 494–506.
- Hevers, W., Hadley, S. H., Luddens, H., and Amin, J. (2008). Ketamine, but not phencyclidine, selectively modulates cerebellar GABA(A) receptors containing alpha6 and delta subunits. *J. Neurosci.* 28, 5383–5393.
- Hindriks, R., and van Putten, M. J. (2012). Meanfield modeling of propofol-induced changes in spontaneous EEG rhythms. *Neuroimage* 60, 2323–2334.
- Hui, T. W., Short, T. G., Hong, W., Suen, T., Gin, T., and Plummer, J. (1995). Additive interactions between propofol and ketamine when used for anesthesia induction in female patients. *Anesthesiology* 82, 641–648.
- Hutt, A. (2011). *Sleep and Anesthesia: Neural Correlates in Theory and Experiment*. New York: Springer.
- Hutt, A., and Longtin, A. (2010). Effects of the anesthetic agent propofol on neural populations. *Cogn. Neurodyn.* 4, 37–59.
- Hutt, A., and Schimansky-Geier, L. (2008). Anesthetic-induced transitions by propofol modeled by non-local neural populations involving two neuron types. *J. Biol. Phys.* 34, 433–440.
- Irifune, M., Katayama, S., Takarada, T., Shimizu, Y., Endo, C., Takata, T., et al. (2007). MK-801 enhances gabaculine-induced loss of the righting reflex in mice, but not immobility. *Can. J. Anaesth.* 54, 998–1005.
- Irifune, M., Shimizu, T., Nomoto, M., and Fukuda, T. (1992). Ketamine-induced anesthesia involves the N-methyl-D-aspartate receptor-channel complex in mice. *Brain Res.* 596, 1–9.
- Jirsa, V. K., and Haken, H. (1996). Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* 77, 960–963.
- Kelland, M. D., Soltis, R. P., Boldry, R. C., and Walters, J. R. (1993). Behavioral and electrophysiological comparison of ketamine with dizocilpine in the rat. *Physiol. Behav.* 54, 547–554.
- Kitamura, A., Marszalec, W., Yeh, J. Z., and Narahashi, T. (2003). Effects of halothane and propofol on excitatory and inhibitory synaptic transmission in rat cortical neurons. *J. Pharmacol. Exp. Ther.* 304, 162–171.
- Kochs, E., Scharein, E., Mollenberg, O., Bromm, B., and Schulte Am Esch, J. (1996). Analgesic efficacy of low-dose ketamine. *Anesthesiology* 85, 304–314.
- Kuizenga, K., Kalkman, C. J., and Henis, P. J. (1998). Quantitative electroencephalographic analysis of the biphasic concentration-effect relationship of propofol in surgical patients during extradural analgesia. *Br. J. Anaesth.* 80, 725–732.

- Kuizenga, K., Wierda, J. M., and Kalkman, C. J. (2001). Biphasic EEG changes in relation to loss of consciousness during induction with thiopental, propofol, etomidate, midazolam or sevoflurane. *Br. J. Anaesth.* 86, 354–360.
- Liley, D. T., and Bojak, I. (2005). Understanding the transition to seizure by modeling the epileptiform activity of general anesthetic agents. *J. Clin. Neurophysiol.* 22, 300–313.
- Liley, D. T., Cadusch, P. J., Gray, M., and Nathan, P. J. (2003). Drug-induced modification of the system properties associated with spontaneous human electroencephalographic activity. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 68, 051906.
- Liley, D. T. J., Bojak, I., Dafilis, M. P., Veen, L., Frascoli, F., and Foster, B. L. (2010). “Bifurcations and state changes in the human alpha rhythm: theory and experiment,” in *Modeling Phase Transitions in the Brain*, eds D. A. Steyn-Ross, and M. Steyn-Ross (New York: Springer), 117–145.
- Liley, D. T. J., Cadusch, P. J., and Dafilis, M. P. (2002). A spatially continuous mean field theory of electrocortical activity. *Network* 13, 67–113.
- Liley, D. T. J., Foster, B. L., and Bojak, I. (2011). “A mesoscopic modelling approach to anaesthetic action on brain electrical activity,” in *Sleep and Anesthesia: Neural Correlates in Theory and Experiment*, ed. A. Hutt (New York: Springer), 139–166.
- Liley, D. T. J., Foster, B. L., and Bojak, I. (2012). “Co-operative populations of neurons: mean field models of mesoscopic brain activity,” in *Computational Systems Neurobiology*, ed. N. Le Novère (Dordrecht: Springer), 315–362.
- Mathew, S. J., Shah, A., Lapidus, K., Clark, C., Jarun, N., Ostermeyer, B., et al. (2012). Ketamine for treatment-resistant unipolar depression: current evidence. *CNS Drugs* 26, 189–204.
- Molae-Ardekani, B., Senhadji, L., Shamsollahi, M. B., Vosoughi-Vahdat, B., and Wodey, E. (2007). Brain activity modeling in general anesthesia: enhancing local mean-field models using a slow adaptive firing rate. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 76, 041911.
- Murrough, J. W. (2012). Ketamine as a novel antidepressant: from synapse to behavior. *Clin. Pharmacol. Ther.* 91, 303–309.
- Murrough, J. W., Perez, A. M., Pillemer, S., Stern, J., Parides, M. K., Aan Het Rot, M., et al. (2012). Rapid and longer-term antidepressant effects of repeated ketamine infusions in treatment-resistant major depression. *Biol. Psychiatry*. doi:10.1016/j.biopsych.2012.06.022
- Nonaka, A., Makino, K., Suzuki, S., Ike-moto, K., Furuya, A., Tamaki, F., et al. (2012). Low doses of ketamine have no effect on bispectral index during stable propofol-remifentanyl anesthesia. *Masui* 61, 364–367.
- Oye, I., Paulsen, O., and Maurset, A. (1992). Effects of ketamine on sensory perception: evidence for a role of N-methyl-D-aspartate receptors. *J. Pharmacol. Exp. Ther.* 260, 1209–1213.
- Phillips, W., Anderson, A., Rosengreen, M., Johnson, J., and Halpin, J. (2010). Propofol versus propofol/ketamine for brief painful procedures in the emergency department: clinical and bispectral index scale comparison. *J. Pain Palliat. Care Pharmacother.* 24, 349–355.
- Robinson, P. A., Rennie, C. J., and Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 56, 826–840.
- Sakai, T., Singh, H., Mi, W. D., Kudo, T., and Matsuki, A. (1999). The effect of ketamine on clinical endpoints of hypnosis and EEG variables during propofol infusion. *Acta Anaesthesiol. Scand.* 43, 212–216.
- Schnoebel, R., Wolff, M., Peters, S. C., Brau, M. E., Scholz, A., Hempelmann, G., et al. (2005). Ketamine impairs excitability in superficial dorsal horn neurones by blocking sodium and voltage-gated potassium currents. *Br. J. Pharmacol.* 146, 826–833.
- Schuttler, J., Stanski, D. R., White, P. E., Trevor, A. J., Horai, Y., Verotta, D., et al. (1987). Pharmacodynamic modeling of the EEG effects of ketamine and its enantiomers in man. *J. Pharmacokin. Biopharm.* 15, 241–253.
- Schwender, D., Daunerer, M., Mulzer, S., Klasing, S., Finsterer, U., and Peter, K. (1996). Spectral edge frequency of the electroencephalogram to monitor “depth” of anaesthesia with isoflurane or propofol. *Br. J. Anaesth.* 77, 179–184.
- Sengupta, S., Ghosh, S., Rudra, A., Kumar, P., Maitra, G., and Das, T. (2011). Effect of ketamine on bispectral index during propofol–fentanyl anesthesia: a randomized controlled study. *Middle East J. Anesthesiol.* 21, 391–395.
- Shafer, S. L., Hendrickx, J. F., Flood, P., Sonner, J., and Eger, E. I. II (2008). Additivity versus synergy: a theoretical analysis of implications for anesthetic mechanisms. *Anesth. Analg.* 107, 507–524.
- Sinner, B., and Graf, B. M. (2008). Ketamine. *Handb. Exp. Pharmacol.* 182, 313–333.
- Steyn-Ross, A., and Steyn-Ross, M. (2010). *Modeling Phase Transitions in the Brain*. New York: Springer.
- Tsuda, N., Hayashi, K., Hagiwara, S., and Sawa, T. (2007). Ketamine, an NMDA-antagonist, increases the oscillatory frequencies of alpha-peaks on the electroencephalographic power spectrum. *Acta Anaesthesiol. Scand.* 51, 472–481.
- White, P. E., Schuttler, J., Shafer, A., Stanski, D. R., Horai, Y., and Trevor, A. J. (1985). Comparative pharmacology of the ketamine isomers. *Br. J. Anaesth.* 57, 197–203.
- Wilson, M. T., Sleight, J. W., Steyn-Ross, D. A., and Steyn-Ross, M. L. (2006). General anesthetic-induced seizures can be explained by a mean-field model of cortical dynamics. *Anesthesiology* 104, 588–593.
- Wolff, K., and Winstock, A. R. (2006). Ketamine: from medicine to misuse. *CNS Drugs* 20, 199–218.

Conflict of Interest Statement: David T. J. Liley is Chief Scientific Officer of Cortical Dynamics Ltd., an unlisted subsidiary of Biopharmica Ltd. (Perth, Australia), which is a medical device company focused on developing an EEG based depth of anesthesia monitor. David T. J. Liley is an inventor on several patent applications filed by Cortical Dynamics Ltd. since 2004 that describe new approaches to monitoring depth of anesthesia using the EEG. None of the IP declared in any published, pending, or granted patent has been licensed.

Received: 07 December 2012; accepted: 10 March 2013; published online: 05 April 2013.

Citation: Bojak I, Day HC and Liley DTJ (2013) Ketamine, propofol, and the EEG: a neural field analysis of HCN1-mediated interactions. *Front. Comput. Neurosci.* 7:22. doi: 10.3389/fncom.2013.00022

Copyright © 2013 Bojak, Day and Liley. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

Table A1 | The 10 selected parameter sets, whose PSDs are shown in Figure 2.

	I	II	III	IV	V	VI	VII	VIII	IX	X
h_e^r (mV)	-68.718	-70.286	-69.774	-78.169	-63.407	-60.745	-67.15	-64.128	-64.061	-60.588
h_i^r (mV)	-71.115	-78.148	-69.325	-79.978	-72.375	-70.488	-79.864	-79.399	-70.777	-69.941
τ_e (ms)	149.16	83.072	125.02	104.53	69.026	86.932	134.75	77.855	96.538	109.63
τ_i (ms)	125.78	122.72	116.58	112.75	118.65	50.200	66.766	137.77	43.662	76.350
h_{ee}^{eq} (mV)	1.8642	-16.433	3.2177	8.7034	-2.5551	1.9520	-18.571	-9.8354	8.4179	0.83573
h_{ei}^{eq} (mV)	-13.716	5.2227	8.2845	-14.231	-17.725	-15.828	-19.572	-4.4417	4.0220	2.4429
h_{ie}^{eq} (mV)	-86.369	-85.969	-86.775	-86.941	-85.466	-83.939	-86.449	-87.012	-82.833	-87.292
h_{ii}^{eq} (mV)	-80.439	-85.348	-77.200	-86.445	-83.047	-79.708	-87.791	-87.906	-78.523	-78.755
Γ_{ee} (mV)	0.22666	0.15856	0.17189	0.11073	0.25964	0.18606	0.31401	0.11187	0.10671	0.15192
Γ_{ei} (mV)	0.72933	1.8661	1.7385	1.8429	1.7030	0.91706	1.7073	1.2797	0.60619	1.7838
Γ_{ie} (mV)	1.9579	1.6800	1.5436	1.6612	1.8285	1.1699	0.53775	1.2751	0.31684	1.2942
Γ_{ii} (mV)	1.0898	0.68575	0.61488	0.53105	0.87001	0.38026	0.10299	0.59535	0.38033	1.2539
γ_{ee} (s ⁻¹)	768.09	979.20	626.91	494.19	399.13	848.11	964.40	795.80	689.20	355.38
γ_{ei} (s ⁻¹)	128.26	399.71	357.24	170.95	246.39	219.24	238.50	191.63	224.39	258.87
γ_{ie} (s ⁻¹)	192.29	178.41	135.36	411.10	251.74	82.043	468.53	221.95	133.21	203.10
γ_{ii} (s ⁻¹)	57.060	58.091	52.773	53.437	49.217	50.113	43.132	57.302	59.581	49.348
N_{ee}^{β}	3393.0	2552.5	2337.7	4557.9	3571.9	2945.1	2002.9	3941.9	2061.2	2718.0
N_{ei}^{β}	4520.7	4183.2	4168.8	4922.0	4290.5	2771.2	3297.8	4833.0	4357.1	4964.4
N_{ie}^{β}	270.31	674.75	566.64	934.52	927.91	520.26	703.66	838.55	835.88	607.11
N_{ii}^{β}	125.69	453.59	594.80	141.53	472.96	658.36	294.98	890.80	314.72	147.19
N_{ee}^{α}	4223.4	2234.3	4974.3	2517.6	2871.9	2230.2	2678.3	2874.7	4781.7	4128.1
N_{ei}^{α}	2892.3	1559.1	2837.9	2412.1	2952.9	1441.3	1693.6	2896.0	2095.3	2078.7
Λ (cm ⁻¹)	0.69280	0.27742	0.2529	0.76041	0.16148	0.22388	0.84812	0.56276	0.83448	0.51625
ν (cm ⁻¹)	116.05	137.60	483.12	158.20	156.22	283.76	483.97	504.36	790.03	325.52
S_e^{\max} (s ⁻¹)	311.08	201.57	474.21	126.59	88.686	103.54	280.25	422.80	190.77	246.41
S_i^{\max} (s ⁻¹)	249.73	280.48	287.93	171.23	227.44	238.74	473.82	294.99	485.67	411.64
$\bar{\mu}_e$ (mV)	-45.365	-49.195	-53.432	-54.004	-44.616	-46.851	-46.811	-43.850	-47.622	-51.391
$\bar{\mu}_i$ (mV)	-49.046	-45.093	-51.480	-44.165	-48.954	-47.996	-51.395	-50.619	-42.285	-44.831
$\hat{\sigma}_e$ (mV)	6.5908	6.7284	5.2051	6.8209	6.9030	5.9824	5.9045	6.6268	5.6959	5.0919
$\hat{\sigma}_i$ (mV)	4.3224	4.7270	4.4501	4.1542	5.4314	3.0605	5.4229	5.8536	5.6195	6.6326
$\bar{\rho}_{ee}$ (s ⁻¹)	7795.9	7344.6	7966.7	5876.2	4496.4	3882.5	6781.8	2649.9	9342.5	2833.6
ρ_{ei} (s ⁻¹)	329.39	2554.0	999.87	2120.3	2188.8	2337.9	1196.8	2063.5	914.37	1339.4



Neural mass modeling of power-line magnetic fields effects on brain activity

J. Modolo^{1,2,3*}, A. W. Thomas^{1,2,3} and A. Legros^{1,2,3,4}

¹ Human Threshold Research Group, Lawson Health Research Institute, London, ON, Canada

² Department of Medical Biophysics, Western University, London, ON, Canada

³ Department of Medical Imaging, Western University, London, ON, Canada

⁴ School of Kinesiology, Western University, London, ON, Canada

Edited by:

Peter B. Graben,
Humboldt-Universität zu Berlin,
Germany

Reviewed by:

Fabrice Wendling, INSERM, France
Flavio Frohlich, University of North
Carolina, USA

*Correspondence:

J. Modolo, Human Threshold
Research Group, Lawson Health
Research Institute, 268 Grosvenor
Street, London, ON N6A4V2,
Canada.
e-mail: jmodolo@lawsonimaging.ca

Neural mass models are an appropriate framework to study brain activity, combining a high degree of biological realism while being mathematically tractable. These models have been used, with a certain success, to simulate brain electric (electroencephalography, EEG) and metabolic (functional magnetic resonance imaging, fMRI) activity. However, concrete applications of neural mass models have remained limited to date. Motivated by experimental results obtained in humans, we propose in this paper a neural mass model designed to study the interaction between power-line magnetic fields (MFs) (60 Hz in North America) and brain activity. The model includes pyramidal cells; dendrite-projecting, slow GABAergic neurons; soma-projecting, fast GABAergic neurons; and glutamatergic interneurons. A simple phenomenological model of interaction between the induced electric field and neuron membranes is also considered, along with a model of post-synaptic calcium concentration and associated changes in synaptic weights. Simulated EEG signals are produced in a simple protocol, both in the absence and presence of a 60 Hz MF. These results are discussed based on results obtained previously in humans. Notably, results highlight that (1) EEG alpha (8–12 Hz) power can be modulated by weak membrane depolarizations induced by the exposure; (2) the level of input noise has a significant impact on EEG power modulation; and (3) the threshold value in MF flux density resulting in a significant effect on the EEG depends on the type of neuronal populations modulated by the MF exposure. Results obtained from the model shed new light on the effects of power-line MFs on brain activity, and will provide guidance in future human experiments. This may represent a valuable contribution to international regulation agencies setting guidelines on MF values to which the general public and workers can be exposed.

Keywords: neural mass models, power-line magnetic fields, electroencephalogram (EEG), synaptic plasticity, brain stimulation

INTRODUCTION

Since the pioneering work of Wilson and Cowan (1973) and Amari (1977), neural field models have been increasingly used, expanded and studied by a developing multidisciplinary community including: mathematicians, physicists, neuroscientists, medical imaging scientists *etcetera*. Neural field models provide a concise, yet insightful description of cortical activity. This theory has not only led to successful reproduction of numerous experimental results, but also to the prediction of a certain number of phenomena that have been observed *in vivo* (for a review of these phenomena, see Modolo et al., 2010). A popular simplification of neural field models consists in neglecting the role of space, consider neural populations present in cortical columns (such as pyramidal neurons) and to consider a connectivity matrix between the different neural masses considered (Wendling et al., 2002; Sotero and Trujillo-Barreto, 2008; Bojak et al., 2010). This approach has been used to build large-scale models of brain activity, including models simulating the electroencephalogram (EEG), or the blood oxygen level-dependent signal (BOLD) reflecting metabolic activity of brain tissue (Wendling et al., 2002;

Sotero and Trujillo-Barreto, 2008; Bojak et al., 2010). Indeed, since neural mass models are tractable and make the link between local variables and observables, these models appear as an excellent compromise between biological realism and computational complexity. Therefore, here we present an application of neural mass model to a specific question that many teams over the world have been tackling for a number of years using various neuroimaging modalities: how do power-line frequency (60 Hz in North America) magnetic fields (MFs) interact with human brain activity (interaction mechanisms), and how does it translate into observable outcomes (neuroimaging data such as EEG, motor/cognitive performance)?

The effects of extremely low-frequency (categorized as being <300 Hz) MF such as power-line MF on human neurophysiology have been studied for several decades. Despite an impressive amount of experimental data *in vitro*, *in vivo*, and in humans, a complete understanding of the interaction mechanisms and associated effects is still to be achieved. One complication in comparing outcomes from these studies is the wide range of MF flux densities, MF exposure setups, and exposure protocols

used (Crasson, 2003). First, significant work has been done has been done *in vivo* and *in vitro* in order to characterize the effects of electric fields in terms of membrane potential perturbation, excitability and neural network oscillations (Jefferys et al., 2003; Bikson et al., 2004; Deans et al., 2007; Fröhlich and McCormick, 2010; Reato et al., 2010). These studies have highlighted the role of neuronal morphology and orientation were critical in understanding the interaction with electric field, but also that membrane depolarization far below the firing threshold can influence the activity of neuronal networks. Second, among the reported effects in humans, let us mention the modulation of pain threshold (Ghione et al., 2005), effects on resting tremor (Legros and Beuter, 2005), modulations in functional brain activity as measured by BOLD (Legros et al., 2011), interference with learning in a short-term memory test (Corbacio et al., 2011). The most established interaction mechanism of MF exposure consists in induced currents, resulting from Faraday's law of induction, stating that a time-varying MF induces a time-varying electric field. This electric field will induce charge movement, creating a current. This is termed as the *induced current mechanism* (National Institute of Environmental Health Sciences of the National Institutes of Health, 1998). The induced current mechanism will be the mechanism considered in this paper. Interestingly, several studies have reported lasting effects associated with ELF MF exposure, i.e., an effect that is still detectable after cessation of the exposure. Using a specific pulsed MF, modulations of the EEG have been observed post-exposure (Cook et al., 2004, 2005). In the case of 60 Hz MF in the millitesla range (1.8 and 3 mT), modulations of the BOLD signal measured in humans during motor (finger tapping) and cognitive (mental rotation) tasks (Legros et al., 2010; Miller et al., 2010) have been found to be modulated post-exposure.

The interest for the interaction between power-line MF and human neurophysiology is twofold. First, there is a growing concern from the general public regarding the possible deleterious effects of power-line MF on human health, even if such negative effects remain to be demonstrated. Second, international regulation agencies such as ICNIRP (International Commission on Non-Ionizing Radiation Protection, <http://www.icnirp.de>) need results from the scientific literature to set their exposure guidelines (ICNIRP, 2010), aiming to protect the general public and workers, that are used by governments. One of the most long awaited data is the threshold in MF flux density at 60 Hz resulting in detectable effects in humans. One well-known effect of ELF MF exposure is magnetophosphenes, the perception of flickering lights in the visual field in the presence of a sufficiently strong MF. ICNIRP states that: "*Since the perception of magnetophosphenes constitutes the most reliable effect of MF exposure on human biology, this serves as a basis for the ICNIRP guidelines*" (ICNIRP, 2010). However, no data is available at 60 Hz regarding threshold values, therefore new approaches that could assist in the interpretation of existing experimental results, but also in the prediction of new results such as an estimation of threshold values that could be tested experimentally, would constitute significant advances in the field.

In order to shed light on the mechanisms involved in lasting effects of 60 Hz MF exposure on human neurophysiology, and also provide an estimate of the threshold value resulting in

detectable changes in EEG caused by 60 Hz MF exposure, we present in this paper a neural mass model aiming to model brain tissue dynamics at different time scales, bridging biophysical mechanisms with changes in observables. Among the panel of brain tissue dynamics models available, the approach initiated by Jansen and Rit (1995), consisting in considering sub-populations of neurons synaptically connected; later extended and improved by Wendling et al. (2002), provides a meaningful and accurate description of cortical dynamics. Indeed, such models have been successfully applied to understand the transition between baseline EEG and epileptic activity (Wendling et al., 2002, 2005; Molaee-Ardekani et al., 2010). First, we present an extension of the neural mass model developed by Sotero and Trujillo-Barreto (2008), by including a population of fast GABAergic neurons as suggested by Wendling et al. (2002). Second, we include a simple biophysical model of interaction between the 60 Hz MF and neuronal activity, along with a model of synaptic plasticity changes related to post-synaptic calcium concentration levels (Shouval et al., 2002a,b). Third, we use this model to investigate *in silico* the effects of 60 Hz MF on cortical dynamics, notably to evaluate the threshold in MF flux density resulting in detectable changes in variables of interest. The role of synaptic input noise, neuronal populations modulated by the exposure, and synaptic plasticity are also explored. Finally, we discuss future directions of research using this modeling approach.

MATERIALS AND METHODS

NEURAL MASS MODEL

In order to develop a biologically grounded model to study the effect of 60 Hz MF on neuronal activity, we have extend the thalamo-cortical model proposed by Sotero and Trujillo-Barreto (2008) by including a population of soma-projecting, fast inhibitory γ -amino-butyric acid (GABA) interneurons to extend the possible dynamical repertoire of the model [as shown by Wendling et al. (2002)]. The proposed modification of the block diagram proposed by Sotero and Trujillo-Barreto (2008) used to describe the thalamo-cortical model is the following:

Model equations are obtained by using the fact that the synaptic response function (Green's function) for a type of synapse i (e.g., glutamatergic) writes as $V_i(t) = A_i \cdot a_i \cdot t \cdot \exp(-a_i \cdot t)$, where A_i is the response amplitude and a_i the response time constant. Considering the temporal operator $\hat{L} = \frac{d^2}{dt^2} + 2a_i \frac{d}{dt} + a_i^2$, and using the fact that the synaptic response is a Greens' function for the temporal operator, we can use $\hat{L}V_i(t) = \delta(t)$ to write the following neural mass equation:

$$\frac{d^2}{dt^2} V_i(t) + 2 \cdot a_i \cdot \frac{d}{dt} V_i(t) + a_i^2 \cdot V_i(t) = A \cdot a \cdot v_i(t) \quad (1)$$

where $v_i(t)$ is the incoming firing rate. It is often practical to write Equation (1) under the form of a system of two first-order differential equations:

$$\begin{aligned} \frac{d}{dt} V_i(t) &= y_i(t) \\ \frac{d}{dt} y_i(t) &= A \cdot a \cdot v_i(t) - 2 \cdot a_i \cdot \frac{d}{dt} y_i(t) + a_i^2 \cdot V_i(t) \end{aligned} \quad (2)$$

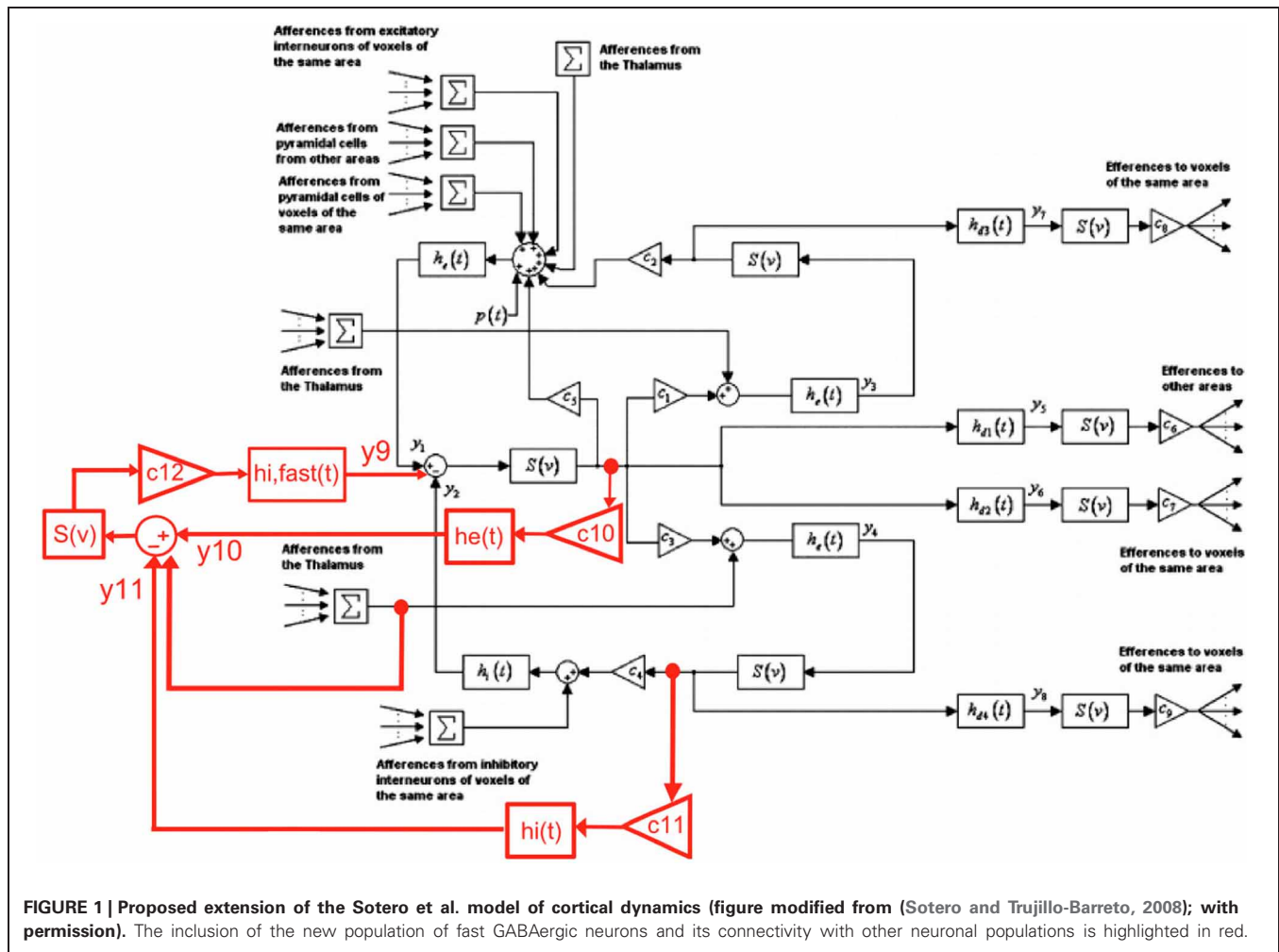


FIGURE 1 | Proposed extension of the Sotero et al. model of cortical dynamics (figure modified from (Sotero and Trujillo-Barreto, 2008); with permission). The inclusion of the new population of fast GABAergic neurons and its connectivity with other neuronal populations is highlighted in red.

Using this principle for each block of EPSP/IPSP presented in **Figure 1**, it is possible to formulate our extended thalamo-cortical model as a system of 22 differential equations (6 new equations corresponding to the new population of fast inhibitory interneurons, its feedback loop with pyramidal neurons, and its inhibitory input from slow inhibitory interneurons) presented below:

$$\begin{aligned}\dot{y}_1^{nj}(t) &= y_{12}^{nj}(t) \\ \dot{y}_2^{nj}(t) &= y_{13}^{nj}(t) \\ \dot{y}_3^{nj}(t) &= y_{14}^{nj}(t) \\ \dot{y}_4^{nj}(t) &= y_{15}^{nj}(t) \\ \dot{y}_5^{nj}(t) &= y_{16}^{nj}(t) \\ \dot{y}_6^{nj}(t) &= y_{17}^{nj}(t) \\ \dot{y}_7^{nj}(t) &= y_{18}^{nj}(t) \\ \dot{y}_8^{nj}(t) &= y_{19}^{nj}(t) \\ \dot{y}_9^{nj}(t) &= y_{20}^{nj}(t)\end{aligned}$$

(3)

$$\begin{aligned}\dot{y}_{10}^{nj}(t) &= y_{21}^{nj}(t) \\ \dot{y}_{11}^{nj}(t) &= y_{22}^{nj}(t) \\ \dot{y}_{12}^{nj}(t) &= A.a. \left\{ c_5.S[y_1^{nj}(t) - y_2^{nj}(t) - y_9^{nj}(t)] \right. \\ &\quad \left. + c_2.S[y_3^{nj}(t)] + K^{th,n}c_3.S[x_4(t)] \right\} \\ &\quad + A.a. \left\{ \sum_{\substack{m=1 \\ m \neq j}}^{M_n} \left(k_{e1}^{mj}.c_7.S[y_6^{mn}(t)] + k_{e2}^{mj}.c_8.S[y_7^{mn}(t)] \right) \right. \\ &\quad \left. + \sum_{\substack{i=1 \\ i \neq n}}^N K^{i,n} \sum_{m=1}^{M_i} c_6.S[y_5^{im}(t)] \right\} \\ &\quad - 2.a.y_{12}^{nj}(t) - a^2.y_{11}^{nj}(t) + A.a.p^{nj}(t)\end{aligned}$$

$$\begin{aligned}
\dot{y}_{13}^{nj}(t) &= B.b. \left\{ c_4.S[y_4^{nj}(t)] + \sum_{\substack{m=1 \\ m \neq j}}^{M_n} \left(k_i^{mj}.c_9.S[y_8^{mj}(t)] \right) \right\} \\
&\quad - 2.b.y_{13}^{nj}(t) - b^2 y_2^{nj}(t) \\
\dot{y}_{14}^{nj}(t) &= A.a. \left\{ c_1.S[y_1^{nj}(t) - y_2^{nj}(t) - y_9^{nj}(t)] + K^{th,n}.c_4^t S[x_5(t)] \right\} \\
&\quad - 2.a.y_{14}^{nj}(t) - a^2 y_3^{nj}(t) \\
\dot{y}_{15}^{nj}(t) &= A.a. \left\{ c_3.S[y_1^{nj}(t) - y_2^{nj}(t) - y_9^{nj}(t)] + K^{th,n}.c_5^t S[x_6(t)] \right\} \\
&\quad - 2.a.y_{15}^{nj}(t) - a^2 y_4^{nj}(t) \\
\dot{y}_{16}^{nj}(t) &= A.a_{d1}. \left\{ S[y_1^{nj}(t) - y_2^{nj}(t) - y_9^{nj}(t)] \right\} \\
&\quad - 2.a_{d1}.y_{16}^{nj}(t) - a_{d1}^2 y_5^{nj}(t) \\
\dot{y}_{17}^{nj}(t) &= A.a_{d2}. \left\{ S[y_1^{nj}(t) - y_2^{nj}(t) - y_9^{nj}(t)] \right\} \\
&\quad - 2.a_{d2}.y_{17}^{nj}(t) - a_{d2}^2 y_6^{nj}(t) \\
\dot{y}_{18}^{nj}(t) &= A.a_{d3}. \left\{ S[y_3^{nj}(t)] \right\} - 2.a_{d3}.y_{18}^{nj}(t) - a_{d3}^2 y_7^{nj}(t) \\
\dot{y}_{19}^{nj}(t) &= B.b_{d4}. \left\{ S[y_4^{nj}(t)] \right\} - 2.b_{d4}.y_{19}^{nj}(t) - b_{d4}^2 y_8^{nj}(t) \\
\dot{y}_{20}^{nj}(t) &= G.g. \left\{ c_{12}.S[y_{10}^{nj}(t) - y_{11}^{nj}(t)] + K^{th,n}.c_5^t S[x_6(t)] \right\} \\
&\quad - 2.g.y_{20}^{nj}(t) - g^2 y_9^{nj}(t) \\
\dot{y}_{21}^{nj}(t) &= A.a. \left\{ c_{10}.S[y_1^{nj}(t) - y_2^{nj}(t) - y_9^{nj}(t)] \right\} \\
&\quad - 2.a.y_{21}^{nj}(t) - a^2 y_{10}^{nj}(t) \\
\dot{y}_{22}^{nj}(t) &= B.b. \left\{ c_{11}.S[y_4^{nj}(t)] \right\} - 2.b.y_{22}^{nj}(t) - b^2 y_{11}^{nj}(t)
\end{aligned}$$

The equations describing the activity of the thalamus, composed of a population of thalamocortical cells and a population of reticular cells are [modified from Sotero and Trujillo-Barreto (2008)]:

$$\begin{aligned}
\dot{x}_1(t) &= x_7(t) \\
\dot{x}_2(t) &= x_8(t) \\
\dot{x}_3(t) &= x_9(t) \\
\dot{x}_4(t) &= x_{10}(t) \\
\dot{x}_5(t) &= x_{11}(t) \\
\dot{x}_6(t) &= x_{12}(t) \\
\dot{x}_7(t) &= A_t.a_t. \left\{ \sum_{i=1}^N K^{th,i} \sum_{m=1}^{M_i} c_6.S[y_5^{im}(t) + p_{th}(t)] \right\} \\
&\quad - 2.a_t.x_7(t) - a_t^2 x_1(t)
\end{aligned}$$

$$\begin{aligned}
\dot{x}_8(t) &= B_t.b_t.c_{2t}.S[c_{1t}x_3(t)] - 2.b_t.x_8(t) - b_t^2 x_2(t) \\
\dot{x}_9(t) &= A_t.a_t.S[x_1(t) - x_2(t)] - 2.a_t.x_9(t) - a_t^2 x_3(t) \\
\dot{x}_{10}(t) &= A_t.a_{d1t}.S[x_1(t) - x_2(t)] - 2.a_{d1t}.x_{10}(t) - a_{d1t}^2 x_4(t) \\
\dot{x}_{11}(t) &= A_t.a_{d2t}.S[x_1(t) - x_2(t)] - 2.a_{d2t}.x_{11}(t) - a_{d2t}^2 x_5(t) \\
\dot{x}_{12}(t) &= A_t.a_{d3t}.S[x_1(t) - x_2(t)] - 2.a_{d3t}.x_{12}(t) - a_{d3t}^2 x_6(t)
\end{aligned} \quad (4)$$

The physical meaning and values of model parameters are detailed in **Table A1** (Appendix section). For more details, the reader can refer to Sotero and Trujillo-Barreto (2008). Overall, the model is composed of 22 differential equations describing cortical dynamics of four different neuronal populations (pyramidal neurons, glutamatergic interneurons, fast/slow GABAergic neurons, and 12 differential equations defining thalamic activity). Therefore, this set of 34 differential equations describes the thalamocortical activity including time delays between cortical areas, connectivity parameters, synaptic responses derived from neurophysiology, and a biologically plausible (even if it is obviously simplified) circuitry between the neuronal populations considered. **Table A1** provided in Appendix summarizes the parameters used in the model, with new parameters added due to the population of fast inhibitory interneurons that have been highlighted.

MODEL OF INTERACTION BETWEEN 60 Hz EXPOSURE AND NEURON MEMBRANES

In order to model the interaction between the electric field induced by 60 Hz MF exposure and neural tissue, we have used the $\ll \lambda \cdot E \gg$ model in order to simulate the modulation of neuron membrane polarization (Molaei-Ardekani et al., 2013). In this model, the membrane depolarization dV in the presence of an electric field E is a function of a constant λ termed “polarization length” (Radman et al., 2009). More precisely, the membrane depolarization is expressed as $dV = \lambda \cdot E$, where λ is a vector oriented along the neuron fibre, and E is the electric field vector. This expression is valid for a static electric field. In the case of a time-varying electric field (such as the electric field induced by 60 Hz MF), a frequency-dependent term needs to be included (Gianni et al., 2006), resulting in:

$$dV = \frac{\lambda \cdot E}{\sqrt{1 + \omega^2 \tau^2}} \quad (5)$$

where $\omega = 2\pi f$, f being the frequency (in our case, $f = 60$ Hz), and τ is the polarization time constant. In order to use Equation (5), describing membrane depolarization induced by the induced electric field, at the level of a neuronal population, we made the following assumptions: (1) the induced 60 Hz electric field is homogeneous in space at the level of the neural mass (i.e., the MF flux density is constant at each point of the cortical column); (2) the MF-induced membrane depolarization is applied to pyramidal neurons only because of their large size compared to other types of neurons in the human cortex; (3) pyramidal neurons in a given neural mass all have the same spatial orientation. Taken

together, assumptions (1) and (3) result in identical dV values for all pyramidal neurons in a neural mass at a given time. These assumptions lead to the use Equation (3) in the context of a neural mass model.

This was achieved by modifying the expression of the total post-synaptic potential at the level of pyramidal neurons:

$$y_1(t) - y_2(t) - y_9(t) \rightarrow y_1(t) - y_2(t) - y_9(t) + dV(t) \quad (6)$$

As it is commonly calculated in neural mass model, the EEG signal was computed as the summation of excitatory and inhibitory post-synaptic potential at the level of pyramidal neurons:

$$EEG(t) = y_1(t) - y_2(t) - y_9(t) + dV(t) \quad (7)$$

Let us mention that, even if the MF-induced depolarization is included in the model as a simple additive perturbation, it has the potential to induce non-linear effects. Indeed, the effective potential at the level of pyramidal neurons (7) is used as an input for other neuronal populations, and is transformed from a potential to a firing rate using a sigmoid function, which is fundamentally non-linear. In the “Results” section, we have used arbitrary values for the field-induced membrane depolarization dV , guided by preliminary simulation results. Based on dV values resulting in significant changes in the EEG alpha power with or without synaptic plasticity in the model, we will provide an estimate of the corresponding level of 60 Hz MF flux density. This will provide us with an order of magnitude of the 60 Hz MF flux density threshold value that should result in effects detectable experimentally in humans.

In our simulations, we focused specifically on the EEG alpha rhythm (8–12 Hz). The reason of this choice is twofold. First, as mentioned in the Introduction, there is converging evidence that extremely low-frequency MF in the millitesla range, such as 60 Hz MF, can induce EEG alpha activity modulation. Second, the model we have developed is basically an extension of the Jansen and Rit model, designed to model EEG alpha activity. It is possible to reproduce other types of EEG rhythms (e.g., beta –13 to 30 Hz), for example by introducing heterogeneity in the time constant of neural populations over different neural masses (Wendling et al., 2002), which exceeds the scope of this paper.

BIOPHYSICAL MODEL OF SYNAPTIC PLASTICITY

In order to investigate the hypothesis that 60 Hz MF exposure might modulate with human neurophysiology by modulating synaptic plasticity, we have implemented a simplified model of synaptic plasticity, based on the biophysical model developed by Shouval et al. (2002a,b). It is now well accepted that the mechanisms of long-term synaptic potentiation and depression (LTP/LTD, respectively) involve changes in post-synaptic calcium concentration and the trafficking of α -amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid (AMPA) glutamate receptors between the intracellular medium and the synapse site. Depending on the calcium concentration, AMPA receptors can either insert into the membrane at the level of the synaptic cleft, or undergo an endocytosis, which is termed *receptor trafficking*

(Collingridge et al., 2004). An increase in the number of AMPA receptors at the synaptic level will increase the number of glutamate molecules that can bind on post-synaptic membranes, thereby increasing membrane depolarization during a synaptic event. Consequently, the number of post-synaptic AMPA receptors is directly proportional to the synaptic weight. The model proposed by Shouval et al. (2002a,b) has been a significant progress in the modeling of the biophysical processes at play during LTP/LTD. This model is based on the “calcium control hypothesis,” according to which the level of post-synaptic calcium is the main factor regulating the exocytosis/endocytosis rate of AMPA receptors, and therefore the dynamics of synaptic plasticity changes.

We have adapted the model by Shouval et al. (2002a,b) to our neural mass model, and despite some simplifications with respects to the original model; our synaptic plasticity model captures some of its essential features. Based on the experimental literature on 60 Hz MF exposure effects on the EEG alpha rhythm, we assume that (1) no qualitative changes of EEG dynamics will occur due to 60 Hz MF exposure, changes will be purely quantitative (i.e., EEG alpha rhythm amplitude/spectral power changes, but no qualitative change in dynamical regime such as a transition toward high-amplitude, low-frequency spiking); (2) the coupling between the synaptic plasticity model and the 60 Hz MF is *via* the equation linking the EEG with the post-synaptic calcium concentration, occurring on long timescales (depending on the opening of N-methyl-D-aspartate (NMDA) glutamate receptors, not represented in the model). Therefore, the model offers the possibility to test the hypothesis that 60 Hz MF exposure can modulate synaptic plasticity by interfering with the calcium fluxes at the level of synapses. However, it does not take into account possible effects of 60 Hz MF exposure on spike timing (see the “Discussion” section).

Let us consider the average calcium post-synaptic concentration in a neural mass. The model proposed by Shouval et al. (2002a,b) links the calcium current at the level of NMDA receptors with the calcium concentration, and finally to a differential equation describing the dynamics of synaptic weight change as a function of two different calcium-dependent functions. The time constant of calcium concentration dynamics is long (on the order of minutes), and the calcium concentration increases with the membrane potential. Therefore, it appears reasonable to approximate calcium dynamics by a low-pass filtering of the mean potential of a given neural mass:

$$\tau_{Ca^{2+}} \frac{d}{dt} [Ca^{2+}] + [Ca^{2+}] = \gamma(y_1 - y_2 - y_9) \quad (8)$$

where $y_1 - y_2 - y_9$ is the “EEG” signal at the level of a neural mass (e.g., summation of post-synaptic potentials at the level of pyramidal neurons as defined previously). Once the post-synaptic dendritic calcium concentration is obtained, it is possible to evaluate the calcium-dependent functions η and Ω present in the Shouval et al. (2002a,b) model, used to express the dynamics of the synaptic weight c_i (i denoting the type of synapse in the neural

mass, e.g., afferent glutamatergic synapses on pyramidal neurons) at the level of a given neural mass:

$$\frac{dc_i(t)}{dt} = \eta(t) \cdot [\Omega_{Ca^{2+}}(t) - c_i(t)] \quad (9)$$

The function Ω was approximated by a combination of piecewise-linear and quadratic functions (see Appendix, **Figure A1** for details) similar to the function proposed in Shouval et al. (2002b). The function Ω used in our model differs quantitatively from the one proposed in Shouval et al. (2002a,b), since the authors were linking with this function the level of post-synaptic calcium concentration with the relative change in synaptic weight, where we directly link the post-synaptic calcium concentration with the synaptic weight itself. Nevertheless, the Ω used in this paper captures the most important qualitative properties proposed by Shouval et al. (2002a,b). In our simulations, we have assumed that was $\eta(t)$ a constant, $[Ca^{2+}(t)]$ being bounded between 0 and $1 \mu\text{M}$. We assumed that the synapses modulated were the synapses terminating on pyramidal neurons, pooled in the constant c_5 , becoming the variable $c_5(t)$ in our model. Numerical implementation for the neural mass model was performed using Matlab 2010 (The Mathworks, USA) on a quad-core Apple iMac (2.66 GHz/CPU) with 8 GB of RAM. The simulation of a neural mass using the complete model during 2 h with a time step of $dt = 1 \text{ ms}$ took typically 8 min.

60 Hz MF EXPOSURE PROTOCOL

In order to study the effects of 60 Hz MF on the simulated EEG, we used the following protocol: the neural mass was simulated during 2 h overall with a 1 ms resolution, which was decomposed

as (1) 30 min without 60 Hz MF exposure (termed “sham,” of sufficient duration to reach a steady state); (2) 60 min with 60 Hz MF exposure (sufficient to reach the new steady state); and (3) 30 min without 60 Hz MF exposure. Previous research in our team using fMRI to image the functional changes in brain activity due to 60 Hz MF exposure involved comparable durations (notably, a 60 min exposure period and fMRI acquisitions performed before and after, see Legros et al., 2010). Simulations were performed both (1) using the synaptic plasticity model, and (2) using a fixed synaptic weight value taken as the steady state value when synaptic plasticity was taken into account. By doing so, we aimed at decomposing the respective contribution of the 60 Hz sinusoidal perturbation in membrane potential one the one hand, and of possible calcium-related synaptic plasticity modulations on the other hand.

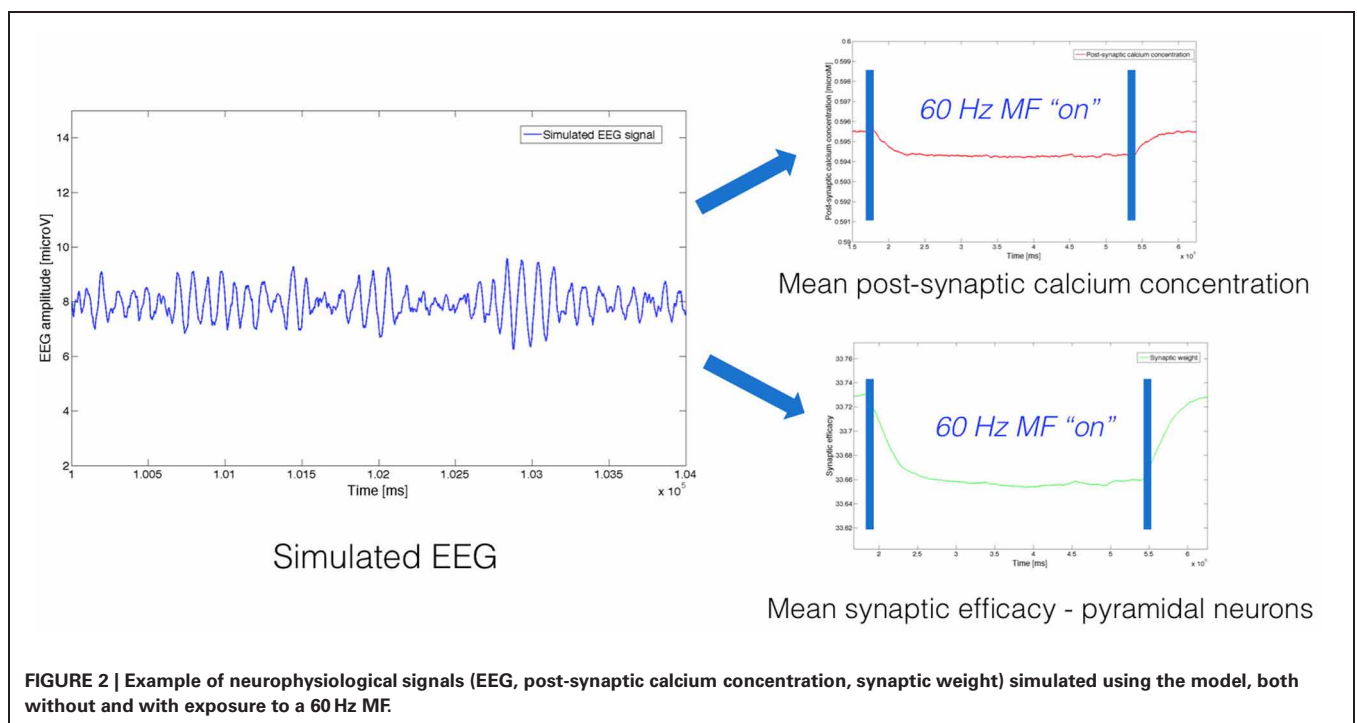
RESULTS

EFFECT OF 60 Hz MF ON THE EEG IN THE MODEL

As an example, we present in **Figure 2** an example of simulated EEG data, and associated mean post-synaptic calcium concentration and synaptic weight obtained by solving Equations (3, 4, 8, 9).

We have investigated the effect of increasing values (125, 250, 500, and $1000 \mu\text{V}$) for the MF-induced membrane depolarization on the EEG alpha power. EEG alpha spectral power was computed before, during and after the 1-h 60 Hz MF exposure. 10 runs of 7200 s were performed for each tested value of dV . The averaged EEG alpha power for each condition (before, during, and after exposure) is presented in **Figure 3**.

From the results presented in **Figure 3**, it appears that increasing values of dV gradually decreases EEG alpha power during



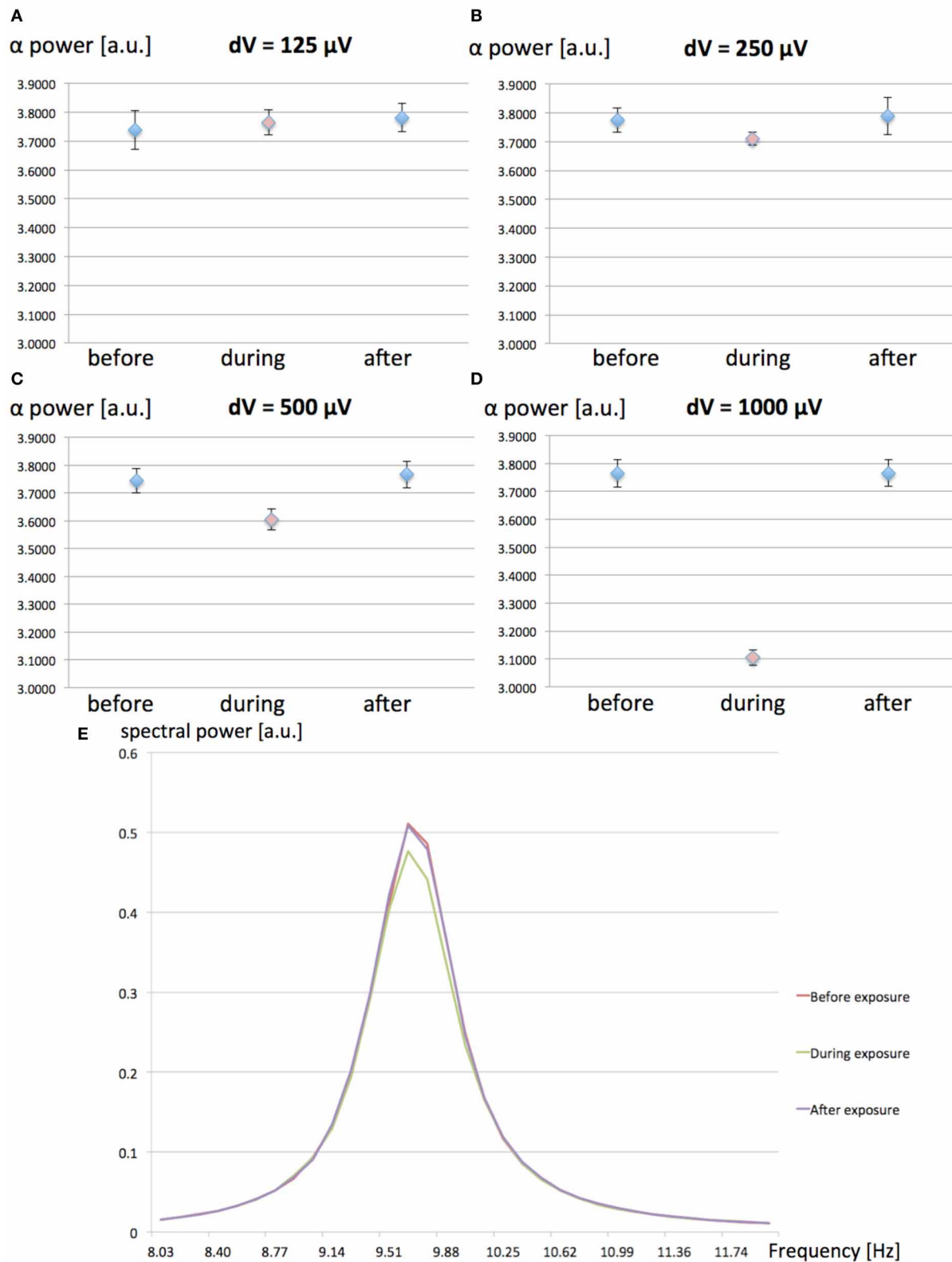


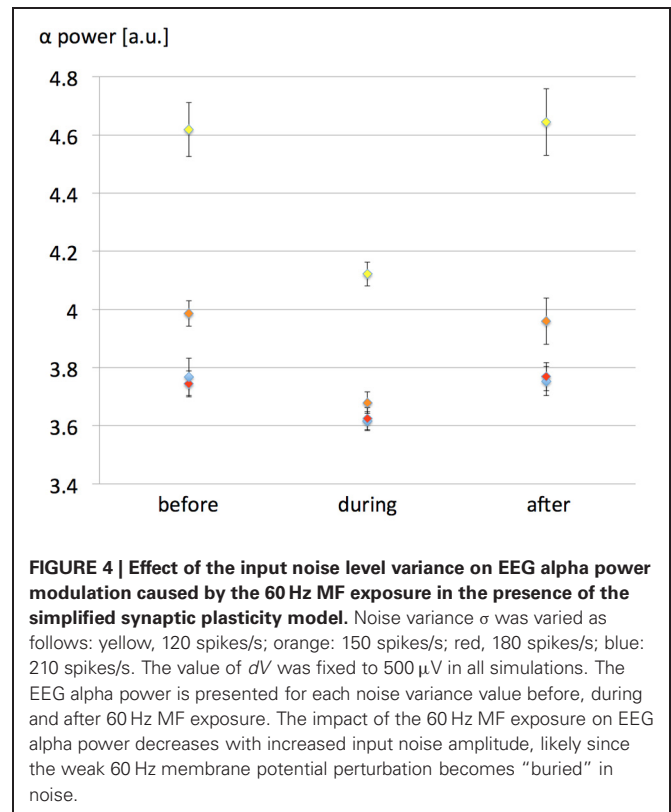
FIGURE 3 | (A–D): Spectral power in the EEG alpha (8–12 Hz) band as a function of the MF-induced membrane polarization dV ; before (blue), during (red) and after (blue) the 1-h 60 Hz MF exposure period. **(A)** $dV = 125 \mu V$; **(B)**, $dV = 250 \mu V$; **(C)** $dV = 500 \mu V$; **(D)** $dV = 1000 \mu V$.

A decrease in EEG alpha power is observed as the value of dV (proportional to the MF flux density) increases. **(E)** Example of average power spectrum before, during and after exposure to the 60 Hz MF, for $dV = 500 \mu V$.

exposure. In order to test the significance of the amplitude of dV on the EEG alpha power during 60 Hz MF exposure, we conducted a statistical analysis of the results. We performed a $4 \times 3 \times 2$ ANOVA for repeated measures (SPSS 21, IBM, USA), respectively testing for the effects of “ dV ” (125, 250, 500, and 1000 μV), “time” (before, during and after), and “plasticity” (with/without synaptic plasticity). The standard p -value of 0.05 (Greenhouse-Geisser) was chosen as the threshold for significance, and p -values were corrected for multiple comparisons. The statistical results reveal a significant decrease of the EEG alpha power for $dV = 500 \mu\text{V}$ as compared to the other values of dV ($p < 0.001$). This indicates that the threshold for a significant decrease of EEG alpha power due to 60 Hz MF exposure lies between induced membrane depolarization values of 250–500 μV . In the next section, we attempt to link the membrane depolarization values to the corresponding MF flux density at 60 Hz.

Due to the possibility that the weak 60 Hz membrane depolarization can be seen as an additive noise, we have tested the influence of the input noise level [$p(t)$ in the model, see Table A1 of the Appendix] variance on EEG alpha power modulation due to the exposure. Since the dynamics of the model itself depend critically on the input noise level, we have indeed investigated the possibility that the 60 Hz MF exposure has an effect of variable amplitude depending on input noise. The interest is that model predictions could be tested experimentally (e.g., in an experimental setting where different levels of visual input would be tested). Therefore, the objective was not to study the influence of the noise level on the model dynamics, but rather how the effects of the 60 Hz MF on model dynamics are dependent on the noise level. Four different values of noise variance were tested ($\sigma = 120, 150, 180$, and 210 spikes/s) for the same maximal dV value of 0.5 mV. 10 runs of 7200 s following the same protocol than previously were run for each noise level value (40 runs total). The influence of the input noise level on EEG alpha power modulation by the 60 Hz MF is presented in Figure 4.

The results presented in Figure 4 highlight the importance of the input noise level of the model. If the input $p(t)$, representing external noisy input to the neural mass, is too high; then the effect induced by the 60 Hz MF on EEG alpha power modulation decreases. This has an immediate consequence on threshold values of MF flux density resulting in detectable effects in brain activity: the MF flux density needed to elicit a response in brain tissue will be lower in the presence of a low level of noise. Interestingly, there is experimental evidence that the visual input can play a role on the effects of MF exposure in humans, with an higher effect when the eyes are closed (Legros et al., 2011). EEG alpha oscillations increase dramatically eyes closed, and decrease in the presence of a visual input, that increases the input noise to the occipital cortex. Therefore, even if there is a considerable gap between the model and human data, it is tempting to make a parallel between smaller effects of 60 Hz MF in the model with high levels of noise, and smaller effects of 60 Hz MF exposure eyes open with an increased input noise level. One advantage of using our neural mass over interpreting experimental results is the possibility to point at precise mechanisms by which the observed decrease in EEG alpha activity occurs due to the 60 Hz MF exposure. From



a physiological point of view, it is relevant to investigate which neuronal pathways are mainly modulated by the exposure. In the model, we observe an immediate decrease in the activity of the loop between pyramidal neurons and slow GABAergic neurons, likely increasing the effect of excitatory input. To complement this observation, it is relevant to note that, using a bifurcation theory analysis of the Jansen and Rit model (the core of our model), Grimbert and Faugeras (2006) have shown that increasing the input noise level at the level of pyramidal neurons in the alpha oscillations regime (corresponding to a Hopf bifurcation) had the effect to decrease the amplitude of alpha oscillations. Therefore, the 60 Hz MF stimulus used in our model seems to have a similar effect than an additive, positive constant membrane depolarization on pyramidal neurons. This results physiologically speaking from an efficiency decrease of the slow inhibitory GABAergic feedback at the pyramidal neurons level. In terms of dynamical systems theory, this seems to be the natural result of increased input level in a specific dynamical system on a Hopf cycle.

In order to distinguish between the contribution of the MF-induced membrane polarization on the one hand, and changes in synaptic plasticity on the other hand, we ran the same simulations than previously for four different values of dV (125, 250, 500, and 1000 μV), with a constant value for the synaptic weight c_5 . The objective was to identify if synaptic plasticity was affecting the direction (increase/decrease of EEG alpha spectral power) or amplitude of the effects. In the following, the value c_5 of was chosen as the steady-state value in the case where synaptic plasticity was considered. 10 simulations of 7200 s were ran for each value of dV . The results are presented in Figure 5.

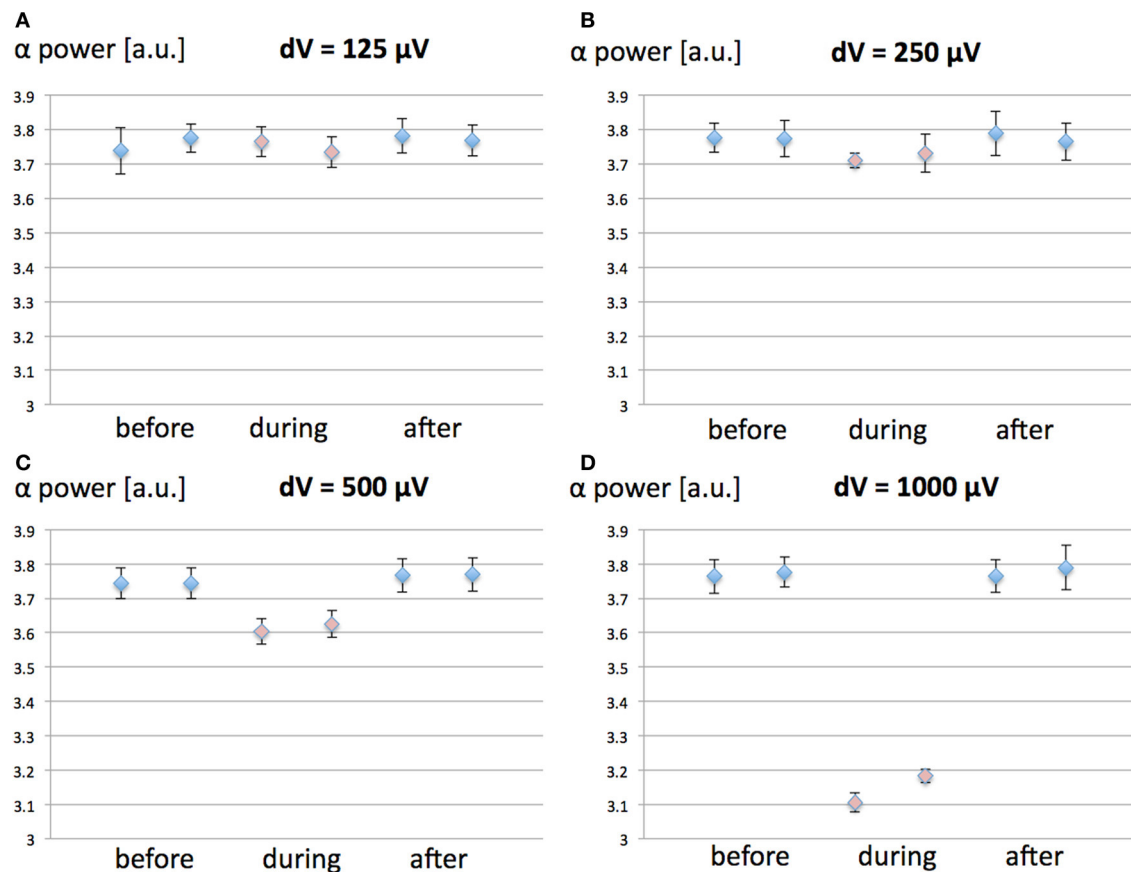
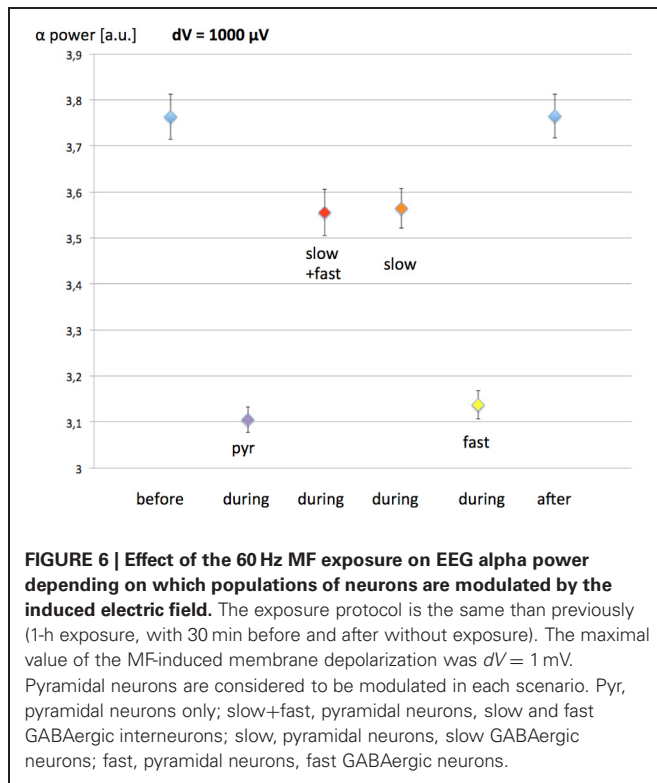


FIGURE 5 | Effect of synaptic plasticity on EEG alpha power modulation by 60 Hz MF exposure compared to the case where synaptic plasticity is not taken into account. The conditions are

Before, During, and After 60 Hz MF exposure, with (1, 3, 5) and without (2, 4, 6) synaptic plasticity. **(A)** $dV = 125 \mu V$; **(B)** $dV = 250 \mu V$; **(C)** $dV = 500 \mu V$; **(D)** $dV = 1000 \mu V$.

From the results in **Figure 5**, it appears that the modulation of post-synaptic calcium concentration and corresponding changes in synaptic weight plays a minimal role in EEG alpha power modulation due to the 60 Hz MF exposure, and does not impact qualitatively the result (the direction of the effects is the same, and the amplitude of the effects is minimally affected). Indeed, the results from the ANOVA shows no significant interaction effect between synaptic plasticity and dV values ($p = 0.253$). This indicates that the presence of the synaptic plasticity mechanisms included in the model does not significantly change the effect of the membrane depolarization. It seems however to induce a non-significant increase the amplitude of the 60 Hz MF exposure effect. Different choices for the function linking the post-synaptic calcium concentration level with the updated synaptic weight Ω lead to similar results (not shown). Therefore, it seems that, if a modulation of synaptic plasticity explains lasting effects of 60 Hz MF exposure, it does not occur primarily by the modulation of post-synaptic calcium currents. However, it is still plausible that receptor trafficking and synaptic plasticity could be impacted by a perturbation of spike timing due to the 60 Hz MF exposure, a mechanism not included in the present model, which we discuss later.

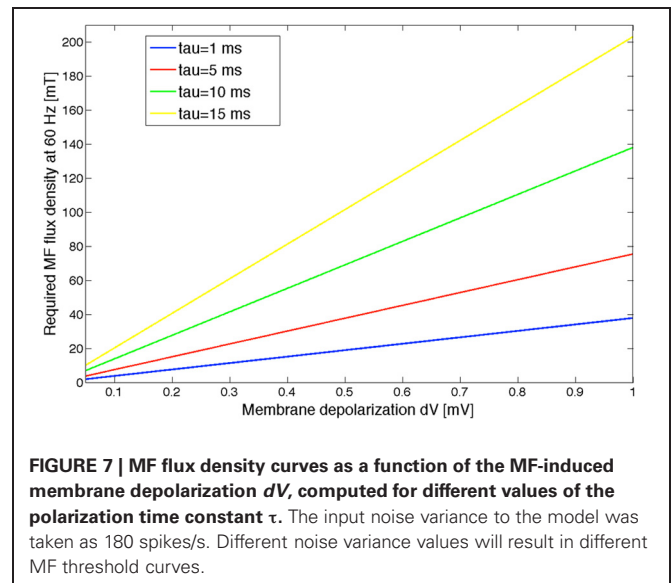
In the model develop by Molaee-Ardekani et al. (2013), investigating the effects of transcranial direct current stimulation (tDCS), the neurons being modulated by the induced field were pyramidal neurons and inhibitory interneurons. Since the simulated EEG could result in different outcomes due to the exposure depending on the neuronal populations simulated, possibly assisting in discriminating between different interaction mechanisms; we have also simulated the EEG in the case of a 60 Hz MF exposure modulating the activity of different populations of inhibitory interneurons. To do so, in a similar fashion to Equation (7) describing the membrane depolarization of pyramidal neurons, we simulated different scenarios: (1) slow and fast GABAergic interneurons are involved, (2) slow GABAergic interneurons are involved, and (3) fast inhibitory neurons are involved. Depending on the scenario, we also added the variable $dV(t)$ to $y_4(t)$ (slow GABAergic interneurons), or to $y_{10}(t)$ - $y_{11}(t)$ (fast GABAergic interneurons), and to both of these quantities for scenario (1). We have simulated the EEG for a similar protocol than previously (30 min without exposure, 1 h of exposure, 30 min without exposure), with a maximal value of $dV = 1000 \mu V$. The resulting EEG alpha power before, during and after exposure for each of these scenarios is presented in **Figure 6**.



The results presented in **Figure 6** show that, if the fast GABAergic interneurons are modulated by the 60 Hz MF in addition to pyramidal neurons, the difference in the EEG is minimal. However, if the slow GABAergic interneurons are modulated, the decrease in EEG alpha power is much smaller, dropping from 17% (pyramidal neurons only) to 5%. Therefore, if the slow GABAergic interneurons are also modulated by the induced electric field due to the 60 Hz MF exposure, the threshold leading to a systematic EEG alpha power modulation will be higher. From the interpretation on the decrease in EEG alpha power when the pyramidal neurons alone are stimulated, where a decrease in the efficiency of the loop between pyramidal neurons and slow inhibitory interneurons is observed, we can speculate that the effect of modulating the slow GABAergic interneurons in addition to pyramidal neurons has an opposite effect of increasing the activity of this loop. Consequently, if slow GABAergic interneurons are modulated by the exposure to the 60 Hz MF exposure, the model suggests that it would result in a compensation of the modulation of pyramidal neurons' activity alone, thereby increasing the threshold in MF flux density leading to a systematic decrease in alpha. In other words, the modulation of pyramidal neurons and slow GABAergic interneurons activity would have competing effects regarding the decrease in EEG alpha power.

THRESHOLD OF 60 Hz MF FLUX DENSITY RESULTING IN DETECTABLE EFFECTS

Based on the expression of the electric field induced by a time varying MF at the level of a sphere of radius R (approximating the brain in that scenario), we can obtain an estimate of



the corresponding MF flux density at 60 Hz. Let us approximate the head as a sphere of radius R , and let us write the 60 Hz MF as $B(t) = B_0 \sin(\omega t + \phi)$. From Maxwell-Faraday's law of magnetic induction, the induced electric field expresses as $E = \frac{R}{2} \frac{dB}{dt} = \pi R f B_0$. By using this expression in Equation (5), we obtain $B = \frac{dV(1 + \omega^2 \tau^2)^{1/2}}{\lambda \pi R f}$, linking the MF-induced membrane depolarization to the MF flux density. We used the following values to estimate the threshold values: $R = 0.15$ m, $\lambda = 10^{-3}$ m, and $f = 60$ Hz. The MF flux density value as a function of the MF-induced membrane depolarization dV , depending on different τ values, is shown in **Figure 7**.

We have shown based on our statistical analysis that a significant EEG alpha power modulation occurs in the model for a dV value between 250 and 500 μ V. Given the uncertainty on the polarization time constant, it is only possible to provide an estimate of the MF flux density threshold at 60 Hz which should result in a significant decrease of EEG alpha activity. Assuming an intermediate dV value as a threshold value (375 μ V), the corresponding threshold MF flux density would range between 15 and 75 mT, for polarization time constants between 1 and 15 ms. Let us also mention that this threshold value is, as shown above, depending on the input noise level. Uncertainties on the values of the polarization time constant and on the input noise level are problematic to estimate a more precise MF flux density threshold value. Depending on the neural elements activated by the induced electric field, this polarization time constant can be very different (higher for the whole soma than for fibers for example). For a neuron soma, the polarization time constant would be of several milliseconds; whereas if the membrane depolarization occurs at the level of Ranvier nodes, then the polarization time constant is considerably smaller, around 20 μ s (Gianni et al., 2006). Therefore, it is reasonable to assume that the polarization time constant is in the low millisecond range. As an example, a polarization time constant between 1 and 5 ms would result in a threshold value between 15 and 25 mT for a 60 Hz MF. The validation of these values will require experimental recordings

performed in humans, which will be performed in the near future in our group (Legros et al., 2012a,b).

DISCUSSION

In this paper, we have developed an innovative application of neural mass models, i.e., the study of how extremely low-frequency MF such as power-line MF interact with brain activity. Indeed, this is the first time that this problem is tackled using neural mass modeling. We have shown that 60 Hz MF exposure can result in a modulation of the EEG alpha rhythm, even for small membrane depolarization values (<1 mV). For reasonable polarization time constant values, the model predicts that 60 Hz MF between 15 and 25 mT could induce a systematic decrease in EEG alpha power. Furthermore, the neural mass model that we have developed includes, in a simplified manner, a contribution of synaptic plasticity processes. To our knowledge, this is the first attempt to include a contribution of calcium-related processes into changes of effective connectivity in neural mass models. The contribution of calcium currents on synaptic weights changes is obviously overly simplified in our model since the calcium concentration is modeled as a low-pass filtered version of the EEG. Nevertheless, it represents a first step that could serve as a basis in future models integrating more biophysically detailed models of synaptic plasticity. Let us mention that Robinson (2011) proposed a neural field model including synaptic plasticity, though in a different manner. Indeed, this model used the relative phase between pre- and post-synaptic neural populations to compute the synaptic weight changes due to an STDP rule. In the present paper, we have intended to provide the bases for mechanism-based neural mass models, on the grounds of a reliable synaptic plasticity model (Shouval et al., 2002a,b). Expanding our model using modeling principles of Robinson (2011) could be a solution so include the effect of spike timing perturbation induced by 60 Hz MF on neural mass activity.

One hypothesis investigated was that changes in post-synaptic calcium concentration could modulate synaptic weights, resulting in a lasting modulation of brain tissue dynamics. Obviously, our model of synaptic plasticity is still simplified and does not explicitly model the voltage-dependence of calcium currents through NMDA receptors. From our results, it appears that, despite a modulation in post-synaptic calcium concentration taking some time to build up, and lasting several minutes after the exposure, these changes are too small to impact neural mass dynamics. However, there is another important mechanism by which receptor trafficking and synaptic plasticity could be modulated by 60 Hz MF exposure. There is indeed a convergence of theoretical (Reato et al., 2010; Stodilka et al., 2011) and experimental (Radman et al., 2007) studies that illustrate the possibility for weak membrane depolarizations induced by electric fields to impact spike timing. Indeed, Radman et al. (2007) have shown that, due to the non-linear properties of neuron membranes, small membrane depolarizations can modulate spike timing. Since post-synaptic calcium currents play the role of “coincidence detector” between pre- and post-synaptic spikes, a perturbation of spike timing could impact receptor trafficking and synaptic weight changes. The challenge is to consider these mechanisms in

neural mass or neural field models, which are rate-coding based and not time-coding based. A recent study proposed how to consider plasticity rules based on spike timing in neural field theory (Robinson, 2011), providing a possibility to investigate the impact of 60 Hz MF perturbation on spike timing at a mesoscopic scale. The perturbation of spike timing by 60 Hz MF exposure will be considered in a future extension of the model presented in this paper, since this synaptic plasticity pathway could be more prone to small membrane perturbations due to the 60 Hz MF exposure, and could induce lasting effects in neural dynamics.

Using our model, we have also investigated the effect of the small 60 Hz membrane depolarization induced by 60 Hz MF exposure depending of the input noise level of the model. Our suggests that the threshold value in MF flux density for which significant changes can be detected in the EEG alpha frequency band is a function of the model input noise level. More specifically, for lower values of input noise level, the decrease in EEG alpha power is higher than for higher input noise level values. However, these results do not imply that stochastic resonance effects are present, which could be tested however by testing many different values of dV , and identify a range of dV values for which the modulation of EEG alpha power would be present. The stochastic resonance mechanism has been already explored in the literature to explain the effects of 60 Hz MF exposure, and constitutes a possibility of future study using our model. This result on the importance of the input noise level has also important implications for the detection of EEG alpha power modulation in humans due to 60 Hz MF exposure. We predict that the threshold in MF flux density at 60 Hz, required to modulate systematically the EEG in the occipital cortex, is lower a condition when the ambient light is low, compared to the effects of the same exposure using a high ambient light level. Indeed, EEG alpha oscillations decrease in the occipital cortex in humans due to the higher input noise level. Therefore, it might be relevant to study a variation of 60 Hz MF threshold values in humans at the level of the occipital cortex using different intensities of ambient light. If such experimental evidence was provided, that would be a precious piece of information that could be of interest to agencies such as ICNIRP. Another indication that the prediction that the level of input is important in the physiological outcome is that the perception of magnetophosphenes and the threshold at which they can be observed is light intensity-dependent (Lövsund et al., 1980). This motivates a future human experimental study using in parallel our neural mass model in order to provide an improved knowledge on the underlying interaction mechanisms. Another valuable insight from the model is the differential effect observed on EEG alpha activity depending on the neuronal populations modulated (pyramidal neurons only, or pyramidal neurons and slow/fast inhibitory interneurons). Since the model predicts a different outcome in the case where slow inhibitory interneurons are also modulated (smaller decrease in alpha activity), this offers a possibility to discriminate in future EEG data acquired in humans which neuronal populations are modulated by the MF exposure. This adds further support for the use of neural mass models to study the effects of power-line MF on human brain activity, since they can offer a deeper insight

into the experimental data in order to clarify the interaction mechanisms involved.

Among the limitations of our approach, let us mention first the absence of ephaptic interactions. It has been indeed demonstrated that post-synaptic potentials can induce in neighboring cells a small but measurable polarization (for a review, see Weiss and Faber, 2010). In the cortex, where the axons of pyramidal axons have a similar and consistent orientation, it is likely that ephaptic interactions could enhance the effect of weak membrane depolarizations. Therefore, the presence of ephaptic interactions should lower the threshold for detectable modulations of neuronal activity, and should be included in future biophysical models studying the effects of low-frequency MF on cortical activity. Second, our model of synaptic plasticity is a significant simplification compared to the biophysically detailed model by Shouval et al. (2002a,b). The present model could be extended by including a more detailed model of the detailed processes underlying receptor trafficking at the synaptic level. Third, the exact orientation of pyramidal axons with respect to the induced electric field was not taken into account, since it was assumed that the orientation was “ideal” (electric field parallel to pyramidal neuron axons). Fourth, we assumed that the power-line MF interacts with brain tissue *via* the induced electric field. However, there is evidence that the MF itself could interact with cellular signaling, and induce biological effects (Pilla, 2012). Studying such phenomena appears however out of reach with our proposed model. Let us note that the possibility that either the MF or the induced electric field modulate neuronal activity is not exclusive, and both mechanisms might even turn out to be complementary and have effects on different cellular components.

Finally, let us mention that, in most studies investigating the effects of low-frequency MF on the human EEG, the data analyzed is from before or after exposure, not during. There is indeed an experimental difficulty in recording the EEG during exposure to low-frequency MF. However, it is possible to compensate using specific signal processing techniques, such as wavelet-based methods (Modolo et al., 2011). We believe that such signal processing techniques applied to EEG acquired during 60 Hz MF

exposure, combined with the neural mass model proposed in this paper, could provide an integrated framework for a thorough understanding of power-line frequency MF on human brain activity.

CONCLUDING REMARKS

We have presented a novel application of neural field models, in the context of brain exposure to 60 Hz MF. The model takes into account different neural populations in cortical tissue, synaptic kinetics proper to each type of synapse considered, synaptic connectivity patterns inspired from neuroanatomy, and a simplified model of synaptic plasticity based on the “calcium-control” hypothesis. The model includes the interaction with the electric field induced by 60 Hz MF exposure, and results in a time-varying membrane depolarization. Using this model, we have shown that membrane depolarization between 250 and 500 μ V at 60 Hz is sufficient to induce a significant decrease in EEG alpha power. We also conclude that the modulation of post-synaptic calcium currents by 60 Hz MF exposure does not appear to predict the lasting effects observed experimentally. Future work should investigate the role of spike timing perturbation by the induced electric field during 60 Hz MF exposure, which is another candidate mechanism to induce plastic changes and lasting changes in neuronal activity. The models provides predictions that can, and will be, tested in experimental protocols during which humans will be exposed to increasing levels of 60 Hz MF exposure up to 50 mT (Legros et al., 2012a,b). Thorough comparison of experimental data with model predictions will constitute a unique opportunity for the validation and calibration of this neural mass model, which might become a relevant tool in the assessment of public and workers exposure to environmental MF, and assisting in the development and evaluation of guidelines developed by ICNIRP.

ACKNOWLEDGMENTS

This work is supported by an Elevate Post-Doctoral Fellowship from Mitacs, Hydro-Québec/Electricité de France/Réseau de Transport d'Electricité, and the Canadian Institutes of Health Research.

REFERENCES

- Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87.
- Bikson, N., Inoue, M., Akiyama, H., Deans, J. K., Fox, J. E., Miyakawa, H., et al. (2004). Effects of uniform extracellular DC electric fields on excitability in rat hippocampal slices *in vitro*. *J. Physiol.* 557(Pt 1), 175–190.
- Bojak, I., Oostendorp, T. F., Reid, A. T., and Kotter, R. (2010). Connecting mean field models of neural activity to EEG and fMRI data. *Brain Topogr.* 23, 139–149.
- Collingridge, G. L., Isaac, J. T., and Wang, Y. T. (2004). Receptor trafficking and synaptic plasticity. *Nat. Rev. Neurosci.* 5, 952–962.
- Cook, C. M., Thomas, A. W., Keenliside, L., and Prato, F. S. (2005). Resting EEG effects during exposure to a pulsed ELF magnetic field. *Bioelectromagnetics* 26, 367–376.
- Cook, C. M., Thomas, A. W., and Prato, F. S. (2004). Resting EEG is affected by exposure to a pulsed ELF magnetic field. *Bioelectromagnetics* 25, 196–203.
- Corbacio, M., Brown, S., Dubois, S., Goulet, D., Prato, F. S., Thomas, A. W., et al. (2011). Human cognitive performance in a 3 mT power-line magnetic field. *Bioelectromagnetics* 23, 620–633.
- Crasson, M. (2003). 50-60 Hz electric and magnetic field effects on cognitive function in humans: a review. *Radiat. Prot. Dosimetry* 106, 333–340.
- Deans, J. K., Powell, A. D., and Jefferys, J. G. (2007). Sensitivity of coherent oscillations in rat hippocampus to AC electric fields. *J. Physiol.* 583(Pt 2), 555–565.
- Fröhlich, F., and McCormick, D. A. (2010). Endogenous electric fields may guide neocortical network activity. *Neuron* 67, 129–143.
- Ghione, S., Seppia, C. D., Mezzalana, L., and Bonfiglio, L. (2005). Effects of 50 Hz electromagnetic fields on electroencephalography alpha activity, dental pain threshold and cardiovascular parameters in humans. *Neurosci. Lett.* 382, 112–117.
- Gianni, M., Liberti, M., Appollonio, F., and D'Inzeo, G. (2006). Modeling electromagnetic fields detectability in a HH-like neuronal system: stochastic resonance and window behavior. *Biol. Cybern.* 94, 118–127.
- Grimbert, F., and Fauergas, O. (2006). Bifurcation analysis of Jansen's neural mass model. *Neural Comput.* 18, 3052–3068.
- ICNIRP—International Commission on Non-Ionizing Radiation Protection. (2010). Guidelines for Limiting Exposure to Time-Varying Electric and Magnetic Fields (1 Hz – 100 kHz). *Health Phys.* 99, 818–836.
- Jansen, B. H., and Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.* 73, 357–366.

- Jefferys, J. G., Deans, J., Bikson, M., and Fox, J. (2003). Effects of weak electric fields on the activity of neurons and neuronal networks. *Radiat. Prot. Dosimetry* 106, 321–323.
- Legros, A., and Beuter, A. (2005). Effect of a low intensity magnetic field on human motor behavior. *Bioelectromagnetics* 26, 657–669.
- Legros, A., Corbacio, M., Beuter, A., Modolo, J., Goulet, D., Prato, F. S., et al. (2012a). Neurophysiological and behavioral effects of a 60 Hz, 1800 microtesla magnetic field in humans. *Eur. J. Appl. Physiol.* 112, 1751–1762.
- Legros, A., Modolo, J., Goulet, D., Plante, M., Souques, M., Deschamps, F., et al. (2012b). “Threshold for a systematic neurophysiological response to 50 and 60 Hz magnetic fields up to 50 millitesla,” *34th Annual Conference of the Bioelectromagnetics Society* (Brisbane, QLD).
- Legros, A., Miller, J. E., Modolo, J., Corbacio, M., Robertson, J. R., Goulet, D., et al. (2010). “Is finger tapping induced brain activation modulated by an exposure to a 60 Hz, 3000 μ T magnetic field?” in *32nd Annual Conference of the Bioelectromagnetics Society* (Seoul).
- Legros, A., Miller, J., Modolo, J., Corbacio, M., Robertson, J., Goulet, D., et al. (2011). Multi-modalities investigation of 60 Hz magnetic field effects on the human central nervous system. *Electra* 256, 14–18.
- Lövsund, P., Oberg, P. A., Nilsson, S. E., and Reuter, T. (1980). Magnetophosphores: a quantitative analysis of thresholds. *Med. Biol. Eng. Comput.* 18, 326–334.
- Miller, J. E., Modolo, J., Robertson, J. R., Corbacio, M., Dubois, S., Goulet, D., et al. (2010). “Effects of a 60 Hz magnetic field exposure on human brain activity during a mental rotation task as measured by fMRI,” *32nd Annual Conference of the Bioelectromagnetics Society* (Seoul).
- Modolo, J., Bhattacharya, B., Edwards, R., Campagnaud, J., Legros, A., and Beuter, A. (2010). Using a virtual cortical modulate implementing a neural field model to modulate brain rhythms in Parkinson's disease. *Front. Neurosci.* 4:45. doi: 10.3389/fnins.2010.00045
- Modolo, J., Juen, N., Robertson, J. A., Thomas, A. W., and Legros, A. (2011). “EEG frequency analysis of 60 Hz magnetic field exposure within the MRI,” *Conference of the Conseil International des Grands Réseaux Electriques (CIGRE)* (Paris).
- Molae-Ardekani, B., Benquet, P., Bartolomei, F., and Wendling, F. (2010). Computational modeling of high-frequency oscillations at the onset of neocortical partial seizures: from “altered structure” to “dysfunction”. *Neuroimage* 52, 1109–1122.
- Molae-Ardekani, B., Márquez-Ruiz, J., Merlet, I., Leal-Campagnario, R., Gruart, A., Sánchez-Campusano, R., et al. (2013). Effects of transcranial direct current stimulation (tDCS) on cortical activity: a computational modeling study. *Brain Stimul.* 6, 25–39.
- National Institute of Environmental Health Sciences of the National Institutes of Health. (1998). *Assessment of Health Effects From Exposure to Power-Line Frequency Electric and Magnetic Fields*, eds C. J. Protier and M. S. Wolfe (Research Triangle Park, NC: NIH publication No. 98–3981).
- Pilla, A. (2012). Electromagnetic fields instantaneously modulate nitric oxide signaling in challenged biological systems. *Biochem. Biophys. Res. Commun.* 426, 330–333.
- Radman, T., Ramos, R. L., Brumberg, J. C., and Bikson, M. (2009). Role of cortical cell type and morphology in subthreshold and suprathreshold uniform electric field stimulation *in vitro*. *Brain Stimul.* 2, 215–228.
- Radman, T., Su, Y., An, J. H., Parra, L. C., and Bikson, M. (2007). Spike timing amplifies the effects of electric fields on neurons: implications for endogenous field effects. *J. Neurosci.* 27, 3030–3036.
- Reato, D., Rahman, A., Bikson, M., and Parra, L. C. (2010). Low-intensity electrical stimulation affects network dynamics by modulating population rate and spike timing. *J. Neurosci.* 30, 15067–15079.
- Robinson, P. A. (2011). Neural field theory of synaptic plasticity. *J. Theor. Biol.* 285, 156–163.
- Shouval, H. Z., Castellani, G. C., Blais, B. S., Yeung, L. C., and Cooper, L. N. (2002a). Converging evidence for a simplified biophysical model of synaptic plasticity. *Biol. Cybern.* 87, 383–391.
- Shouval, H. Z., Bear, M. F., and Cooper, L. N. (2002b). A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 99, 10831–10836.
- Stodilka, R. Z., Modolo, J., Prato, F. S., Robertson, J. A., Cook, C., Patrick, J., et al. (2011). Pulsed magnetic field exposure induces lasting changes in neural network dynamics. *Neurocomputing* 74, 2164–2175.
- Sotero, R. C., and Trujillo-Barreto, N. J. (2008). Biophysical model for integrating neuronal activity, EEG, fMRI and metabolism. *Neuroimage* 39, 290–309.
- Weiss, S. A., and Faber, D. S. (2010). Field effects in the CNS play functional roles. *Front. Neural Circuits* 4:15. doi: 10.3389/fncir.2010.00015
- Wendling, F., Bartolomei, F., Bellanger, J. J., and Chauvel, P. (2002). Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *Eur. J. Neurosci.* 15, 1499–1508.
- Wendling, F., Hernandez, A., Bellanger, J. J., Chauvel, P., and Bartolomei, F. (2005). Interictal to ictal transition in human temporal lobe epilepsy: insights from a computational model of intracerebral EEG. *J. Clin. Neurophysiol.* 22, 343–356.
- Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 November 2012; accepted: 25 March 2013; published online: 11 April 2013.

Citation: Modolo J, Thomas AW and Legros A (2013) Neural mass modeling of power-line magnetic fields effects on brain activity. *Front. Comput. Neurosci.* 7:34. doi: 10.3389/fncom.2013.00034

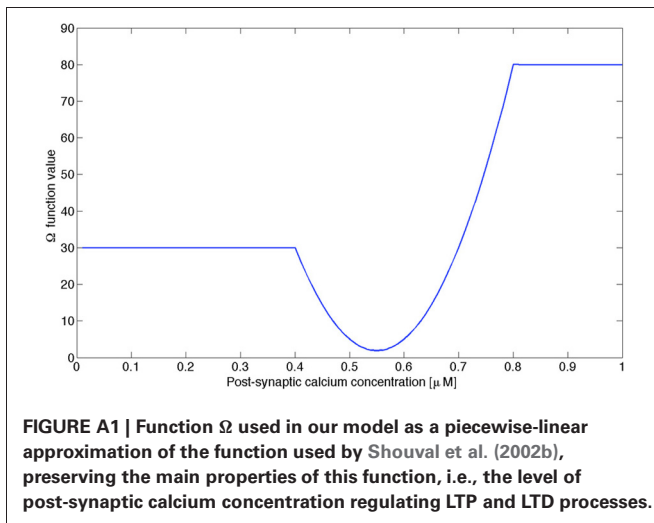
Copyright © 2013 Modolo, Thomas and Legros. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

Table A1 | Summary of the parameters used in the model with their physiological meaning and their values.

Parameter	Physiological significance	Value
c_1	Number of synaptic contacts from pyramidal to exc. Interneurons	135
c_2	Number of synaptic contacts from exc. interneurons to exc. interneurons	$0.8 \times c_1$
c_3	Number of synaptic contacts from pyramidal to slow inh. Interneurons	$0.25 \times c_1$
c_4	Number of synaptic contacts from slow inh. interneurons to pyramidal cells	$0.25 \times c_1$
c_5	Number of synaptic contacts between pyramidal cells in the same cortical column	$0.25 \times c_1$
c_6	Number of synaptic contacts from pyramidal cells to other areas	200
c_7	Number of synaptic contacts from pyramidal cells to pyramidal cells in the same brain area	100
c_8	Number of synaptic contacts from exc. interneurons to pyramidal cells in the same brain area	100
c_9	Number of synaptic contacts from slow inh. interneurons to pyramidal cells in the same brain area	100
c_{10}	Number of synaptic contacts from pyramidal cells to fast inh. interneurons	$0.3 \times c_1$
c_{11}	Number of synaptic contacts from slow inh. interneurons to fast inh. interneurons	$0.8 \times c_1$
c_{12}	Number of synaptic contacts from fast inh. interneurons to pyramidal cells	$0.1 \times c_1$
r	Slope of the sigmoidal function	0.56
e_0	Maximum of the sigmoid function	2.5 Hz
v_0	Threshold of the sigmoid function	6 mV
A	Amplitude of glutamatergic EPSPs	3.25 mV
B	Amplitude of GABA _a slow and GABA _b IPSPs	22 mV
G	Amplitude of GABA _a fast IPSPs	10 mV
a	Synaptic time constant for excitatory connections	100/s
b	Synaptic time constant for slow inhibitory connections	50/s
g	Synaptic time constant for fast inhibitory connections	350/s
a_{d1}	Time constant of efferent connections	33/s
a_{d2}	Time constant of efferent connections	100/s
a_{d3}	Time constant of efferent connections	100/s
b_{d4}	Time constant of efferent connections	40/s
c_{1t}	Number of synaptic contacts from TC to RE cells	50/s
c_{2t}	Number of synaptic contacts from RE to TC cells	50/s
c_{3t}	Number of synaptic contacts from TC to pyramidal cells	80/s
c_{4t}	Number of synaptic contacts from TC to exc. Interneurons	100/s
c_{5t}	Number of synaptic contacts from TC to inh. Interneurons	80/s
A_t	Amplitude of thalamic excitatory EPSPs (TC cells)	3.25 mV
A_t	Amplitude of thalamic excitatory EPSPs (TC cells)	3.25 mV
B_t	Amplitude of thalamic inhibitory IPSPs (RE cells)	22 mV
a_t	Synaptic time constant of TC cells	200/s
b_t	Synaptic time constant RE cells	40/s
$p(t)$	Input noise to pyramidal cells	120 pulses/s
K	Coupling constant between the thalamic and cortical modules of the model	0.2
K_{Cx}	Coupling constant between the cortical and thalamic modules of the model	10
K	Coupling constant between different areas	3.33
τ_{Ca}	Time constant of calcium current dynamics	0.05 s
γ	Sensitivity of calcium concentration to EEG changes	0.055
η	Time constant of synaptic weight changes	0.054 ms

Novel parameters compared to Sotero and Trujillo-Barreto (2008) are highlighted in red.





Stability constraints on large-scale structural brain networks

Richard T. Gray^{1*} and Peter A. Robinson^{2,3}

¹ The Kirby Institute, The University of New South Wales, Sydney, NSW, Australia

² School of Physics, University of Sydney, Sydney, NSW, Australia

³ Brain Dynamics Center, Sydney Medical School – Western, University of Sydney, Westmead, NSW, Australia

Edited by:

Dimitris Pinotsis, University College London, UK

Reviewed by:

Ingo Bojak, University of Birmingham, UK

Viktor Jirsa, Aix-Marseille University, France

Paul C. Bressloff, University of Utah, USA

*Correspondence:

Richard T. Gray, The Kirby Institute, The University of New South Wales, Sydney, NSW 2052, Australia.
e-mail: rgray@kirby.unsw.edu.au

Stability is an important dynamical property of complex systems and underpins a broad range of coherent self-organized behavior. Based on evidence that some neurological disorders correspond to linear instabilities, we hypothesize that stability constrains the brain's electrical activity and influences its structure and physiology. Using a physiologically-based model of brain electrical activity, we investigated the stability and dispersion solutions of networks of neuronal populations with propagation time delays and dendritic time constants. We find that stability is determined by the spectrum of the network's matrix of connection strengths and is independent of the temporal damping rate of axonal propagation with stability restricting the spectrum to a region in the complex plane. Time delays and dendritic time constants modify the shape of this region but it always contains the unit disk. Instabilities resulting from changes in connection strength initially have frequencies less than a critical frequency. For physiologically plausible parameter values based on the corticothalamic system, this critical frequency is approximately 10 Hz. For excitatory networks and networks with randomly distributed excitatory and inhibitory connections, time delays and non-zero dendritic time constants have no impact on network stability but do effect dispersion frequencies. Random networks with both excitatory and inhibitory connections can have multiple marginally stable modes at low delta frequencies.

Keywords: brain networks, stability, network spectra, random matrices, mean-field modeling

INTRODUCTION

The brain is possibly the most complicated example of a system of interacting dynamical units whose activity self-organizes to produce complex global behavior. The human brain performs cognitive functions through the transmission of action potentials within a vast structurally dynamic network consisting of approximately 10^{11} neurons and up to 10^{15} synaptic interconnections (Kandel et al., 2000; Koch, 2004; Sporns et al., 2005). The aggregate of all neural firings within this network results in large-scale coherent electrical activity and the performance of high-level cognitive functions. Understanding the structure and physiology of the brain thus gives insight into its overall behavior.

At large scales the excitatory and inhibitory neurons in the brain are organized into a complex large-scale network of distinct anatomical and functional structures (Sporns et al., 2004, 2005; Bullmore and Sporns, 2009; van den Heuvel and Sporns, 2011). We can represent this structure as a complex network—the structure of which has been studied extensively in recent times with a number of experimental cortical connection networks determined for the cat and the macaque monkey (Felleman and van Essen, 1991; Scannell et al., 1995; Jouve et al., 1998; Hilgetag et al., 2000a,b; Sporns, 2011). These networks have a modular hierarchical structure with the small-world properties of high local clustering and short path length between structures (Hilgetag et al., 2000a,b; Sporns et al., 2000, 2004, 2005; Young, 2000;

Sporns and Zwi, 2004; Bassett and Bullmore, 2006; Bullmore and Sporns, 2009, 2012).

Reasons for why the brain has evolved this particular large-scale structure are currently unknown. A number of investigations have concentrated on the effect of physical constraints on brain structure. Such constraints include brain volume, wiring length, and energy consumption or metabolic demands (Laughlin et al., 1998; Attwell and Laughlin, 2001; Lennie, 2003). Other studies have looked at functional constraints such as minimizing the conduction delay or processing steps for a signal to travel from one neuron to another (Wen and Chklovskii, 2005). Alternatively, the dynamics of the brain's electrical activity may constrain the brain's structure. If the physiology and structural characteristics of the brain produce adverse electrical activity resulting in seizures, tremors, or other neurological disorders then it is likely the structural characteristics of the brain will be constrained to limit these disorders.

One of the most important dynamical properties of complex systems such as the brain is stability. It has been associated with pattern formation (Turing, 1952; Murray, 2002), synchronized activity (Kuramoto, 1984; Pecora and Carroll, 1998; Jirsa and Ding, 2004; Acebrón et al., 2005; Feng et al., 2006), the complexity and diversity of ecosystems (May, 1972, 1974; Hogg et al., 1989; McCann, 2000; Allesina and Tang, 2012), the functioning of biological systems (Murray, 2002; Taverna and

Goldstein, 2002; Steuer, 2007), and the generation of coherent self-organized behavior. Stability is also an important aspect of the design and control of advanced technological systems (Bechhoefer, 2005).

A common approach to studying the large-scale dynamics of the brain's electrical activity is to use a continuum mean-field approximation for neural activity. This approach has been extensively studied over the past 30 years producing numerous models for the electrical activity within the brain (Wilson and Cowan, 1973; da Silva et al., 1974; Nunez, 1974, 1995; Freeman, 1975; Steriade et al., 1990; Jirsa and Haken, 1996; Wright and Liley, 1996; Robinson et al., 1997, 2003a; Wright et al., 2001; Robinson, 2005). This work has been reviewed recently (Deco et al., 2008; Bressloff, 2011). These models have been used extensively to perform stability analysis and understand pattern formation, oscillations and waves in the brain's electrical activity (Deco et al., 2008; Bressloff, 2011).

Measurements of brain activity suggest that the brain operates close to marginal stability, permitting a wide range of flexible, adaptable, and complex behavior (Stam et al., 1999; Robinson et al., 2001b; Breakspear, 2002; Breakspear et al., 2003). Physiological modeling also suggests that linear instabilities in the brain's electrical activity correspond to neurological disorders, such as epilepsy (Robinson et al., 1998, 2002; Breakspear et al., 2006; Kim and Robinson, 2007; Deco et al., 2008). It is therefore possible that stability is a dynamical property that imposes constraints on the brain's physiology and structure.

In previous work we have used a simplified version of the Robinson, Rennie, Wright (RRW) physiologically-based continuum (mean-field) model (Robinson et al., 1997, 1998, 2003a; Wright et al., 2001; Rennie et al., 2002; Robinson, 2003, 2005) to study the dynamics of the electrical activity in large-scale structural brain networks. Based on the hypothesis that stability is a dynamical constraint on the structure and physiology of structural brain networks we have investigated the effect of stability of large-scale structural brain networks (Gray and Robinson, 2006, 2008, 2009b,a; Robinson et al., 2009).

Our previous work ignored the dendritic time constants of neurons and the propagation time delays for signals to travel between neural populations. However, time delays due to axonal propagation affect stability and the possible physiology of neuronal networks (Atay and Hutt, 2004; Jirsa and Ding, 2004; Coombes, 2005; Coombes et al., 2007; Qubbaj and Jirsa, 2007, 2009; Venkov et al., 2007; Jirsa, 2009). This previous work has generally used integro-differential neural field equations with connectivity within a neural mass described by homogeneous or heterogeneous kernels. Our approach here is to focus on the temporal dynamics of the overall electrical activity of arbitrarily connected large-scale structural brain networks, ignoring the spatial spread and propagation of electrical activity within individual neuronal populations.

In this study, we increase the physiological realism of our structural brain network model by allowing propagation time delays and non-zero dendritic time constants. After reviewing the stability of structural brain networks, we aim to investigate how these physiological features might affect the dynamics and stability of networks of neuronal populations where the connection

patterns between populations are arbitrary—ignoring the spatial and geometric placement of the populations and simply focusing on which populations are inter-connected.

We also investigate the dispersion frequencies of marginally stable modes of electrical activity using plausible physiological parameters. The incorporation of non-zero dendritic time constants generalizes the work in (Jirsa and Ding, 2004) which, by including time delays, extended May's original analysis on the stability of complex systems (May, 1972, 1974).

METHODS

A structural brain network of n neural populations is represented by a directed graph N whose vertices and edges represent specific neural populations and inter-population connections, respectively. Neural populations within a network are collections of neurons with an assumed effective range and of sufficient number for a mean-field approximation to be valid. For example, a neural population can represent all the neurons in a distinct region or nuclei of the brain (e.g., cortical area, thalamus), a particular neuron type (e.g., interneuron, pyramidal cell), or a particular neurotransmitter type (e.g., glutamate, GABA, dopamine). Neurons in one population do not have to be separated geometrically or physically within the brain and can be intermixed with the neurons of another population (e.g., excitatory and inhibitory neurons in the cortex).

The structure of N is represented by a connection matrix $C(N) = [C_{ab}]$; where $C_{ab} = 1$ if there is a connection from population b to population a , $C_{ab} = 0$ otherwise. If $C_{ab} = C_{ba}$ for all a and b , the network is symmetric; otherwise it is asymmetric. Self-connections in structural brain networks correspond to non-zero diagonal entries in $C(N)$. The connection matrix simply records whether one neuronal population sends neural signals to another neuronal population. Properties of connections are not included in $C(N)$.

PHYSIOLOGICALLY-BASED STRUCTURAL BRAIN NETWORK DYNAMICS

In this section we outline the physiological model used to describe the dynamics of a brain network. If a neural population contains a sufficient number of neurons a continuum approximation can be used, whereby the properties of population neurons are averaged over. This approximation is valid for length scales greater than a few tenths of a millimeter and is thus suitable for investigating the dynamics of large-scale structural brain networks.

The continuum approximation allows the use of a previously developed model, the RRW model, for the brain's electrical activity (Robinson et al., 1997; Wright et al., 2001; Rennie et al., 2002; Robinson et al., 2003b, 2004; Robinson, 2005). This continuum model incorporates and describes three features of neural dynamics: (1) the synapto-dendritic dynamics resulting in the cell body potential; (2) from the mean cell body potential an average firing rate is determined via a non-linear sigmoid function; and (3) the population firing rate generates a neural pulse forming a field $\phi(t)$ that propagates along the populations outgoing connections. The field within a population is temporally described using a damped wave equation. Implicitly the neurons in each population are assumed to have an effective range which

gives a rate at which spikes reach axonal terminals and cease existence.

This model has been extensively used to model the corticothalamic system with the linear version having been shown to produce excellent agreement with EEG spectra, ERP, and other neurophysiological phenomena (Robinson et al., 1997, 2001a,b; Rennie et al., 2002; Robinson, 2003). To apply this continuum model to brain networks we previously used a number of simplifying assumptions. In particular, we assumed that all neural populations have instantaneous dendritic response times and there is no time delay for a signal to be sent from one population to the other (Gray and Robinson, 2006, 2008, 2009b,a; Robinson et al., 2009). For this study we relax some of these assumptions.

Firstly we assume time delays τ for a signal to be sent from one population to another are equal. Secondly, we assume each population in a network has the same dendritic decay rate α and rise rate β . The values of $1/\alpha$ and $1/\beta$ equal the dendritic decay and rise time constants, respectively; instantaneous rise and decay times imply $1/\alpha = 1/\beta = 0$. These assumptions are unrealistic for real structural brain networks but improve our previous analysis and allow us to analytically determine stability. Though for some structural brain networks these assumptions may be good approximations of the networks physiology. Another weakness of these assumptions is self-connections have the same time delay as connections between distinct populations—if $\tau \neq 0$ then self-connections in a network also involve a delay. Generally, self-connections represent interconnections within a neural population and would be expected to have zero time delay. However, for cortical networks self-connections can be used to represent feedback from underlying structures such as the thalamus. Time delayed self-connections would be appropriate for this type of feedback. We will generalize the assumption of equal time delays in future work.

The neurophysics and neurophysiology incorporated into the general RRW model and the equations for the linear perturbations of the neural field ϕ_a , for each neural population a , are described and derived in detail elsewhere (Robinson et al., 1997; Wright et al., 2001; Rennie et al., 2002; Robinson et al., 2003b, 2004; Robinson, 2005). This study uses the notation and equations derived in (Robinson, 2005). Under the assumptions used here the RRW equations describing linear perturbations of the neural field ϕ_a of population a about the assumed steady state in Fourier space reduce to

$$(1 - i\omega/\gamma)^2 \phi_a(\omega) = \sum_b L(\omega) G_{ab} e^{i\omega\tau} \phi_b(\omega), \quad (1)$$

$$= L(\omega) e^{i\omega\tau} \sum_b G_{ab} \phi_b(\omega), \quad (2)$$

where ω is the angular frequency and

$$L(\omega) = \frac{\alpha\beta}{(\alpha - i\omega)(\beta - i\omega)} = \frac{1}{(1 - i\omega/\alpha)(1 - i\omega/\beta)}. \quad (3)$$

The gain G_{ab} is a dimensionless quantity describing the effect of changes in the firing rate of neurons in population b on the

neurons of population a . Physiologically, G_{ab} is the number of extra action potentials produced in a per extra action potential incident from b . Hence, G_{ab} is a measure of how sensitive and responsive a is to changes in b 's activity. In the general RRW model γ is a damping rate equal to the velocity of the ϕ 's propagation within a neural population divided by the characteristic range of the axons that carry it. In the spatially uniform case used here, γ represents a temporal damping rate.

Letting $\mathbf{G} = [G_{ab}]$ be the matrix of gains and setting

$$D(\omega) = [L(\omega)]^{-1} (1 - i\omega/\gamma)^2 e^{-i\omega\tau}, \quad (4)$$

$$= (1 - i\omega/\alpha)(1 - i\omega/\beta)(1 - i\omega/\gamma)^2 e^{-i\omega\tau}, \quad (5)$$

which is a complex analytic function. Equation (1) can be written in matrix form as

$$\mathbf{\Delta}(\omega) \mathbf{\Phi}(\omega) = \mathbf{G} \mathbf{\Phi}(\omega), \quad (6)$$

where $\mathbf{\Phi}$ is a column vector of the ϕ_a and $\mathbf{\Delta}(\omega) = D(\omega) \mathbf{I}$, where \mathbf{I} is the identity matrix. Setting $\mathbf{A} = \mathbf{G} - \mathbf{\Delta}$, Equation (6) can be simplified to

$$\mathbf{A}(\omega) \mathbf{\Phi}(\omega) = 0. \quad (7)$$

The linear stability of a network is then determined by the solutions ω of the dispersion relation,

$$\det[\mathbf{A}(\omega)] = 0. \quad (8)$$

The gain matrix $\mathbf{G} = [G_{ab}]$ encodes all of the information in \mathbf{C} , since $G_{ab} \neq 0$ implies $C_{ab} \neq 0$, as well as the strength of connections between populations. No assumptions (such as homogeneity or isotropy) are made for the characteristics of a connection and any attenuation or phase shifting of an incoming signal due to time delays are reflected in the exponential term of Equation (1). However, the model implicitly assumes an effective range for neurons within a population. If $G_{ab} > 0$ then the connection is excitatory and if $G_{ab} < 0$ the connection is inhibitory. Note that Equation (5) shows that if the values of α and β are exchanged, the brain network has the same dynamics and stability.

REALISTIC PARAMETER VALUES FOR LARGE-SCALE STRUCTURAL BRAIN NETWORKS

Physiologically plausible parameter values for γ , α , β , and τ are shown in **Table 1**. These values are based on the parameters used in the corticothalamic model (Robinson et al., 2003a, 2004), with the specific values taken from (Robinson et al., 2004). The nominal values in **Table 1** are the default model parameters used to illustrate our results under our assumptions α and β are the same value for all populations and τ is the same for all connections. The values for γ are based on the cortical excitatory neurons which form the long range connections within the cortex. Inhibitory inter-neurons in the cortex are short range (Nunez, 1995) and therefore have $\gamma \approx \infty$. Under the assumptions used here all neural populations are given the same γ value.

In real structural brain networks dendritic time constants and propagation time delays may vary. The spatial distribution and

Table 1 | Physiologically plausible ranges and nominal values of parameters.

Parameter	Range	Nominal value	Unit
γ	30–220	100	s^{-1}
α	5–200	60	s^{-1}
β	17–2500	240	s^{-1}
β/α	1–10	4	–
τ	0–50	10	ms

Based on the corticothalamic model parameters in Robinson et al. (2004). The nominal values are the default model parameters we use in this work.

physical separation of structures within the brain will lead to distinct time delays. The values of τ in **Table 1** are physiologically plausible values for the time delays based on corticothalamic modeling. The nominal value is a realistic value for the average delay between the large-scale neural populations (or areas) in the cerebral cortex.

RANDOMLY CONNECTED LARGE-SCALE STRUCTURAL BRAIN NETWORKS

To illustrate our results we investigate randomly connected structural brain networks where neural populations are connected randomly with probability p . The size n and probability of connection we use is based on experimentally determined cortical connection networks for animals. These have been analyzed with graph-theoretical methods and all of these networks have less than 100 neural populations with a connection density (percentage of existing connections out all possible connections) of 20–40% (Felleman and van Essen, 1991; Scannell et al., 1995; Hilgetag et al., 2000a; Sporns et al., 2000, 2004; Bullmore and Sporns, 2009; Rubinov and Sporns, 2010; Sporns, 2011). We use random networks with $n = 50$ and $p = 0.5$ to illustrate our results, allowing comparisons with real cortical networks. These values of p ensure the networks are strongly connected (Bollobás, 1985) and all populations have at least one input and one output with high probability; i.e., there are no sources or sinks of electrical activity.

The specific random networks we investigate are the same random networks we have previously investigated (Gray and Robinson, 2006, 2008, 2009a,b). These networks consist of excitatory and inhibitory connections. The probability that a connection is inhibitory is given by p_i and such a connection has a negative gain. Excitatory gains are given values from a normal distribution with a mean $\mu_e > 0$ and variance σ_e^2 . Similarly, inhibitory connections have a gain sampled from a normal distribution with $\mu_i < 0$ and variance σ_i^2 . In terms of the gain matrix \mathbf{G} all positive entries are sampled from $\mathcal{N}(\mu_e, \sigma_e^2)$ and all negative entries are sampled from $\mathcal{N}(\mu_i, \sigma_i^2)$.

Based on these parameters we investigated the stability of three types of networks: random networks with fixed excitatory gains (RENS), random connection networks (RCNs) with excitatory and inhibitory connections distributed randomly within the network (Gray and Robinson, 2009b), and random population networks (RPNs) (Gray and Robinson, 2009a). RPNs represent random networks with excitatory and inhibitory populations of

neurons, this implies the outgoing connections of a given populations are all excitatory or all inhibitory. The gain matrix of RPNs consists of columns with either all entries ≤ 0 or all entries ≥ 0 .

We determine the dispersion solutions for these structural brain networks numerically using a FORTRAN program called CROOT (Botten et al., 1983). This program finds dispersion solutions by implementing a recursive algorithm that employs Cauchy's integral formula (Mitrinović and Kečkić, 1984) within a specified annulus or disk.

RESULTS

Our results describe the stability of structural brain networks by determining the criteria for a network to stable—starting from simple excitatory networks and then adding time delays and dendritic time constants. For our network model we show that stability is determined by the eigenvalues of the gain matrix with stability constraining the eigenvalues to a specific zone in the complex plane. The first subsections translate the results from previous work into the current context. In particular, when time delays are included we produce a similar tear-drop shaped stability zone found by others (Jirsa and Ding, 2004; Feng et al., 2006; Qubbaj and Jirsa, 2007; Jirsa, 2009; Qubbaj and Jirsa, 2009). However, we show the addition of dendritic time constants modifies the shape of the stability zone. Finally, we use our results to assess how stability constrains the physiology of randomly connected networks with excitatory and inhibitory connections.

STABILITY OF STRUCTURAL BRAIN NETWORKS

The solutions ω of the dispersion relation Equation (8) determine the linear stability of a network. Setting $\lambda = D(\omega)$, the dispersion relation is

$$\det(\mathbf{G} - \lambda \mathbf{I}) = 0. \quad (9)$$

Therefore, network stability is determined by the spectrum of \mathbf{G} , which we denote $\text{Sp}(\mathbf{G})$. All the dispersion solutions ω of the network can be obtained by solving

$$\lambda - D(\omega) = 0 \quad (10)$$

for each λ in $\text{Sp}(\mathbf{G})$. If all the λ in $\text{Sp}(\mathbf{G})$ have corresponding ω [given by Equation (10)] with $\text{Im}\omega < 0$ then the network is stable. However, if there exists one λ which has a corresponding dispersion solution with $\text{Im}\omega \geq 0$ then the network is unstable. The set of dispersion solutions of a brain network is termed the *dispersion spectrum*. Taking the complex conjugate of Equations (4, 10) show that if $\omega_1 = \text{Re}\omega + i\text{Im}\omega = \omega_r + i\omega_i$ is solution for λ then $\omega_2 = -\omega_r + i\omega_i$ is a solution for the complex conjugate $\bar{\lambda}$ of λ . Therefore, since both λ and $\bar{\lambda}$ are in $\text{Sp}(\mathbf{G})$, the dispersion spectrum is symmetric about the real axis.

Solving Equation (10) for ω is equivalent to solving $\lambda - D(\varpi) = 0$ for ϖ where

$$D(\varpi) = D(\gamma\varpi) = (1 - i\varpi\gamma/\alpha)(1 - i\varpi\gamma/\beta)(1 - i\varpi)^2 e^{-i\varpi\gamma\tau}, \quad (11)$$

and $\varpi = \omega/\gamma$ is a dimensionless frequency parameter. From the ϖ solutions the dispersion solutions for the network are

$\omega = \gamma\varpi$. Since D and \mathcal{D} are analytic, the dispersion spectrum can be obtained by numerically solving Equation (10) for each λ .

Boundary between unstable and stable states

As the stability of a network is determined by $\text{Sp}(\mathbf{G})$ we are interested in the zone of the complex plane where the all the eigenvalues of \mathbf{G} must lie for the network to be stable. If a dispersion solution has $\text{Im}\omega = 0$ (i.e., ω is real and marginally stable) then the λ corresponding to ω lies on the critical boundary between the unstable and stable zones in the complex plane. Therefore, the stability boundary is given by $D(\omega)$ for real ω ranging from $-\infty$ to ∞ . To describe the stability boundary we use $\mathcal{D}(\varpi)$ with real ϖ . This function traces out a continuous curve in the complex plane as ϖ ranges from $-\infty$ to ∞ . Points on this curve represent λ in $\text{Sp}(\mathbf{G})$ with real dispersion solutions $\omega = \gamma\varpi$. The stability boundary of a network is a segment of this curve since, in general, $\mathcal{D}(\varpi)$ can intersect itself to form loops.

Since \mathcal{D} is a re-parameterization of D using $\varpi = \omega/\gamma$, a brain network with model parameters $\gamma = \gamma'$, $\alpha = \alpha'$, $\beta = \beta'$, and $\tau = \tau'$ has the same stability boundary as a network with $\gamma = 1$, $\alpha = \alpha'/\gamma'$, $\beta = \beta'/\gamma'$, and $\tau = \gamma'\tau$. However, the dispersion spectra of these networks will differ. If an initially stable network becomes unstable due to changes in its connection gains, its spectrum initially lies in the stability zone before at least one eigenvalue moves across the stability boundary from the stable to unstable zones. The λ that crosses the boundary is an instability with frequency $\omega/2\pi = \gamma\varpi/2\pi$, where $\lambda = \mathcal{D}(\varpi)$ on the boundary.

We now describe some general properties of the stability boundary for structural brain networks with time delays and non-zero dendritic time constants. In the following sections we investigate particular cases. Firstly, $\mathcal{D} \neq 0$ for all real $-\infty < \varpi < \infty$, thus $\lambda = 0$ does not lie on the stability boundary. If $\lambda = 0$ is substituted into Equation (11) then

$$(1 - i\varpi)^2 = 0, (1 - i\varpi\gamma/\alpha)^2 = 0, \text{ or } (1 - i\varpi\gamma/\beta)^2 = 0, \quad (12)$$

for complex ϖ . The only solutions to Equation (12) are $\varpi = -i$, $\varpi = -i\alpha/\gamma$, or $\varpi = -i\beta/\gamma$ giving the dispersion solution $\omega = -i\gamma$, $\omega = -i\alpha$, and $\omega = -i\beta$. These solutions all lie in the lower half plane for $\alpha, \beta, \gamma > 0$ and hence $\lambda = 0$ is a stable eigenvalue. If $\omega = \varpi = 0$ then $\mathcal{D}(\varpi) = 1$, since $L(0) = 1$, and $\lambda = 1$ lies on the stability boundary. This implies that if $\lambda = 1$ is in $\text{Sp}(\mathbf{G})$ then the network has a zero frequency marginally stable dispersion solution. Furthermore, if we consider

$$|\lambda| = |L(\varpi)|^{-1} |e^{-i\varpi\gamma\tau}| |1 - i\varpi|^2, \quad (13)$$

$$= |1 - i\varpi\gamma/\alpha| |1 - i\varpi\gamma/\beta| |e^{-i\varpi\gamma\tau}| |1 - i\varpi|^2, \quad (14)$$

then $|\lambda| < 1$ implies at least one of the factors on the right of Equation (14) is less than 1. Since α, β, γ , and τ are all positive, this condition can only be satisfied if $\text{Im}\omega < 0$. Therefore, the unit disk is always contained in the stability zone and if all the eigenvalues of a gain matrix lie in the unit disk the network is stable independent of α, β, γ , and τ .

Stability of excitatory networks

If all the connections in a structural brain network are excitatory then $G_{ab} \geq 0$ and \mathbf{G} is a non-negative matrix. The Perron–Frobenius theorem (Horn and Johnson, 1985; Cvetković et al., 1995) then implies that \mathbf{G} has a real eigenvalue λ_p such that $|\lambda_i| \leq \lambda_p$ for all λ_i in $\text{Sp}(\mathbf{G})$. Therefore, an excitatory brain network is stable if and only if $\lambda_p < 1$; i.e., all the eigenvalues are in the unit disk. This stability criteria follows from the discussion at the end of the previous section and the fact that if $\lambda_p = 1$ then $\omega = 0$ is a solution to Equation (11). Since 0 lies on the stability boundary this implies if $\lambda_p \geq 1$ then $\text{Im}\omega \geq 0$. This means the stability of an excitatory brain network is independent of α, β, γ , and τ (as described in the previous section). In general, structural brain networks have inhibitory connections and the Perron–Frobenius theorem does not apply. This implies the presence of inhibitory connections allows the stability zone to extend beyond the unit disk.

IMPACT OF TIME DELAYS AND DENDRITIC TIME CONSTANTS ON NETWORK STABILITY

Stability of networks with no time delays and instantaneous rise and decay times

For networks with no time delays and instantaneous dendritic rise and decay times (i.e., $1/\alpha = 1/\beta = \tau = 0$)

$$\mathcal{D}(\varpi) = (1 - i\varpi)^2, \quad (15)$$

and hence, for each λ in $\text{Sp}(\mathbf{G})$ there are two dispersion solutions given by

$$\omega = \gamma\varpi = -\gamma(i \pm i\sqrt{\lambda}). \quad (16)$$

Taking the imaginary part of Equation (16) we obtain

$$\text{Im}\omega = -\gamma(1 \pm \text{Re}\sqrt{\lambda}). \quad (17)$$

The stability condition $\text{Im}\omega < 0$ implies that $-\gamma(1 \pm \text{Re}\sqrt{\lambda})$ or $\text{Re}\sqrt{\lambda} \leq 1$ since $\gamma > 0$. Thus the stability of a network with $1/\alpha = 1/\beta = \tau = 0$ is independent of γ and all λ must satisfy $\text{Re}\sqrt{\lambda} \leq 1$ or alternatively $\text{Re}\lambda + |\lambda| \leq 2$ [since $(\text{Re}\sqrt{\lambda})^2$ equals $(\text{Re}\lambda + |\lambda|)/2$, with equality corresponding to the stability boundary].

If $\lambda_r = \text{Re}\lambda$ and $\lambda_i = \text{Im}\lambda$ then the stability zone is a parabolic zone in the complex plane given by

$$\lambda_i^2 \leq 4 - 4\lambda_r \quad (18)$$

with equality giving the stability boundary. The axis of the parabolic boundary is along the real axis with a turning point at $(\lambda_r, \lambda_i) = (1, 0)$ and imaginary axis intercepts at $\lambda_i = \pm 2$. This stability zone is the light gray zone in **Figure 1** and is the stability region described previously in (Gray and Robinson, 2008, 2009a,b).

Stability of networks with equal time delays

We now consider the effect of time delays on the stability of networks with $1/\alpha = 1/\beta = 0$. In this section we determine how the addition of a time delay to structural brain networks modifies the

parabolic stability zone described by Equation (18) and describe the characteristics of the stability boundary. As noted previously, under our assumptions any self-connection in a brain network has the same time delay as connections between populations.

In this case $L(\omega) = 1$, $\tau \neq 0$, and

$$\mathcal{D}(\omega) = (1 - i\omega)^2 e^{-i\omega\gamma\tau}. \quad (19)$$

Due to the exponential in Equation (19), $\lambda - \mathcal{D}(\omega) = 0$ has an infinite number of solutions for each eigenvalue of the gain matrix. If $\lambda = \text{Re}\lambda + i\text{Im}\lambda = \lambda_r + i\lambda_i$ then $\lambda = \mathcal{D}(\omega)$ (for real ω) implies

$$\lambda_r = \text{Re}\mathcal{D}(\omega) = (1 - \omega^2) \cos(\omega\gamma\tau) - 2\omega \sin(\omega\gamma\tau), \quad (20)$$

and

$$\lambda_i = \text{Im}\mathcal{D}(\omega) = (\omega^2 - 1) \sin(\omega\gamma\tau) - 2\omega \cos(\omega\gamma\tau), \quad (21)$$

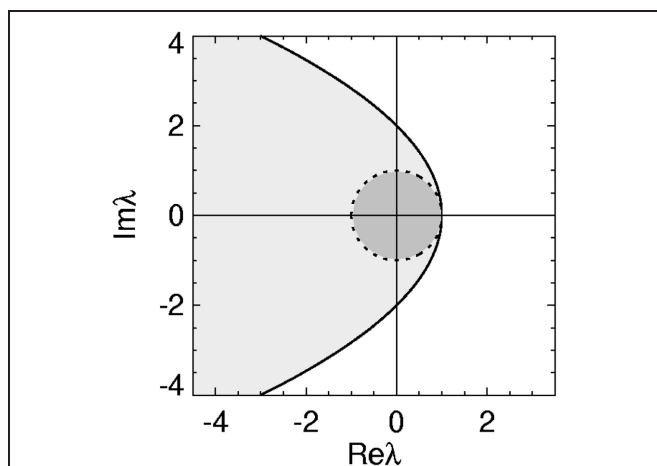


FIGURE 1 | Stability zone for a brain network with $1/\alpha = 1/\beta = \tau = 0$. The gray zone is where all the eigenvalues λ of the brain network must lie for the network to be stable. The dark region within this zone is the unit disk.

these equations can be combined giving

$$\lambda_i = -\tan(\omega\gamma\tau)[\lambda_r + 2\omega \text{cosec}(\omega\gamma\tau)], \quad (22)$$

which gives the stability boundary in the complex plane. For $\omega \geq 0$, Equation (22) describes a spiral curve in the complex plane starting at the point $(\lambda_r, \lambda_i) = (1, 0)$, centered on the origin, and spiraling in a clockwise direction as ω increases, as seen in **Figure 2A**. For $\omega < 0$, Equation (22) describes a similar counterclockwise spiral curve, corresponding to a reflection of the curve in **Figure 2A** about the imaginary axis.

Using Equations (20–22) we now describe the characteristics of the stability boundary and the resulting stability zone. The $\mathcal{D}(\omega)$ curve crosses the real axis when $\lambda_i = 0$. Substituting $\lambda_i = 0$ into Equation (21) and (22) gives

$$(\omega^2 - 1)\sin(\omega\gamma\tau) = 2\omega \cos(\omega\gamma\tau), \quad (23)$$

and

$$\lambda_r = -2\omega \text{cosec}(\omega\gamma\tau), \quad (24)$$

respectively. The values of ω for which $\lambda_i = 0$ can be obtained by solving Equation (23) numerically. Note that Equation (24) is only valid if $\omega\gamma\tau \neq \pm m\pi/2$ for integers m . Due to the periodicity of the sine and cosine functions there is an infinite number of ω that satisfy these equations.

Thus the stability boundary consists of two spiral curves produced by bending the arms of the parabola described by Equation (18) inwards. These spiral curves intersect an infinite number of times on the real axis enclosing larger and larger regions of the complex plane as $|\omega\gamma\tau|$ increases. The intersection of all these enclosed regions, corresponding to the innermost zone, represents the stability zone for the network. The reason eigenvalues outside this innermost zone are instabilities is because they correspond to eigenvalues outside the stability zone described by Equation (18) when transformed through multiplication by $e^{i\omega\gamma\tau}$ (which removes the effect of the time delay).

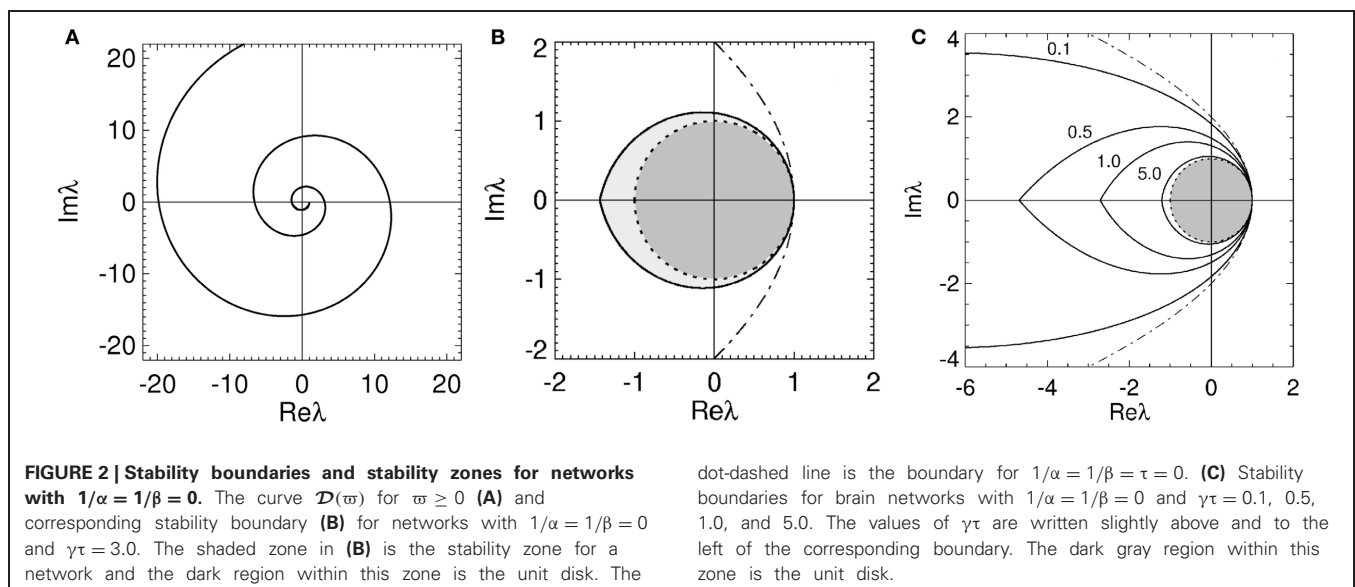


FIGURE 2 | Stability boundaries and stability zones for networks with $1/\alpha = 1/\beta = 0$. The curve $\mathcal{D}(\omega)$ for $\omega \geq 0$ (A) and corresponding stability boundary (B) for networks with $1/\alpha = 1/\beta = 0$ and $\gamma\tau = 3.0$. The shaded zone in (B) is the stability zone for a network and the dark region within this zone is the unit disk. The

dot-dashed line is the boundary for $1/\alpha = 1/\beta = \tau = 0$. (C) Stability boundaries for brain networks with $1/\alpha = 1/\beta = 0$ and $\gamma\tau = 0.1, 0.5, 1.0$, and 5.0 . The values of $\gamma\tau$ are written slightly above and to the left of the corresponding boundary. The dark gray region within this zone is the unit disk.

We define the smallest $\omega = \gamma\varpi > 0$ that gives $\lambda_i = 0$ the critical ω value. This critical value is denoted by ω_c with the corresponding ϖ and λ_r denoted ϖ_c and λ_r^c , respectively. The stability zone is defined by

$$|\operatorname{Im}\lambda| < |\operatorname{Im}[\mathcal{D}(\varpi)]|, \quad (25)$$

for $0 \leq \varpi \leq \varpi_c$. This zone has a boundary defined by Equation (19) for $-\varpi_c \leq \varpi \leq \varpi_c$. The real axis intercepts of the stability boundary are given by $\lambda_r = 1$ and $\lambda_r = \lambda_r^c = -2\varpi_c \operatorname{cosec}(\varpi_c \gamma \tau)$. As seen in **Figure 2B** the stability zone has a teardrop shape containing the unit disc. Note that Equation (25) also defines the stability zone in **Figure 1** with $\varpi_c = \infty$. As $\gamma\tau \rightarrow 0$ then $\lambda_r \rightarrow (1 - \varpi^2)$ in Equation (20), $\lambda_i \rightarrow -2\varpi$ in Equation (21), and the stability boundary converges to the parabola $\lambda_i^2 = 4 - 4\lambda_r$, as expected.

The effect of increasing τ on the stability boundary is shown in **Figure 2C**. As $\gamma\tau$ increases, the stability boundary converges to the unit circle with the stability zone converging to the unit disk, shaded dark gray. The values of ϖ_c in **Figure 2C** for the four values of $\gamma\tau$ shown are 3.0 ($\gamma\tau = 0.1$), 1.92 ($\gamma\tau = 0.5$), 1.31 ($\gamma\tau = 1.0$), and 0.46 ($\gamma\tau = 5.0$), respectively. The corresponding λ_r^c are -9.4 (intersection not seen), -4.6 , -2.7 , and -1.2 . For $\gamma = 100 \text{ s}^{-1}$ and $\tau = 0.01 \text{ s}$ (the nominal values in **Table 1**) $\omega_c/2\pi \approx 30 \text{ Hz}$.

Overall, the presence of time delays bends the parabola described by Equation (18) inward, forming a teardrop-shaped stability zone containing the unit disk. As $\tau \rightarrow \infty$ the stability boundary wraps around the unit circle an infinite number of times and the stability zone converges to the unit disk, restricting the critical frequency.

Stability of networks with non-zero dendritic rise and decay time constants and no time delays

The teardrop shaped zone and the time-delay effects on stability described in the previous section produce similar results to those seen in other studies (Marcus and Westervelt, 1989; Jirsa and Ding, 2004; Feng et al., 2006). However, our model also incorporates dendritic rise and decay time constants. In the next two sections we describe the stability of structural brain networks with non-zero dendritic time constants.

We first investigate brain networks with dendritic time constants and no propagation time delays. In this case $\tau = 0$, $\alpha \neq 0$, and $\beta \neq 0$. Hence $L(\omega) \neq 1$ and

$$\mathcal{D}(\varpi) = (1 - i\varpi\gamma/\alpha)(1 - i\varpi\gamma/\beta)(1 - i\varpi)^2. \quad (26)$$

From Equation (26) the stability boundary is given by

$$\lambda_r = 1 - [1 + 2\gamma(1/\alpha + 1/\beta) + \gamma^2/(\alpha\beta)]\varpi^2 + \gamma^2/(\alpha\beta)\varpi^4, \quad (27)$$

and

$$\lambda_i = -[2 + \gamma(1/\alpha + 1/\beta)]\varpi + [\gamma(1/\alpha + 1/\beta) + 2\gamma^2/(\alpha\beta)]\varpi^3. \quad (28)$$

where $\lambda_r = \operatorname{Re}\lambda$, $\operatorname{Im}\lambda = \lambda_i$, and ϖ is real. In this case $\lambda - \mathcal{D}(\varpi) = 0$ only has a finite number of solutions since $\mathcal{D}(\varpi)$ is a polynomial of degree four. From Equation (28) the values of ϖ_c and ω_c are given by

$$\varpi_c^2 = \frac{2\alpha\beta + \gamma(\alpha + \beta)}{2\gamma^2 + \gamma(\alpha + \beta)}, \quad (29)$$

and

$$\omega_c = \sqrt{\frac{2\alpha\beta\gamma + \gamma^2(\alpha + \beta)}{\alpha + \beta + 2\gamma}}, \quad (30)$$

respectively. An equation similar to Equation (30) was previously derived to describe gamma resonances produced by a similar mechanism (Robinson, 2005). If $\alpha\beta < \gamma^2$ then Equation (30) implies $\omega_c < \gamma$.

Unlike the case for time delays in the previous section, the $\mathcal{D}(\varpi)$ curve only crosses the real axis once for real $\varpi > 0$; an example of such a $\mathcal{D}(\varpi)$ curve is shown in **Figure 3A**. However, the region enclosed by $\mathcal{D}(\varpi)$ for $-\varpi_c \leq \varpi \leq \varpi_c$ is again the stability zone and defined by Equation (25). The corresponding stability zone for the \mathcal{D} curve in **Figure 3A** is shown in **Figure 3B**.

The stability zone in **Figure 3B** contains the unit disk and has a similar teardrop shape to the zone in **Figure 2B** with the arms of the parabola given by Equation (18) bent inward. In terms of stability, this implies non-zero $1/\alpha$ and $1/\beta$ have similar effects to a propagation time delay. This is consistent with previous work on the corticothalamic model and highlights the low-pass filter effect of $L(\omega)$ (Robinson et al., 1997, 2001a,b; Rennie et al., 2002). The effective time delay resulting from γ/α and γ/β can be obtained by solving

$$e^{-i\varpi_c\gamma\tau} = (1 - i\varpi_c\gamma/\alpha)(1 - i\varpi_c\gamma/\beta), \quad (31)$$

for τ .

In the remainder of this section we explore the effect of varying γ , α , and β on the stability zone. We illustrate these effects by setting β/α to a positive constant. Firstly, for fixed β/α and large γ/α , Equations (27, 29) imply

$$\lambda_r \approx 1 + (\gamma/\alpha)^2\varpi^2(\varpi^2 - 1), \quad (32)$$

and

$$\varpi_c^2 \approx \alpha/\gamma, \quad (33)$$

respectively. From Equation (33), $\omega_c \approx 0$ for large γ/α and substituting Equation (33) into (32) shows that $\lambda_r^c \rightarrow -\infty$ as $\gamma/\alpha \rightarrow \infty$. These results are illustrated in **Figures 3C, 4**.

The change in the stability zone for fixed β/α and varying γ/α is shown in **Figure 3C**. All the stability zones contain the unit circle, and for $\gamma/\alpha = 0.1$ and 1.7 the stability zone is contained within the parabolic zone defined by Equation (18). However, the zone for $\gamma/\alpha = 100$ has expanded so that its boundary intersects the parabolic boundary at $\lambda_r \approx -11$. The ϖ_c in **Figure 3C** for $\gamma/\alpha = 100, 0.1$, and 1.7 are 0.15 , 4.0 , and 1.08 , respectively. The corresponding λ_r^c are -66 (intersection not seen), -19 , and -4.9 . For $\gamma/\alpha = 1.7$, the nominal value from **Table 1**, the critical frequency is $\omega_c/2\pi \approx 17 \text{ Hz}$.

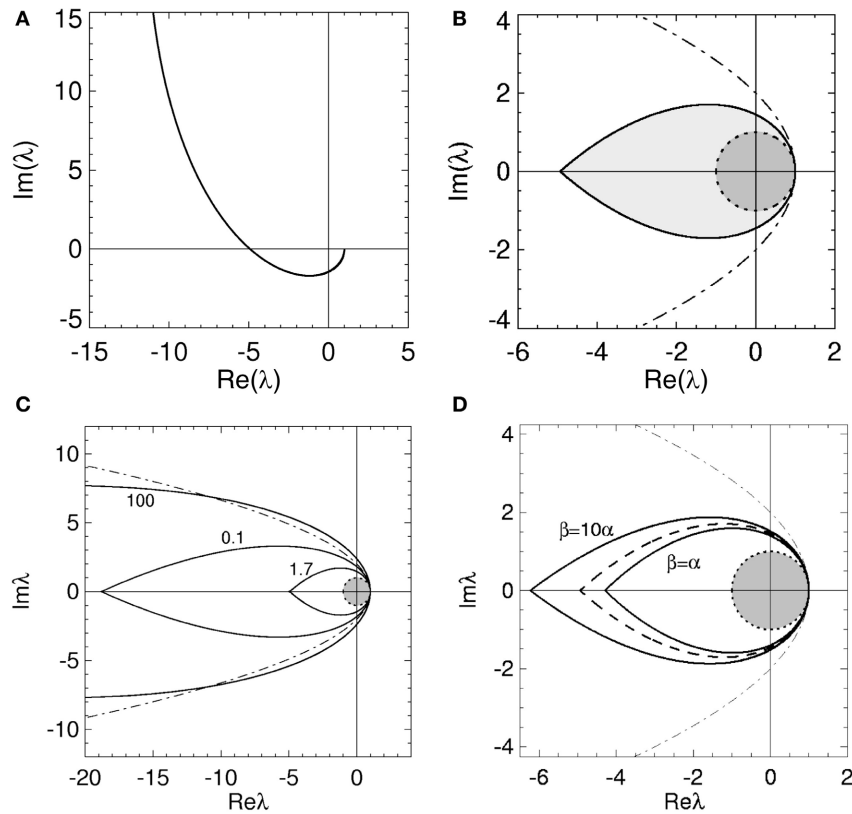


FIGURE 3 | Stability boundaries and stability zones for networks with non-zero dendritic rise and decay time constants and no time delays. The curve $\mathcal{D}(\omega)$ for real $\omega \geq 0$ (A) and corresponding stability zone and boundary (B) for networks with $\alpha = 60 \text{ s}^{-1}$, $\beta = 240 \text{ s}^{-1}$ and $\gamma = 100 \text{ s}^{-1}$ and $\tau = 0 \text{ s}$. The stability zone is shaded gray and the dark region within this zone is the unit disk. The dot-dashed line is the

boundary for $1/\alpha = 1/\beta = \tau = 0$. (C) Stability boundaries for $\beta = 4\alpha$ and $\gamma/\alpha = 100, 0.1$, and 1.7 . The values of γ/α are written next to their corresponding boundary. (D) Stability boundaries for $\gamma/\alpha = 1.7$ and $\beta/\alpha = 1, 4$ (dashed curve), and 10 . The dot-dashed line in (C,D) is the boundary for $1/\alpha = 1/\beta = \tau = 0$ and the dark gray region in (C,D) is the unit disk.

The effect of changing β/α on the stability zone, while γ/α remains fixed, is shown in **Figure 3D**. The $\omega_c(\lambda_r^c)$ in **Figure 3D** for $\beta/\alpha = 1$ and $\beta/\alpha = 10$ are 0.77 ($\lambda_r^c = -4.3$) and 1.26 ($\lambda_r^c = -6.2$), respectively. Note that if $\beta/\alpha < 1$, then the values of α and β can be swapped and the results in **Figure 3** are reproduced. This is because exchanging γ/α and γ/β in Equation (4) has no effect on the dynamics and stability of a network. This implies the smallest stability zone with the minimum λ_r^c occurs when $\alpha = \beta$. Note that experimental measurements of dendritic time constants in the brain give $\beta/\alpha \approx 4\text{--}10$ (Robinson et al., 2003a, 2004), the upper range in **Table 1**, and hence, a larger stability zone for brain activity.

Figure 3 suggests that as β/α increases the stability zone expands in a similar way to decreasing τ . However, as $\tau \rightarrow 0$ the stability zone converges to the parabolic zone (Equation 18), this is not the case for $\gamma/\beta \rightarrow \infty$ and fixed γ/α . If γ/α is fixed to a positive constant and $\gamma/\beta \gg \gamma/\alpha$ then, from Equation (29), $\omega_c^2 \lesssim 1$ and therefore

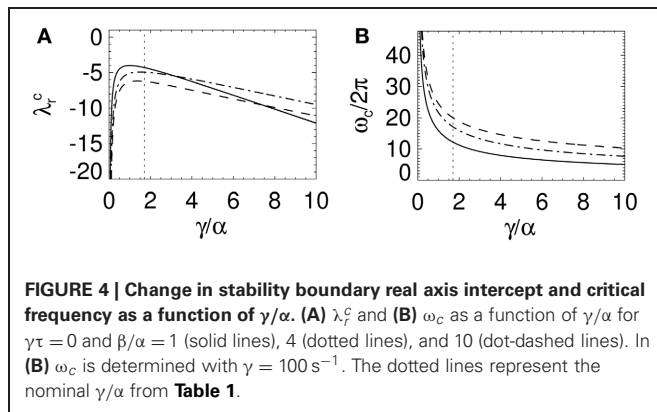
$$\lambda_r \approx 1 - \gamma\omega^2/\beta, \quad (34)$$

and

$$\lambda_i \approx -\gamma\omega/\beta. \quad (35)$$

When $\lambda_r = 0$, Equation (34) implies $\omega \approx \sqrt{\beta/\gamma}$ and $\lambda_i \approx -\sqrt{\gamma/\beta}$. Therefore as $\gamma/\beta \rightarrow \infty$, the imaginary axis intercepts converge to $\pm\infty$ and the stability zone expands to cover the entire region of the complex plane defined by $\text{Re}\lambda < 1$. Note that in this case, even though the eigenvalues can lie anywhere to the left of $\text{Re}\lambda = 1$, the dispersion solutions have an angular frequency $\omega < \gamma$. These results explain the intersection of the stability boundary with the parabolic boundary in **Figure 3C**.

With **Figure 3C** these convergence results suggest that as γ/α increases from 0, the stability zone contracts toward the unit circle, and then expands again. **Figure 4** shows the values of λ_r^c and ω_c as a function of γ/α and β/α . As γ/α increases from 0, λ_r^c rapidly increases from $-\infty$ to a maximum value and then slowly decreases back to $-\infty$; this decrease is greatest for the $\beta/\alpha = 1$ curve which intersects the other two curves in **Figure 4A**. The maximum turning point for λ_r^c occurs when $\gamma \approx \alpha$ for each β/α , with maximum λ_r^c decreasing as β/α increases. When $\beta/\alpha = 4$ (which is the nominal value in **Table 1**), λ_r^c is approximately constant for $1 \lesssim \gamma/\alpha \lesssim 3$, with a maximum at $\gamma/\alpha \approx 1.7$, the nominal value. The corresponding ω_c curves in **Figure 4B** all show similar monotonic decreases as γ/α increases. The curves do not intersect or have a turning point as in **Figure 4A**. This decrease



from ∞ is initially very rapid, before gradually decreasing to 0 as $\gamma/\alpha \rightarrow \infty$. This change approximately occurs at the nominal γ/α value in **Table 1**, where $\omega_c/2\pi \lesssim 20 \text{ Hz}$ for each β/α . Note that when β/α is a fixed constant, Equation (33) shows that $\omega_c \rightarrow 0$. **Figure 4B** also shows that increasing γ/α and decreasing γ/β results in a decreased value for ω_c .

In this section we have shown that physiologically realistic dendritic time constants have an effect on network stability similar to that of propagation time delays restricting, the critical frequency and the stability zone to a teardrop-shaped zone in the complex plane. However, unlike τ , for particular values of γ/α and γ/β the stability zone can expand to enclose an area outside the parabolic region described by Equation (18).

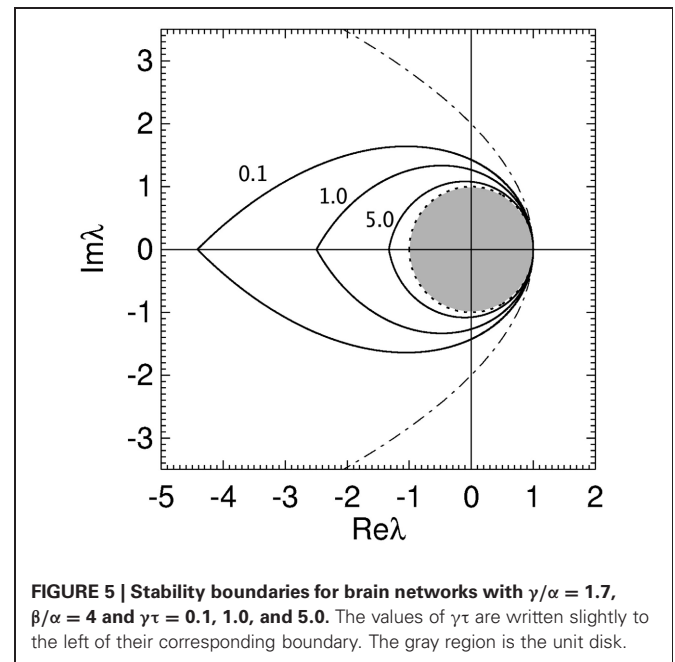
Stability of networks with time delays and non-zero dendritic time constants

The effect of having both time delays and non-zero dendritic time constants on stability is now described. In this case each of the parameters γ , $1/\alpha$, $1/\beta$, and τ are non-zero and \mathcal{D} is given by Equation (11). As in the previous section, the stability boundary for these networks is defined by $\lambda_r = \text{Re}\mathcal{D}(\varpi)$ and $\lambda_i = \text{Im}\mathcal{D}(\varpi)$ for $-\infty < \varpi < \infty$. Analysis of these equations gives the properties of the stability boundary and zone. However, the effects on stability of having equal time delays and dendritic time constants are easily understood qualitatively as a combination of our previous results. Beginning with a network that has $\gamma/\alpha = \gamma/\beta = \gamma\tau = 0$ and a parabolic stability zone, given by Equation (18), adding a time delay τ contracts the stability zone toward the unit circle, by “pulling in” the parabolic boundary, forming a teardrop-shaped zone within the original parabolic region. Adding dendritic rise and decay constants then, depending on their value, expands or contracts this stability zone. In all cases the stability zone is defined by

$$|\text{Im}\lambda| < |\text{Im}[\mathcal{D}(\varpi)]|, \quad (36)$$

for $0 \leq \varpi \leq \varpi_c$. Note that for large $\gamma\tau$, γ/α , and γ/β the contraction caused by $\gamma\tau$ is greater than the expansion effects due to γ/α and γ/β because of the exponential in Equation (11).

In **Figure 5** the stability zone of a brain network with plausible time delays and dendritic time constants (from **Table 1**) is shown. This shows that for realistic parameter values brain networks have



a teardrop-shaped stability zone completely within the parabolic zone (Equation 18). For large realistic τ the stability zone is only slightly larger than the unit disc. The ϖ_c in **Figure 5** for increasing $\gamma\tau$ are 1.02, 0.72, and 0.35, respectively. The corresponding critical frequencies $\omega_c/2\pi$ for $\gamma = 100$ are then 16, 11.3, and 5.5 Hz, respectively.

STABILITY AND DISPERSION SOLUTIONS OF RANDOMLY CONNECTED NETWORKS

We now investigate the stability and dispersion solutions of the randomly connected structural brain networks defined in the “Methods” section. For these networks we fix the model parameters γ , α , β , and τ to their nominal values in **Table 1**. The stability and dispersion solutions of this networks are determined from $\text{Sp}(\mathbf{G})$. If all the eigenvalues λ in $\text{Sp}(\mathbf{G})$ satisfy Equation (25) then the network is stable. The corresponding dispersion solutions are obtained numerically by solving Equation (10) for each λ using CROOT (Botten et al., 1983) (as described in “Methods”).

The spectrum of a RENs consists of one eigenvalue at $n\mu_e$ with the other $n - 1$ eigenvalues uniformly distributed in a disc of radius $\mu_e\sqrt{np(1-p)} < n\mu_e$. The spectrum of RCNs and RPNs with maximum μ_e and μ_i allowed by stability is distributed within the unit disk, see (Gray and Robinson, 2009a), with multiple eigenvalues near the stability boundary. Therefore, stability constrains the spectrum of random brain networks to the unit disk and the stability of random brain networks is independent of the γ , α , β , and τ . However, the frequencies of the dispersion solutions do depend on the model parameters.

In **Figure 6** the spectrum and dispersion solutions for a REN, RCN, and RPN with the nominal model parameters are shown. The parameters of each network are set so that the networks are marginally stable and μ_e and $|\mu_i|$ are as large as possible while maintaining stability. Note the larger values of μ_e and $|\mu_i|$ for the

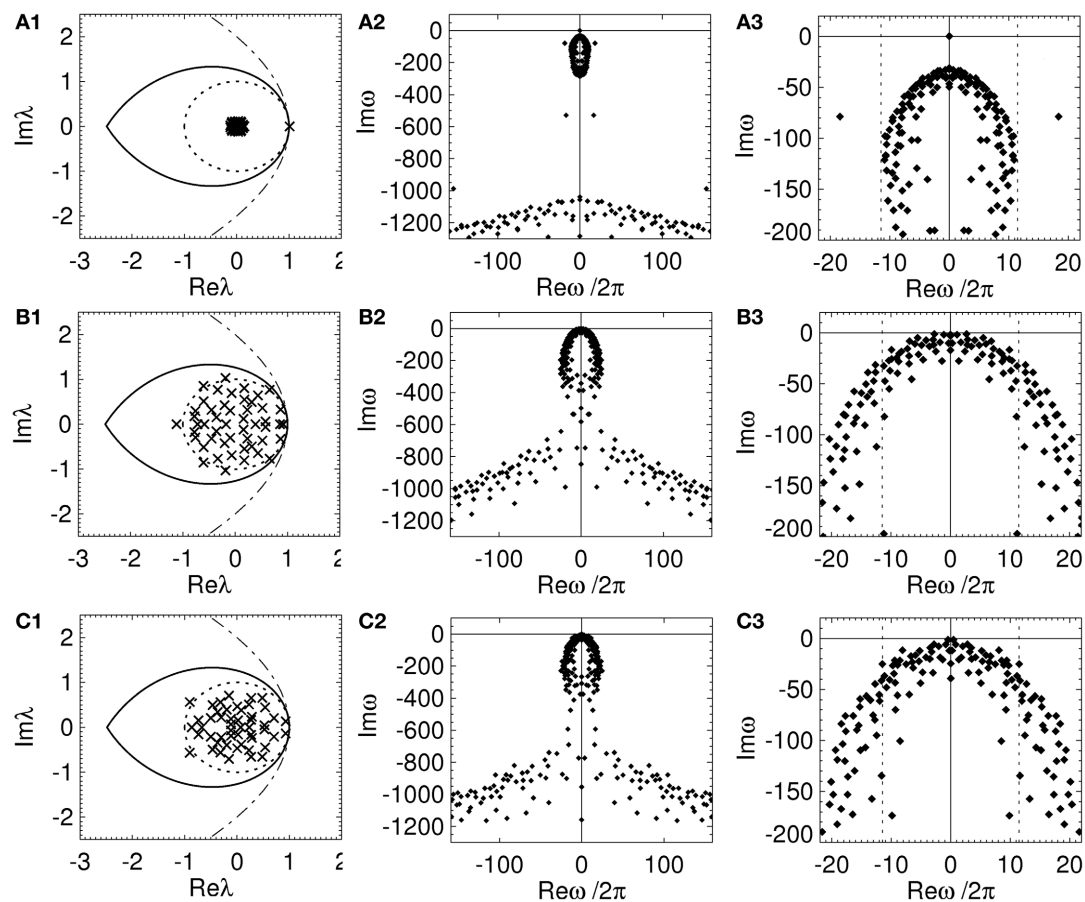


FIGURE 6 | Spectrum and dispersion solutions for a REN, a RCN, and a RPN. Each network has $n = 50$, $p = 0.5$, and the nominal parameters in **Table 1**, other parameters are set so that the network is close to marginal stability. The left column is the spectrum (crosses), the middle column shows the dispersion solutions

(diamonds), and the right column shows an expanded view of the dispersion solutions near the imaginary axis with dotted lines representing the critical frequencies $\pm \omega_c/2\pi$. **(A)** REN with $\mu_e = 0.04$, **(B)** RCN with $p_i = 0.5$, $\mu_e = -\mu_i = 0.2$, $\sigma_e = \sigma_i = 0$, and **(C)** RPN with $p_i = 0.5$, $\mu_e = -\mu_i = 0.22$, $\sigma_e = \sigma_i = 0$.

RPN, compared to the RCN. This highlights that RPNs can have larger μ_e and $|\mu_i|$, and hence be more responsive, before becoming almost certainly unstable as shown in (Gray and Robinson, 2009a). This suggests stability may have an effect on the arrangement of inhibitory and excitatory neurons and their physiology in structural brain networks.

Each network has an infinite number of dispersion solutions because $\tau \neq 0$. The second column of **Figure 6** shows the dispersion spectrum is symmetrically placed around the real axis. Each network has a qualitatively similar dispersion spectrum, with a finite cluster of solutions near the origin and a broad “arrowhead” of solutions for $\text{Im}\omega \lesssim 800 \text{ s}^{-1}$; this arrowhead has an infinite number of solutions with $\text{Im}\omega$ decreasing as $|\text{Re}\omega|$ increases. The third column shows the dispersion solutions near the imaginary axis. In **Figure 6A3** one solution, $\omega = \omega_1$, lies on the origin, separate from the other dispersion solutions. This solution corresponds to the eigenvalue at $\lambda \approx 1$ that is separated from the rest of the spectrum in **Figure 6A1**. This implies the dynamics of an REN will be dominated by a zero-frequency mode. Since the rest of the dispersion solutions

have $\text{Im}\omega \ll 0$ all other modes rapidly decay to zero amplitude. However, the RCN and RPN in **Figure 6** have a very similar dispersion spectrum with multiple dispersion solutions near the imaginary axis. The solutions closest to the imaginary axis have small frequencies $\lesssim 5 \text{ Hz}$. This shows that the presence of inhibitory connections allows random networks to have multiple marginally stable low frequency modes (Gray and Robinson, 2008, 2009a,b).

DISCUSSION

We increased the physiological realism of a structural brain network model we studied previously in (Gray and Robinson, 2006, 2008, 2009a,b; Robinson et al., 2009) by allowing the network to have equal time delays τ for propagation between neuronal populations and non-zero dendritic rise $1/\beta$ and decay $1/\alpha$ time constants. Under these assumptions the stability of an arbitrarily connected network of neural populations is determined by the network’s gain matrix. The addition of time delays changed the stability zone in the complex plane from a parabolic region to a teardrop-shaped zone, dependent on α , β , τ , and the temporal

damping rate γ . Our results are similar to previous work investigating the effect of time delays on stability of electrical activity within spatially continuous networks of neural tissue (Marcus and Westervelt, 1989; Jirsa and Ding, 2004; Feng et al., 2006; Qubbaj and Jirsa, 2007, 2009; Jirsa, 2009). This previous work has generally used integro-differential neural field equations with connectivity within a neural mass described by homogeneous or heterogeneous kernels. While in principle this work could be applied to large-scale connection topologies of discrete neural masses, as we have investigated here, this has generally not been done as it is difficult to incorporate arbitrary connectivity patterns (Qubbaj and Jirsa, 2009). In this work we have investigated the temporal dynamics of the overall electrical activity of arbitrarily connected structural brain networks, ignoring the spatial spread and propagation of electrical activity within individual neuronal populations.

In terms of stability the effect of non-zero dendritic time constants is similar to a time delay, further suggesting that dendrites act as a low-pass filter on synaptic inputs (Robinson et al., 1997, 2001a,b; Rennie et al., 2002). However, dendritic time constants can change the shape of the stability zone even allowing it to expand and enclose areas outside the parabolic stability region for networks with zero time delays and instantaneous dendritic rise and decay times. For all values of γ , α , β , and τ the stability zone contains the unit disk. This result implies that the stability criteria originally derived by May (1972, 1974) (that a network is stable if all its eigenvalues lie in the unit disk) is a sufficient condition for the stability of structural brain networks.

We also explored the dispersion solutions and frequencies of structural brain networks. If an initially stable brain network becomes unstable through a change in its connection gains, then at least one eigenvalue moves across the stability boundary and the network has an instability at a frequency given by the eigenvalue's corresponding dispersion solution. For networks with time delays and non-zero dendritic time constants there is a maximum frequency, the critical frequency $\omega_c/2\pi$, at which initially stable networks will become unstable. For example, if all the gain matrix eigenvalues of a networks are initially inside the stability zone but then move across the stability boundary (e.g., due to changes in connection gains) then the frequency of these instabilities will be less than the critical frequency.

Measurements of brain activity (Stam et al., 1999; Robinson et al., 2001b; Breakspear, 2002; Breakspear et al., 2003) suggest the brain operates near marginal stability allowing the brain to have rich dynamics and a wide range of complex behavior. A network near marginal stability has eigenvalues near the stability boundary with corresponding modes that are the slowest to decay back to the steady state dominating the network's dynamics. These modes have a frequency less than the critical frequency. Using physiologically plausible parameter values in our structural network model (see **Table 1**) we would expect the electrical dynamics to be dominated by frequencies $\lesssim 100$ Hz. When the nominal parameter values of α , β , γ , and τ are used the critical frequency is approximately 10 Hz and decreases as $\gamma\tau$ increases (see **Figures 4, 6**).

For the randomly connected structural brain networks we investigated previously (Gray and Robinson, 2006, 2008, 2009a,b)

the spectrum of the gain matrix is almost certainly contained in a disk centered on the origin with a radius dependent on the network's architecture and the average values of its excitatory and inhibitory gains. Thus, the stability zone of these networks is the unit disk and their stability is independent of time delays and dendritic time constants. Therefore, the results of that work remain valid in the more general case studied here. However, for the critical frequency is dependent on dendritic time constants, temporal damping rate, and time delays. We showed marginally stable randomly connected networks with inhibitory connections have multiple marginally stable low frequency dispersion solutions.

The primary goal of this and our previous work on structural brain networks is to understand how stability potential constrains the structure physiology of networks. For the randomly connected networks studied in (Gray and Robinson, 2006, 2008, 2009a,b), we have shown that time delays and non-zero dendritic time constants have minimal effect on their stability. One network type whose stability could be affected by these physiological properties is networks with inhibitory self-connections. The spectrum of these networks, even if they are randomly connected, is no longer restricted to a disc by stability but can have eigenvalues distributed within the teardrop-shaped region. Such networks could have marginally stable modes with frequencies (up to the critical frequency) in the alpha, beta, and gamma ranges. This will be explored in future work.

LIKELY EFFECTS OF DISTRIBUTED OR VARYING PARAMETERS ON STABILITY

Assuming structural brain networks have equal γ , α , and β for each neural population and equal τ for each connection is unrealistic. Different neuronal populations in the brain have different parameter values; for example, excitatory cortical neurons have $\gamma \approx 100 \text{ s}^{-1}$ while for inhibitory cortical neurons $\gamma \approx 10^3\text{--}10^4 \text{ s}^{-1}$ (Robinson et al., 2004). Also, the time delay in real cortical networks is expected to vary from $\tau = 0$ for self-connections to a large value for areas physically far apart. A realistic model of a structural brain network would therefore allow the model parameters γ , α , β , and τ to vary across neural populations. This variation could possibly be represented as a distribution.

The effect of distributed time delays on network stability has been studied using general models for network activity (Yi and Tan, 2002; Atay, 2003; Jirsa and Ding, 2004; Feng et al., 2006). In (Feng et al., 2006) and (Jirsa and Ding, 2004) networks with a distributed time delay with mean τ were shown to have a stability zone that contained the stability zone of networks with a constant delay equal to τ . These results are applicable to the brain networks studied here, since our model without dendritic time constants can be described as a specific case of the model studied in (Jirsa and Ding, 2004; Feng et al., 2006). This implies brain networks with distributed delay are more stable than networks with equal time delays; in the sense that a stable network with a distribution of delays could be unstable if its delays were replaced with a constant delay equal to the distribution mean. Hence, the equal time delay case is the least stable case and yields a bound on the stability of a structural brain network.

Here we have shown that plausible dendritic time constants have similar effects on stability as a time delay. This suggests that similar results to those found in (Jirsa and Ding, 2004; Feng et al., 2006) will likely be observed for distributed α and β . Also, we have shown that γ only affects the dispersion frequencies of a network, not its stability. A distributed γ is therefore expected to have no effect on the stability, only on the dispersion frequencies but this needs to be confirmed numerically.

Given the previous results investigating distributed parameter values we argue our results are still informative. But the exact effect of distributed γ , α , and β on structural brain network dynamics needs to be determined, particularly to understand the critical frequency and the dynamics of marginally stable modes in a network's electrical activity. To fully understand the stability and dynamics of structural brain networks with varying time delays and dendritic time constants requires a numerical approach. This will be investigated using CROOT (Botten et al., 1983) in future work.

REFERENCES

- Acebrón, J. A., Bonilla, L. L., Vicente, C. J. P., Ritort, F., and Spigler, R. (2005). The kuramoto model: a simple paradigm for synchronization phenomena. *Rev. Mod. Phys.* 77, 137–185.
- Allesina, S., and Tang, S. (2012). Stability criteria for complex ecosystems. *Nature* 483, 205–208.
- Atay, F. M. (2003). Distributed delays facilitate amplitude death of coupled oscillators. *Phys. Rev. Lett.* 91:094101. doi: 10.1103/PhysRevLett.91.094101
- Atay, F. M., and Hutt, A. (2004). Stability and bifurcations in neural fields with finite propagation speed and general connectivity. *SIAM J. Appl. Math.* 65, 644–666.
- Attwell, D., and Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *J. Cereb. Blood Flow Metab.* 21, 1133–1145.
- Bassett, D. S., and Bullmore, E. (2006). Small-world brain networks. *Neuroscientist* 12, 1–12.
- Bechhoefer, J. (2005). Feedback for physicists: a tutorial essay on control. *Rev. Mod. Phys.* 77, 783–836.
- Bollobás, B. (1985). *Random Graphs*. London: Academic Press.
- Botten, L. C., Craig, M. S., and McPhedran, R. C. (1983). Complex zeros of analytic functions. *Comput. Phys. Commun.* 29, 245–259.
- Breakspear, M. (2002). Nonlinear phase desynchronization in human electroencephalographic data. *Hum. Brain Mapp.* 15, 175–198.
- Breakspear, M., Roberts, J. A., Terry, J. R., Rodrigues, S., Mahant, N., and Robinson, P. A. (2006). A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cereb. Cortex* 16, 1296–1313.
- Breakspear, M., Terry, J. R., and Friston, K. J. (2003). Modulation of excitatory synaptic coupling facilitates synchronization and complex dynamics in a nonlinear model of neuronal dynamics. *Neurocomputing* 52–54, 151–158.
- Bressloff, P. (2011). Spatiotemporal dynamics of continuum neural fields. *J. Phys. A Math. Theor.* 45:033001. doi: 10.1088/1751-8113/45/3/033001
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198.
- Bullmore, E., and Sporns, O. (2012). The economy of brain network organization. *Nat. Rev. Neurosci.* 13, 336–349.
- Coombes, S. (2005). Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* 93, 91–108.
- Coombes, S., Venkov, N. A., Shiau, L., Bojak, I., Liley, D. T., and Laing, C. R. (2007). Modeling electrocortical activity through improved local approximations of integral neural field equations. *Phys. Rev. E* 76:051901. doi: 10.1103/PhysRevE.76.051901
- Cvetković, D., Doob, M., and Scahs, H. (1995). *Spectra of Graphs*. 3rd Edn. Heidelberg: Johann Ambrosius Barth Verlag.
- da Silva, F. H. L., Hoeks, A., Smits, A., and Zetterberg, L. H. (1974). Model of brain rhythmic activity: the alpha-rhythm of the thalamus. *Kybernetik* 15, 27–37.
- Deco, G., Jirsa, V., Robinson, P., Breakspear, M., and Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* 4:e1000092. doi: 10.1371/journal.pcbi.1000092
- Felleman, D. J., and van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47.
- Feng, J., Jirsa, V. K., and Ding, M. (2006). Synchronization in networks with random interactions: theory and applications. *Chaos* 16:015109. doi: 10.1063/1.2180690
- Freeman, W. J. (1975). *Mass Action in the Nervous System: Examination of the Neurophysiological Basis of Adaptive Behavior Through the EEG*. New York, NY: Academic Press.
- Gray, R., and Robinson, P. (2006). Stability and spectra of randomly connected excitatory cortical networks. *Neurocomputing* 70, 1000–1012.
- Gray, R., and Robinson, P. (2008). Stability and synchronization of random brain networks with a distribution of connection strengths. *Neurocomputing* 71, 1373–1387.
- Gray, R. T., and Robinson, P. A. (2009a). Stability and structural constraints of random brain networks with excitatory and inhibitory neural populations. *J. Comput. Neurosci.* 27, 81–101.
- Gray, R. T., and Robinson, P. A. (2009b). Stability of random brain networks with excitatory and inhibitory connections. *Neurocomputing* 72, 1849–1858.
- Hilgetag, C.-C., Burns, G. A. P. C., O'Neill, M. A., Scannell, J. W., and Young, M. P. (2000a). Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat. *Philos. Trans. R. Soc. Lond. B* 355, 91–110.
- Hilgetag, C.-C., O'Neill, M. A., and Young, M. P. (2000b). Hierarchical organization of macaque and cat cortical sensory systems explored with a novel network processor. *Philos. Trans. R. Soc. Lond. B* 355, 71–89.
- Hogg, T., Huberman, B. A., and McGlade, J. M. (1989). The stability of ecosystems. *Proc. R. Soc. B* 237, 43–51.
- Horn, R. A., and Johnson, C. R. (1985). *Matrix Analysis*. New York, NY: Cambridge University Press.
- Jirsa, V. (2009). Neural field dynamics with local and global connectivity and time delay. *Philos. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci.* 367, 1131–1143.
- Jirsa, V. K., and Ding, M. (2004). Will a large complex system with time delays be stable? *Phys. Rev. Lett.* 93:070602, 1–4. doi: 10.1103/PhysRevLett.93.070602
- Jirsa, V. K., and Haken, H. (1996). Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* 77, 960–963.
- Jouve, B., Rosenstiehl, P., and Imbert, M. (1998). A mathematical approach to the connectivity between the cortical visual areas of the macaque monkey. *Cereb. Cortex* 8, 28–39.
- Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (2000). *Principles of Neural Science*. 4th Edn. New York, NY: McGraw-Hill.
- Kim, J. W., and Robinson, P. A. (2007). Compact dynamical model of brain

CONCLUSIONS

We investigated the stability of discrete networks of neuronal populations using a simplified physiologically-based mean-field model of brain electrical activity. Incorporating time delays and non-zero dendritic time constants affects the stability of arbitrarily connected structural brain networks by constraining the eigenvalues of the gain matrix to a teardrop-shaped region in the complex plane. The stability of randomly connected networks of excitatory and inhibitory neuronal populations is unaffected by time delays and dendritic time constants; as stability constrains the gain matrix eigenvalues to the unit circle. However, the dispersion frequencies of instabilities are affected by network physiology. Randomly connected brain networks with the largest average excitatory and inhibitory gains allowed by stability can have multiple marginally stable low-frequency modes. Such networks would be highly responsive and adaptable to external stimuli while remaining stable, and have a wide range of flexible, adaptable, and complex behavior.

- activity. *Phys. Rev. E* 75:031907. doi: 10.1103/PhysRevE.75.031907
- Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Englewood: Roberts and Company.
- Kuramoto, Y. (1984). *Chemical Oscillations, Waves, and Turbulence*. New York, NY: Dover.
- Laughlin, S. B., de Ruyter van Stevenick, R. R., and Anderson, J. C. (1998). The metabolic cost of neural information. *Nat. Neurosci.* 1, 36–41.
- Lennie, P. (2003). The cost of cortical computation. *Cogn. Brain Res.* 13, 493–497.
- Marcus, C. M., and Westervelt, R. M. (1989). Stability of analog neural networks with delay. *Phys. Rev. A* 39, 347–359.
- May, R. M. (1972). Will a large complex system be stable? *Nature* 238, 413–414.
- May, R. M. (1974). *Stability and Complexity in Model Ecosystems*. 2nd Edn. Princeton, NJ: Princeton University Press.
- McCann, K. S. (2000). The diversity-stability debate. *Nature* 405, 228–233.
- Mitrinović, D. S., and Kečkić, J. D. (1984). *The Cauchy Theory of Residues: Theory and Applications*. Dordrecht: D. Reidel Publishing Company.
- Murray, J. D. (2002). *Mathematical Biology*. 3rd Edn. New York, NY: Springer.
- Nunez, P. L. (1974). Wave-like properties of the alpha rhythm. *IEEE Trans. Biomed. Eng.* 21, 473–482.
- Nunez, P. L. (1995). *Neocortical Dynamics and Human EEG Rhythms*. New York, NY: Oxford University Press.
- Pecora, L. M., and Carroll, T. L. (1998). Master stability functions for synchronized coupled systems. *Phys. Rev. Lett.* 80, 2109–2112.
- Qubbaj, M., and Jirsa, V. (2007). Neural field dynamics with heterogeneous connection topology. *Phys. Rev. Lett.* 98:238102. doi: 10.1103/PhysRevLett.98.238102
- Qubbaj, M., and Jirsa, V. (2009). Neural field dynamics under variation of local and global connectivity and finite transmission speed. *Physica D* 238, 2331–2346.
- Rennie, C. J., Robinson, P. A., and Wright, J. J. (2002). Unified neurophysical model of EEG spectra and evoked potentials. *Biol. Cybern.* 86, 457–471.
- Robinson, P. A. (2003). Neurophysical theory of coherence and correlations of electroencephalographic and electrocorticographic signals. *J. Theor. Biol.* 222, 163–175.
- Robinson, P. A. (2005). Propagator theory of brain dynamics. *Phys. Rev. E* 72:0119404. doi: 10.1103/PhysRevE.72.0119404
- Robinson, P. A., Henderson, J. A., Matar, E., Riley, P., and Gray, R. T. (2009). Dynamical reconnection and stability constraints on cortical network architecture. *Phys. Rev. Lett.* 103, 1–4.
- Robinson, P. A., Loxley, P. N., O'Connor, S. C., and Rennie, C. J. (2001a). Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. *Phys. Rev. E* 63:041909. doi: 10.1103/PhysRevE.63.041909
- Robinson, P. A., Rennie, C. J., Wright, J. J., Bahramali, H., Gordon, E., and Rowe, D. L. (2001b). Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E* 63:021903. doi: 10.1103/PhysRevE.63.021903
- Robinson, P. A., Rennie, C. J., and Rowe, D. L. (2002). Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Phys. Rev. E* 65:041924. doi: 10.1103/PhysRevE.65.041924
- Robinson, P. A., Rennie, C. J., Rowe, D. L., and O'Connor, S. C. (2004). Estimation of multiscale neurophysiologic parameters by electroencephalographic means. *Hum. Brain Mapp.* 23, 53–72.
- Robinson, P. A., Rennie, C. J., Rowe, D. L., O'Connor, S. C., Wright, J. J., Gordon, E., et al. (2003a). Neurophysical modeling of brain dynamics. *Neuropsychopharmacology* 28, S74–S79.
- Robinson, P. A., Whitehouse, R. W., and Rennie, C. J. (2003b). Nonuniform corticothalamic continuum model of electroencephalographic spectra with application to split-alpha peaks. *Phys. Rev. E* 68:021922. doi: 10.1103/PhysRevE.68.021922
- Robinson, P. A., Rennie, C. J., and Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* 56, 826–840.
- Robinson, P. A., Rennie, C. J., Wright, J. J., and Bourke, P. D. (1998). Steady states and global dynamics of electrical activity in the cerebral cortex. *Phys. Rev. E* 58, 3557–3571.
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069.
- Scannell, J. W., Blakemore, C., and Young, M. P. (1995). Analysis of connectivity in the cat cerebral cortex. *Int. J. Neural Syst.* 15, 1463–1483.
- Sporns, O. (2011). The human connectome: a complex network. *Ann. N.Y. Acad. Sci.* 1224, 109–125.
- Sporns, O., Chialvo, D. R., Kaiser, M., and Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *TRENDS Neurosci.* 8, 418–425.
- Sporns, O., Tononi, G., and Edelman, G. M. (2000). Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cereb. Cortex* 10, 127–141.
- Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Comput. Biol.* 1:e42. doi: 10.1371/journal.pcbi.0010042
- Sporns, O., and Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics* 2, 145–161.
- Stam, C., Pijn, J. P. M., Suffczynski, P., and da Silva, F. H. L. (1999). Dynamics of the human alpha rhythm: evidence for non-linearity? *Clin. Neurophysiol.* 110, 1801–1813.
- Steriade, M., Gloor, P., Linas, R. P., da Silva, F. H. L., and Mesulam, M. M. (1990). Basic mechanisms of cerebral rhythmic activities. *Electroencephalogr. Clin. Neurophysiol.* 76, 481–508.
- Steuer, R. (2007). Computational approaches to the topology, stability and dynamics of metabolic networks. *Phytochemistry* 68, 2139–2151.
- Taverna, D. M., and Goldstein, R. A. (2002). Why are proteins marginally stable? *Proteins* 46, 105–109.
- Turing, A. M. (1952). The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. B* 237, 37–72.
- van den Heuvel, M., and Sporns, O. (2011). Rich-club organization of the human connectome. *J. Neurosci.* 31, 15775–15786.
- Venkov, N. A., Coombes, S., and Matthews, P. C. (2007). Dynamic instabilities in scalar neural field equations with space-dependent delays. *Physica D* 232, 1–15.
- Wen, Q., and Chklovskii, D. B. (2005). Segregation of the brain into gray and white matter: a design minimizing conduction delays. *PLoS Comput. Biol.* 1:e78. doi: 10.1371/journal.pcbi.0010078
- Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80.
- Wright, J. J., and Liley, D. T. J. (1996). Dynamics of the brain at global and microscopic scales: neural networks and the EEG. *Behav. Brain Sci.* 19, 285–320.
- Wright, J. J., Robinson, P. A., Rennie, C. J., Gordon, E., Bourke, P. D., Chapman, C. L., et al. (2001). Towards an integrated continuum model of cerebral dynamics: the cerebral rhythms, synchronous oscillation and cortical stability. *Biosystems* 63, 71–88.
- Yi, Z., and Tan, K. K. (2002). Dynamic stability conditions for Lotka-Volterra recurrent neural networks with delay. *Phys. Rev. E* 66:011910, 1–8. doi: 10.1103/PhysRevE.66.011910
- Young, M. P. (2000). The architecture of visual cortex and inferential processes in vision. *Spat. Vis.* 13, 137–146.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 October 2012; accepted: 24 March 2013; published online: 12 April 2013.

Citation: Gray RT and Robinson PA (2013) Stability constraints on large-scale structural brain networks. *Front. Comput. Neurosci.* 7:31. doi: 10.3389/fncom.2013.00031

Copyright © 2013 Gray and Robinson. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Top-down influences on local networks: basic theory with experimental implications

Ramesh Srinivasan^{1,2,3*}, Samuel Thorpe³ and Paul L. Nunez⁴

¹ Department of Cognitive Sciences, University of California, Irvine, CA, USA

² Department of Biomedical Engineering, University of California, Irvine, CA, USA

³ Institute for Mathematical Behavioral Sciences, University of California, Irvine, CA, USA

⁴ Cognitive Dissonance, LLC, San Diego, CA, USA

Edited by:

Dimitris Pinotsis, University College London, UK

Reviewed by:

William C. Gaetz, The Children's

Hospital of Philadelphia, USA

Dimitris Pinotsis, University College

London, UK

*Correspondence:

Ramesh Srinivasan, Department of Cognitive Sciences, University of California, 3151 SSPA, Irvine, CA 92617, USA.

e-mail: r.srinivasan@uci.edu

The response of a population of cortical neurons to an external stimulus depends not only on the receptive field properties of the neurons, but also the level of arousal and attention or goal-oriented cognitive biases that guide information processing. These top-down effects on cortical neurons bias the output of the neurons and affect behavioral outcomes such as stimulus detection, discrimination, and response time. In any physiological study, neural dynamics are observed in a specific brain state; the background state partly determines neuronal excitability. Experimental studies in humans and animal models have also demonstrated that slow oscillations (typically in the alpha or theta bands) modulate the fast oscillations (gamma band) associated with local networks of neurons. Cross-frequency interaction is of interest as a mechanism for top-down or bottom up interactions between systems at different spatial scales. We develop a generic model of top-down influences on local networks appropriate for comparison with EEG. EEG provides excellent temporal resolution to investigate neuronal oscillations but is space-averaged on the cm scale. Thus, appropriate EEG models are developed in terms of population synaptic activity. We used the Wilson–Cowan population model to investigate fast (gamma band) oscillations generated by a local network of excitatory and inhibitory neurons. We modified the Wilson–Cowan equations to make them more physiologically realistic by explicitly incorporating background state variables into the model. We found that the population response is strongly influenced by the background state. We apply the model to reproduce the modulation of gamma rhythms by theta rhythms as has been observed in animal models and human ECoG and EEG studies. The concept of a dynamic background state presented here using the Wilson–Cowan model can be readily applied to incorporate top-down modulation in more detailed models of specific cortical systems.

Keywords: EEG, population dynamics, neural mass models, top-down control, ECoG

INTRODUCTION

A fundamental question in any neurophysiological study is whether observed modulations of neural responses in cortex by cognitive processes are the result of the action of a *local* network or due to the interactions between this local network and the rest of the brain in *global* networks. This conceptual framework of local and global networks interacting in cognitive processes is salient to the interpretation of physiological signals obtained from the brain with any technique – EEG, MEG, fMRI, LFPs, or unit activity and to models of the underlying cognitive processes. That is, even when signals are recorded from a small number of neurons (or even just one neuron) the observed dynamics result both from the intrinsic properties of the local network and from the influence of other neurons located in nearby or even distant cortex (Mountcastle, 1997). This simple distinction can be understood in terms of behavior – the response of neurons to inputs depends not only on the receptive field of the neurons but also on the level of arousal, typically by the action of neuromodulators, and attention or

goal-oriented cognitive biases that guide information processing. The latter are sometimes called top-down effects (Engel et al., 2001), which bias the output of the neurons and affect behavioral outcomes such as stimulus detection, short term memory, and reaction time.

The objective of this paper is to develop a model of local networks with which we can investigate the effect background brain state or top-down signaling on the local network. In order to develop this model, we have to make choices of spatial scale and physiological detail to incorporate into the model. Very detailed models have the potential to provide more information about specific neural systems, e.g., details models of the visual system (Lumer et al., 1997). However, detailed model parameters are not available in humans, where the competition/interaction between global and local dynamics is expected to be the most robust (Nunez, 1995, 2000; Nunez and Srinivasan, 2006). In addition, detailed models may not lead to generalized principles that can potentially guide experimental studies in a variety of behavioral contexts; detailed cellular models are not easily compared

to electrocorticogram (ECoG) and electroencephalogram (EEG) data obtained in humans.

Electroencephalogram is uniquely positioned to differentiate local and global processes and to examine their interactions in human subjects. EEG provides excellent temporal resolution allowing us to separate processes at different time scales at electrodes over cortex while allowing for sufficient spatial coverage to investigate interactions of sensory neurons with neural processes in other areas of the brain. The main limitation of EEG is spatial resolution; EEG signals are space-averaged on the cm scale (Nunez, 1981; Nunez and Srinivasan, 2006) by volume conduction through the tissues of the head. An active area of research is to improve our understanding of the structure of cortical sources and connectivity from EEG data (Pinotsis et al., 2012). ECoG in humans combines the temporal dynamics of EEG with the greater spatial detail and (depending on the patient) partial coverage of the cortex (Schalk and Leuthardt, 2011). Although ECoG is only available in limited cases in patients with intractable epilepsy, these data are a useful source of information on the nature of dynamics of localized population of neurons.

EEG signals span a frequency range of 1–50 Hz while ECoG signals span a broader frequency range of 1–150 Hz (Canolty et al., 2010; Schalk and Leuthardt, 2011). The lower portion of this spectrum (below 20 Hz) has strongly global properties with spatial distribution across the brain that depends strongly on the frequency (von Stein and Sarnthein, 2000; Nunez et al., 2001; Nunez and Srinivasan, 2006). For example human alpha rhythms, which are quite robust in alert subjects, may be recorded over nearly all of the upper scalp or cortex with a visible peak in the power spectrum near 10 Hz. Alpha rhythm power and phase synchronization (usually measured as coherence) are modulated in specific large-scale cortical networks by a wide variety of different cognitive processes including attention (Thut et al., 2006; Thorpe et al., 2012) and working memory (Sarnthein et al., 1998; Sauseng et al., 2005). Consistent with this “global” picture of low frequency EEG signals are studies using periodic visual input to elicit steady-state visual evoked potentials (SSVEPs). SSVEPs are responses to visual flicker at the flicker frequency (and harmonics). Low frequency (<20 Hz) SSVEPs elicit “resonant” responses in large-scale networks whose spatial distribution depends strongly on the input temporal frequency (Ding et al., 2006; Srinivasan et al., 2006). These large-scale networks have both distinct characteristic frequencies and functional properties (Ding et al., 2006; Bridwell and Srinivasan, 2012).

At higher frequencies (>30 Hz) the spatial distribution of EEG and ECoG signals is (apparently) localized at the cm scale. EEG studies have shown task dependent modulations of gamma networks in networks localized in sensory and motor cortex. These studies were inspired by single-unit and LFP studies in animal models, most notably by Singer and colleagues (Engel and Singer, 2001; Fries et al., 2007) that demonstrate localized networks synchronizing at gamma band frequencies. This local view of the origin of gamma rhythms is supported by ECoG studies that show relatively low coherence between electrodes at gamma band frequencies (Menon et al., 1996). SSVEP data at gamma band frequencies are consistent with this localized picture of fast EEG

rhythms – γ -SSVEPs appear to be local processes in the visual cortex (Thorpe et al., 2011).

The distinct spatial and dynamical property of EEG oscillations in low (<20 Hz) and high (>20 Hz) frequency bands suggests the need for different types of models to explain these phenomena. Given any unknown physical or biological system that produces oscillations at some preferred (or resonant) frequency $f = \omega/2\pi$, a reasonable starting point for developing a model is the origin of the implied underlying time delay τ roughly estimated as

$$\tau \sim \omega^{-1} \quad (1)$$

The implied physiological time scale for the (8–13 Hz) alpha rhythm is $\tau = 12$ –20 ms. More generally, the most robust human EEG rhythms recorded from the scalp (1–20 Hz) correspond to time delays $\tau = 8$ –160 ms. How does this delay range compare with mammalian physiology? Whereas early studies of membrane time constants in mammalian cortex were very short, typically less than 10 ms, more modern studies with improved recording methods report a wider range up to 100 ms (Koch et al., 1996). While synaptic delays (PSP rise and decay times) lie in a general range (within a factor of perhaps 5 or 10) that might account for dominant EEG frequencies, claims of close agreement between the *details* of observed EEG spectra and dynamic theories based on membrane time constants are not by themselves a critical validation of a model. Model parameters can always be chosen to “match” EEG data, which, in any case, varies widely between brain states.

Local network theories refers to models of cortical or thalamo-cortical interactions in which signal propagation delays in axons are neglected. For example, coupled non-linear oscillators interact without any transmission delay in a local theory. In contrast, models that incorporate the spatial extent of the cortex and the transmission delays between neural populations are *global* theories. Global theories predict spatially coherent oscillations over the surface of the cortex with wave-like properties that depend primarily on the transmission delays between cortical populations and the size (surface area) of the cortex (Nunez, 1981, 1995, 2000). The dominant modes of these spatially distributed oscillations are predicted to lie below 15 Hz in the theta and alpha bands. While both global and local network theories have been developed independently, their interaction across spatial and temporal scales is less well understood. Previous studies have focused on how local networks influence global networks (Jirsa and Haken, 1996; Nunez, 2000), and a recent study investigates the interaction between local connectivity and long-range interactions (Pinotsis et al., 2013). In this paper we consider how global network dynamics may influence local networks.

The underlying time scales in local network theories are typically postsynaptic potential rise and decay times due to membrane capacitive-resistive properties (Wilson and Cowan, 1972, 1973). Local theories typically predict EEG signals with frequencies above 20 Hz. These results are consistent with more detailed studies of spiking neuron models (Izhikevich, 2006; Izhikevich and Edelman, 2008) that predict fast frequency oscillations in cortical populations unless coupled with delays as in a global network. Physiologically realistic compartment models incorporating

the interactions between excitatory and inhibitory populations in cortex give rise to fast oscillations at gamma band frequencies (Bush and Sejnowski, 1996; Traub et al., 1997; Whittington et al., 2000). In these types of model, the dynamics are determined primarily by the synaptic rise and decay times and the strength of excitatory and inhibitory synaptic connections. More specific local models in sensory systems incorporate the essential spiking dynamics and connectivity of thalamocortical networks (Lumer et al., 1997) also giving rise to gamma band oscillations. While the physiological detailed models are useful to compare to data in animal models, comparisons to EEG and ECoG require model development in macroscopic variables that describe synaptic mass action.

The Wilson–Cowan model is one of the earliest and most often cited dynamic models based on local (PSP rise and decay) delays (Wilson and Cowan, 1972, 1973). The Wilson–Cowan model produces either sustained (limit cycle) or damped oscillations over a broad range of physiologically realistic parameter space in response to a step function input to the excitatory population. The oscillations in all parts of the network are highly correlated, as there is no independent noise in each population. The rate of damping of the oscillations is largely determined by the ratio of excitatory to inhibitory weights with higher inhibition leading to damped oscillations. The frequency of the oscillation is determined primarily by the membrane time constants and connectivity strength.

In this paper, we will make use of the Wilson–Cowan model to investigate how properties of high frequency (gamma band) oscillations generated by a local network in response to input is influenced by modulation of the background state by top-down influences. Our objective here is to formalize the general principles by which local networks in cortex are influenced by modulatory signals. For this purpose, we have modified the Wilson–Cowan equations to make more physiologically realistic by incorporating background state parameters into the model. In any physiological study neural dynamics are observed in a specific brain state (e.g., asleep, awake, alert, attentive, etc.) determined partly by neuromodulatory action at much longer time scales. As brain state changes, the background state partly determines the excitability of the network (Fellous and Linster, 1998; Romei et al., 2008). Experimental studies in humans and animal models have also demonstrated that top-down influences in cognitive processes involve the action of slower oscillations typically in the alpha or theta bands which appear to reflect the coherent behavior of global networks distributed across the cortex. We believe that the mostly likely underlying time scale for such global oscillations is transmission delays in corticocortical axons, and we have proposed a specific global model that predicts global standing waves with frequencies in the general range at the slower end of the EEG spectrum (Nunez, 1995, 2000; Nunez and Srinivasan, 2006). Our analysis here depends only on the existence of such global, low frequency oscillations as has been commonly observed for almost 100 years with scalp EEG and not any specific global field theory of EEG. Using the modified Wilson–Cowan model, we identify cross-frequency coupling as an EEG or ECoG signature of the effects of background state changes by top-down signals on local network dynamics.

MATERIALS AND METHODS

THE MODIFIED WILSON–COWAN MODEL

Wilson and Cowan (1972) derived a model neural population containing both excitatory and inhibitory neurons with dynamics described by a set of coupled, non-linear differential equations, herein labeled WC. The solution of these equations gives the proportion of cells in each subpopulation (excitatory/inhibitory) that become active per unit time. The cells comprising the population are assumed to be in close spatial proximity, with interconnections dense enough so that any two cells within it are path-connected. Furthermore, the model assumes that local interactions between neurons within the population are largely random, but that this local randomness gives rise to structure at larger spatial scales. The situation is analogous to an example taken from thermodynamics, in which a fluid with a macroscopically structured flow can be observed to be undergoing stochastic Brownian motion at the molecular level. The same framework set forth by Wilson and Cowan has been extended in a number of straightforward ways to models with more general connectivity, and an arbitrary number of spatially distinct neural populations (Campbell and Wang, 1996; Borisjuk et al., 2000). Extensions of the WC framework have been developed to model interacting thalamic (reticular formation) and cortical structures involved in the generation of spindle oscillations (7–14 Hz) in early sleep stages (Yousif and Denham, 2005). Jirsa and Haken (1997), used a WC model interacting with a global model to interpret MEG data in a syncopated tapping audio-motor task. Other model developments related to the WC model have incorporated spatially extended models with axonal delays and more detailed physiological parameters (Jirsa and Haken, 1996, 1997; Robinson et al., 1997; Liley et al., 1999).

Here we adopt a modified version of WC to make it more physiologically realistic as outlined in the Appendix. The basic dependent variables are the *fractions of excitatory and inhibitory active cells* (action potential densities) $E(t)$, $I(t)$, which can evidently exhibit high frequency jitter not treated in this analysis. Rather, the WC equations are expressed in terms of coarse grained excitatory $\langle E(t) \rangle$ and inhibitory $\langle I(t) \rangle$ action potential densities. The basic model is illustrated in **Figure 1**. We introduce the new dependent variables $X_E(t)$, $X_I(t)$, which provide perturbations about the critical (equilibrium) point (E_0, I_0) , which we have interpreted as the background brain state which is controlled by various neuromodulators or top-down signaling. Thus, we express

$$\begin{aligned}\langle E(t) \rangle &= E_0 + X_E(t) \\ \langle I(t) \rangle &= I_0 + X_I(t)\end{aligned}\quad (2)$$

Since the excitatory action potential densities are defined as fractions of the total cell populations, we require

$$\begin{aligned}0 &\leq E_0 + X_E(t) \leq 1 \\ 0 &\leq I_0 + X_I(t) \leq 1\end{aligned}\quad (3)$$

The basic WC equations then become

$$\begin{aligned}\frac{dX_E}{dt} &= -E_0 - X_E + (1 - E_0 - X_E) S_E(X_E, X_I, P) \\ A \frac{dX_I}{dt} &= -I_0 - X_I + (1 - I_0 - X_I) S_I(X_E, X_I)\end{aligned}\quad (4)$$

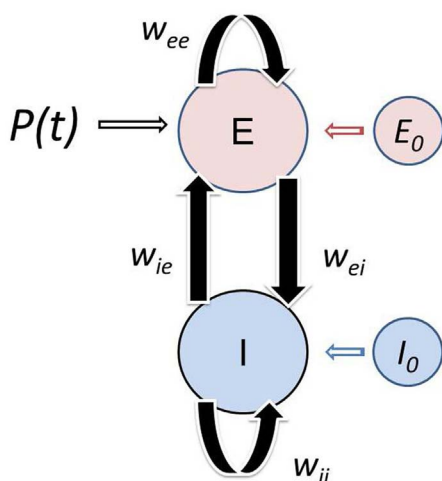


FIGURE 1 | Schematic of the modified Wilson-Cowan model. The localized population consists of excitatory and inhibitory neurons that interact with each other with negligible transmission delays. The population receives afferent input $P(t)$. Simultaneously, the population is subject to influences from both nearby and distant cortex. This top-down modulation of the neural population is the result of feedback from large-scale networks and/or global synaptic fields spanning the cortex. The model consists of an Excitatory (E) and Inhibitory (I) subpopulations with membrane time constants τ_E and τ_I , which interact with each other via the connection weights w_{EI} and w_{IE} . The neurons within each subpopulation also interact with each other, reflected in the self-excitation w_{EE} and self-inhibition w_{II} weights. The influence of other cortical areas on the population is reflected in the background state of the excitatory E_0 and inhibitory I_0 subpopulation. The Wilson-Cowan model was modified to incorporate the background state variables (see Appendix).

Here $A = \frac{\tau_I}{\tau_E}$ is the ratio of inhibitory to excitatory time constants, and $P(t)$ is an excitatory external (driving) input from another cortical population or potentially input to the population from the thalamus. The set of parameters (w_{EE} , w_{IE} , w_{EI} , w_{II}) are gain parameters that give the strength of connections between the excitatory and inhibitory populations as indicated in **Figure 1**. As shown in the Appendix for the special case $P(t) = 0$, the sigmoid functions S_E , S_I in Eq. 4 then take the forms

$$S_E = \frac{1}{1 + \left(\frac{1}{E_0} - 2\right) \exp(-w_{EE}X_E + w_{IE}X_I - P)} \quad E_0 < \frac{1}{2} \quad (5)$$

$$S_I = \frac{1}{1 + \left(\frac{1}{I_0} - 2\right) \exp(-w_{EI}X_E + w_{II}X_I)} \quad I_0 < \frac{1}{2} \quad (6)$$

PARAMETER CHOICES

For our simulations the main parameters of interest are the background state variables E_0 and I_0 , which we will vary as described in the following sections. The free parameters in our analysis are the set of connection weights (w_{EE} , w_{EI} , w_{IE} , w_{II}) which are determined by the following physiological considerations: (1) In the cortex, excitatory connections are estimated to be 4–5 times more common than inhibitory connections (Bush and Sejnowski, 1996) and (2) Inhibitory connections are more typically found on the cell body possibly increasing their effectiveness in comparison to

excitatory connections on dendritic trees (Mountcastle, 1997). Taking these two points into consideration we first fixed the two parameters $w_{EI} = 50$ and $w_{IE} = 15$. We set the self-inhibition $w_{II} = 0$, as we found little practical effect for the small values of this parameter, other than to increase damping in the system, and shift the critical point for transition from a damped oscillation to a limit cycle regime.

Equation 4 produce stable limit cycle solutions about the critical point (E_0 , I_0) for a wide range of the parameters. For example, setting $A = 1$ and $E_0 = I_0$ the necessary condition for oscillatory solutions about E_0 , I_0 is

$$w_{EE} < 2\sqrt{w_{IE}w_{EI}} \quad (7)$$

This oscillatory solution is unstable (e.g., an unstable spiral allowing for a stable limit cycle) if

$$w_{EE} > \frac{2}{E_0(1 - 2E_0)} \quad (8)$$

From Eq. 8 we were always able to find the critical value of w_{EE} below which the system produced damped oscillations in response to a step function input, while above this value the system produced limit cycle oscillations.

TOP-DOWN (GLOBAL) INFLUENCES ON A LOCAL WC NETWORK

We explicitly consider two types of top-down influences on the local WC network developed in section “The Modified Wilson-Cowan Model”: (1) the effect of neuromodulators setting the background state (E_0 , I_0) of the population. For the purpose of the analysis here we consider this effect on the background state to be static as it takes place at very long time scales as compared to the frequency of the oscillations and (2) the effect of dynamic modulation of the background state of the local network (top-down) by oscillations in larger scale networks that incorporate the cells that constitute the local network. For simplicity of analysis we presume that the larger scale networks (or global synaptic fields) generate oscillations at frequency ω_α that modulate the background state of the WC oscillator; that is

$$\begin{aligned} E_0 &\rightarrow E_0 + \alpha_E \cos(\omega_\alpha t) \\ I_0 &\rightarrow I_0 + \alpha_I \cos(\omega_\alpha t + \phi_\alpha) \end{aligned} \quad (9)$$

Here the amplitudes (α_E , α_I) of the background modulations are constrained to be less than the constant background (E_0 , I_0). We introduce a phase offset ϕ_α to allow for differences in local processing of the modulatory input by the excitatory and inhibitory subpopulations, as might occur if they have different membrane time constants.

SIMULATIONS AND DATA ANALYSIS

All of the simulations carried out here were performed using the built in ode solver in MATLAB (Natick, MA, USA), *ode23*. We considered several types of inputs $P(t)$ – step function, impulse, sinusoidal, and random noise and found the essential characteristics of the system response were represented by the step function input. The spectrum of the model output was analyzed using a

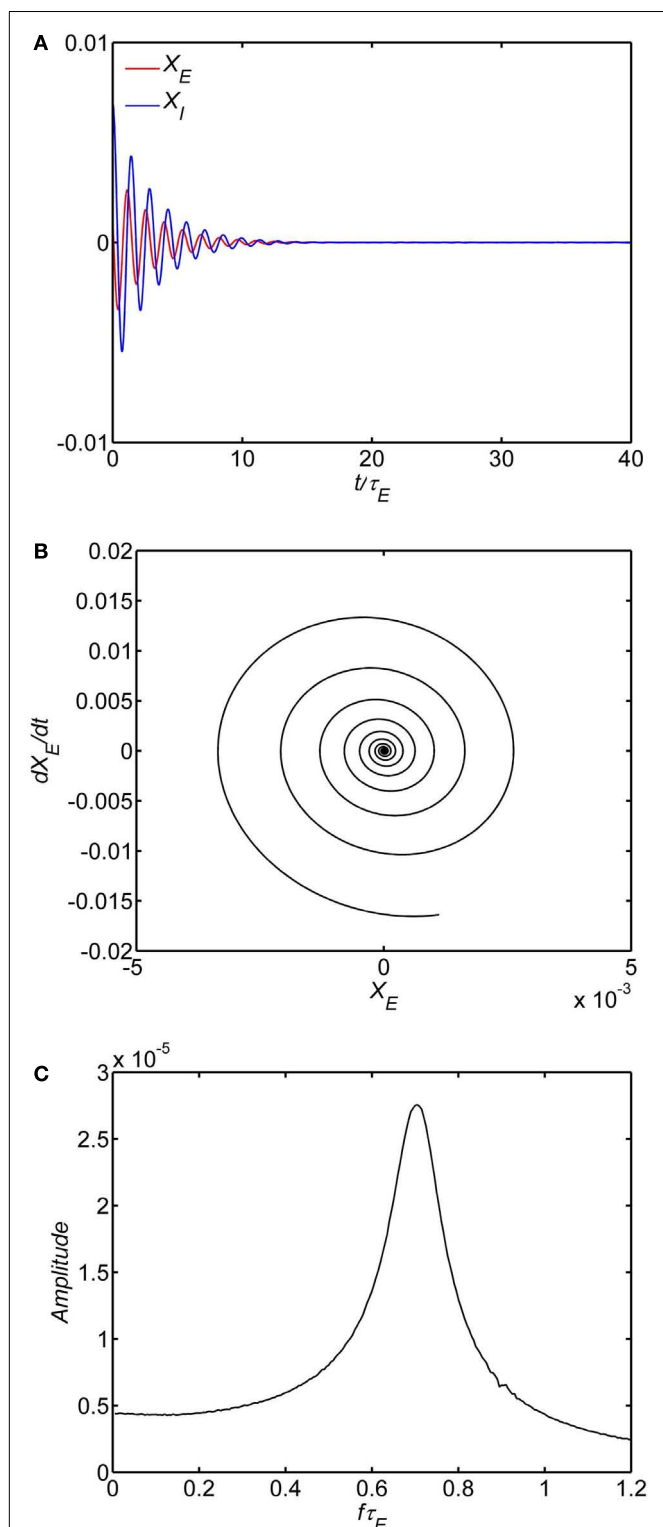


FIGURE 2 | Damped oscillation regime of the model. In all of the simulations, the following parameters are fixed: (1) The ratio of time constants $A = \tau_I/\tau_E = 1$, (2) the background state $E_0(t) = I_0(t) = 0.25$, (3) the connection weights are $(w_{EI}, w_{IE}, w_{II}) = (50, 15, 0)$ and (4) the input $P(t) = 0.1$ is a step function at time 0. Damped oscillation observed with
(Continued)

FIGURE 2 | Continued

self-excitation $w_{EE} = 12$. The time series of the excitatory and inhibitory subpopulations are shown in (A). In these plots time is normalized by excitatory membrane time constant τ_E . Phase-plane plots for the excitatory subpopulation are shown in (B). Amplitude spectra obtained by the FFT are shown in (C). Normalized frequency is $f\tau_E$. If $\tau_E = 20$ ms, a normalized frequency of 1 corresponds to 50 Hz.

FFT in MATLAB (Mathworks, Natick, MA, USA). For sustained oscillations in the limit cycle regime, we also analyzed the model outputs either by using Hilbert Transforms to estimate the frequency and amplitude of the oscillation or by a complex Morlet wavelet transform. For the damped oscillations, we fit the oscillation to a damped sinusoid $\exp(j2\pi ft(1 + j\gamma))$ where f is the frequency of the oscillation and γ is the damping coefficient. We obtained direct estimates of frequency using zero crossings and estimated the damping coefficient by fitting an exponential to the decay of amplitude across cycles of the oscillation.

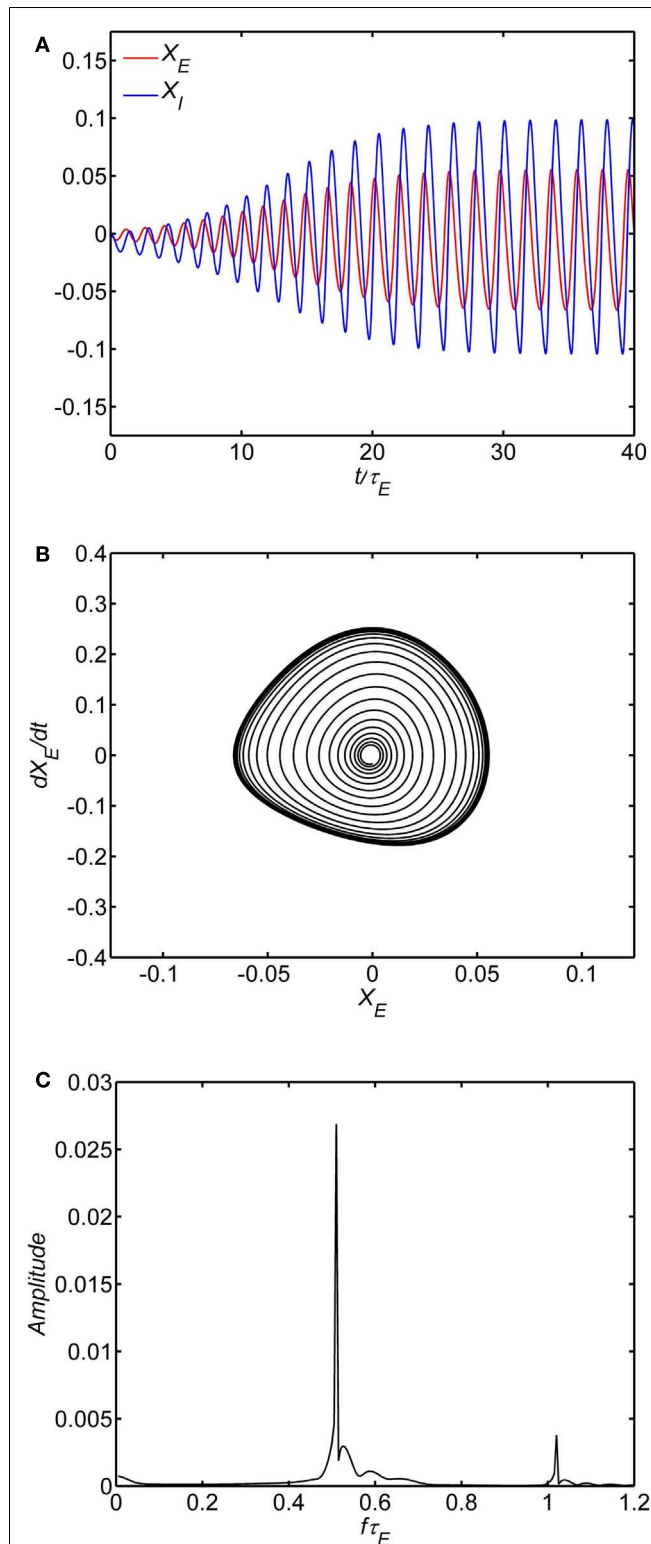
RESULTS

BASIC RESPONSE PROPERTIES OF WC OSCILLATOR

We first examined the behavior of the system with identical excitatory and inhibitory time constants ($A = \tau_I/\tau_E = 1$) and a fixed background state ($E_0 = I_0 = 0.25$). The specific value of w_{EE} separating limit cycle from damped oscillations depends on the background excitability as in Eq. 8; with $E_0 = I_0 = 0.25$ the critical value is $w_{EE} = 15$. The limit cycle is observed if the self-excitation is sufficiently large ($w_{EE} > 15$); smaller values lead to damped oscillations of higher frequency. An example of numerical solutions for the model with the self-excitation (w_{EE}) parameter in the damped oscillation range is shown in Figure 2. In the time series plot (Figure 2A), the time variable is normalized with respect to the excitatory membrane time constant τ_E and in the amplitude spectra (Figure 2C) the frequency variable f is normalized as $f\tau_E$. For example, if τ_E falls in the range of 10–20 ms range, the damped oscillation corresponds to a gamma band oscillation in the 35–70 Hz band. The limit cycle is observed if the self-excitation is sufficiently large (Figure 3). The limit cycle has a lower fundamental frequency as shown in the spectrum in Figure 3C; if τ_E falls in the range of 10–20 ms range, the dominant frequency is in the 25–50 Hz range and also exhibits harmonics (Second harmonic shown).

The ratio of inhibitory to excitatory time constants A influences both the frequency and damping of the oscillations. Figures 4A,B show the oscillation frequency and damping coefficient in the damped oscillation regime for self-excitation in the damped oscillation range ($w_{EE} = 12$). When the inhibitory time constant is smaller than the excitatory time constant ($A < 1$) the oscillations are highly damped, but if the inhibitory time is constant is larger than the excitatory time constant ($A > 1$) the oscillations are weakly damped. Thus, in order to observe the damped oscillations, it must be the case that inhibitory time constants are longer than the excitatory time constant. As the ratio A increases further the system will eventually transition to a limit cycle oscillation.

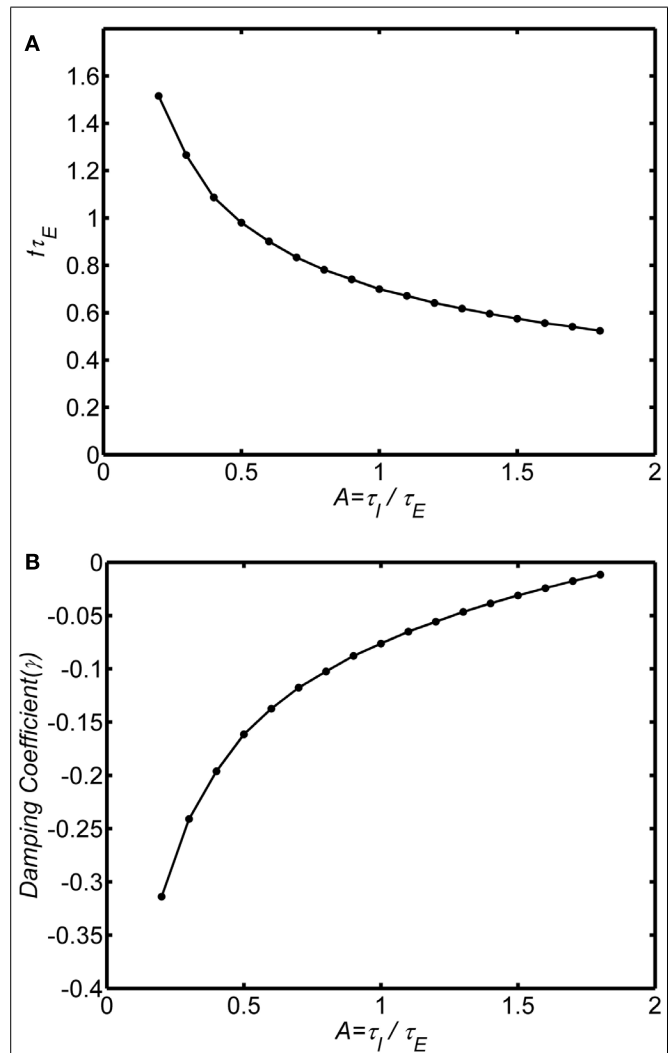
Figures 5A,B shows an example with the self-excitation parameter in the limit cycle regime ($w_{EE} = 18$). As A increases the frequency of the oscillation decreases and the amplitude increases



(Continued)

FIGURE 3 | Continued

time 0. Limit cycle oscillation observed with self-excitation $w_{EE} = 20$. The time series of the excitatory and inhibitory subpopulations are shown in (A). In these plots time is normalized by excitatory membrane time constant τ_E . Phase-plane plots for the excitatory subpopulation are shown in (B). Amplitude spectra obtained by the FFT are shown in (C). Normalized frequency is $f\tau_E$. If $\tau_E = 20$ ms, a normalized frequency of 1 corresponds to 50 Hz.



consistent with reduced damping in the system. Essentially, for any level of self-excitation w_{EE} , as inhibitory time constant increases, damping is reduced, and frequency decreases. The frequency range

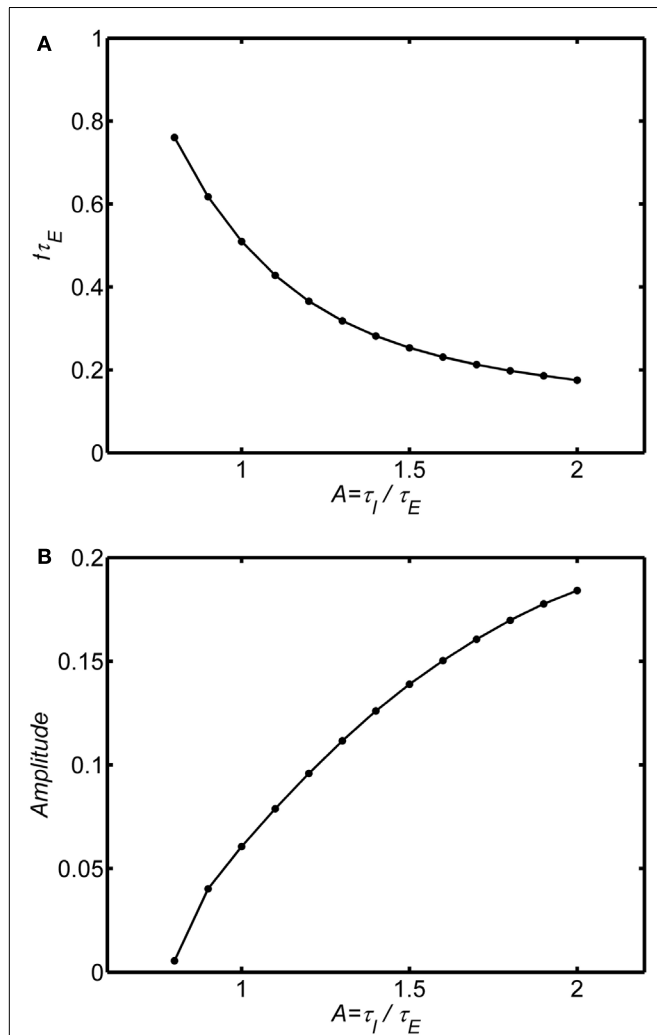


FIGURE 5 | Effect of the ratio of inhibitory to excitatory time constants in the limit cycle regime. The background state is fixed as $E_0(t) = I_0(t) = 0.25$. Input is a step function of magnitude $P(t) = 0.1$. System connection weights are fixed as $(w_{EI}, w_{IE}, w_{II}) = (50, 15, 0)$ and $w_{EE} = 18$ to fix the system in the limit cycle regime. Frequency and amplitude estimated using a Hilbert Transform of the period from of $100\tau_E$ to $200\tau_E$. Normalized frequency is f_{τ_E} . If $\tau_E = 10$ ms, normalized frequency of 1 is 100 Hz. For A smaller than the range shown for each plot the limit cycle transitions to a damped oscillation. The main result is that increasing τ_E relative to τ_I reduces damping (increasing amplitude) and lowers the frequency of the oscillation. **(A)** Normalized frequency **(B)** Amplitude.

of the limit cycle oscillations is much lower than the damped oscillation. For $\tau_E = 10$ ms the damped oscillations are in the high gamma band frequency range (60–100 Hz) when A ranges from 0.5–1.5. In contrast, for the same range of τ_E and A ranging from 1–2 the limit cycle oscillations range is in the lower gamma frequency range (20–50 Hz).

EFFECT OF BACKGROUND STATE

The frequency and damping of the WC oscillator is strongly influenced by the background state. **Figures 6A,B** shows the frequency and damping as a function of background state variables $E_0 = I_0$

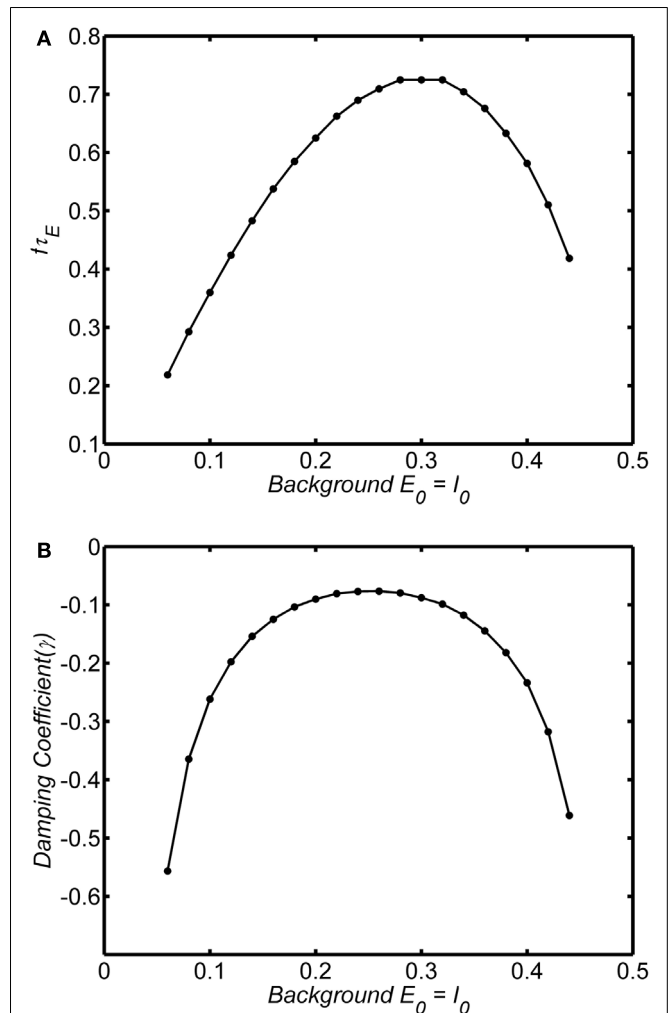


FIGURE 6 | Dependence of frequency and damping on background state for the damped oscillations. System connection weights are fixed as $(w_{EI}, w_{IE}, w_{II}, w_{EE}) = (50, 15, 0, 15)$. Input is a step function of magnitude 0.1. We have verified that the curves are the same for step functions up to 0.3. Frequency was estimated by analyzing zero crossings. Damping was estimated by fitting an exponential decay to the peaks of a rectified (absolute value) of the time series. Normalized frequency is f_{τ_E} . If $\tau_E = 20$ ms, normalized frequency of 1 corresponds to a 50 Hz oscillation. **(A)** Normalized frequency **(B)** Damping coefficient.

for an example in the damped oscillation regime ($w_{EE} = 12$; $A = 1$). At very low levels of background activity the system exhibits low frequency rapidly damped oscillations. As the background activity increases above $E_0 = I_0 = 0.1$ the oscillations become weakly damped and frequency increases as damping decreases. Damping reaches a minimum at $E_0 = I_0 = 0.25$ and the frequency of the oscillation reaches a peak at $E_0 = I_0 = 0.3$. At higher levels of background activity, the oscillations decrease in frequency and are again highly damped. Thus only at the center of the range, at around 0.2–0.3 can we observe high frequency weakly damped oscillations.

If the system is in the limit cycle regime the same essential damping behavior is observed as shown for an example in

Figures 7A,B ($w_{EE} = 18$; $A = 1$). The limit cycle amplitude is suppressed above and below $E_0 = I_0 = 0.25$; at very large or very small values of the background state the limit cycle disappears and is replaced by a damped oscillation. The range of background activity over which the limit cycle is observed can be expanded by increasing the self-excitation parameter (w_{EE}). However, in contrast to the damped oscillation regime, in the limit cycle the frequency of the oscillations is much less dependent on the background activity level, remaining stable over the range of background states with the high amplitude.

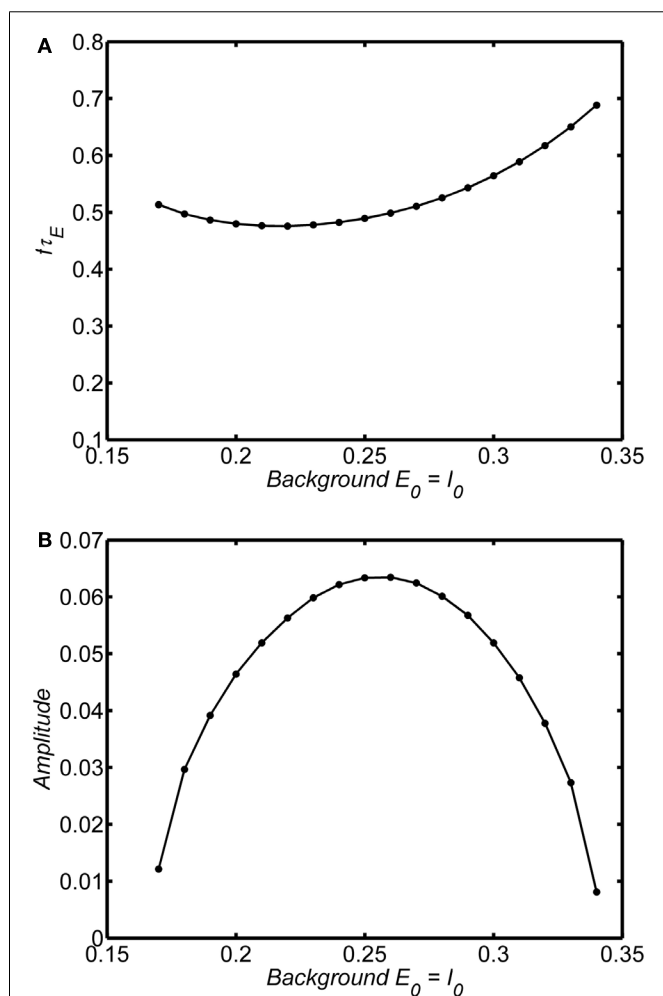


FIGURE 7 | Dependence of frequency and amplitude of limit cycle oscillations on the background state. System connection weights are fixed as (w_{EI} , w_{IE} , w_{II} , w_{EE}) = (50, 15, 0, 20) (dashed line) or (w_{EI} , w_{IE} , w_{II} , w_{EE}) = (50, 15, 0, 25) (solid line). Input is a step function of magnitude 0.1. Frequency and amplitude estimated using a Hilbert Transform of the period from of $100\tau_E$ – $200\tau_E$. Normalized frequency is $f\tau_E$. If $\tau_E = 20$ ms, normalized frequency of 1 corresponds to a 50 Hz oscillation. Outside the range of background states shown there are no limit cycle oscillation for each level of w_{EE} and the oscillations damp out. The main point is that the amplitude of the limit cycle oscillations depend strongly on the background state. Compared to the damped oscillations (Figure 6) the limit cycle frequency does not depend strongly on the background state. (A) Normalized frequency (B) Amplitude.

EFFECTS OF TOP-DOWN SIGNALING

It is increasingly appreciated that neural populations are subject to top-down signals reflected in oscillations in large-scale cortical networks. We modified the WC system equations to incorporate dynamic modulation of background state as in Eq. A12. For simplicity we modulated the background state variables (E_0, I_0) with a sinusoidal signal of fixed frequency; these modulatory frequencies are much slower than the intrinsic frequencies of the WC oscillator. The presence of the modulatory signal alone was not sufficient to drive the system – excitatory input $P(t)$ was always required.

Figure 8 shows some example simulations of the model with the self-excitation parameter set in the limit cycle regime ($w_{EE} = 18$). The main effect of the dynamic modulation of background state is to modulate the amplitude and frequency of the oscillation. Figures 8A,B show the time course of a modulatory signal and the oscillation in the WC model. In this example, the modulatory signal is an oscillation about a background state $E_0 = I_0 = 0.2$ with normalized frequency $f\tau_E = 0.03$. The oscillation in population activity can be seen to modulate in amplitude at the rate of the modulatory signal, with higher amplitude when background activity increases. Thus the phase of the modulatory signal modulates the amplitude of the oscillation. Figures 8D,E show another example where the modulatory signal is an oscillation about a background state $E_0 = I_0 = 0.3$. Here a different phase relationship is evident with higher amplitude when the background activity decreases. For each example, the temporal evolution of the spectrum obtained with wavelet transform is shown in Figures 8C,F. In these examples, the population oscillates at roughly $f\tau_E = 0.5$ with amplitude modulated at $f\tau_E = 0.03$. If $\tau_E = 10$ ms, the underlying oscillation frequency is in the gamma band at approximately 50 Hz and the modulation is in the theta band at 3 Hz. In both cases, during each cycle of the modulatory signal as the amplitude of the population activity increases the frequency decreases.

We carried out simulations over a broad range of parameters to determine if we could produce the apparent effect of amplitude modulation of the intrinsic limit cycle oscillation by adding a sinusoidal modulation at low frequencies to the input $P(t)$. In no case were we able to reproduce the amplitude modulation shown in Figure 8, and the limit cycle show stable amplitude and frequency.

DISCUSSION

In this paper we have revised the Wilson–Cowan model of the interactions within a population of excitatory and inhibitory neurons in order to investigate the impact of background activity on the dynamics of neural populations. In our model this background state is determined statically at very long time scales (presumably by neuromodulator systems) and dynamically at faster time scales by the activity of other cortical systems that exert top-down control on the neural population. We find that our model formalizes the mechanisms by which background state can influence local population dynamics consistent with observations in experimental studies in different behavioral contexts and recording methods.

DYNAMICS OF THE WILSON–COWAN SYSTEM

The basic response properties of the system around a fixed background state indicate that the system produced damped oscillations or sustained limit cycle oscillations depending on the

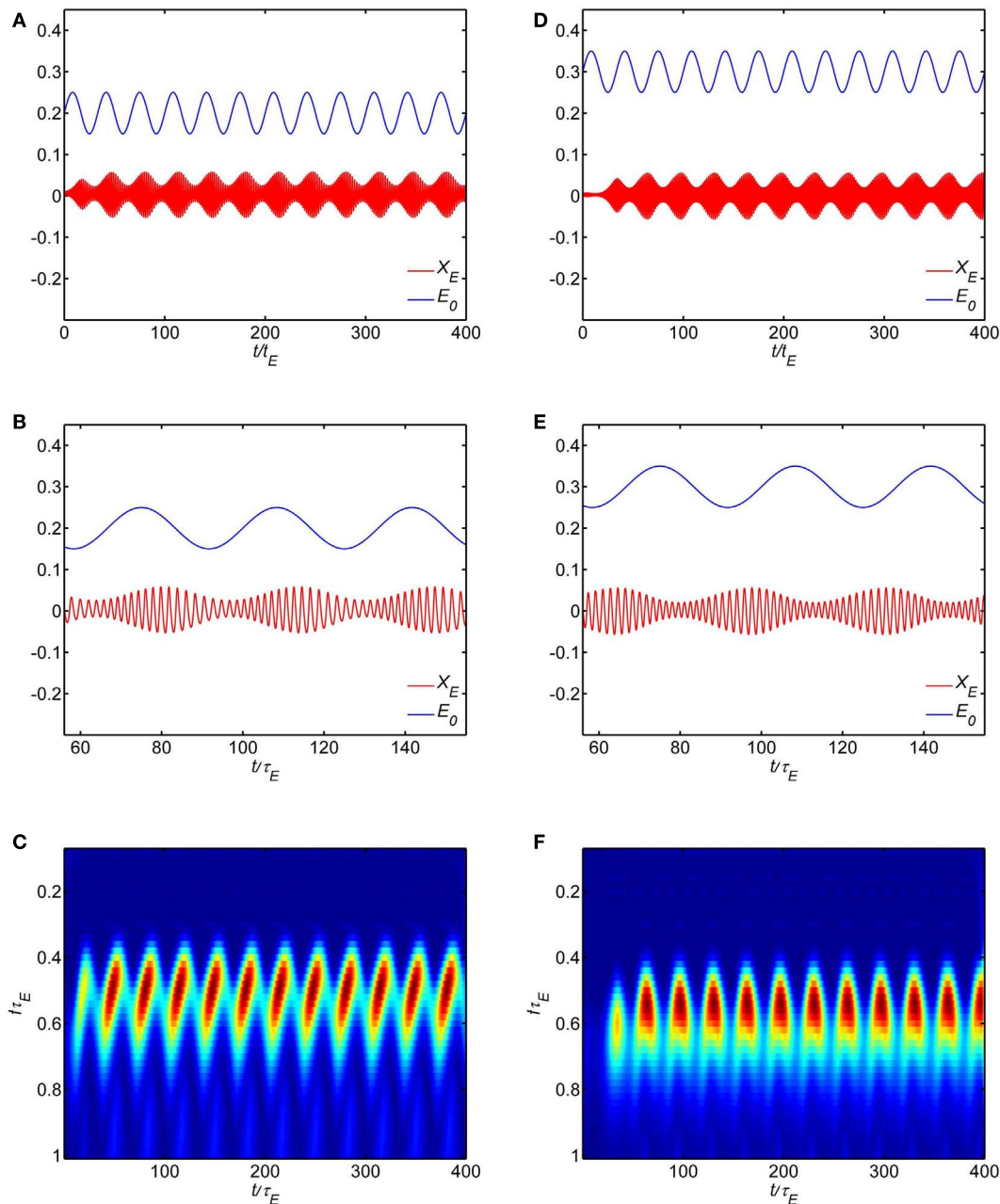


FIGURE 8 | Dynamic modulation of background state produces cross-frequency coupling in the model. Model parameters are in the limit cycle regime (w_{EI} , w_{IE} , w_{II} , w_{EE}) = (50, 15, 0, 18). (A–C) correspond to a

background state $E_0 = I_0 = 0.2$ while (D–F) correspond to a background state $E_0 = I_0 = 0.3$. (A,B,D,E) show the time series while (C,F) show the wavelet spectrum.

level of self-excitation of the excitatory neurons and the relative value of excitatory and inhibitory time constants. Linear analysis about critical (equilibrium, fixed) points indicates that for lower values of self-excitation damped oscillations will be observed while at higher values of self-excitation limit cycle oscillations are observed. In addition, sustained oscillations are more likely to be observed within local populations with longer inhibitory time constants than excitatory time constants. When excitatory time constants are longer than the inhibitory time constants,

rapidly damped high frequency oscillations are observed in the system. Inhibitory time constants that are longer than the excitatory time constants in the 10 ms range result in higher amplitude lower frequency oscillations in the gamma band (30–100 Hz) and generally support limit cycle oscillations rather than damped oscillations. In cases where the inhibitory time constants are much longer these sustained oscillations can be produced at even lower frequencies in the beta (13–30 Hz) and alpha (8–12 Hz) ranges.

Whereas early studies of membrane time constants in mammalian cortex were very short, typically less than 10 ms, more modern studies with improved recording methods report the wide range up to 100 ms (Koch et al., 1996). In particular, very long inhibitory time constants have been reported for thalamic and cortical populations. Thus, we can reasonably expect the ratio of inhibitory to excitatory time constants to be significantly larger than one, supporting the existence of linear instability and corresponding limit cycle oscillations over a broad range of model parameters. Such self-sustained dynamics may contribute locally to the generation of spontaneous EEG rhythms.

EFFECTS OF BACKGROUND STATE ON POPULATION DYNAMICS

The dynamics of the model population depend very strongly on the background state in both limit cycle and damped oscillation ranges of the parameters. At very low levels of background activity ($E_0, I_0 < 0.2$), the population does not respond to the external input. As background activity increases the system responds to external input. In the limit cycle regime, the *amplitude* of the oscillation depends strongly on the background activity while in the damped oscillation regime the *frequency* of the oscillations depends strongly on the background activity.

The overall level of background activity is determined by the neuromodulatory systems that control the sleep-wake cycles, level of arousal, and the response to reward and/or threat. Although there are variations in the densities of neuromodulator inputs to different cortical areas, almost all cortical areas receive neuromodulatory input (Goldman-Rakic et al., 1990). Changes in these states occur over very long time scales; in an awake, behaving animal (or human) these tonic influences are generally considered constant. Thus in general, we can expect that the background level is constant over a responsive level of the population, although there are also phasic modulations of the cholinergic and dopaminergic neuromodulator systems that may play a role in stimulus response and reward seeking behavior (Sarter and Bruno, 1997; Chuhma et al., 2004; Zhang and Sulzer, 2004), and can be expected to have dynamic influences on local networks.

MODULATION OF POPULATION DYNAMICS BY TOP-DOWN SIGNALS

The dynamic modulation of background state creates amplitude and frequency modulation of the intrinsic gamma oscillations of the WC system. The most salient effect is amplitude modulation by the phase of slow modulations of the background states. In the experimental literature this phenomenon is explained as dynamic modulation of the excitability of the population (Jensen and Colgin, 2007). Our model captures this essential behavior, and provides a plausible mechanism to incorporate these types of effects in computational models. In our model, the specific phase/amplitude relationship was influenced by the overall activity level; the specific phase of the modulation that produced robust oscillations was arbitrary. The experimental literature is consistent with this picture, with different studies reporting different phases of the modulation signal for peak amplitude of the local oscillation as shown in **Figure 1** of (Lisman and Buzsaki, 2008).

This essential phenomenon of cross-frequency coupling has been observed in animal models in a number of experimental contexts (Buzsaki and Draguhn, 2004; Scheffzuck et al., 2011) and

human ECoG recordings (Canolty and Knight, 2010; Voytek et al., 2010). These findings have since been confirmed in human EEG where the phase of the theta rhythm is shown to modulate the amplitude of the gamma rhythm (Demiralp et al., 2007). It has long been known that the phase of the alpha rhythm at stimulus onset influences amplitude and phase of the evoked potential (Dustman and Beck, 1965; Jansen and Brandt, 1991; Gruber et al., 2005; Hanslmayr et al., 2007). Moreover, there are a number of studies that have shown that the state of the cortex (as measured by EEG oscillations) can predict the perception of a sensory stimulus, presumably by modulating the sensory evoked response (Haig and Gordon, 1998; Hanslmayr et al., 2007).

LOCAL-GLOBAL INTERACTIONS

We have proposed a conceptual framework in which local networks (cell assemblies) are embedded in a global environment that produces standing waves due to propagation in the corticocortical (white matter) fibers and periodic boundary conditions (Nunez, 1995, 2000; Nunez and Srinivasan, 2006). That is, the neocortex and underlying white matter are modeled as a closed loop or spherical shell. In this paper, we have proposed a method to model the top-down influences of such systems on a local network. These top-down influences may be the result of feedback from global network. In our analysis we have isolated the local network from the global system, and only analyzed the local network dynamics. Similarly, global models typically assume that the local network is sufficiently localized such that its (bottom up) influence on the global dynamics may be neglected to first approximation. This condition might be satisfied in the eyes closed resting state, for example. On the other hand, the (eyes open) processing of substantial visual input or complex cognitive functions may involve multiple local thalamocortical networks that act (bottom up) to modify the global networks that are influencing the local networks. Future work must explicitly consider in more detail how the local networks and global networks interact.

Our study suggests that models of these local networks must incorporate the idea that the response properties of the networks can be modified by modulatory inputs. In our modified WC model, the addition of oscillatory afferent input does not modify the system dynamics. We explicitly incorporated dynamic modulation of the system properties by making the background state an explicit part of the model. In most models (including the original WC) the background state of the neurons is mathematically removed, and the dynamics of the system is studied without further consideration of the background state. This approximation is limiting; modification of the background state may be an important mechanism of top-down signaling in the cortex, especially in the control of goal-oriented behavior such as attention. Local networks in the cortex experience dynamic background states which can be readily incorporated into most model formulations. This may also have importance in specific models that seek to make a distinction between feedforward and feedback connections in sensory systems (Mountcastle, 1997; Lamme and Roelfsema, 2000).

CONCLUSIONS

Since the first human recording in the early 1920s the physiological bases for the wide range of rhythmic EEG activity has been

somewhat of a mystery. As one important “window on the mind,” EEG has long provided a critical tool in pursuit of connecting neural dynamics to cognitive processes. Human brains produce a proverbial “spectral zoo” that is closely correlated to behavior and cognition. A major obstacle in this quest is a shortage of robust and widely appreciated theoretical support for EEG’s dynamic behavior in time and spatial location over the scalp. The conceptual framework facilitated by such theory could have a substantial influence on the design of new EEG-cognitive experiments. In this

paper, we propose an approach to incorporate global (top-down) influences on local networks. The essence of our approach is to immerse the local network in a dynamic background state. These dynamics could be generated by a global model of interactions across the cortex; they could also be modeled from experimental EEG data. This approach is sufficiently general to be applied to other theoretical formulations of population dynamics in neural populations and to models of specific cognitive influences on local circuit dynamics.

REFERENCES

- Borisjuk, R., Denham, M., Hoppensteadt, F., Kazanovich, Y., and Vinogradova, O. (2000). An oscillatory neural network model of sparse distributed memory and novelty detection. *BioSystems* 58, 265–272.
- Bridwell, D. A., and Srinivasan, R. (2012). Distinct attention networks for feature enhancement and suppression in vision. *Psychol. Sci.* 23, 1151–1158.
- Bush, P., and Sejnowski, T. (1996). Inhibition synchronizes sparsely connected cortical neurons within and between columns in realistic network models. *J. Comput. Neurosci.* 3, 91–110.
- Buzsaki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929.
- Campbell, S., and Wang, D. (1996). Synchronization and desynchronization in a network of locally coupled Wilson–Cowan oscillators. *IEEE Trans. Neural Netw.* 7, 541–554.
- Canolty, R. T., Ganguly, K., Kennerley, S. W., Cadieu, C. F., Koepsell, K., Wallis, J. D., et al. (2010). Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17356–17361.
- Canolty, R. T., and Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends Cogn. Sci. (Regul. Ed.)* 14, 506–515.
- Chuhma, N., Zhang, H., Masson, J., Zhuang, X., Sulzer, D., Hen, R., et al. (2004). Dopamine neurons mediate a fast excitatory signal via their glutamatergic synapses. *J. Neurosci.* 24, 972–981.
- Demiralp, T., Bayraktaroglu, Z., Lenz, D., Junge, S., Busch, N. A., Maess, B., et al. (2007). Gamma amplitudes are coupled to theta phase in human EEG during visual perception. *Int. J. Psychophysiol.* 64, 24–30.
- Ding, J., Sperling, G., and Srinivasan, R. (2006). Attentional modulation of SSVEP power depends on the network tagged by the flicker frequency. *Cereb. Cortex* 16, 1016–1029.
- Dustman, R. E., and Beck, E. C. (1965). Phase of Alpha Brain Waves. *Electroencephalogr. Clin. Neurophysiol.* 18, 433–440.
- Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* 2, 704–716.
- Engel, A. K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci. (Regul. Ed.)* 5, 16–25.
- Fellous, J. M., and Linster, C. (1998). Computational models of neuromodulation. *Neural Comput.* 10, 771–805.
- Fries, P., Nikolic, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci.* 30, 309–316.
- Goldman-Rakic, P. S., Lidow, M. S., and Gallager, D. W. (1990). Overlap of dopaminergic, adrenergic, and serotonergic receptors and complementarity of their subtypes in primate prefrontal cortex. *J. Neurosci.* 10, 2125–2138.
- Gruber, W. R., Klimesch, W., Sauseng, P., and Doppelmayr, M. (2005). Alpha phase synchronization predicts P1 and N1 latency and amplitude size. *Cereb. Cortex* 15, 371–377.
- Haig, A. R., and Gordon, E. (1998). Prestimulus EEG alpha phase synchronicity influences N100 amplitude and reaction time. *Psychophysiology* 35, 591–595.
- Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., and Bauml, K. H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage* 37, 1465–1473.
- Izhikevich, E. M. (2006). Polychronization: computation with spikes. *Neural Comput.* 18, 245–282.
- Izhikevich, E. M., and Edelman, G. M. (2008). Large-scale model of mammalian thalamocortical systems. *Proc. Natl. Acad. Sci. U.S.A.* 105, 3593–3598.
- Jansen, B. H., and Brandt, M. E. (1991). The effect of the phase of prestimulus alpha activity on the averaged visual evoked response. *Electroencephalogr. Clin. Neurophysiol.* 80, 241–250.
- Jensen, O., and Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends Cogn. Sci. (Regul. Ed.)* 11, 267–269.
- Jirsa, V. K., and Haken, H. (1996). Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* 77, 960–963.
- Jirsa, V. K., and Haken, H. (1997). A derivation of a macroscopic field theory of the brain from the quasiscale microscopic neural dynamics. *Physica D* 99, 503–526.
- Koch, C., Rapp, M., and Segev, I. (1996). A brief history of time (constants). *Cereb. Cortex* 6, 93–101.
- Lamme, V. A., and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579.
- Liley, D. T. J., Cadusch, P. J., and Wright, J. J. (1999). A continuum theory of electrocortical activity. *Neurocomputing* 26, 795–800.
- Lisman, J., and Buzsaki, G. (2008). A neural coding scheme formed by the combined function of gamma and theta oscillations. *Schizophr. Bull.* 34, 974–980.
- Lumer, E. D., Edelman, G. M., and Tononi, G. (1997). Neural dynamics in a model of the thalamocortical system. *Cereb. Cortex* 7, 207–227.
- Menon, V., Freeman, W. J., Cuttillo, B. A., Desmond, J. E., Ward, M. F., Bressler, S. L., et al. (1996). Spatiotemporal correlations in human gamma band electrocorticograms. *Electroencephalogr. Clin. Neurophysiol.* 98, 89–102.
- Mountcastle, V. B. (1997). *Perceptual Neuroscience: The Cerebral Cortex*. Cambridge: Harvard University Press.
- Nunez, P. L. (1974). The brain wave equation: a model for the EEG. *Math. Biosci.* 21, 279–297.
- Nunez, P. L. (1981). *Electric Fields of the Brain: The Neurophysics of EEG*. New York: Oxford University Press.
- Nunez, P. L. (1995). *Neocortical Dynamics and Human EEG Rhythms*. New York: Oxford University Press.
- Nunez, P. L. (2000). Toward a quantitative description of large-scale neocortical dynamic function and EEG. *Behav. Brain Sci.* 23, 371–398; discussion 399–437.
- Nunez, P. L., and Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*, 2nd Edn. New York: Oxford University Press.
- Nunez, P. L., Wingeier, B. M., and Silberstein, R. B. (2001). Spatial-temporal structures of human alpha rhythms: theory, microcurrent sources, multiscale measurements, and global binding of local networks. *Hum. Brain Mapp.* 13, 125–164.
- Pinotsis, D. A., Hansen, E., Friston, K. J., and Jirsa, V. K. (2013). Anatomical connectivity and the resting state activity of large cortical networks. *Neuroimage* 65, 127–138.
- Pinotsis, D. A., Schwarzkopf, D. S., Litvak, V., Rees, G., Barnes, G., and Friston, K. J. (2012). Dynamic causal modelling of lateral interactions in the visual cortex. *Neuroimage* 66C, 563–576.
- Robinson, P. A., Rennie, C. J., and Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* 55, 826–840.
- Romei, V., Rihs, T., Brodbeck, V., and Thut, G. (2008). Resting electroencephalogram alpha-power over posterior sites indexes baseline visual cortex excitability. *Neuroreport* 19, 203–208.
- Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G. L., and Von Stein, A. (1998). Synchronization between prefrontal and posterior association cortex during human working memory. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7092–7096.
- Sarter, M., and Bruno, J. P. (1997). Cognitive functions of cortical acetylcholine: toward a unifying hypothesis. *Brain Res. Brain Res. Rev.* 23, 28–46.
- Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R., and Hanslmayr, S. (2005). EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Hum. Brain Mapp.* 26, 148–155.

- Schalk, G., and Leuthardt, E. C. (2011). Brain-computer interfaces using electrocorticographic signals. *IEEE Rev. Biomed. Eng.* 4, 140–154.
- Scheffzuck, C., Kukushka, V. I., Vysotski, A. L., Draguhn, A., Tort, A. B., and Brankack, J. (2011). Selective coupling between theta phase and neocortical fast gamma oscillations during REM-sleep in mice. *PLoS ONE* 6:e28489. doi:10.1371/journal.pone.0028489
- Srinivasan, R., Bibi, F. A., and Nunez, P. L. (2006). Steady-state visual evoked potentials: distributed local sources and wave-like dynamics are sensitive to flicker frequency. *Brain Topogr.* 18, 167–187.
- Thorpe, S., D'zmura, M., and Srinivasan, R. (2012). Lateralization of frequency-specific networks for covert spatial attention to auditory stimuli. *Brain Topogr.* 25, 39–54.
- Thorpe, S. D. S., Garcia, J. O., Lee, R. R., Huang, M., and Srinivasan, R. (2011). Spatial attention enhances steady-state visual evoked potentials in the gamma band. *Int. J. Bioelectromagn.* 13, 233–238.
- Thut, G., Nietzel, A., Brandt, S. A., and Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.* 26, 9494–9502.
- Traub, R. D., Jefferys, J. G., and Whittington, M. A. (1997). Simulation of gamma rhythms in networks of interneurons and pyramidal cells. *J. Comput. Neurosci.* 4, 141–150.
- von Stein, A., and Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int. J. Psychophysiol.* 38, 301–313.
- Voytek, B., Canolty, R. T., Shestyuk, A., Crone, N. E., Parvizi, J., and Knight, R. T. (2010). Shifts in gamma phase-amplitude coupling frequency from theta to alpha over posterior cortex during visual tasks. *Front. Hum. Neurosci.* 4:191. doi:10.3389/fnhum.2010.00191
- Whittington, M. A., Traub, R. D., Kopell, N., Ermentrout, B., and Buhl, E. H. (2000). Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int. J. Psychophysiol.* 38, 315–336.
- Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24.
- Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80.
- Yousif, N. A., and Denham, M. (2005). A population-based model of the nonlinear dynamics of the thalamocortical feedback network displays intrinsic oscillations in the spindling (7–14 Hz) range. *Eur. J. Neurosci.* 22, 3179–3187.
- Zhang, H., and Sulzer, D. (2004). Frequency-dependent modulation of dopamine release by nicotine. *Nat. Neurosci.* 7, 581–582.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 December 2012; accepted: 19 March 2013; published online: 18 April 2013.

Citation: Srinivasan R, Thorpe S and Nunez PL (2013) Top-down influences on local networks: basic theory with experimental implications. *Front. Comput. Neurosci.* 7:29. doi: 10.3389/fncom.2013.00029

Copyright © 2013 Srinivasan, Thorpe and Nunez. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

THE BASIC WC ANALYSIS

The classical WC model (Wilson and Cowan, 1972, 1973) applies to populations of interacting excitatory and inhibitory neurons in some local neocortical region as indicated in **Figure 1**. An important WC assumption is that axon propagation delays are negligible; that is, all WC delays are due to PSP rise and decay times. WC is then a strictly local model and represents the opposite limiting case to global models in which delays are axonal, especially in the longer corticocortical axons forming most of human white matter (Nunez, 1974; Nunez and Srinivasan, 2006). These distinct local and global models have been shown to be fully compatible and may be combined into local/global models (Jirsa and Haken, 1996; Nunez, 2000).

The basic WC dependent variables are the *fractions of excitatory and inhibitory active cells* (dimensionless action potential densities) $E(t)$, $I(t)$, which can evidently exhibit very high frequency jitter not treated in the analysis. Rather, the WC equations are expressed in terms of coarse grained (in time) excitatory $\langle E(t) \rangle$ and inhibitory $\langle I(t) \rangle$ action potential densities, where the critical point $\langle E(t) \rangle, \langle I(t) \rangle = (0,0)$ is considered by WC to be an equilibrium background state occurring when the external (afferent) driving function $P(t) = 0$. Thus, the variables $\langle E(t) \rangle, \langle I(t) \rangle$ are allowed to take on negative values by WC, an inaccurate and (as we show here) unnecessary approximation to their physiological interpretation as fractions of active cells. The WC equations (1.3.1 and 1.3.2 from the 1973 paper) are

$$\begin{aligned} \tau_E \frac{d\langle E(t) \rangle}{dt} &= -\langle E(t) \rangle + [1 - r_E \langle E(t) \rangle] S_E[P(t) \\ &\quad + w_{EE} \langle E(t) \rangle - w_{IE} \langle I(t) \rangle] \\ \tau_I \frac{d\langle I(t) \rangle}{dt} &= -\langle I(t) \rangle + [1 - r_I \langle I(t) \rangle] S_I[w_{EI} \langle E(t) \rangle \\ &\quad - w_{II} \langle I(t) \rangle] \end{aligned} \quad (\text{A1})$$

Here we drop the spatial dependence x of all variables since axon speeds are assumed to be infinite implying that neural spatial separations have no effect on dynamic behaviors in this approximation. We also distinguish between the excitatory and inhibitory membrane time constants τ_E, τ_I . In addition, we only allow excitatory afferent input $P(t)$.

SOME ISSUES WITH THE ORIGINAL WC ANALYSIS

1. WC write the *proportion of sensitive excitatory cells* (neurons currently firing or in their refractory periods r_E) as $R_E(t) = 1 - \int_{t-r_E}^t E(t') dt'$. This is not dimensionally correct and leads to the dimensionally incorrect Eq. A1. The WC equation also yields the incorrect result $R_E(t) \rightarrow 1$ as the refractory period $r_E \rightarrow 0$. The correct expression is

$$\begin{aligned} R_E(t) &= 1 - \frac{1}{r_E} \int_{t-r_E}^t E(t') dt' \\ R_E(t) &\rightarrow 1 - E(t) \text{ when } r_E \rightarrow 0 \end{aligned} \quad (\text{A2})$$

In this limit, the sensitive population $R_E(t)$ consists of all cells not firing at time t . The excitatory and inhibitory integrals should have been divided by the refractory times r_E, r_I , equivalent to setting $r_E, r_I = 1$, in Eq. A1, a simple corrective step often adopted by others using the WC model.

2. Negative values of $\langle E(t) \rangle, \langle I(t) \rangle$ are not realistic physiologically for essentially the same reason. During times when $E(t), I(t) < 0$, the fractions of sensitive cells $R_E(t), R_I(t) > 1$, which is also inconsistent with realistic physiology. By contrast, the modified WC model presented here forces

$$\langle E(t) \rangle, \langle I(t) \rangle \geq 0, \text{ for } -\infty < t < +\infty.$$

3. WC choose sigmoid response functions such that $S_E(0) = S_I(0) = 0$ in order to force $(0,0)$ to be a critical (equilibrium) point where the variable time derivatives equal zero. Such critical points may be either state or unstable. If a critical point is stable, any brain dynamic state coming sufficiently close to this point will become fixed (forever static or “brain dead”). Of more interest to us are unstable critical points associated with on-going oscillations (limit cycles), possibly underlying EEG. In WC, the tissue response functions can become negative, a physiological impossibility; thus, in our following modified analysis, the WC conditions are replaced by

$$\begin{aligned} S_E[P(t), \langle E(t) \rangle, \langle I(t) \rangle] &\geq 0 \\ S_I[\langle E(t) \rangle, \langle I(t) \rangle] &\geq 0 \end{aligned} \quad (\text{A3})$$

4. WC interpret the WC parameter μ as a single membrane time constant. Based on the classic solution of the cable equation and the distribution of excitatory synapses on dendrites with inhibitory synapses typically near cell bodies and our increased appreciation of wide ranges of excitatory and inhibitory time constants τ_E, τ_I , we consider cases $\tau_E \neq \tau_I$.

A MODIFIED VERSION OF THE WC ANALYSIS

Define the non-dimensional time $t_1 = \frac{t}{\tau_E}$ and time constant ratio $A = \frac{\tau_I}{\tau_E}$ so that with rescaled variables Eq. A1 become

$$\begin{aligned} \frac{d\langle E \rangle}{dt_1} &= -\langle E \rangle + (1 - \langle E \rangle) S_E(P + w_{EE} \langle E \rangle - w_{IE} \langle I \rangle) \\ A \frac{d\langle I \rangle}{dt_1} &= -\langle I \rangle + (1 - \langle I \rangle) S_I(w_{EI} \langle E \rangle - w_{II} \langle I \rangle) \end{aligned} \quad (\text{A4})$$

Here the non-dimensional parameter A may range from somewhat less than one to as high as perhaps 5. For convenience we drop the subscript 1 on the non-dimensional time variable. We introduce the new dependent variables $X_E(t), X_I(t)$, which provide perturbations about the critical point (E_0, I_0)

$$\begin{aligned} \langle E(t) \rangle &= E_0 + X_E(t) \\ \langle I(t) \rangle &= I_0 + X_I(t) \end{aligned} \quad (\text{A5})$$

We assume that the input function $P(t)$ is exclusively excitatory such that $P(t) \geq 0$ at all times. The following conditions follow from the variable definitions

$$\begin{aligned} 0 &\leq E_0 + X_E(t) \leq 1 \\ 0 &\leq I_0 + X_I(t) \leq 1 \end{aligned} \quad (\text{A6})$$

Equations (A4) then yield

$$\begin{aligned} \frac{dX_E}{dt} &= -E_0 - X_E + (1 - E_0 - X_E) S_E [P + w_{EE}(E_0 + X_E) \\ &\quad - w_{IE}(I_0 + X_I)] \\ A \frac{dX_I}{dt} &= -I_0 - X_I + (1 - I_0 - X_I) S_I [w_{EI}(E_0 + X_E) \\ &\quad - w_{II}(I_0 + X_I)] \end{aligned} \quad (\text{A7})$$

The sigmoid functions S_E, S_I are chosen here such that (E_0, I_0) is a critical point when $P(t) = 0$. Thus, we choose the following alternate sigmoid tissue response functions.

$$S_I = \frac{1}{1 + \exp[-w_{EI}(E_0 + X_E) + w_{II}(I_0 + X_I) + K_I]} \quad (\text{A8})$$

$$S_E = \frac{1}{1 + \exp[-w_{EE}(E_0 + X_E) + w_{IE}(I_0 + X_I) + K_E]} \quad (\text{A9})$$

The constants K_E, K_I are added to the original WC sigmoid functions, thereby determining the response range during oscillations about fixed points. This choice of the forms of the sigmoid response functions insures that (1) (E_0, I_0) or $(X, Y) = (0, 0)$ is a critical point and (2) $0 \leq S_E \leq 1$ and $0 \leq S_I \leq 1$. Substitution of Eqs A8 and A9 into Eq. A7 and setting $X_E(t) = X_I(t) = 0$ yields the sigmoid constants in terms of the critical point

$$K_I = w_{EI}E_0 - w_{II}I_0 + \text{Log} \left[\frac{1}{I_0} - 2 \right] \quad (\text{A10})$$

$$K_E = w_{EE}E_0 - w_{IE}I_0 + \text{Log} \left[\frac{1}{E_0} - 2 \right] \quad (\text{A11})$$

Substitution of Eqs A10 and A11 into Eqs A8 and A9 yields

$$\begin{aligned} S_I [w_{EI}(E_0 + X_E) - w_{II}(I_0 + X_I)] \\ = \frac{1}{1 + \left(\frac{1}{I_0} - 2 \right) \exp(-w_{EI}X_E + w_{II}X_I)} \quad I_0 < \frac{1}{2} \end{aligned} \quad (\text{A12})$$

$$\begin{aligned} S_E [w_{EE}(E_0 + X_E) - w_{IE}(I_0 + X_I)] \\ = \frac{1}{1 + \left(\frac{1}{E_0} - 2 \right) \exp(-w_{EE}X_E + w_{IE}X_I)} \quad E_0 < \frac{1}{2} \end{aligned} \quad (\text{A13})$$

PHASE-PLANE ANALYSIS

We assume $w_{II} \cong 0$ in all of the following analyses based on our preliminary studies: In simulations with non-zero w_{II} the effect is only to set the activity level of the system and has no significant influence on the dynamics. The first step in the analysis of Eq. A7 is to find the nature of the critical point $(0, 0)$. To accomplish this we expand the functions $F(X_E, X_I)$ and $G(X_E, X_I)$ about $(0, 0)$, where these functions are the expressions on the right sides of Eq. A7, that is

$$\begin{aligned} F(X_E, X_I) &= -E_0 - X_E + (1 - E_0 - X_E) S_E \\ G(X_E, X_I) &= [-I_0 - X_I + (1 - I_0 - X_I) S_I]/A \end{aligned} \quad (\text{A14})$$

Taylor expansion about the critical point $(0, 0)$ yields equations of the general form

$$\begin{aligned} \frac{dX_E}{dt} &= F(X_E, X_I) \cong F(0, 0) + \left(\frac{\partial F}{\partial X_E} \right)_0 X_E + \left(\frac{\partial F}{\partial X_I} \right)_0 X_I \\ &\equiv aX_E + bX_I \\ \frac{dX_I}{dt} &= G(X_E, X_I) \cong G(0, 0) + \left(\frac{\partial G}{\partial X_E} \right)_0 X_E + \left(\frac{\partial G}{\partial X_I} \right)_0 X_I \\ &\equiv cX_E + dX_I \end{aligned} \quad (\text{A15})$$

Here we have forced $F(0, 0) = G(0, 0) = 0$ by proper choice of the constants K_E, K_I in Eqs A10 and A11. The partial derivatives are evaluated at $(0, 0)$ yielding the parameters (a, b, c, d) . Eq. A15 then consist of two first order linear equations governing the dynamic behavior of the non-linear system close to $(0, 0)$. Define the parameters $\beta = a + d$ and $\gamma = ad - bc$; the eigenvalues $\lambda_{1,2}$ of the linear system satisfy $\lambda^2 - \beta\lambda + \gamma = 0$ with solution

$$\lambda_{1,2} = \frac{1}{2} (\beta \pm \sqrt{\beta^2 - 4\gamma}) \quad (\text{A16})$$

We are mainly interested in oscillatory solutions about the critical point $(0, 0)$; that is, stable limit cycle solutions. These are expected when $(0, 0)$ is an *unstable* spiral, which occurs when $\beta > 0$ and $\beta^2 < 4\gamma$. By contrast, stable spirals result in damped oscillations. Saddle points and nodes result in non-oscillatory solutions (stable or unstable) that are of minimal interest here.

Consider the following example with $E_0 = I_0$. The critical point (E_0, I_0) is unstable if the following condition is met

$$w_{EE} > \frac{A + 1}{AE_0(1 - 2E_0)} \quad 0 < E_0 < \frac{1}{2} \quad (\text{A17})$$

Note that $A = \frac{\tau_I}{\tau_E}$ so if inhibitory time constants are much shorter than excitatory time constants, larger values of w_{EE} are required to produce linear instability. For the physiologically interesting range $0.1 < E_0 < 0.4$ and $A = 1$, all $w_{EE} > 25$ cause the fixed point to be unstable.

For the case $E_0 = I_0$ and $A = 1$, the fixed point is a spiral if

$$w_{EE} < 2\sqrt{w_{IE}w_{EI}} \quad (\text{A18})$$

By combining Eqs A17 and A18 we find the necessary (but possibly not sufficient) conditions for a stable limit cycle about (E_0, E_0) when $A = 1$, that is

$$\frac{2}{E_0(1 - 2E_0)} < w_{EE} < 2\sqrt{w_{IE}w_{EI}} \quad (\text{A19})$$

If this condition is met, the corresponding spiral frequency is

$$\omega_{\text{spiral}} = N(E_0) \sqrt{4w_{IE}w_{EI} - w_{EE}^2} \quad (\text{A20})$$

Here the numerical factor lies in the range $0.00790 < N(E_0) < 0.0294$ if $0.1 < E_0 < 0.4$. An *unstable* spiral point at (E_0, I_0)

suggests a likely *stable limit cycle*, but the limit cycle frequency will not generally equal the (linear) spiral frequency. If w_{EE} exceeds the upper limit in Eq. A19 an unstable node or saddle point will occur. In this case the solutions $X_E(t)$, $X_I(t)$ are likely to grow beyond physiologically realistic ranges, implying that the basic WC equations are no longer valid.



Cortical information flow in Parkinson's disease: a composite network/field model

Cliff C. Kerr^{1,2,3*}, Sacha J. Van Albada⁴, Samuel A. Neymotin^{1,5}, George L. Chadderton¹, P. A. Robinson^{2,3} and William W. Lytton^{1,6}

¹ Department of Physiology and Pharmacology, State University of New York Downstate Medical Center, Brooklyn, NY, USA

² School of Physics, University of Sydney, NSW, Australia

³ Brain Dynamics Centre, Westmead Millennium Institute, Westmead, NSW, Australia

⁴ Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6), Jülich Research Centre and JARA, Jülich, Germany

⁵ Department of Neurobiology, Yale University, New Haven, CT, USA

⁶ Department of Neurology, Kings County Hospital, Brooklyn, NY, USA

Edited by:

Dimitris Pinotsis, University College London, UK

Reviewed by:

Petra Ritter, Charité University Medicine Berlin, Germany
Stephan Van Gils, Universiteit Twente, Netherlands

*Correspondence:

Cliff C. Kerr, Department of Physiology and Pharmacology, State University of New York Downstate Medical Center, 450 Clarkson Avenue, Po Box 31, Brooklyn, NY 11203, USA.
e-mail: ckerr@physics.usyd.edu.au

The basal ganglia play a crucial role in the execution of movements, as demonstrated by the severe motor deficits that accompany Parkinson's disease (PD). Since motor commands originate in the cortex, an important question is how the basal ganglia influence cortical information flow, and how this influence becomes pathological in PD. To explore this, we developed a composite neuronal network/neural field model. The network model consisted of 4950 spiking neurons, divided into 15 excitatory and inhibitory cell populations in the thalamus and cortex. The field model consisted of the cortex, thalamus, striatum, subthalamic nucleus, and globus pallidus. Both models have been separately validated in previous work. Three field models were used: one with basal ganglia parameters based on data from healthy individuals, one based on data from individuals with PD, and one purely thalamocortical model. Spikes generated by these field models were then used to drive the network model. Compared to the network driven by the healthy model, the PD-driven network had lower firing rates, a shift in spectral power toward lower frequencies, and higher probability of bursting; each of these findings is consistent with empirical data on PD. In the healthy model, we found strong Granger causality between cortical layers in the beta and low gamma frequency bands, but this causality was largely absent in the PD model. In particular, the reduction in Granger causality from the main "input" layer of the cortex (layer 4) to the main "output" layer (layer 5) was pronounced. This may account for symptoms of PD that seem to reflect deficits in information flow, such as bradykinesia. In general, these results demonstrate that the brain's large-scale oscillatory environment, represented here by the field model, strongly influences the information processing that occurs within its subnetworks. Hence, it may be preferable to drive spiking network models with physiologically realistic inputs rather than pure white noise.

Keywords: neural field model, spiking neural networks, Parkinson's disease, thalamus, cortex, basal ganglia, Granger causality, interlaminar processing

1. INTRODUCTION

Parkinson's disease (PD) is a multiscale phenomenon, encompassing pathology at the level of single neurons, local networks, large neuronal ganglia, and the complex interactions between these ganglia and the cortex. PD is caused by the degeneration of dopaminergic neurons in the substantia nigra pars compacta, with the damage later spreading to dopaminergic neurons in the ventral tegmental area (Cools, 2006). The loss of dopaminergic input alters the dynamics of the striatum, which then affects the dynamics of large portions of the thalamus and cortex, which in turn affects the spinal cord and muscles (Bolam et al., 2002). Striatal dynamics are crucial to several large-scale projection pathways, including the well-characterized direct and indirect pathways. Dopaminergic input to the striatum increases

transmission in D1-expressing striatal neurons involved in the direct pathway. These neurons inhibit the globus pallidus internal segment (GPi). Dopaminergic input also decreases input to D2-expressing striatal neurons involved in the indirect pathway. These neurons inhibit the globus pallidus external segment (GPe), which in turn inhibits the GPi. Thus, alterations to the direct and indirect pathways in PD are both thought to increase the firing rate of the GPi, which in turn inhibits the thalamus. There is also a hyperdirect pathway from the cortex to the GPi via the subthalamic nucleus (STN), as well as other lesser pathways (Figure 1).

Numerous models of PD and the basal ganglia have been proposed, using either field or network approaches. Van Albada and Robinson (2009) developed a field-based model of the basal

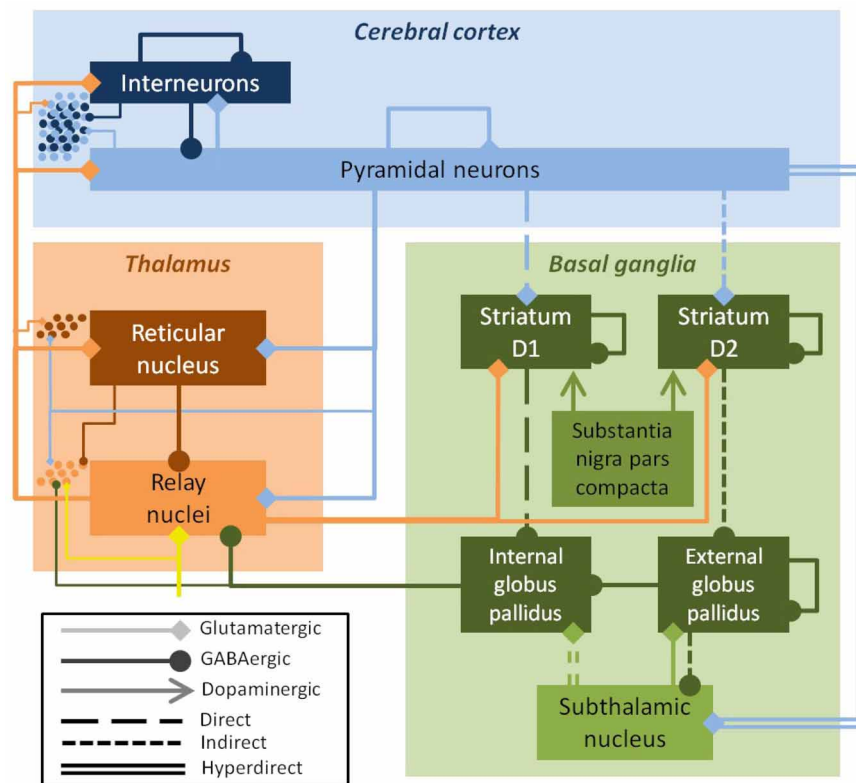


FIGURE 1 | Schematic of the field model, showing excitatory populations and connections (light colors, diamond arrows) and inhibitory ones (dark colors, round arrows). The key efferent nucleus of the basal ganglia is the internal globus pallidus (GPi), which receives cortical input via direct, indirect, and hyperdirect pathways. The field model drives a

spiking network model, shown here schematically (dots at left); the inputs from the field model to the spiking model are indicated by the thin lines. The substantia nigra pars compacta modulates parameters, but is not explicitly modeled. Inputs to the thalamus (yellow arrow) were modeled as white noise.

ganglia/thalamocortical system. This model was shown to reproduce realistic firing rates of each neuronal population in both healthy and PD states. One early network model was that of Terman et al. (2002), which represented a small network of neurons in the GPe and STN. A considerably larger and more complex (non-spiking) network model was developed by Leblois et al. (2006). This model explored both basal ganglia and thalamocortical cell populations, looking at competition between the direct and hyperdirect pathways. They suggested that PD disrupted this competitive balance, resulting in loss of the network's ability to select motor programs. Another network model focusing on motor-selection abilities was developed by Humphries et al. (2006), who also found that decreased dopamine interfered with the basal ganglia's capacity for selecting actions. Network models have also been used to analyze and predict the effects of deep brain stimulation on basal ganglia nuclei (Hahn and McIntyre, 2010; Guo and Rubin, 2011; Dovzhenok et al., 2013).

Previous neuronal network models of PD have either not included a cortex at all (Terman et al., 2002; Rubchinsky et al., 2003; Park et al., 2011), approximated it as a random Poisson process (Humphries et al., 2006), or considered it as a single layer with a single cell type (Leblois et al., 2006). The thalamus has also either been omitted or treated as a single population. In this work,

we sought to fill this gap by exploring the interactions of the large-scale dynamics of basal ganglia, represented by a field model, with a far smaller but more spatially detailed network model of the thalamus and six-layered cortex.

1.1. COMPOSITE MODEL

The primary aim of this paper is to determine how the large-scale dynamics of the brain affect the information flow in small networks of neurons. Most previous brain modeling efforts have been directed at one of these two scales, rather than their interaction. These efforts have consisted of either (1) neural field models that describe the dynamics of the whole brain, without explicitly modeling the activity of individual neurons (Nunez, 1974; Jirsa and Haken, 1996; Robinson et al., 1997; Destexhe and Sejnowski, 2009), or (2) spiking neuronal network models that capture individual neurons' dynamics, but are many orders of magnitude smaller than the brains of even the simplest vertebrates (Lumer et al., 1997; Neymotin et al., 2011b). Several large network models have also been published that have roughly as many "neurons" as the full mammalian brain (Izhikevich and Edelman, 2008; Ananthanarayanan et al., 2009). However, these models have not yet reproduced large-scale dynamics with the same degree of fidelity as neural field models. For example, the model of

Izhikevich and Edelman (2008) showed simultaneous peaks in the delta and alpha bands, whereas experimentally these peaks are characteristic of sleep and wakefulness, respectively, and are hence rarely observed simultaneously (Niedermeyer and Lopes da Silva, 1999). Such infidelity may be because the enormous computational resources required to run these models makes it impractical to constrain their parameters by fitting their dynamics to experimental data.

Recently, both Deco and Jirsa (2012) and Wilson et al. (2012) described approximations that allow small and large spatial scales to be spanned at a mesoscopic level of description, allowing large-scale dynamics (e.g., BOLD signals) to be related to small-scale network properties (e.g., criticality). Robinson and Kim (2012) took a different approach: they described the theoretical basis of combining spiking network and neural field components into a single model. The fundamental challenge in combining these two modeling approaches is to create a common representation of neuronal activity, since individual spikes are used in network models, while field models use average firing rates. Converting individual spikes into an average firing rate is a straightforward reduction of dimensionality: one simply needs to average over multiple neurons in the model. In contrast, converting an average population firing rate into individual spikes in multiple neurons requires an increase in dimensionality. This is a degenerate problem, so additional assumptions must be made. One approach, described in Robinson and Kim (2012), is to treat each neuron as a phase oscillator. The average firing rate then represents the instantaneous rate of phase change, with a given neuron firing whenever its phase advances by 2π radians. However, here we used an alternative approach, in which the average firing rate is taken as the instantaneous rate for an ensemble of Poisson processes. These are then used to generate individual spike times (Dayan and Abbott, 2001; Leblois et al., 2006; Chadderdon et al., 2012). This approach produces variability in spike timings even with a constant average firing rate, as is seen in real neuronal populations.

2. METHODS

The model we used consisted of a network of spiking neurons that was “embedded” in a neural field model. The embedding consisted of having the field model generate spikes (via an ensemble of Poisson processes) that were used to drive the network model. Except where otherwise noted, all analyses were performed on the network model. The complete model is publicly available via ModelDB: <https://senselab.med.yale.edu/modeldb/ShowModel.asp?model=147366>.

2.1. NEURAL FIELD MODEL

The neural field model was based on the work of Van Albada and Robinson (2009) and Van Albada et al. (2009). The neuronal populations and connections that constitute this model are shown in **Figures 1** and **3A** respectively. The basal ganglia nuclei modeled were the striatum, internal and external pallidal segments, and STN. The internal pallidal population can be thought of as including the substantia nigra pars reticulata, which has very similar connections and properties. The substantia nigra pars compacta was not explicitly modeled, except through its effects on the other nuclei. The thalamus was modeled as two populations: the

inhibitory thalamic reticular nucleus (TRN) and the excitatory thalamocortical relay nuclei (TCR). The cortex was also modeled as two populations, representing inhibitory interneurons and excitatory pyramidal neurons. Since together these neuronal populations comprise a large portion of the brain, a network formulation would be computationally intractable. Except for a unitless normalization constant, all parameter values were based on anatomical and physiological data, as listed in Table 2 of Van Albada and Robinson (2009).

In neural field models, neuronal properties are spatially averaged. The dynamics are then governed by a set of equations relating the mean firing rates of populations of neurons to changes in mean cell-body potential, which are in turn triggered by mean rates of incoming spikes. The neural field model used here was based on a previously published model of the electrophysiology of the thalamocortical system (Robinson et al., 1997, 2001, 2002, 2005; Rennie et al., 1999), which in turn was based on earlier field models (Wilson and Cowan, 1973; Nunez, 1974; Freeman, 1975; Steriade et al., 1990; Wright and Liley, 1996).

The first component of the model is the description of the average response of populations of neurons to changes in mean cell-body potential. The mean firing rate Q_a of each population a is the maximum attainable firing rate Q_a^{\max} times the proportion of neurons with a membrane potential V_a above the mean threshold potential θ_a . This can be approximated by the sigmoid function

$$Q_a(\mathbf{r}, t) = \frac{Q_a^{\max}}{1 + e^{-[V_a(\mathbf{r}, t) - \theta_a]/\sigma'}}, \quad (1)$$

where \mathbf{r} is the spatial coordinate, t is time, and σ' is $\sqrt{3}/\pi$ times the standard deviation of the distribution of firing thresholds (Wright and Liley, 1995). This function increases smoothly from 0 to Q_a^{\max} as V_a changes from $-\infty$ to ∞ .

The change in the mean cell-body potential due to afferent activity depends on the mean number of synapses N_{ab} from neurons of population b to neurons of population a (note that the direction of projection $b \rightarrow a$ follows the conventions of control theory and matrix multiplication). The change in potential also depends on s_{ab} , the time-integrated change in cell-body potential per incoming spike. Defining $v_{ab} = N_{ab}s_{ab}$, the change in the mean cell-body potential in neurons of population a is (Robinson et al., 2004).

$$D_{\alpha\beta}(t)V_a(t) = \sum_b v_{ab}\phi_b(t - \tau_{ab}), \quad (2)$$

$$D_{\alpha\beta}(t) = \frac{1}{\alpha\beta} \frac{d^2}{dt^2} + \left(\frac{1}{\alpha} + \frac{1}{\beta}\right) \frac{d}{dt} + 1. \quad (3)$$

Here, $\phi_b(t - \tau_{ab})$ is the incoming firing rate, τ_{ab} represents the axonal time delay for signals traveling from population b to population a neurons, and α and β are the decay and rise rates of mean cell-body potential. The differential operator $D_{\alpha\beta}(t)$ represents dendritic and synaptic integration of incoming signals (Robinson et al., 1997; Rennie et al., 2000). The synapses and dendrites form an effective low-pass filter with a cut-off frequency between $1/\alpha$ and $1/\beta$.

In this model, neuronal activity spreads along the cortex in a wavelike fashion. This reflects previous models (Nunez, 1995; Jirsa and Haken, 1996; Bressloff, 2001) as well as experimental observations of such waves following cortical stimulation (Burns, 1951; Nunez, 1974; Rubino et al., 2006). Estimates of characteristic axonal ranges and propagation speeds suggest that these waves are significantly damped on the scale of the human cortex (Robinson et al., 2001, 2004; Wright and Liley, 1995). Assuming that the range distribution of corticocortical fibers decays exponentially at large distances, activity propagates according to a 2D damped-wave equation of the form (Robinson et al., 1997)

$$Q_a(\mathbf{r}, t) = \left[\frac{1}{\gamma_a^2} \frac{\partial^2}{\partial t^2} + \frac{2}{\gamma_a} \frac{\partial}{\partial t} + 1 - r_a^2 \nabla^2 \right] \phi_a(\mathbf{r}, t), \quad (4)$$

where $\gamma_a = v_a/r_a$ is the damping rate, consisting of the average axonal transmission speed v_a ($\simeq 10 \text{ m}\cdot\text{s}^{-1}$) and the characteristic axonal range r_a . In practice, most types of axons are short enough to justify setting $\gamma_a = \infty$, which has been termed the local interaction approximation (Robinson et al., 2004). We therefore take only γ_e , the damping rate of cortical

pyramidal neurons, to be finite. This turns all wave equations except the cortical one into delayed one-to-one mappings. The model was implemented on a 5×5 grid of nodes with coupling to nearest-neighbor nodes via this damped-wave equation.

2.2. SPIKING NETWORK MODEL

The spiking network was based on several previous models developed by our group (Lytton and Stewart, 2005; Lytton et al., 2008b; Neymotin et al., 2011b; Kerr et al., 2012; Song et al., 2013). It consisted of 4950 event-driven integrate-and-fire neurons. These were divided into three types (excitatory pyramidal cells *E*, fast-spiking inhibitory interneurons *I*, and low-threshold spiking interneurons *IL*), which were in turn distributed across the six layers of the cortex, plus two thalamic cell populations (excitatory thalamocortical relay *TCR* and inhibitory thalamic reticular *TRN*), for 15 distinct neuronal populations in total. The numbers and locations of each neuronal population are illustrated in **Figure 2**, and were as follows: E2 (i.e., excitatory pyramidal neurons of layer 2/3), 1500; I2, 250; IL2, 150; E4, 300; I4, 200; IL4, 150; E5R, 650; E5B, 150; I5, 250; IL5, 150; E6, 600; I6, 250; IL6, 150; TCR, 100; and TRN, 100. The pyramidal neurons

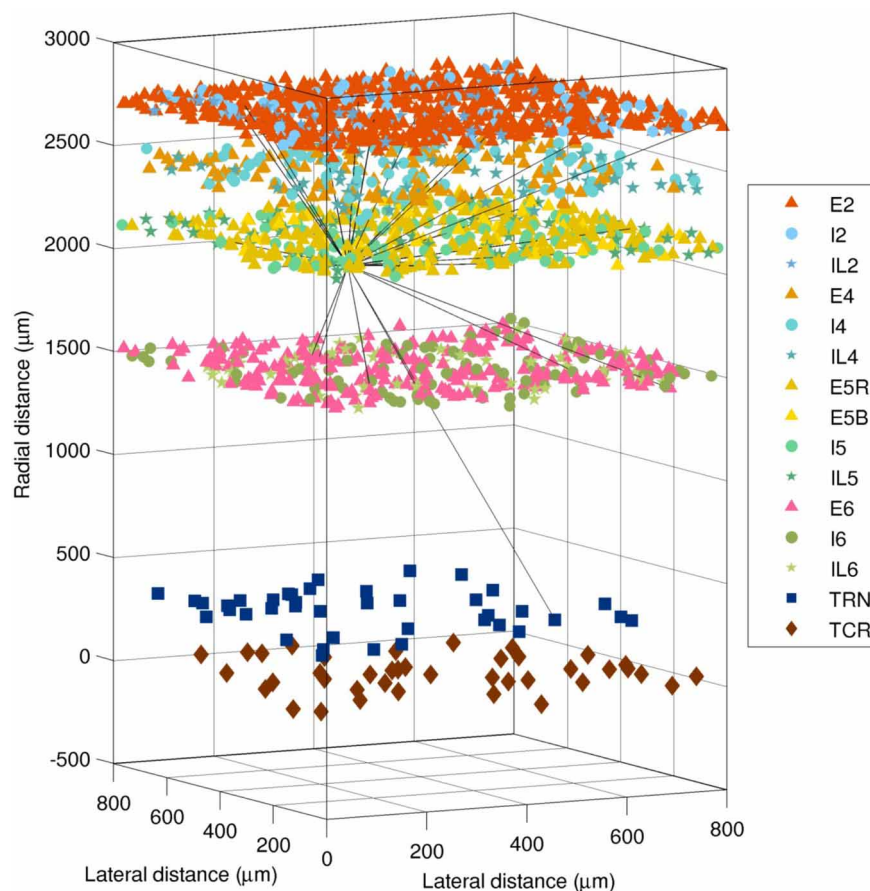


FIGURE 2 | Layout of the 4950 neurons in the spiking network model (1980 cells shown). Shapes show type (triangle = excitatory pyramidal, E; circle = fast-spiking interneuron, I; star = low-threshold spiking interneuron, IL;

square = thalamic reticular, TRN; diamond = thalamocortical relay, TCR). The 28 efferent connections from a single layer 5 pyramidal neuron are shown (black lines). The distance from the thalamus to the cortex is not shown to scale.

in layer 5 are divided into two populations, *R* (regular firing) and *B* (bursting), since these have different cellular properties and connectivity patterns.

Connectivity (shown in **Figure 3B**) and the relative numbers of neurons per layer were based on published models (Traub et al., 2005; Neymotin et al., 2011a,b) and anatomical studies (Thomson et al., 2002; Binzegger et al., 2004; Song et al., 2005; Lefort et al., 2009; Adesnik and Scanziani, 2010). Connectivity was strongest between populations within a given layer, as seen from the four clusters visible along the diagonal of **Figure 3B**. Overall, excitatory neurons had more projections than inhibitory ones, but inhibitory projections were typically stronger. This balanced excitation and inhibition such that the overall gain of the system (the number of additional output spikes for every additional input spike) was close to unity. Such balance is necessary for avoiding the stable but undesirable states of seizure

(pathologically high firing) and quiescence (pathologically low firing).

Individual neurons were modeled as event-driven, rule-based units. Since computing resources are finite, a tradeoff must be made between the complexity of neurons vs. the complexity of the network. The neuron model used was complex enough to replicate key features found in real neurons, including adaptation, bursting, depolarization blockade, and voltage-sensitive NMDA conductance (Lytton and Stewart, 2005, 2006; Lytton and Omurtag, 2007; Lytton et al., 2008a,b; Neymotin et al., 2011b), yet was simple enough to connect into large (10^3 – 10^6 neuron) networks.

Each neuron had a membrane voltage state variable (V_m) with a baseline value determined by a resting membrane potential parameter (V_{RMP} , set at -65 mV for pyramidal neurons and low-threshold-spiking interneurons, and at -63 mV for fast-spiking interneurons). This membrane voltage was updated by one of three events: synaptic input, threshold spike generation, and refractory period. These events are described briefly below; further detail can be found in the papers and code cited above.

2.2.1. Synaptic input

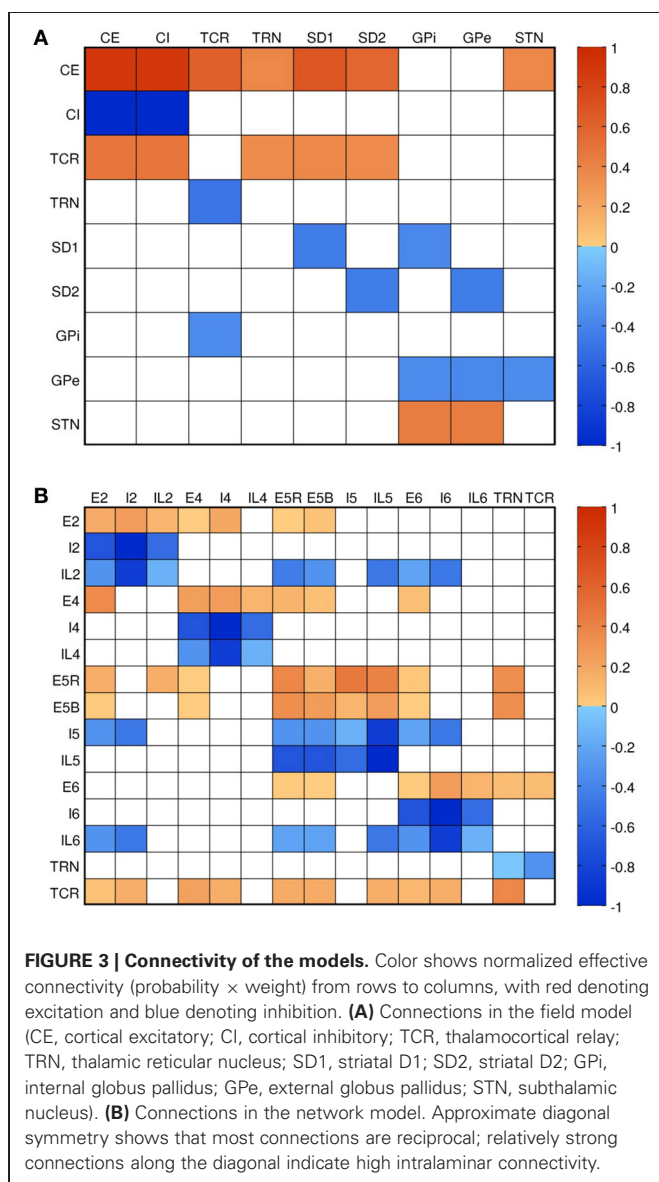
The response of the membrane voltage to synaptic input was modeled as an instantaneous rise and exponential decay: $V_n(t) = V_n(t_0) + w_s(1 - V_n(t_0)/E_i)e^{-\frac{t-t_0}{\tau_i}}$, where V_n is the membrane voltage of neuron n ; t_0 is the synaptic event time (i.e., $t - t_0$ is the time since the event); w_s is the weight of synaptic connection s ; E_i is the reversal potential of ion channel i , relative to resting membrane potential (where i = AMPA, NMDA, or GABA_A; and $E_{AMPA} = 65$ mV, $E_{NMDA} = 90$ mV, and $E_{GABA_A} = -15$ mV); and τ_i is the receptor time constant for ion channel i (where $\tau_{AMPA} = 20$ ms; $\tau_{NMDA} = 30$ ms; and $\tau_{GABA_A} = 10$ or 20 ms for somatic and dendritic GABA_A, respectively).

2.2.2. Action potentials

A neuron fires an action potential at time t if $V_n(t) > T_n(t)$ and $V_n(t) < B_n$, where V_n , T_n , and B_n are the membrane voltage, threshold voltage (-40 mV for pyramidal neurons and fast-spiking interneurons, -47 mV for low-threshold-spiking interneurons), and blockade voltage (-10 mV for interneurons and -25 mV for pyramidal neurons), respectively, for neuron n . Action potentials arrive at target neurons at time $t_2 = t_1 + l(n_1, n_2)/v + \tau_s$, where t_1 is the time the first neuron fired, τ_s is the delay due to synaptic conduction effects, $l(n_1, n_2)$ is the axon length between neurons n_1 and n_2 , and v is the axonal conduction velocity (≈ 1 m·s $^{-1}$, which is smaller than in the field model, since long-range fibers tend to be more heavily myelinated).

2.2.3. Refractory period

After firing, a neuron cannot fire during the absolute refractory period, τ_A (10 ms for interneurons and 50 ms for pyramidal neurons). Firing is reduced during the relative refractory period by two effects: first, an increase in threshold potential, $T_n(t) = \left(1 + Re^{-\frac{t-t_0}{\tau_R}}\right) T_n(t_0)$, where R is the fractional increase in threshold voltage due to the relative refractory period (0.25 for interneurons and 0.75 for pyramidal neurons) and τ_R is its



time constant (1.5 ms for interneurons and 8 ms for pyramidal neurons); and second, by hyperpolarization, $V_n(t) = V_n(t_0) - He^{-\frac{t-t_0}{\tau_H}}$, where H is the amount of hyperpolarization (0.5 mV for interneurons and 1 mV for pyramidal neurons) and τ_H is its time constant (50 ms for interneurons and 400 ms for pyramidal neurons).

Local field potentials (LFPs) were computed for each cortical layer as the average membrane voltage across all neurons in that layer; after baseline removal and normalization, this approach is roughly equivalent to summing over all synaptic currents (Mazzoni et al., 2010). While this approach does not take into consideration synaptic and dendritic geometry, this is not possible in the event-driven point-neuron model used here.

Simulations were run in NEURON 7.3 (Hines and Carnevale, 2001; Carnevale and Hines, 2006) on a Linux workstation with an Intel Xeon 2.7 GHz CPU; each 20 s simulation took approximately 10 min to run on a single core. To avoid edge effects, the first and last 2 s of simulated data were discarded. All analyses were performed on the remaining 16 s of simulated data. Since the model is at steady-state and does not incorporate plasticity effects, longer runs produced similar results (data not shown). Model parameters were tuned manually (within physiological limits) to match experimentally observed firing rates, dynamics, and information-theoretic properties, as described in Song et al. (2013).

2.3. INPUT DRIVE

The composite model consisted of the spiking network model being driven by ("embedded in") the activity of the field model. Since the field model represents a brain region much larger than the network model, the field causally influences the network, but not vice versa. The key methodological novelty of this work is that the spiking network model is thus embedded in an environment with physiologically realistic dynamics (as provided by the field model), rather than the white noise environment such models are typically embedded in.

To obtain realistic firing rates in the network model, the input spiking rate each neuron receives must be bounded. Hence, the firing rate from each neuronal population in the field model was normalized so that the minimum and maximum input spiking rates were 225 and 1125 s^{-1} for excitatory neurons and 30% lower for inhibitory neurons. The input drive was obtained by treating each of these normalized instantaneous firing rates as the rate of an ensemble of Poisson processes for generating spikes. These spikes were then used to drive each population of spiking neurons, using the same connections as used in the field model itself (e.g., excitatory cortical neurons in the network model received input from the excitatory cortical field, the inhibitory cortical field, and the thalamic field); relative connection weights were also set to match those of the field model. Thus, each neuron belonging to a given population in the network model receives the same average rate of input from the field model, but from a separate Poisson process, thereby avoiding artificial correlations in input spike times between neurons.

Four different inputs were explored in this work. First, all neurons in the network were driven by spikes drawn from a spectrally

white distribution ("WN", the white noise model). This represents the control condition, and is identical to the approach used in previous work with the network model (Neymotin et al., 2011b). Second, neurons were driven by the thalamocortical version of the field model ("TC", the thalamocortical model); i.e., connection strengths to and from the basal ganglia neuronal populations were set to zero. Third, neurons were driven by the full basal ganglia/thalamocortical model described above ("BG", the healthy basal ganglia model). Finally, neurons were driven by the full basal ganglia/thalamocortical model, using parameter values shown by Van Albada and Robinson (2009) to best match the electrophysiological changes associated with the degeneration of dopaminergic projections to the striatum ("PD", the Parkinson's disease model).

2.4. ANALYSIS

2.4.1. Fano factor

The Fano factor, a common measure of spiking variability (Churchland et al., 2010), is the ratio of the variance to the mean of the spike rate:

$$F = \frac{\sigma^2}{\mu}, \quad (5)$$

where σ is the standard deviation and μ is the mean of the time series of binned spiking activity across all neurons. To explore spiking variability on a range of different time scales, the time bin size was varied from 1 ms (resulting in 16,000 bins, with an average of roughly 10 spikes per bin) to 8 s (resulting in 2 bins, with roughly 80,000 spikes per bin).

2.4.2. Population burst probability

A population burst (Benayoun et al., 2010) was defined as ≥ 2 neurons firing within a given 10 ms time bin. The probability of a burst of size N was defined as the number of time bins with N cells firing divided by the total number of time bins. The relative burst probability was calculated by dividing the observed number of bursts of each size by the number of bursts of that size expected from uncorrelated activity, which in turn was determined via the observed firing rate (averaged over the entire simulation) and the binomial probability distribution.

2.4.3. Spectral granger causality

Information flow was quantified in terms of spectral Granger causality, also called the directed transfer function (Kaminski et al., 2001). Although many alternative tools for inferring causality exist, such as directed transfer entropy (Lizier et al., 2011), no others allow the spectral properties of the signals to be analyzed in detail.

As in standard Granger causality analysis, spectral Granger causality of $\alpha(f) \rightarrow \beta(f)$ is non-zero if prior knowledge of variable α at frequency f reduces error in the prediction of β at frequency f . The directionality of the causation arises from the fact that Granger causality quantifies how much the history of time series α can be used to predict the future of time series β : if α has a strong causal influence on β , then the prediction error will be reduced.

Spectral Granger causality is calculated by Fourier transforming the multivariate autoregressive model used in standard

Granger causality. Hence, the spectral Granger causality from time series $\alpha(t)$ to time series $\beta(t)$ is defined as (Cui et al., 2008)

$$G_{\alpha \rightarrow \beta}(f) = -\log \left(1 - \frac{\left(N_{\alpha\alpha} - \frac{N_{\beta\alpha}^2}{N_{\alpha\alpha}} \right) |H_{\beta, \alpha}(f)|^2}{S_{\beta, \beta}(f)} \right), \quad (6)$$

where N is the noise covariance, $H(f)$ is the transfer function, and $S(f)$ the spectral matrix, as derived from the bivariate autoregressive model of $\alpha(t)$ and $\beta(t)$. This analysis was performed in Matlab 2012a using code based on the BSMART toolbox, available via <http://www.brain-smart.org>.

3. RESULTS

The neural field model results were similar to those reported previously (Van Albada and Robinson, 2009; Van Albada et al., 2009), and are briefly presented here for completeness. We then present the overall dynamics of the spiking network model (Kerr et al., 2012), comparing its dynamics for each of the four drives (white noise, the thalamocortical model, the healthy basal ganglia model, and the PD model). Finally, we focus more closely on the alterations that occur in the PD-driven model and their implications. We have split the results into these sections in order to better accomplish our dual goals of (1) presenting the new composite model, and (2) applying this model to help understand the pathophysiology of PD.

3.1. FIELD MODEL DYNAMICS

Firing rates in each neuronal population were similar to those reported previously (Van Albada and Robinson, 2009; Van Albada et al., 2009). Because the drive from the field model to the network model was normalized to a range that provided realistic firing rates in the latter, tonic firing rates had negligible effect on the simulations.

Changes in coherence are a commonly reported finding in PD. In the PD model, coherence between the GPe and the GPi was lost, and high frequency power (>10 Hz) in the GPi was reduced (Figure 4). In the healthy state, activity in the GPe and GPi is strongly correlated ($r^2 = 0.9$). Following dopamine loss, this correlation is substantially reduced ($r^2 = 0.3$). This is because the GPe and GPi are both mostly influenced by the striatum in the healthy state, whereas the GPi is strongly driven by the STN in the parkinsonian state, resulting in strong coherence between the GPi and STN in the PD model. Increased STN-GPi coherence at frequencies up to about 35 Hz has indeed been found in PD off levodopa compared to the on-levodopa condition (Brown et al., 2001). Since the GPi is the only nucleus of the basal ganglia that projects to the thalamus or cortex (Figure 1), all changes observed in the network model in the healthy versus PD cases are due to the altered dynamics of the GPi.

To characterize the overall dynamics of the different field models, we looked at their power spectra. In the absence of the basal ganglia, cortical excitatory neurons had a strong alpha peak (10 Hz), and a weaker harmonic in the beta range (20 Hz), as shown in Figure 4B. Cortical inhibitory neurons were driven strongly by thalamocortical cells, evident both in the phase

locking between the two populations (Figure 4A), and in the similarity of their power spectra below 70 Hz (Figure 4B). The addition of the basal ganglia (Figure 4B, middle panel) reduced the strength of the alpha peak in cortical excitatory neurons and reduced the slope of the power law spectral fall-off at high frequencies; in cortical excitatory neurons, this slope changed from $P(f) \propto f^{-5.3}$ to $P(f) \propto f^{-4.3}$. Reduced dopamine corresponding to PD reduced the power of higher frequencies (>10 Hz) relative to lower frequencies (<10 Hz) in the cortical, thalamic, and GPi spectra. For example, the GPi showed a 2% decrease in power at 10 Hz compared to a 76% decrease at 20 Hz. In contrast, reduced dopamine increased power in the STN at frequencies >10 Hz (e.g., 2.2 times larger at 20 Hz), a result also reported experimentally (Brown et al., 2001; Cassidy et al., 2002; Priori et al., 2004).

3.2. NETWORK MODEL DYNAMICS

The field drive into the network model strongly modulated its spiking activity (Figure 5). Firing rates varied from near zero during the troughs of input activity to >10 Hz during the peaks (Figure 5A). The temporal structure of the spiking activity depended strongly on the type of input drive used (Figure 5B). As a control, white noise produced no consistent temporal structure. The TC-driven model input produced some structure, with a characteristic time scale below 500 ms. The BG-driven model added some features on longer time scales (of order 1 s) to the activity produced by the TC-only field model. Variability in firing rate, as measured by the Fano factor, was lowest in the WN-driven model (Figure 5D)—as would be expected since the white noise had the lowest variability of the four inputs. On time scales <1 s, the PD-driven model had the greatest variability, while the BG-driven model had the greatest variability on scales >1 s.

The power spectra of the network model, shown in Figure 6A, were broadly similar to those of the input drives, but with several interesting differences. The basic filter properties of the network model are apparent from the shape of spectrum of the WN-driven model; to a first approximation, the network acts like a low-pass filter, with $P(f) \propto f^{-4.0}$ for $f > 20$ Hz. However, actual afferent activity in the brain is already low-pass filtered due to dendritic properties, so a more realistic input (the thalamocortical drive) results in even greater low-pass filtering. For example, the WN-driven model predicts 5.4 times more power at 10 Hz than the TC-driven model. Both BG- and PD-driven models differed markedly from the TC-driven model in the 20–30 Hz band, where many basal ganglia nuclei have their peak power. Interestingly, this peak was much sharper in the network model than in the input drive, demonstrating a resonance effect (compare Figure 4A with Figure 6B).

To quantify synchrony in the model on a population level, we used population burst size (Benayoun et al., 2010). All of the field-driven models showed substantially higher population bursting than the WN-driven model (Figure 6B). This is because the field drive applies a global modulatory signal to the network, which organizes the firing of its neurons into up and down states (as evident from the bands of spikes in Figure 5A); in contrast, the WN-driven model has a constant, intermediate level of activation.

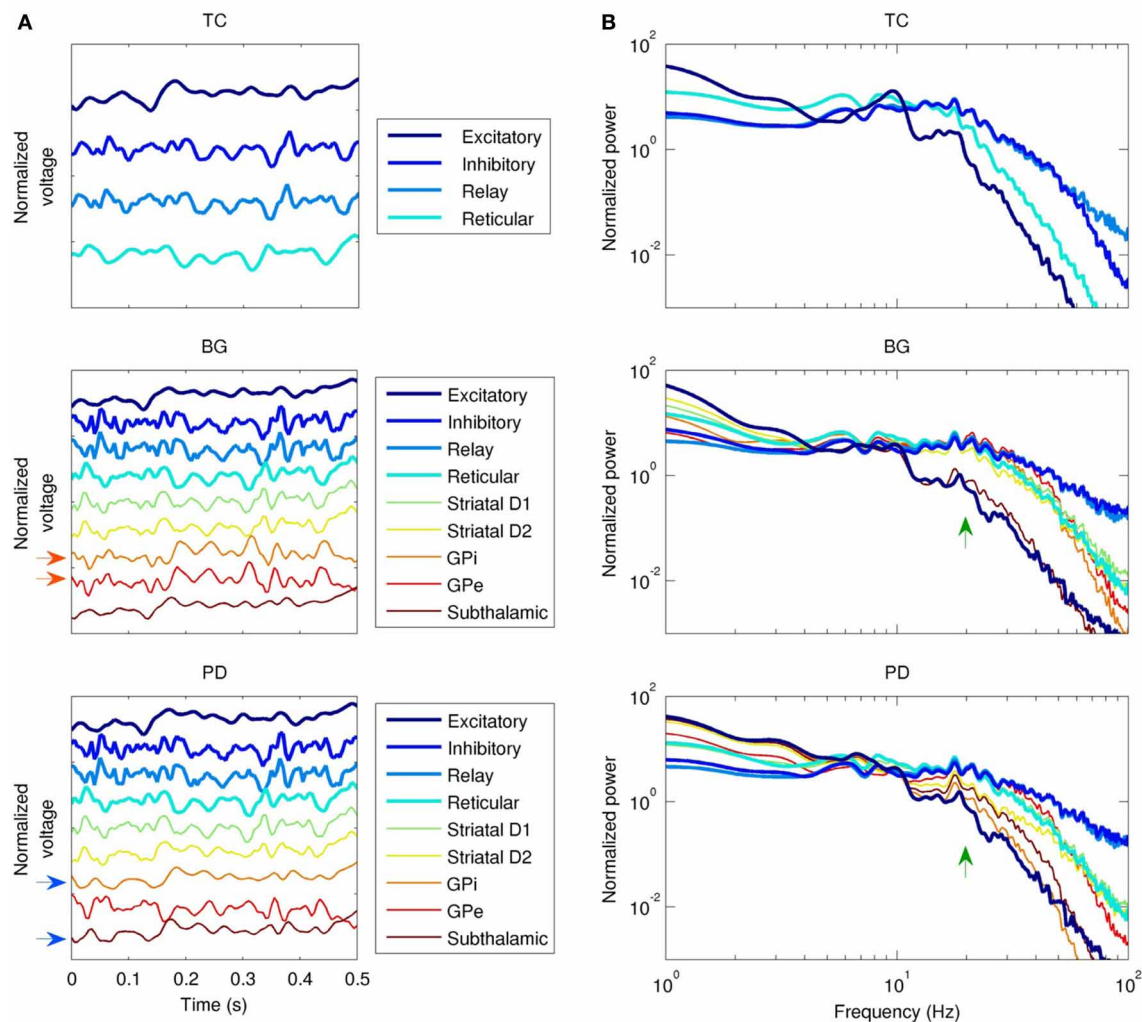


FIGURE 4 | Dynamics of the three field models (without the network model). TC, thalamocortical field model; BG, healthy basal ganglia model; PD, Parkinson's disease model; white noise model not shown. "Excitatory" and "inhibitory" refer to cortical subpopulations. (A) Local field potential (LFP) time series, showing phase relationships between populations. Activity in the globus pallidus internal (GPi) and external

(GPe) segments is normally in phase (red arrows), but this relationship is lost in PD, since the GPi entrains to the subthalamic nucleus instead (blue arrows). (B) LFP spectra. Except for the subthalamic nucleus, healthy basal ganglia nuclei spectra are similar to the spectrum of the thalamic relay nuclei from 10–40 Hz. This is disrupted in PD (green arrows), especially in the GPi.

3.3. DYNAMICAL CHANGES IN THE PARKINSON'S DISEASE MODEL

The Parkinson's disease model (PD-driven model) showed a number of changes that suggest possible mechanisms underlying the clinical dysfunctions of the disease. Compared to the healthy control (BG-driven model), the PD-driven model showed a shift in the LFP spectrum toward lower frequencies, with higher delta power and a lower beta peak frequency (Figure 6), consistent with clinical findings (Stoffers et al., 2007). These changes were also readily apparent looking at the LFP time series, which showed a flattening of activity between the slow, high-amplitude features (Figure 5B). Soikkeli et al. (1991) noted such slowing in 10 out of 18 non-demented PD patients, as well as in all 18 demented PD patients studied [see Figure 1 in Soikkeli et al. (1991)].

The PD-driven model showed an $18 \pm 2\%$ decrease in firing rates compared to the healthy model (Figure 5C), consistent with changes in fMRI indicators of activity (Monchi et al., 2007). The PD-driven model also showed greater firing variability than the healthy model on most time scales. For example, with a bin size of 1 ms, the Fano factor was 41% higher in the PD-driven model (Figure 5D). However, it showed less variability on very long time scales: with a bin size of 8 s, the Fano factor was 2.4 times higher in the BG-driven model. The increased variability in the PD-driven model on all but the longest timescales is consistent with the enhanced oscillations and synchrony associated with PD (Goldberg et al., 2002). Note that maximal dynamical richness does not necessarily correspond to maximal variability in firing rates: for example, tonic firing will have low dynamical richness

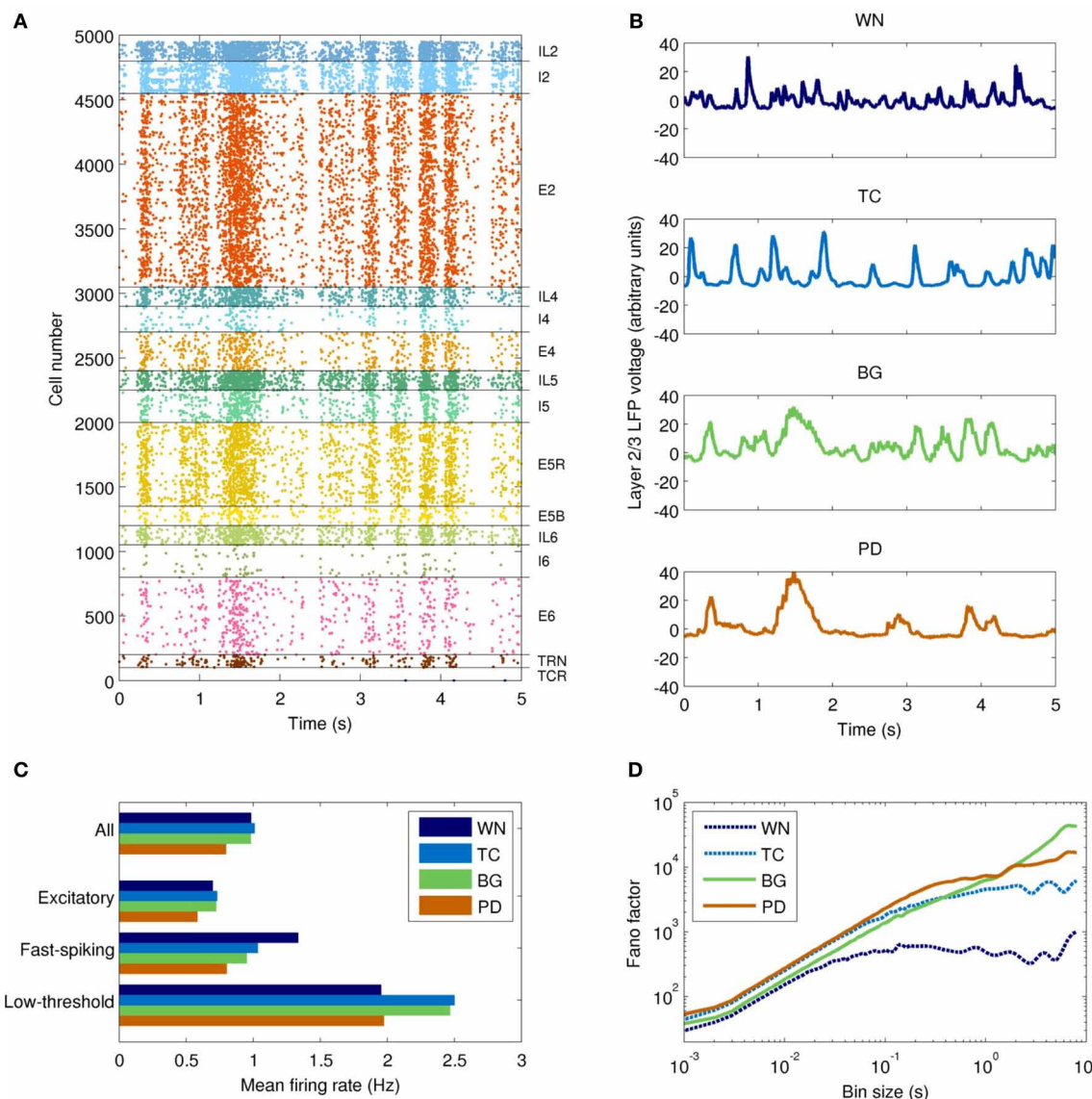


FIGURE 5 | Temporal dynamics of the network model with each type of input drive (WN, white noise; TC, thalamocortical; BG, healthy thalamocortical/basal ganglia; PD, Parkinson's disease). (A) Example spike raster from the BG-driven model. Low-frequency oscillations are clearly visible. **(B)** LFPs from layer 2/3 of each model. The BG case corresponds to the raster shown in **(A)**; peaks in voltage are correlated with peaks in spiking activity. **(C)** Mean firing rates by cell type (averaged over both cortical and

thalamic populations). Overall, the PD-driven model had considerably lower firing rates, which result from excessive inhibition of the thalamic nuclei.

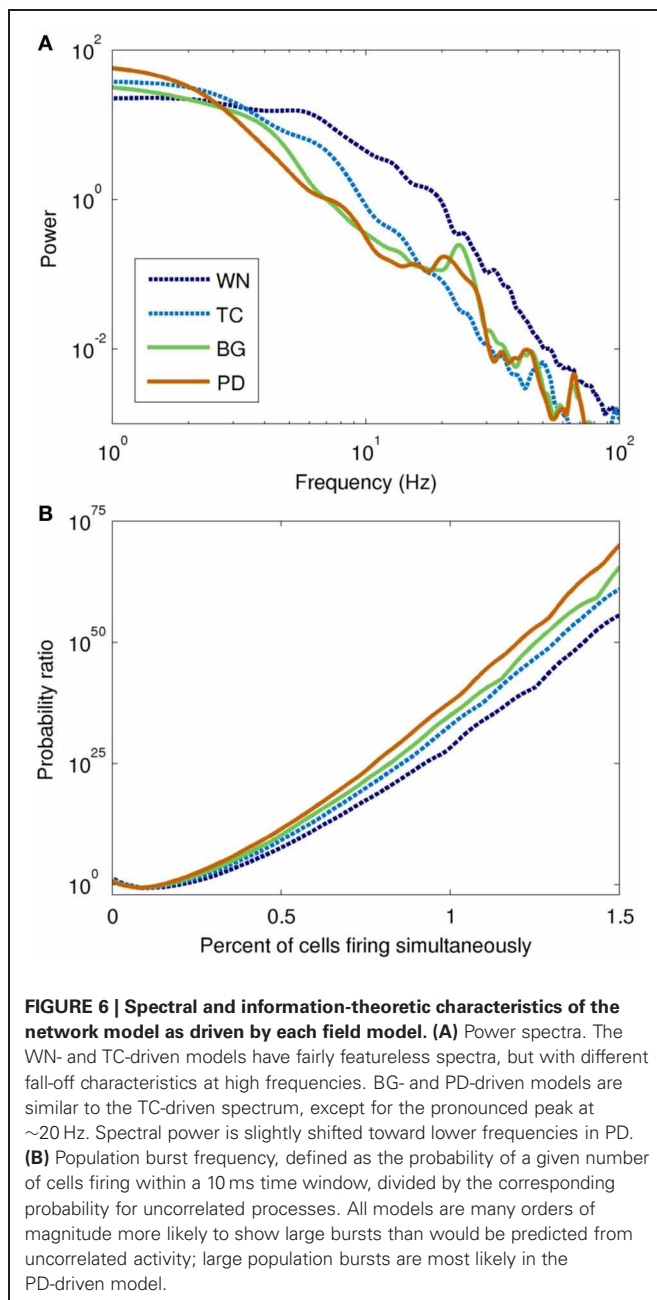
(D) Variability in neuronal firing rates on different time scales. The PD- and BG-driven models (which receive the most highly structured input) show the most variability on short and long time scales, respectively; the WN-driven model (which receives the least structured input) shows the least variability on all scales.

and low variability on all time scales, while strong, seizure-like oscillations will also have low dynamical richness, despite very high variability (at least on the time scale of the oscillation).

The concentration of activity in large population bursts was a prominent feature of the PD-driven model. For example, bursts consisting of 40 neurons were 60% more common in the PD-driven model than in the healthy model, while 70-neuron bursts were three orders of magnitude more common. (Population bursts smaller than 30 neurons were more common in the healthy model, a result of its higher firing rate.) Although it is tempting

to consider these large population events in the context of parkinsonian tremor, we did not note a clear periodicity in their occurrence.

A crucial question in PD is the mechanism by which information flow is disrupted from higher cortical areas (e.g., those involved in motor planning) to primary areas (e.g., those involved in motor execution). Although information flow between cortical layers is bidirectional, a dominant direction of information flow is suggested by both anatomical and functional studies (Bollimunta et al., 2008). This dominant information pathway is believed to



stream from thalamic inputs to layer 4 (or upper layer 5 in agranular motor cortices), up to layer 2/3 for processing, and thence to layer 5, which in turn produces outputs to multiple sites including the thalamus, basal ganglia, and brainstem. We hypothesized that damage to this dominant pathway would represent a pathology with major functional consequences. We therefore used Granger causality to quantify information flow between the cortical layers that comprise this pathway.

Overall, interlaminar spectral Granger causality was highest in the BG-driven model, and lowest in the WN-driven model (Figure 7). Most notably, the BG-driven model showed a prominent peak in causality in the high-beta/low-gamma band (20–35 Hz). This peak was almost entirely absent in the

PD-driven model; for example, peak causality from layer 4 to layer 5 in this frequency range was only half that of the BG-driven model (0.23 and 0.45 for PD- and BG-driven models, respectively), even though these models had similar spectral power (Figure 6A). As shown in Figure 7, similar results were seen in other layer pairs (e.g., 4 → 2/3, 2/3 → 5, and 6 → 2/3).

4. DISCUSSION

We have explored the effects of driving a spiking network model with several different types of input, including those corresponding to the healthy brain and to PD. Many of the differences between the healthy and PD models accord with prior experimental findings. For example, we found a modest but consistent reduction in firing rates of cortical neurons in PD. Although there are no direct studies of cortical firing rates during PD in humans, several indirect measures from functional imaging suggest such a decrease (Jenkins et al., 2004; Monchi et al., 2004, 2007). We also found a shift toward lower LFP frequencies, a finding consistent with PD electroencephalography (Soikkeli et al., 1991; Bosboom et al., 2006; Stoffers et al., 2007). We found increased synchrony between neurons in our PD model, as measured by population burst size and probability; increased synchrony among basal ganglia neurons is a commonly reported finding in PD (Raz et al., 1996), and increased synchrony among cortical neurons has also been reported (Goldberg et al., 2002).

Our major finding was the loss of Granger causality between cortical layers in the high-beta/low-gamma band. The Granger causality for the PD-driven model was more similar to the TC- and WN-driven models than to the BG-driven model, suggesting that the dynamical properties of the basal ganglia that facilitate cortical information flow are almost entirely lost in PD. The frequency range of this disrupted information flow is thought to be crucial for encoding motor commands, especially limb movements (Van Der Werf et al., 2008; Muthukumaraswamy, 2010). Gamma has also been implicated in many cognitive processes (Fries et al., 2007), including the perceptual binding underlying sensorimotor coordination (Lee et al., 2003) and consciousness (Llinas et al., 1998). Hence, our observation of disrupted causality might also partially account for some of the cognitive symptoms of PD, including bradyphrenia and planning deficits (Morris et al., 1988; Chaudhuri and Schapira, 2009).

The fact that Granger causality was disrupted in the PD-driven model (Figure 7) while the power spectrum was nearly unchanged in the same frequency band (Figure 6A) shows that the changed input drive has reorganized the dynamics of the network in complex ways. Since the GPi does not project directly to the cortex, these changes are entirely mediated by the thalamus; indeed, thalamic lesions alone are sufficient for producing parkinsonian symptoms in rats (Oehrns et al., 2007). Since the thalamus projects differentially to the different layers of the cortex, a major change in thalamic input is sufficient explanation for why the causality would shift so dramatically. Specifically, the thalamus normally projects strongly to layer 4; the peak in causality at 20–35 Hz is consistent with thalamic modulation by the GPi. In PD, inhibition to the thalamus is increased, which results in weaker drive to the cortex and thus a loss of information flow.

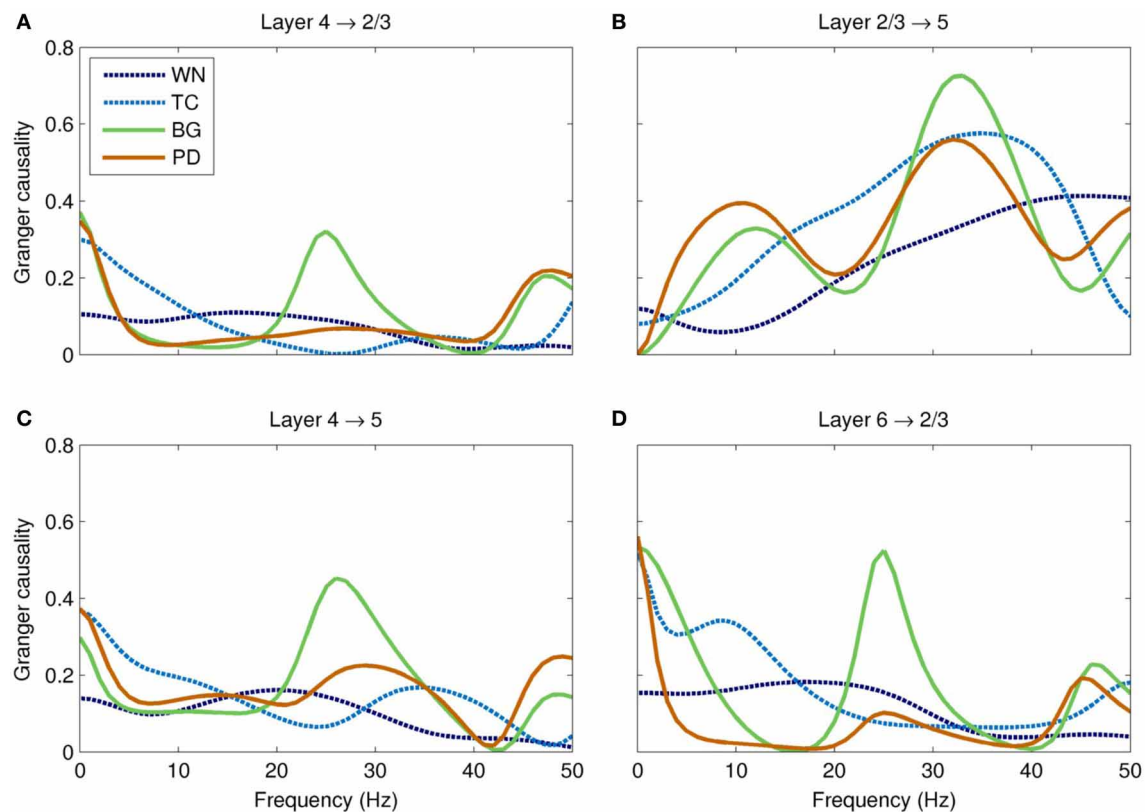


FIGURE 7 | Spectral Granger causality between cortical layers in each of the models. (A) The BG-driven model shows strong causality from layer 4 to 2/3 in the delta (<5 Hz) and high-beta/low-gamma (20–35 Hz) bands; causality in the latter band is almost entirely lost in Parkinson's disease. **(B)** The causality from layer 2/3 to layer 5 is slightly reduced in this band in Parkinson's disease. **(C)** These two effects combine to

significantly reduce the total Granger causality from layer 4 to layer 5 in PD, especially in the high-beta/low-gamma band. **(D)** Similar reductions of Granger causality in this band were seen in other layer pairs, such as layer 6 to layer 2/3. In each case, the high-beta/low-gamma band Granger causality is significantly higher in the BG-driven model than in any of the other models.

Our findings suggest that therapeutic interventions, such as deep brain stimulation (Deuschl et al., 2006), may be more effective if they restore both the dynamics and the tonic level of activity of the GPi, rather than just the latter.

Several of our findings are qualitatively consistent with experimental results pointing to a loss of complexity in EEG time series from patients with a variety of cognitive disorders, including PD (Stam et al., 1994, 1995; Vaillancourt and Newell, 2002). For example, in the healthy model, the slope of the Fano factor increases roughly linearly on time scales from 1 ms to 10 s, indicating dynamical structure across a wide range of time scales (Figure 5D). This result can be seen qualitatively in the LFP time series of the healthy model, which appeared to show meaningful structure over a broader range of time scales than any of the other models (Figure 5B). We speculate that these properties may reflect the number of possible states that the network can assume, which may in turn be related to the number of different motor programs that can be implemented by the network. This principle is closely related to the concept of ϕ , defined as “the repertoire of causal states available to a system as a whole” (Balduzzi and Tononi, 2008). While ϕ cannot be easily computed for moderately large networks such as ours, we

expect that it will be manifested in terms of the network's ability to perform real motor tasks—a topic we will explore in future work. Specifically, we predict that the BG-driven model will perform better on simulated reaching tasks than the WN-, TC-, or PD-driven models.

Beta-band activity (15–30 Hz) was predominantly generated by the thalamic and inhibitory cortical neuronal populations in our model (Figure 4B, top panel), in agreement with previous experimental and modeling studies (Brown and Williams, 2005; Hahn and McIntyre, 2010). Most empirical studies of beta activity in PD have focused on the basal ganglia nuclei, with increased power in the STN being a commonly reported finding (Brown and Williams, 2005; Kühn et al., 2005; Weinberger et al., 2006). In our model, we found that beta power in the STN was indeed enhanced in PD (Figure 4B, middle and bottom panels), which may reflect an idling or antikinetic state (Brown and Williams, 2005; Engel and Fries, 2010).

4.1. LIMITATIONS

Several experimentally observed features of PD, such as increased coherence among neurons in the STN, can only be explicitly represented using a neuronal network model of the basal ganglia

(Terman et al., 2002)—a major benefit of that modeling approach. However, it is not known whether these phenomena are causally linked to parkinsonian symptoms. Hence, in the present context, the benefits of using a neural field model for the basal ganglia outweigh the drawbacks of this approach. In future, a spiking network model of the basal ganglia would be desirable in order to account for these and other phenomena, such as reinforcement learning. An explicit representation of dopamine in such a mode—rather than the implicit representation used here—would also allow the effects of pharmacological interventions to be modeled directly.

Due to the eloquence of the motor system, movement disorders are the most obvious symptoms of PD. Yet the pathophysiology of the disease is widespread; even the retina is affected (Hajee et al., 2009). We stress that the spiking network model used here was designed as a model of association cortex, not primary motor cortex; for example, our model includes layer 4 cells, which are absent from the latter. However, since the thalamus and striatum have broad projections to the cortex, we expect the dynamical and information-theoretic changes in PD (such as increased synchrony and reduced complexity) to extend to motor areas as well. In the future, we will explore the effects of PD in a model of primary motor cortex controlling a virtual arm (Chadderdon et al., 2012), with the aim of directly demonstrating classical parkinsonian motor symptoms. By incorporating sensory feedback into this model, the white noise that was used to drive the neural field component can be replaced with more realistic input, thereby addressing another obvious limitation of the method used here.

4.2. MULTISCALE DYNAMICS IN A COMPOSITE MODEL

To our knowledge, this work represents the first composite spiking network/neural field model of the brain. This is a multiscale model that spans spatial scales from 10 μm to 30 cm and temporal scales from 1 ms to tens of seconds. The composite method provides a way of linking two types of models that provide access to different spatial scales—a network model that spans scales from individual neurons (10 μm) to a cortical column (600 μm), and a field model encompassing the whole diencephalon (30 cm).

REFERENCES

- Adesnik, H., and Scanziani, M. (2010). Lateral competition for cortical space by layer-specific horizontal circuits. *Nature* 464, 1155–1160.
- Ananthanarayanan, R., Esser, S. K., Simon, H. D., and Modha, D. S. (2009). “The cat is out of the bag: cortical simulations with 10^9 neurons, 10^{13} synapses,” in *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, SC '09* (New York, NY: ACM), 1–12.
- Balduzzi, D., and Tononi, G. (2008). Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS Comput. Biol.* 4:e1000091. doi: 10.1371/journal.pcbi.1000091
- Benayoun, M., Cowan, J. D., Van Drongelen, W., and Wallace, E. (2010). Avalanches in a stochastic model of spiking neurons. *PLoS Comput. Biol.* 6:e1000846. doi: 10.1371/journal.pcbi.1000846
- Binzegger, T., Douglas, R. J., and Martin, K. A. (2004). A quantitative map of the circuit of cat primary visual cortex. *J. Neurosci.* 24, 8441–8453.
- Bolam, J. P., Hanley, J. J., Booth, P. A. C., and Bevan, M. D. (2002). Synaptic organisation of the basal ganglia. *J. Anat.* 196, 527–542.
- Bollimunta, A., Chen, Y., Schroeder, C. E., and Ding, M. (2008). Neuronal mechanisms of cortical alpha oscillations in awake-behaving macaques. *J. Neurosci.* 28, 9976–9988.
- Bosboom, J. L., Stoffers, D., Stam, C. J., Van Dijk, B. W., Verbunt, J., Berendse, H. W., et al. (2006). Resting state oscillatory brain dynamics in Parkinson's disease: an MEG study. *Clin. Neurophysiol.* 117, 2521–2531.
- Bressloff, P. C. (2001). Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Physica D* 155, 83–100.
- Brown, P., Oliviero, A., Mazzone, P., Insola, A., Tonalì, P., and Di Lazzaro, V. (2001). Dopamine dependency of oscillations between subthalamic nucleus and pallidum in Parkinson's disease. *J. Neurosci.* 21, 1033–1038.
- Brown, P., and Williams, D. (2005). Basal ganglia local field potential activity: character and functional significance in the human. *Clin. Neurophysiol.* 116, 2510–2519.
- Burns, B. D. (1951). Some properties of isolated cerebral cortex in the unanaesthetized cat. *J. Physiol.* 112, 156–175.
- Carnevale, N. T., and Hines, M. L. (2006). *The NEURON Book*. Cambridge, UK: Cambridge University Press.
- Cassidy, M., Mazzone, P., Oliviero, A., Insola, A., Tonalì, P., Di Lazzaro, V., et al. (2002). Movement-related changes in synchronization in the human basal ganglia. *Brain* 125, 1235–1246.
- Chadderdon, G. L., Neymotin, S. A., Kerr, C. C., and Lytton, W. W. (2012). Reinforcement learning of targeted movement in a spiking neuronal model of motor

Temporally, both network and field models are valid over many orders of magnitude (approximately 10^{-3} – 10^4 s).

The mechanism used here to couple the field and network models is just one of several alternatives (Wilson et al., 2012). In the present case, the coupling was unidirectional; the network model did not affect the dynamics of the field model. While this can be easily justified in terms of the effective size of each model, an alternative approach generates the neural field based on the dynamics of the network model, using the new neuron-in-cell approach of Robinson and Kim (2012). Because spiking network models are still limited in their capacity to generate accurate dynamics on a large scale, this approach cannot yet be used in place of neural field models. However, this may change if scientific advances and improved computing facilities enable the development of larger and more realistic spiking network models.

Many spiking network models that are too small to show self-sustaining activity are driven by white noise (Hill and Tononi, 2005; Vogels and Abbott, 2005; Oswald et al., 2009; McDonnell et al., 2011; Volman et al., 2011; Kerr et al., 2012; Muller and Destexhe, 2012; Vijayan and Kopell, 2012). Here we demonstrated that using physiologically realistic input instead of white noise has a major impact on multiple measures of network activity, including power spectra, spiking variability, burst probability, and Granger causality. Thus, white-noise-driven spiking network models are an abstraction away from the physiological environment, and should perhaps be considered as being analogous to artificially driven slice preparations rather than *in vivo* activity.

ACKNOWLEDGMENTS

Cliff C. Kerr, Samuel A. Neymotin, George L. Chadderdon, and William W. Lytton were supported by DARPA grant N66001-10-C-2008; P. A. Robinson was supported by the Australian Research Council and the Westmead Millennium Institute; and Sacha J. van Albada was supported by EU FP7 Grant 269921 (BrainScaleS) and the Helmholtz Alliance on Systems Biology. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

- cortex. *PLoS ONE* 7:e47251. doi: 10.1371/journal.pone.0047251
- Chaudhuri, K. R., and Schapira, A. H. (2009). Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment. *Lancet Neurol.* 8, 464–474.
- Churchland, M. M., Byron, M. Y., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* 13, 369–378.
- Cools, R. (2006). Dopaminergic modulation of cognitive function: implications for L-DOPA treatment in Parkinson's disease. *Neurosci. Biobehav. Rev.* 30, 1–23.
- Cui, J., Xu, L., Bressler, S. L., Ding, M., and Liang, H. (2008). BSMAT: a Matlab/C toolbox for analysis of multichannel neural time series. *Neural Netw.* 21, 1094–1104.
- Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience*, Vol. 31. Cambridge, MA: MIT Press.
- Deco, G. and Jirsa, V. K. (2012). Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *J. Neurosci.* 32, 3366–3375.
- Destexhe, A., and Sejnowski, T. J. (2009). The Wilson–Cowan model, 36 years later. *Biol. Cybern.* 101, 1–2.
- Deuschl, G., Schade-Brittinger, C., Krack, P., Volkmann, J., Schäfer, H., Bötzel, K., et al. (2006). A randomized trial of deep-brain stimulation for parkinson's disease. *N. Engl. J. Med.* 355, 896–908.
- Dovzhenok, A., Park, C., Worth, R. M., and Rubchinsky, L. L. (2013). Failure of delayed feedback deep brain stimulation for intermittent pathological synchronization in Parkinson's disease. *PLoS ONE* 8:e58264. doi: 10.1371/journal.pone.0058264
- Engel, A. K., and Fries, P. (2010). Beta-band oscillations – signalling the status quo? *Curr. Opin. Neurobiol.* 20, 156–165.
- Freeman, W. (1975). *Mass Action in the Nervous System*. New York, NY: Academic Press.
- Fries, P., Nikolic, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci.* 30, 309–316.
- Goldberg, J. A., Boraud, T., Maraton, S., Haber, S., Vaadia, E., and Bergman, H. (2002). Enhanced synchrony among primary motor cortex neurons in the 1-methyl-4-phenyl-1, 2, 3, 6-tetrahydropyridine primate model of Parkinson's disease. *J. Neurosci.* 22, 4639–4653.
- Guo, Y., and Rubin, J. E. (2011). Multi-site stimulation of subthalamic nucleus diminishes thalamocortical relay errors in a biophysical network model. *Neural Netw.* 24, 602–616.
- Hahn, P. J., and McIntyre, C. C. (2010). Modeling shifts in the rate and pattern of subthalamopallidal network activity during deep brain stimulation. *J. Comput. Neurosci.* 28, 425–441.
- Hajee, M. E., March, W. F., Lazzaro, D. R., Wolintz, A. H., Shrier, E. M., Glazman, S., et al. (2009). Inner retinal layer thinning in Parkinson disease. *Arch. Ophthalmol.* 127, 737–741.
- Hill, S., and Tononi, G. (2005). Modeling sleep and wakefulness in the thalamocortical system. *J. Neurophysiol.* 93, 1671–1698.
- Hines, M. L., and Carnevale, N. T. (2001). NEURON: a tool for neuroscientists. *Neuroscientist* 7, 123–135.
- Humphries, M. D., Stewart, R. D., and Gurney, K. N. (2006). A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *J. Neurosci.* 26, 12921–12942.
- Izhikevich, E. M., and Edelman, G. M. (2008). Large-scale model of mammalian thalamocortical systems. *Proc. Natl. Acad. Sci. U.S.A.* 105, 3593–3598.
- Jenkins, I. H., Fernandez, W., Playford, E. D., Lees, A. J., Frackowiak, R. S. J., Passingham, R. E., et al. (2004). Impaired activation of the supplementary motor area in Parkinson's disease is reversed when akinesia is treated with apomorphine. *Ann. Neurol.* 32, 749–757.
- Jirsa, V. K., and Haken, H. (1996). Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* 77, 960–963.
- Kaminski, M., Ding, M., Truccolo, W. A., and Bressler, S. L. (2001). Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biol. Cybern.* 85, 145–157.
- Kerr, C. C., Neymotin, S. A., Chadderdon, G. L., Fietkiewicz, C. T., Francis, J. T., and Lytton, W. W. (2012). Electrostimulation as a prosthesis for repair of information flow in a computer model of neocortex. *IEEE Trans. Neural Syst. Rehabil. Eng.* 20, 153–160.
- Kühn, A. A., Trottenberg, T., Kivi, A., Kupsch, A., Schneider, G.-H., and Brown, P. (2005). The relationship between local field potential and neuronal discharge in the subthalamic nucleus of patients with parkinson's disease. *Exp. Neurol.* 194, 212–220.
- Leblois, A., Boraud, T., Meissner, W., Bergman, H., and Hansel, D. (2006). Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *J. Neurosci.* 26, 3567–3583.
- Lee, K. H., Williams, L. M., Breakspear, M., and Gordon, E. (2003). Synchronous gamma activity: a review and contribution to an integrative neuroscience model of schizophrenia. *Brain Res. Rev.* 41, 57–78.
- Lefort, S., Tómm, C., Floyd Sarria, J.-C., and Peterson, C. C. H. (2009). The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. *Neuron* 61, 301–316.
- Lizier, J., Heinze, J., Horstmann, A., Haynes, J.-D., and Prokopenko, M. (2011). Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. *J. Comput. Neurosci.* 30, 85–107.
- Llinas, R., Ribary, U., Contreras, D., and Pedraarena, C. (1998). The neuronal basis for consciousness. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1841–1849.
- Lumer, E. D., Edelman, G. M., and Tononi, G. (1997). Neural dynamics in a model of the thalamocortical system. I. Layers, loops and the emergence of fast synchronous rhythms. *Cereb. Cortex* 7, 207–227.
- Lytton, W. W., Neymotin, S. A., and Hines, M. L. (2008a). The virtual slice setup. *J. Neurosci. Methods* 171, 309–315.
- Lytton, W. W., Omurtag, A., Neymotin, S. A., and Hines, M. L. (2008b). Just-in-time connectivity for large spiking networks. *Neural Comput.* 20, 2745–2756.
- Lytton, W. W., and Omurtag, A. (2007). Tonic-clonic transitions in computer simulation. *J. Clin. Neurophysiol.* 24, 175–181.
- Lytton, W. W., and Stewart, M. (2005). A rule-based firing model for neural networks. *Int. J. Bioelectromagnetism* 7, 47–50.
- Lytton, W. W., and Stewart, M. (2006). Rule-based firing for network simulations. *Neurocomputing* 69, 1160–1164.
- Mazzoni, A., Whittingstall, K., Brunel, N., Logothetis, N. K., and Panzeri, S. (2010). Understanding the relationships between spike rate and delta/gamma frequency bands of LFPs and EEGs using a local cortical network model. *Neuroimage* 52, 956–972.
- McDonnell, M. D., Mohan, A., Stricker, C., and Ward, L. M. (2011). Input-rate modulation of gamma oscillations is sensitive to network topology, delays and short-term plasticity. *Brain Res.* 1434, 162–177.
- Monchi, O., Petrides, M., Doyon, J., Postuma, R. B., Worsley, K., and Dagher, A. (2004). Neural bases of set-shifting deficits in Parkinson's disease. *J. Neurosci.* 24, 702–710.
- Monchi, O., Petrides, M., Mejia-Constain, B., and Strafella, A. P. (2007). Cortical activity in Parkinson's disease during executive processing depends on striatal involvement. *Brain* 130, 233–244.
- Morris, R. G., Downes, J. J., Sahakian, B. J., Evenden, J. L., Heald, A., and Robbins, T. W. (1988). Planning and spatial working memory in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 51, 757–766.
- Muller, L., and Destexhe, A. (2012). Propagating waves in thalamus, cortex and the thalamocortical system: experiments and models. *J. Physiol. Paris* 106, 222–238.
- Muthukumaraswamy, S. D. (2010). Functional properties of human primary motor cortex gamma oscillations. *J. Neurophysiol.* 104, 2873–2885.
- Neymotin, S. A., Jacobs, K. M., Fenton, A. A., and Lytton, W. W. (2011a). Synaptic information transfer in computer models of neocortical columns. *J. Comput. Neurosci.* 30, 69–84.
- Neymotin, S. A., Lee, H., Park, E., Fenton, A. A., and Lytton, W. W. (2011b). Emergence of physiological oscillation frequencies in a computer model of neocortex. *Front. Comput. Neurosci.* 5:19. doi: 10.3389/fncom.2011.00019
- Niedermeyer, E., and Lopes da Silva, F. (1999). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Baltimore, MD: Urban and Schwarzenberg.
- Nunez, P. L. (1974). Wave-like properties of the alpha rhythm. *IEEE Trans. Biomed. Eng.* 21, 473–482.
- Nunez, P. L. (1995). *Neocortical Dynamics and Human EEG Rhythms*. Oxford: Oxford University Press.
- Oehrn, C., Allbutt, H., and Henderson, J. (2007). Effect of ventrolateral thalamic nucleus lesions in the unilateral 6-hydroxydopamine rat model. *Behav. Brain Res.* 183, 67–77.
- Oswald, A. M. M., Doiron, B., Rinzel, J., and Reyes, A. D. (2009). Spatial profile and differential recruitment of GABA_B modulate oscillatory activity in auditory cortex. *J. Neurosci.* 29, 10321–10334.
- Park, C., Worth, R. M., and Rubchinsky, L. L. (2011). Neural

- dynamics in parkinsonian brain: the boundary between synchronized and nonsynchronized dynamics. *Phys. Rev. E* 83:042901. doi: 10.1103/PhysRevE.83.042901
- Priori, A., Foffani, G., Pesenti, A., Tamma, F., Bianchi, A. M., Pellegrini, M., et al. (2004). Rhythm-specific pharmacological modulation of subthalamic activity in Parkinson's disease. *Exp. Neurol.* 189, 369–379.
- Raz, A., Feingold, A., Zelanskaya, V., Vaadia, E., and Bergman, H. (1996). Neuronal synchronization of tonically active neurons in the striatum of normal and parkinsonian primates. *J. Neurophysiol.* 76, 2083–2088.
- Rennie, C. J., Robinson, P. A., and Wright, J. J. (1999). Effects of local feedback on dispersion of electrical waves in the cerebral cortex. *Phys. Rev. E* 59, 3320–3329.
- Rennie, C. J., Wright, J. J., and Robinson, P. A. (2000). Mechanisms of cortical electrical activity and emergence of gamma rhythm. *J. Theor. Biol.* 205, 17–35.
- Robinson, P. A., Loxley, P. N., O'Connor, S. C., and Rennie, C. J. (2001). Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. *Phys. Rev. E* 63:041909. doi: 10.1103/PhysRevE.63.041909
- Robinson, P. A., and Kim, J. W. (2012). Spike, rate, field, and hybrid methods for treating neuronal dynamics and interactions. *J. Neurosci. Methods* 205, 283–294.
- Robinson, P. A., Rennie, C. J., and Rowe, D. L. (2002). Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Phys. Rev. E* 65:041924. doi: 10.1103/PhysRevE.65.041924
- Robinson, P. A., Rennie, C. J., Rowe, D. L., and O'Connor, S. C. (2004). Estimation of multiscale neurophysiologic parameters by electroencephalographic means. *Hum. Brain Mapp.* 23, 53–72.
- Robinson, P. A., Rennie, C. J., Rowe, D. L., O'Connor, S. C., and Gordon, E. (2005). Multiscale brain modelling. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1043–1050.
- Robinson, P. A., Rennie, C. J., and Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* 56, 826–840.
- Rubchinsky, L. L., Kopell, N., and Sigvardt, K. A. (2003). Modeling facilitation and inhibition of competing motor programs in basal ganglia subthalamic nucleus-pallidal circuits. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14427–14432.
- Rubino, D., Robbins, K. A., and Hatsopoulos, N. G. (2006). Propagating waves mediate information transfer in the motor cortex. *Nat. Neurosci.* 9, 1549–1557.
- Soikkeli, R., Partanen, J., Soininen, H., Pääkkönen, A., and Riekkinen, P. (1991). Slowing of EEG in Parkinson's disease. *Electroencephalogr. Clin. Neurophysiol.* 79, 159–165.
- Song, S., Sjöström, P., Reigl, M., Nelson, S., and Chklovskii, D. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* 3:e68. doi: 10.1371/journal.pbio.0030068
- Song, W., Kerr, C. C., Lytton, W. W., and Francis, J. T. (2013). Cortical plasticity induced by spike-triggered microstimulation in primate somatosensory cortex. *PLoS ONE* 8:e57453. doi: 10.1371/journal.pone.0057453
- Stam, C. J., Jelles, B., Achtereekte, H. A. M., Rombouts, S., Slaets, J. P. J., and Keunen, R. W. M. (1995). Investigation of EEG non-linearity in dementia and Parkinson's disease. *Electroencephalogr. Clin. Neurophysiol.* 95, 309–317.
- Stam, K. J., Tavy, D. L. J., Jelles, B., Achtereekte, H. A. M., Slaets, J. P. J., and Keunen, R. W. M. (1994). Non-linear dynamical analysis of multi-channel EEG: clinical applications in dementia and Parkinson's disease. *Brain Topogr.* 7, 141–150.
- Steriade, M., Gloor, P., Llinás, R. R., Lopes da Silva, F. H., and Mesulam, M. M. (1990). Basic mechanisms of cerebral rhythmic activities. *Electroencephalogr. Clin. Neurophysiol.* 76, 481–508.
- Stoffers, D., Bosboom, J. L. W., Deijen, J. B., Wolters, E. C., Berendse, H. W., and Stam, C. J. (2007). Slowing of oscillatory brain activity is a stable characteristic of Parkinson's disease without dementia. *Brain* 130, 1847–1860.
- Terman, D., Rubin, J. E., Yew, A. C., and Wilson, C. J. (2002). Activity patterns in a model for the subthalamic network of the basal ganglia. *J. Neurosci.* 22, 2963–2976.
- Thomson, A., West, D., Wang, Y., and Bannister, A. (2002). Synaptic connections and small circuits involving excitatory and inhibitory neurons in layers 2–5 of adult rat and cat neocortex: triple intracellular recordings and biocytin labelling *in vitro*. *Cereb. Cortex* 12, 936–953.
- Traub, R. D., Contreras, D., Cunningham, M. O., Murray, H., LeBeau, F. E., Roopun, A., et al. (2005). Single-column thalamo-cortical network model exhibiting gamma oscillations, sleep spindles, and epileptogenic bursts. *J. Neurophysiol.* 93, 2194–2232.
- Vaillancourt, D. E., and Newell, K. M. (2002). Changing complexity in human behavior and physiology through aging and disease. *Neurobiol. Aging* 23, 1–11.
- Van Albada, S. J., Gray, R. T., Drysdale, P. M., and Robinson, P. A. (2009). Mean-field modeling of the basal ganglia-thalamocortical system. II. Dynamics of parkinsonian oscillations. *J. Theor. Biol.* 257, 664–688.
- Van Albada, S. J., and Robinson, P. A. (2009). Mean-field modeling of the basal ganglia-thalamocortical system. I. Firing rates in healthy and parkinsonian states. *J. Theor. Biol.* 257, 642–663.
- Van Der Werf, J., Jensen, O., Fries, P., and Medendorp, W. P. (2008). Gamma-band activity in human posterior parietal cortex encodes the motor goal during delayed prosaccades and antisaccades. *J. Neurosci.* 28, 8397–8405.
- Vijayan, S., and Kopell, N. J. (2012). Thalamic model of awake alpha oscillations and implications for stimulus processing. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18553–18558.
- Vogels, T. P., and Abbott, L. F. (2005). Signal propagation and logic gating in networks of integrate-and-fire neurons. *J. Neurosci.* 25, 10786–10795.
- Volman, V., Bazhenov, M., and Sejnowski, T. J. (2011). Pattern of trauma determines the threshold for epileptic activity in a model of cortical deafferentation. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15402–15407.
- Weinberger, M., Mahant, N., Hutchison, W. D., Lozano, A. M., Moro, E., Hodaie, M., et al. (2006). Beta oscillatory activity in the subthalamic nucleus and its relation to dopaminergic response in Parkinson's disease. *J. Neurophysiol.* 96, 3248–3256.
- Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80.
- Wilson, M. T., Robinson, P. A., O'Neill, B., and Steyn-Ross, D. A. (2012). Complementarity of spike-and-rate-based dynamics of neural systems. *PLoS Comput. Biol.* 8:e1002560. doi: 10.1371/journal.pcbi.1002560
- Wright, J. J., and Liley, D. T. J. (1995). Simulation of electrocortical waves. *Biol. Cybern.* 72, 347–356.
- Wright, J. J., and Liley, D. T. J. (1996). Dynamics of the brain at global and microscopic scales: neural networks and the EEG. *Behav. Brain Sci.* 19, 285–294.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 January 2013; accepted: 02 April 2013; published online: 25 April 2013.

Citation: Kerr CC, Van Albada SJ, Neymotin SA, Chadderdon GL, Robinson PA and Lytton WW (2013) Cortical information flow in Parkinson's disease: a composite network/field model. *Front. Comput. Neurosci.* 7:39. doi: 10.3389/fncom.2013.00039

Copyright © 2013 Kerr, Van Albada, Neymotin, Chadderdon, Robinson and Lytton. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



The mesoscopic modeling of burst suppression during anesthesia

David T. J. Liley* and Matthew Walsh

Brain and Psychological Sciences Research Centre, Faculty of Life and Social Sciences, Swinburne University of Technology, Hawthorn, VIC, Australia

Edited by:

Peter Beim Graben,
Humboldt-Universität zu Berlin,
Germany

Reviewed by:

Axel Hutt, INRIA CR Nancy, France
Moira Steyn-Ross, University of
Waikato, New Zealand

*Correspondence:

David T. J. Liley, Brain and
Psychological Sciences Research
Centre, Faculty of Life and Social
Sciences, Swinburne University of
Technology, P.O. Box 218, Hawthorn,
VIC 3122, Australia.
e-mail: dliley@swin.edu.au

The burst-suppression pattern is well recognized as a distinct feature of the mammalian electroencephalogram (EEG) waveform. Consisting of alternating periods of high amplitude oscillatory and isoelectric activity, it can be induced in health by deep anesthesia as well as being evoked by a range of pathophysiological processes that include coma and anoxia. While the electroencephalographic phenomenon and clinical implications of burst suppression have been studied extensively, the physiological mechanisms underlying its emergence remain unresolved and obscure. Because electroencephalographic bursting phenomenologically resembles the bursting observed in single neurons, it would be reasonable to assume that the theoretical insights developed to understand bursting at the cellular ("microscopic") level would enable insights into the dynamical genesis of bursting at the level of the whole brain ("macroscopic"). In general action potential bursting is the result of the interplay of two time scales: a fast time scale responsible for spiking, and a slow time scale that modulates such activity. We therefore hypothesize that such fast-slow systems dynamically underpin electroencephalographic bursting. Here we show that a well-known mean field dynamical model of the electroencephalogram, the Liley model, while unable to produce burst suppression unmodified, is able to give rise to a wide variety of burst-like activity by the addition of one or more slow systems modulating model parameters speculated to be major "targets" for anesthetic action. The development of a physiologically plausible theoretical framework to account for burst suppression will lead to a more complete physiological understanding of the EEG and the mechanisms that serve to modify ongoing brain activity necessary for purposeful behavior and consciousness.

Keywords: burst suppression, anesthesia, electroencephalogram, mean field model, neuronal hyperexcitability

1. INTRODUCTION

Prior to the development of the modern intensive care unit in the early 1960s, that featured intubation, artificial respiration, and comprehensive physiological monitoring, reports of the electroencephalographic pattern of burst suppression (BS) were confined to animal studies involving deep anesthesia and the occasional case of psychosurgery (Niedermeyer, 2009). Since then the burst-suppression pattern has become well recognized as a major diagnostic feature of the EEG waveform that is encountered in a range of encephalopathic conditions, in addition to its appearance in health during deep anesthesia. Typically the BS pattern consists of bursts of high amplitude slow, sharp, or spiking electroencephalographic activity separated by periods of electroencephalographic suppression (isoelectricity). The oscillatory features of the bursts, together with their duration and the duration of suppressed periods show a high degree of variability (see **Figure 1** for examples) that presumably reflects its myriad of initiating causes. First identified during deep anesthesia with tribromoethanol in cats (Derbyshire et al., 1936), labeled burst-suppression pattern by Swank and Watson (1949) during barbiturate and ether anesthesia in dogs, it is now associated with cortical deafferentation (Henry and Scoville, 1952), cerebral anoxia and hypoxia, various types of intracortical lesions (Fischer-Williams and Cooper, 1963), deep

coma, various infantile encephalopathies, the final stages of deteriorated status epilepticus (Treiman et al., 1990), hypothermia, and high levels of many sedative and anesthetic agents (Schwartz et al., 1989; Akrawi et al., 1996).

Burst suppression in the absence of anesthesia is in general associated with a very poor prognosis. For example in neonates (Grigg-Damberger et al., 1989) the appearance of BS, even if transient, is a portent of death or severe neurodevelopmental disability. In contrast, in adult populations while an anoxic/hypoxic BS pattern signals a serious pathophysiological event the outcome is not necessarily fatal and recovery with or without severe neurological damage is possible (Niedermeyer, 2009). Consistent with this are results of experimental work with EEG monitoring in rats revealing that animals with greater rates of high amplitude bursts have a better survival and neurological outcome compared to those with lower rates of low amplitude bursts (Geocadin et al., 2002).

While the electroencephalographic phenomenon and clinical implications of BS have been studied extensively (Brenner, 1985; Niedermeyer, 2009) the physiological mechanisms underlying its emergence remain in general unresolved and obscure.

Burst suppression is typically thought to be spatially homogeneous with burst onset and termination reported to occur near simultaneously across the entire scalp (Brenner, 1985; An et al.,

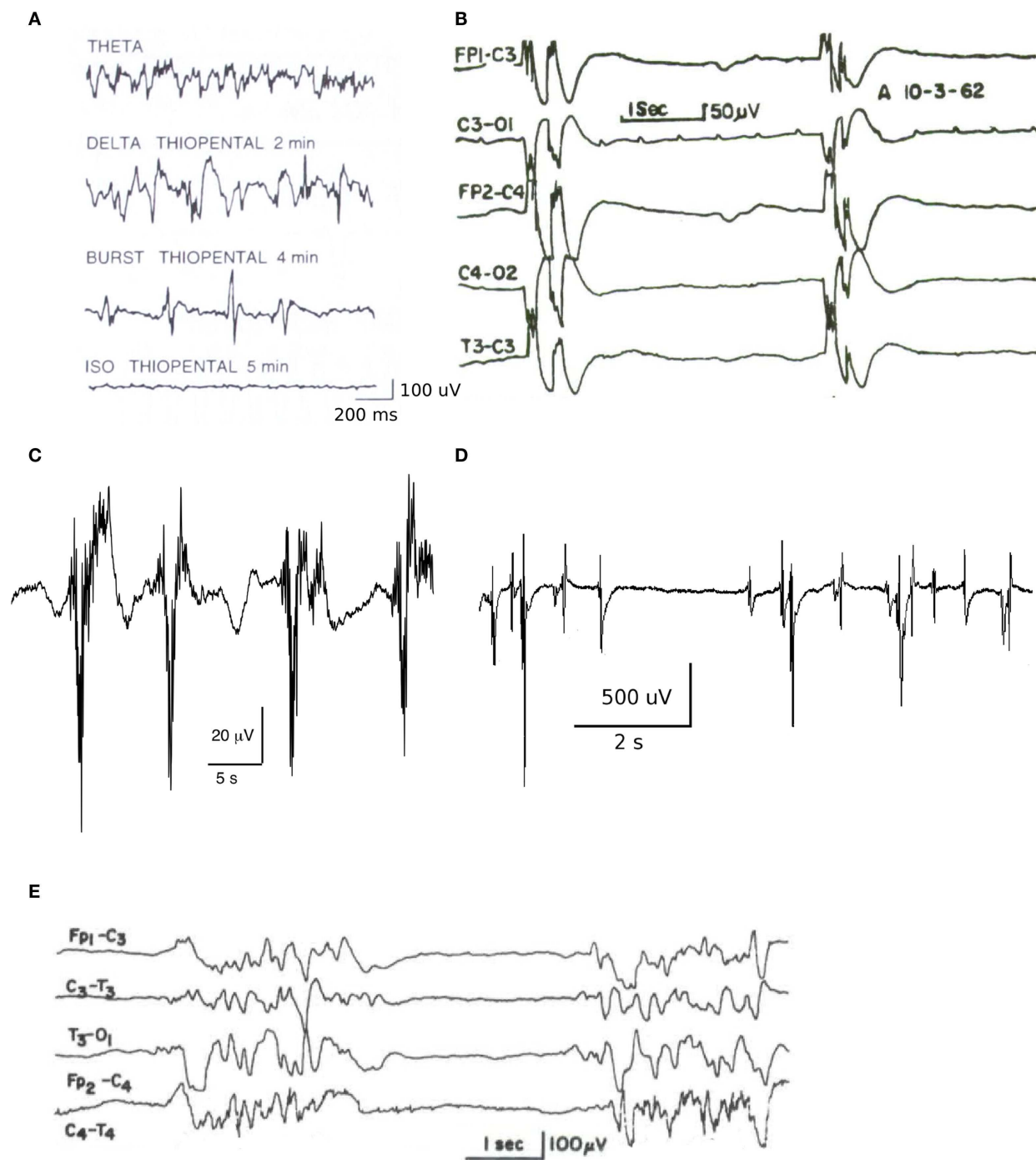


FIGURE 1 | Example traces of electroencephalogram and electrocorticogram illustrating the heterogeneity of BS patterns. (A) Changes in neocortical electroencephalogram in the rat, recorded using dural surface electrodes, in response to a 5 mg/kg/min thiopental infusion [figure reproduced with permission from Lukatch and MacIver (1996)]. (B) Electroencephalogram recorded in acute anoxia showing a clear burst-suppression pattern with grouped spikes [figure reproduced with permission from Hockaday et al. (1965)]. (C) BS pattern during closed loop target controlled propofol infusion at a target level of

approximately 15 μ g/ml (data courtesy of Professor Michel Struys, Groningen). Note the bursts consist of fast activity (>10 Hz) on a slow wave background. (D) Electrocorticogram obtained from an adult merino sheep during deep enflurane anesthesia, demonstrating high amplitude spikes interspersed with isoelectric periods of variable length [figure reproduced with permission from Voss et al. (2006)]. (E) Electroencephalogram recorded from a 3-month-old infant suffering from infantile myoclonic encephalopathy [reproduced with permission from Niedermeyer (2005)].

1996; Ching et al., 2012), indicating that low level subcortical mechanisms may be playing a decisive role. However arguing against this is the fact that this pattern persists subsequent to cortical deafferentation (Lukatch and MacIver, 1996), indicating that it probably represents an intrinsic, though physiologically abnormal, dynamical mode of cortex. Indeed the phenomenal resemblance of the patterns of BS to disorders of neuronal hyperexcitability suggests the involvement of similar physiological mechanisms. For example the bursting during burst suppression is often associated with myoclonic jerks resembling those seen during epileptic fits. Like generalized epileptiform activity, bursts are recorded simultaneously at multiple electrode derivations, implying the wide synchronization of neuronal activity.

At the cellular level a commonly reported finding is that hyperpolarization of the membrane potential of cortical neurons reliably precedes any overt electroencephalographic activity of BS (Steriade et al., 1994). Such hyperpolarization, which has been attributed to an increase in neuronal membrane potassium conductance (Steriade et al., 1994), has been hypothesized to play a major role in the induction of BS. This implied importance of inhibition in the genesis of BS is further supported by results involving rat neocortical brain slice micro-EEG preparations in which the application of a direct acting GABA_A agonist, muscimol, readily induces BS. However contradicting this result is work reporting that inhibition is diminished during isoflurane-induced BS, in an *in vivo* feline preparation, as evidenced by increases in cortical neuronal input resistance and extracellular chloride concentration (Ferron et al., 2009). Of course it may be that slow periodic modulations in inhibition, rather than singular increases or decreases in inhibition, underpin BS. In support of this view is the recent model of Ching et al. (2012), in which alterations in brain metabolism, due to the effects of hypoxia or anesthesia, parametrically regulate an activity dependent slow modulation of an adenosine triphosphate-gated potassium channel conductance to give rise to BS. However modulations in inhibitory activity alone may not be sufficient to account for BS and more consideration might need to be given to other mechanisms. For example Kroeger and Amzica (2007) present empirical evidence suggesting that modulations in excitatory synaptic efficiency, due to the progressive depletion of interstitial calcium during the periods of high amplitude electroencephalographic activity and its recovery during isoelectric periods, might account for BS. Consistent with this are reports involving laboratory slice preparations in which burst suppression induced by thiopental, propofol, and isoflurane is abolished by the application of glutamate receptor antagonists (Lukatch and MacIver, 1996). Whatever the pathophysiology of BS it is reasonably clear that it is unlikely to be accounted for by a unitary physiological perturbation. That the physiological factors identified to date in BS all lead to a single well defined state suggests the possibility of an unifying dynamical mechanism. Thus the best hope for progress in understanding the phenomena of BS may be theoretical.

How might we theoretically approach BS? The well studied dynamical mechanisms of action potential bursting (Izhikevich, 2007) may be able to provide vital insights into the mechanisms of bursting in the EEG. In general the dynamical mechanisms underlying bursting can be divided into two broad classes (i) fast-slow

bursters in which there is a clear separation of the underlying time scales, with a fast system responsible for the fast spiking, and a slow system its slow modulation, and (ii) “hedgehog” bursters (Izhikevich, 2000) in which there is no clear separation of time scales. In terms of developing a theory of BS the former might represent the preferable starting point as the little empirical evidence that is available (Ching et al., 2012), at least in humans, suggests that alphoid activity, indicative of normal resting EEG, is preserved during the bursts of BS. Thus a theoretical starting point to understanding BS might be to consider the slow modulation of a dynamical system developed to describe the resting EEG.

One such dynamical system is the mesoscopic electrocortical model of Liley et al. (Liley et al., 1999, 2002, 2011; Bojak and Liley, 2005; Frascoli et al., 2011). This model is capable of accounting for a range of resting electroencephalographic phenomena that includes the alpha rhythm (Liley et al., 2002), the modulation of resting activity by sedative and anesthetic action (Bojak and Liley, 2005) as well as the proconvulsant properties of the latter (Liley and Bojak, 2005), all within a physiologically plausible/admissible parameter space. This model is therefore well suited as a foundation from which to explore the physiological and dynamical genesis of BS. However, because in this model rhythmogenesis emerges from a strong coupling between cortical excitatory and inhibitory population activity, in its present form it has a restricted ability to exhibit BS through the parametric separation of time scales, either through the simulated actions of anesthetics or through other parametric routes. Here we show that BS can however emerge in this model by the addition of a slow system driven by one or more of the originally defined mean fields. We speculate that such a slow system represents a mathematical ansatz for the slow neuro-modulation of activity by a variety of intracortical, inter-cortical, and subcortical systems that include thalamus and the ascending neurotransmitter modulatory systems.

2. MATERIALS AND METHODS

2.1. MESOSCOPIC MEAN FIELD MODELING OF ELECTROCORTICAL ACTIVITY

The electroencephalogram and electrocorticogram arise out of the cooperative activity of many thousands of neurons. A single electroencephalographic electrode records the synaptically induced currents of well over a 100,000 neurons (Nunez and Srinivasan, 2005) and thus detailing each neurons contribution to this summed activity would appear superfluous. For this reason it is preferable instead to model the activity of populations of neurons. One general way of achieving this, in which known stochastic fluctuations can be included, is to dynamically evolve the probability distributions associated with the states of the neuronal ensemble. While in principle providing a rigorous way forward the formulation of such stochastic equations of motion entails a great deal of physiological uncertainty. For this and other reasons (Deco et al., 2008) a more resolute path is to dynamically evolve some average quantity such as the mean soma membrane potential or the mean firing rate of some suitably defined neuronal ensemble. In this manner a mesoscopic level model can be developed which acts as a bridge between cellular (or microscopic) level activity and whole brain (or macroscopic) level behavior. While the current mathematical approach for formulating the equations of motion for

Thus the response of the mean soma membrane potential h_k ($k = e, i$) at position r on a two-dimensional cortical sheet, in response to induced post synaptic activity I_{lk} ($l = \text{source}$, $k = \text{target population}$) is given by

$$\tau_k \frac{\partial h_k(r, t)}{\partial t} = h_k^I - h_k(r, t) + \sum_{l=e, i} \frac{h_{lk}^{eq} - h_k(r, t)}{|h_{lk}^{eq} - h_k^r|} I_{lk}(r, t) \quad (1)$$

The postsynaptic response to a single pre-synaptic action potential (at $t = 0$) is modeled by the well-known synaptic alpha function of cable theory as $\Gamma_{lk} \gamma_{lk} t \exp(1 - \gamma_{lk} t) \Theta(t)$ where Γ_{lk} is peak amplitude (occurring at $t = t_{\text{peak}} = 1/\gamma_{lk}$) of the respective excitatory ($l = e$) or inhibitory ($l = i$) postsynaptic potential (PSP), and $\Theta(t)$ is the Heaviside step function. Thus we assume that the time course of the synaptically induced excitatory and inhibitory currents is described by a critically damped oscillator driven respectively by the mean rate of incoming excitatory and inhibitory axonal pulses:

$$\left(\frac{\partial}{\partial t} + \gamma_{lk}\right)^2 I_{lk}(r, t) = \exp(1) \Gamma_{lk} \gamma_{lk} A_{lk}(r, t), \quad (2)$$

with

$$A_{ek}(r, t) = N_{ek}^\beta S_e[h_e(r, t)] + \phi_{ek} + p_{ek}(r, t), \quad (3)$$

and

$$A_{ik}(r, t) = N_{ik}^\beta S_i[h_i(r, t)] + p_{ik}(r, t), \quad (4)$$

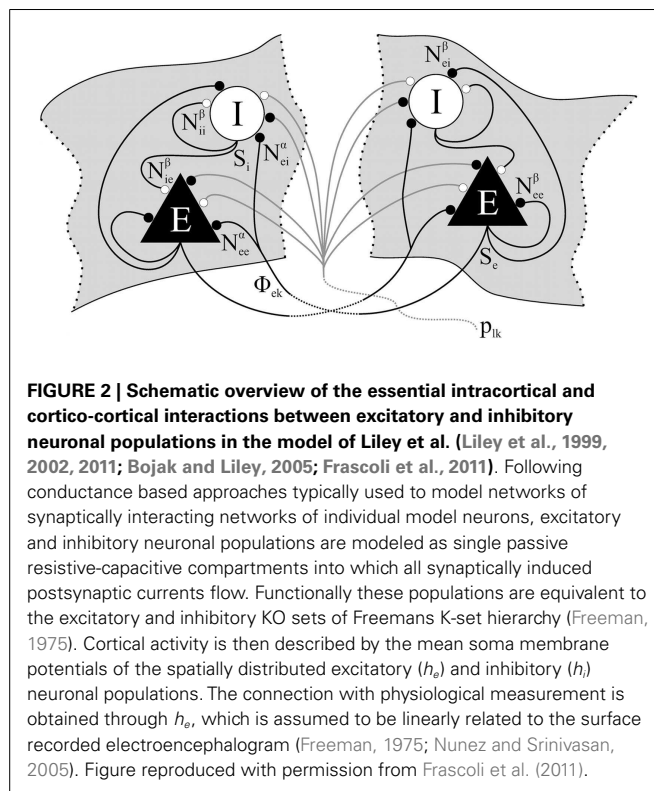


FIGURE 2 | Schematic overview of the essential intracortical and cortico-cortical interactions between excitatory and inhibitory neuronal populations in the model of Liley et al. (Liley et al., 1999, 2002, 2011; Bojak and Liley, 2005; Frascoli et al., 2011). Following conductance based approaches typically used to model networks of synaptically interacting networks of individual model neurons, excitatory and inhibitory neuronal populations are modeled as single passive resistive-capacitive compartments into which all synaptically induced postsynaptic currents flow. Functionally these populations are equivalent to the excitatory and inhibitory KO sets of Freeman's K-set hierarchy (Freeman, 1975). Cortical activity is then described by the mean soma membrane potentials of the spatially distributed excitatory (h_e) and inhibitory (h_i) neuronal populations. The connection with physiological measurement is obtained through h_e , which is assumed to be linearly related to the surface recorded electroencephalogram (Freeman, 1975; Nunez and Srinivasan, 2005). Figure reproduced with permission from Frascoli et al. (2011).

where A_{lk} comprises the different sources of pre-synaptic spikes: $N_{lk}^{\beta} S_i$ (input from local cortical neuronal populations), ϕ_{ek} (input from long-range excitatory cortico-cortical fibers), and p_{lk} (extra-cortical sources). While the present consensus is that extra-cortical sources (thalamo-cortical afferents) are purely excitatory in nature and thus $p_{lk} = 0$, we choose to retain these terms as when time independent they can be utilized to include the effects of tonic inhibition that are known to be induced by anesthetic action. The time courses of the synaptically induced currents, $1/\gamma_{lk}$ are taken to describe the time course of “fast” excitatory [$l = e$: α -amino-3-hydroxy-5-methyl-4-isoxazole-propionate (AMPA) and kainate] and inhibitory [$l = i$: γ -amino-butyric-acid type A (GABA_A)] neurotransmitter kinetics. Thus each type of PSP (excitatory, inhibitory) is described by two parameters Γ_{lk} , γ_{lk} . However, as we will describe later, a parametrically more flexible description of the PSP is required to meaningfully model the effects of anesthetics in which we can independently vary peak amplitude, rise (t_{peak}) and decay times. Mean neuronal population firing rates, S_b , are assumed to instantaneous sigmoid functions of the mean soma membrane potential i.e.,

$$S_l[h_l(r, t)] = S_l^{\max} / \{1 + \exp[\sqrt{2}(h_l(r, t) - \bar{\mu}_l/\sigma_l)]\} \quad (5)$$

The axonal pulses ϕ_{ek} propagated by the exclusively excitatory long-range cortico-cortical fiber system is in the simplest case described by the following two-dimensional telegraph equation,

$$\begin{aligned} & \left(\frac{\partial}{\partial t} + v_{ek} \Lambda_{ek} \right)^2 \phi_{ek}(r, t) - \frac{3}{2} v_{ek}^2 \nabla^2 \phi_{ek}(r, t) \\ & = v_{ek}^2 \Lambda_{ek}^2 N_{ek}^\alpha S_e[h_e(r, t)] \end{aligned} \quad (6)$$

where N_{ek}^α is the total number of excitatory connections formed by long-range cortico-cortical axons on long on local population k , and assumes a single axonal conduction velocity v_{ek} and an exponential fall off with distance (characteristic scale $= 1/\Lambda_{ek}$) of the strength of cortico-cortical connectivity. For simplicity, and given the fact that at least in anesthesia BS appears to have a degree of spatial uniformity, we chose to only study the spatially homogeneous case, i.e., $\nabla^2 = 0$.

Equations (1)–(4) and (6) represent a system of 8 coupled non-linear partial differential equations that typically defines the Liley model of electrocortical rhythmogenesis, which is capable of reproducing the main features of spontaneous human electroencephalogram (alpha resonance, “ $1/f$ ” activity). **Table 1** summarizes all model parameters, their definitions, and approximate ranges.

2.2. MODEL PARAMETERIZATION: GENERATION OF NORMATIVE PARAMETER SETS

Because BS activity (at least that induced by anesthetic and sedative action) is assumed to ultimately arise out of a background of normal electroencephalographic activity it is important to define parametrically normative states in order to study how they may be perturbed during health and disease. We therefore chose to

Table 1 | List of spatially averaged parameters for different types $k, l \in \{e, i\}$ of neuronal target populations in the electrocortical model of Liley et al. (Liley et al., 1999, 2002, 2011; Bojak and Liley, 2005; Frasca et al., 2011), with typical ranges that are assumed to be physiologically admissible.

	Definition	Min	Max	Units
h_k^r	Resting membrane potential	−80	−60	mV
τ_k	Passive membrane decay time	5	150	ms
h_{ek}^{eq}	Excitatory reversal potential	−20	10	mV
h_{ik}^{eq}	Inhibitory reversal potential	−90	$h_k^r - 5$	mV
Γ_{ek}	EPSP peak amplitude	0.1	2.0	mV
Γ_{ik}	IPSP peak amplitude	0.1	2.0	mV
$1/\gamma_{ek}$	EPSP rise time to peak	1	10	ms
$1/\gamma_{ik}$	IPSP rise time to peak	2	100	ms
N_{ek}^α	Number of excitatory cortico-cortical synapses	1000	5000	–
N_{ek}^β	Number of excitatory intracortical synapses	2000	5000	–
N_{ik}^β	Number of inhibitory intracortical synapses	100	1000	–
v_{ek}	Axonal conduction velocity	0.1	1	mm ms ^{−1}
$1/\Lambda_{ek}$	Decay scale cortico-cortical connectivity	10	100	mm
S_k^{\max}	Maximum firing rate	0.05	0.5	ms ^{−1}
$\bar{\mu}_k$	Mean firing threshold	−55	−40	mV
σ_k	Firing threshold standard deviation	2	7	mV
p_{lk}	Extra-cortical synaptic input rate	0	10	ms ^{−1}

Table adapted from Liley et al. (2011). EPSP, excitatory PSP; IPSP, inhibitory PSP.

utilize previously defined parameter sets (Bojak and Liley, 2005) that have the following properties: (i) are confined to the physiologically admissible parameter space (see **Table 1**), (ii) give rise to electroencephalographically and physiologically plausible activity (“ $1/f$ ” decay at low frequencies plus a relatively sharp peak at alpha frequencies, 8–13 Hz; mean excitatory/inhibitory neuronal firing rates $< 20 \text{ s}^{-1}$) and (iii) that exhibit transient increases in total power and monotonic reductions in mean frequency with respect to modeled anesthetic action (see below). In general such sets can be found by randomly searching the high dimensional physiologically admissible (and plausible) parameter space. For further details see Bojak and Liley (2005).

2.3. MODELING ANESTHETIC ACTION

The range of molecular and cellular targets identified to date as sites of anesthetic action is so varied that a unitary biological mechanism for anesthetic effect seems unlikely. Nevertheless, at least functionally, at the level of cortex anesthetics seem to act principally by enhancing the actions of inhibitory activity (Liley et al., 2011). Indeed from the perspective of the mean field model we have described many of its parameters can be related in a fairly straightforward way to these alterations in inhibitory activity and other identified sites of anesthetic action in cortex (see **Table 2**). However a parametrically more flexible description of the PSP, than is presently incorporated, is required to meaningfully model the effects of anesthetics in which we can independently vary peak amplitude, rise (t_{peak}) and decay times. For example isoflurane, a volatile halogenated anesthetic, has been shown to prolong the decay time of the unitary IPSP without altering its time to peak. Fortunately a simple modification of the equation describing the dynamics of the PSP enables independent adjustment of the peak amplitude, rise (t_{peak}) and decay times. By defining I_{lk} to satisfy

$$\left[\frac{\partial}{\partial t} + \gamma_{lk}(\varepsilon_{lk}) \right] \left[\frac{\partial}{\partial t} + \tilde{\gamma}_{lk}(\varepsilon_{lk}) \right] I_{lk}(r, t) = \tilde{\gamma}_{lk}(\varepsilon_{lk}) \exp[\gamma_{lk}(\varepsilon_{lk})/\gamma_{lk}^0] \Gamma_{lk} A_{lk}(r, t), \quad (7)$$

$$\gamma_{lk}(\varepsilon_{lk}) = \varepsilon_{lk} \gamma_{lk}^0 / (e^{\varepsilon_{lk}} - 1), \quad \tilde{\gamma}_{lk} = \gamma_{lk}(\varepsilon) e^{\varepsilon_{lk}} \quad (8)$$

Table 2 | Relationship between major experimentally identified sites of cortical anesthetic action and parameters of the electrocortical model of Liley et al. (Liley et al., 1999, 2002; Bojak and Liley, 2005).

Site of action	Main anesthetic effect	Parameters
2PK channels and extrasynaptic GABA _A	Increase in tonic inhibition	p_{ik}, h_k^r
nACh receptors	Reduction in tonic excitation	p_{ek}, h_k^r
Synaptic GABA _A	Increase of IPSPs	γ_{ik}, Γ_{ik}
AMPA/kainate receptors and NMDA receptors*	Reduction of EPSPs	γ_{ek}, Γ_{ek}
Myelinated axons	Slowdown of conduction [†]	v_{ek}
Na channels	Alteration of neuronal firing	$S_k^{\max}, \bar{\mu}_k, \sigma_k$

*Parameters will depend on membrane potential in this case.

[†] Effect demonstrated in periphery, speculative in cortex (Swindale, 2003).

where $1/\gamma_{lk}^0$ defines the time to peak, we can control the decay of the unitary PSP by altering $\varepsilon_{lk} > 0$. Increasing ε_{lk} will monotonically increase the decay time of the tail of the unitary PSP (see lower left panel, **Figure 6**). Empirically it is found that increasing the aqueous concentration of a range of GABAergic anesthetic agents leads to a progressive increase in the decay time of the unitary inhibitory PSP (e.g., Banks and Pearce, 1999) and thus ε_{lk} will be a monotonic function of anesthetic concentration c , i.e., $\varepsilon_{lk}(c)$. Liley et al. (2011), based on a range of empirical evidence, have numerically estimated $\varepsilon_{lk}(c)$ for the volatile anesthetic agent isoflurane. However because $\varepsilon_{lk}(c)$ is not currently known for other GABAergic anesthetic agents we will assume that in general $\varepsilon_{lk} \propto c$.

It is worth noting that equation (7) reduces to equation (2) as $\varepsilon_{lk} \rightarrow 0$. Further details regarding this formulation can be found in Bojak and Liley (2005).

2.4. DEFINING A THEORETICAL BASIS FOR BURSTING

We call a bursting system *fast-slow* if it can be written in the following form

$$\dot{x} = f(x, y) \quad (\text{fast oscillatory system}) \quad (9)$$

$$\dot{y} = \mu g(x, y) \quad (\text{slow modulatory system}) \quad (10)$$

where $x \in \mathbb{R}^m$ describes the m-dimensional system responsible for the fast oscillatory (spiking in single neuron models) dynamics and $y \in \mathbb{R}^n$ the n-dimensional slow system that modulates the fast oscillations (or spiking behavior). The parameter μ represents the ratio of the time scales between the slow and fast system. It is typically assumed that $\mu \ll 1$. Because μ can be made as small as we like equations (9) and (10) represent a singularly perturbed system.

We will assume that equations (1)–(4) and (6), which putatively describe the genesis of the “fast oscillatory” resting EEG, correspond to the m-dimensional fast system. To motivate the slow n-dimensional slow modulatory system we will make a plausible appeal to the biophysics of synaptic resource depletion and recovery during periods of sustained neuronal population activity. During periods of high firing neural activity a variety of factors come into play to diminish synaptic efficiency. The most important of these are receptor desensitization and synaptic vesicle depletion. Tsodyks and Markram (1997) developed a model to account for the biophysics of such activity dependent short term synaptic depression estimating that its onset is rapid, of the order of milliseconds, but that its recovery is quite slow, of the order of 800 ms. Given that such a time scale is approximately at least an order of magnitude greater than the characteristic time scales associated with resting EEG activity, this may represent a candidate slow EEG modulatory system. On this basis we choose to include this activity dependent short term synaptic depression using the following two different formulations, referred respectively to as SS1 and SS2,

$$\dot{\Gamma}_{lk} = \mu_l[\theta_l - k_l S_l(h_l)], \quad (\text{SS1}) \quad (11)$$

$$\dot{\Gamma}_{lk} = \mu_l[\Gamma_{lk}^0/(1 + \exp[\kappa_l(h_l - \xi_l)]) - \Gamma_{lk}], \quad (\text{SS2}) \quad (12)$$

where $1/\mu_l$ is the characteristic time scale of the respective slow modulatory system. Based on Tsodyks and Markram (1997) we will fix $1/\mu_l$ to 1000 ms. The advantage of the first formulation is that the rates of synaptic recovery ($\mu_l\theta_l$) and depletion ($\mu_l k_l$) can be separately specified. The advantage of the second formulation is that Γ_{lk} remains bounded between Γ_{lk}^0 (the resting value) and zero decreasing monotonically with increasing mean soma membrane potential h_l , with Γ_{lk}^0 remaining as a free parameter. For low levels of the respective neuronal activity (h_l) there is very little decrease in Γ_{lk} until a threshold ξ_l is reached, with the parameter κ_l controlling the sensitivity of the change at this threshold to variations in neuronal activity. This formulation has previously been used by Tabak and Rinzel (2005) in their mean field model for spontaneous electrical bursting activity in embryonic chick spinal cord.

2.5. COMPUTATIONAL METHODS

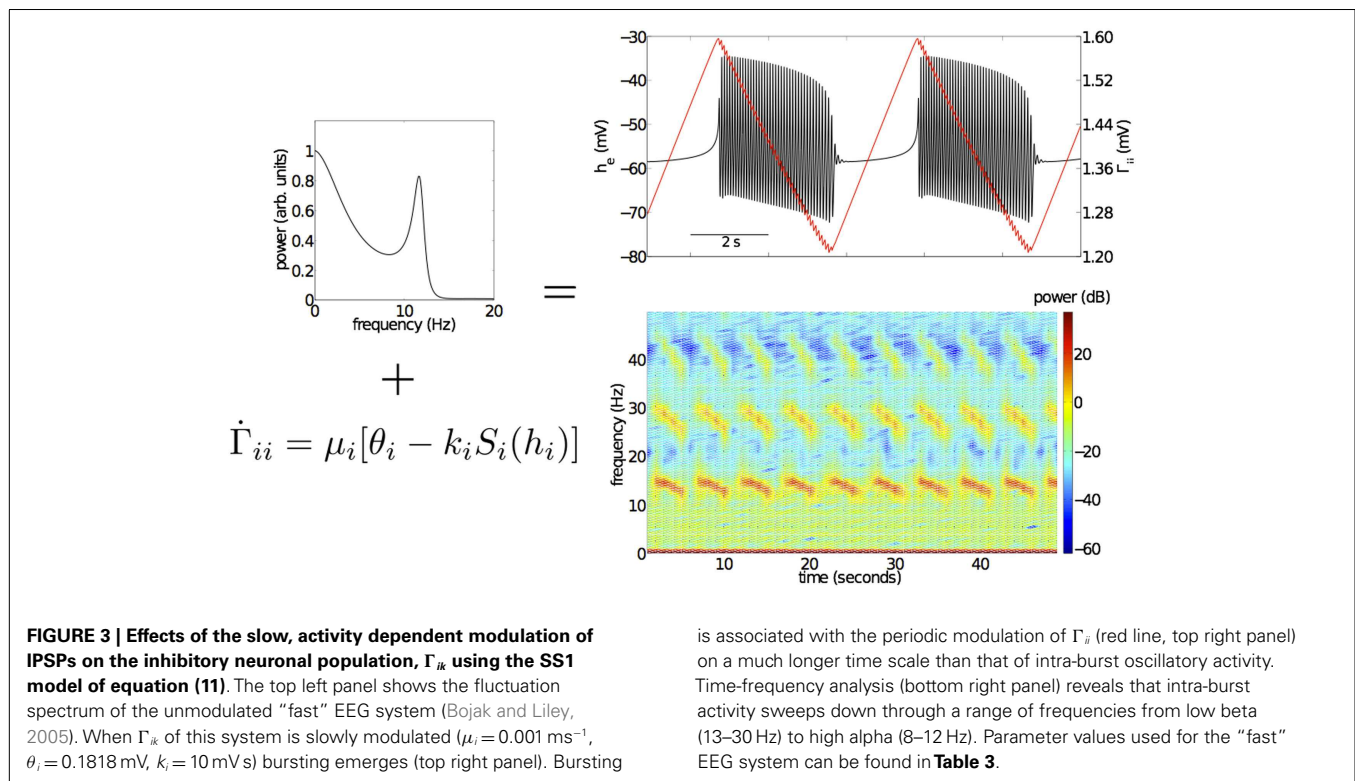
All numerical integrations and one-dimensional dynamical continuations were performed using the XPPAUT package (Ermentrout, 2002). A 4-th order Runge-Kutta scheme with a time step of 0.1 ms was used to numerically integrate the differential equations. Because of the multiple time scales our differential system may suffer from stiffness and numerical solutions may not converge. In these cases we have used the recommended “stiff” integrator CVODE (Cohen and Hindmarsh, 1994) as implemented in XPPAUT.

3. RESULTS

All numerical simulations were performed using a single model parameter set having a physiologically plausible white noise fluctuation spectrum (see top left panel **Figure 3**) and a single stable fixed point. This parameter set was found using the methods described in section 2.2. The parameters used, all within the physiologically admissible domain, can be found in the **Table 3**.

Figure 3 shows the effects of the activity dependent modulation of Γ_{ii} on simulated mean field EEG activity using SS1 [equation (11)]. Bursts emerge periodically, with intervening near isoelectric intervals, apparently driven by slow variations in Γ_{ii} (red line, top right panel). A spectrogram of a sufficiently long simulated time series reveals that the frequency of the model EEG activity decreases from low beta ($\cong 15$ Hz) to high alpha ($\cong 12$ Hz) over the period of the bursts. Such intra-burst “chirping” is a common feature of many of the topologically identified single neuronal bursters (Izhikevich, 2007). Of interest are the multiple harmonics of this dominant oscillatory activity.

Figure 4 illustrates how we might dynamically account for the transition to, and cessation from, bursting and follows the now standard method of the dissection of neural bursting pioneered by Rinzel (1985). Here we have set $\mu_i = 0$ and consider how the “fast” EEG system responds. **Figure 4** shows a one-dimensional bifurcation diagram of this “fast” EEG system with Γ_{ii} as the bifurcation parameter. Thick black lines show the fixed points as a function of Γ_{ii} . For small values of Γ_{ii} there is a single stable fixed point. As Γ_{ii} is increased this fixed point loses stability by a super-critical Hopf bifurcation, thus signaling the onset of limit cycle activity. Periodic continuations of this low amplitude activity reveals that it, and a stable fixed point, co-exist with a higher amplitude limit



cycle, thus suggesting that an activity dependent hysteresis drives the system between a fixed point and a high amplitude oscillation, thus giving rise to the bursting activity observed. This can be better seen by superimposing on this diagram the trajectory of a single burst (thin solid black line). Here we can see that the burst terminates through a fold-limit cycle bifurcation. At this stage it is not clear what bifurcation accounts for the emergence of the burst.

A well described feature of anesthetic action is the reduction in cerebral blood flow and metabolism (Kaisti et al., 2003). Therefore during anesthetic action it would be reasonable to assume that the recovery of pre-synaptic neurotransmitter levels will be impaired. In particular as the anesthetic level increases then the rate of synaptic recovery should decrease. **Figure 5** shows the effects of systematic reductions in the synaptic recovery rate for the SS1 model. As θ_i (synaptic recovery) is decreased the burst duration decreases and the period of non-oscillatory isoelectric activity increases until the model EEG becomes fully isoelectric. This trend is also observed clinically during anesthesia.

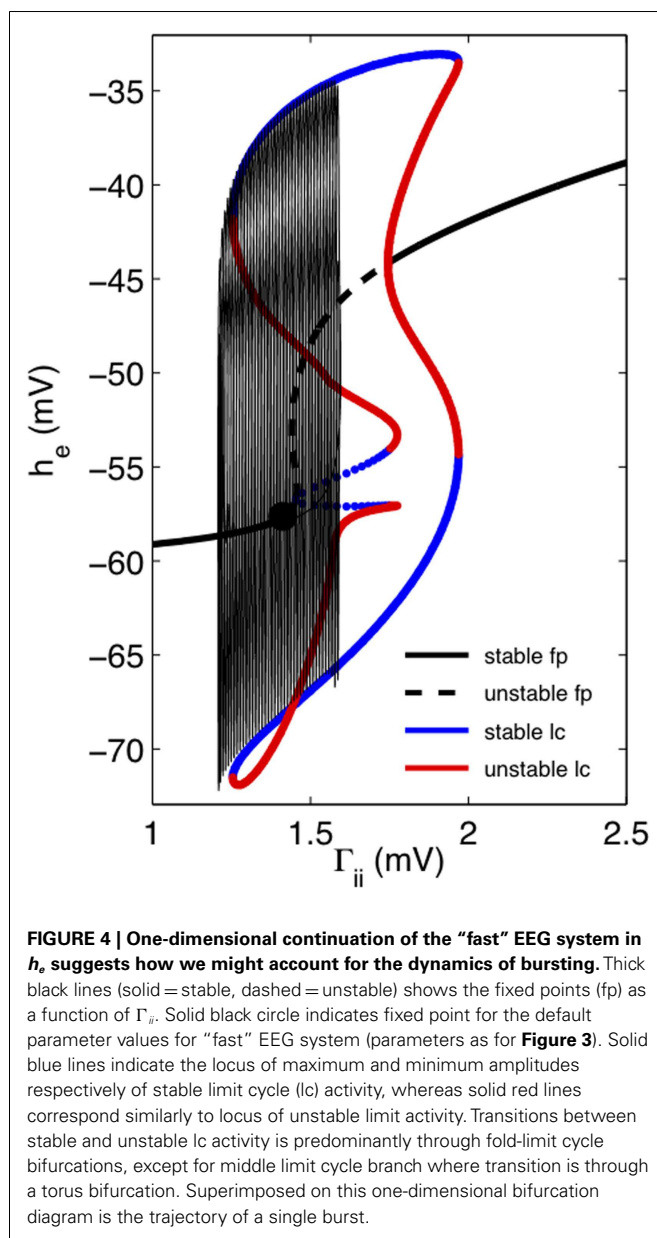
Short term synaptic depression would be expected to affect all synapses, though the depression would not be expected to be uniform. So far we have assumed that the synaptic depression would principally affect inhibitory synapses between inhibitory neurons. Will such bursting survive when all types of “fast” synaptic activity is subject to the biological forces of short term synaptic depression? **Figure 6** reveals that bursting does occur when both excitatory and inhibitory synaptic activity undergoes activity dependent short term synaptic depression. Further, the bursting that emerges is strongly modulated by parameters of the “fast”

Table 3 | Model parameter set used in simulations of Figures 3–8.

Parameter	Value	Parameter	Value
h_e^{rest} (mV)	−68.1355	N_{ee}^{α}	4994.4860
h_i^{rest} (mV)	−77.2602	N_{ei}^{α}	2222.9060
h_{ee}^{eq} (mV)	−15.8527	N_{ee}^{β}	4582.0661
h_{ei}^{eq} (mV)	7.4228	N_{ei}^{β}	4198.1829
h_{ie}^{eq} (mV)	−85.9896	N_{ie}^{β}	989.5281
h_{ii}^{eq} (mV)	−84.5363	N_{ii}^{β}	531.9419
τ_e (ms)	138.3660	v_{ek} (cm ms ^{−1})	0.1714
τ_i (ms)	89.3207	Λ_{ek} (cm ^{−1})	0.2433
Γ_{ee} (mV)	0.3127	S_e^{max} (ms ^{−1})	0.2801
Γ_{ei} (mV)	0.9426	S_i^{max} (ms ^{−1})	0.1228
Γ_{ie} (mV)	0.4947	μ_e (mV)	−47.1364
Γ_{ii} (mV)	1.4122	μ_i (mV)	−45.3751
γ_{ee} (ms ^{−1})	0.4393	σ_e (mV)	2.6120
γ_{ei} (ms ^{−1})	0.2350	σ_i (mV)	2.8294
γ_{ie} (ms ^{−1})	0.0791	p_{ee} (ms ^{−1})	3.6032
γ_{ii} (ms ^{−1})	0.0782	p_{ei} (ms ^{−1})	0.3639

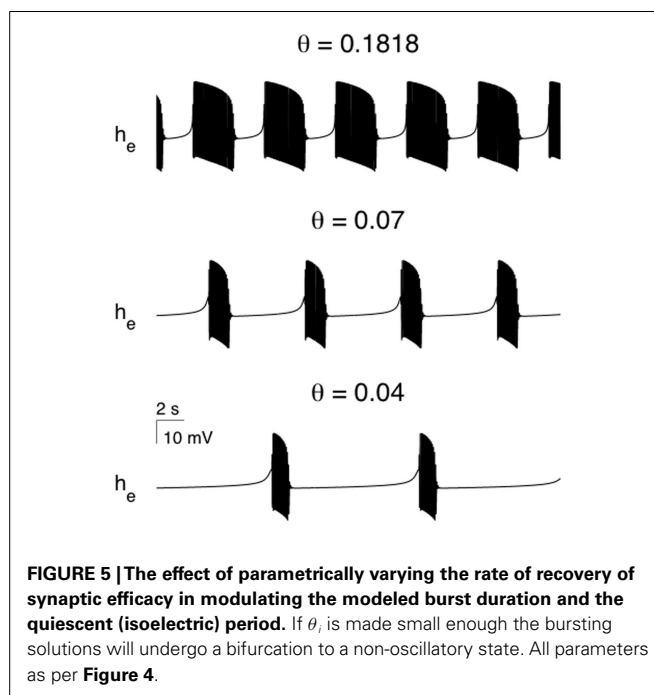
Parameter set taken from Bojak and Liley (2005). Because parameters were obtained as part of a numerical search their full precision had been detailed, however their sensitivity to perturbation is much less than the precision reported (Bojak and Liley, 2005).

EEG system (see **Table 2**) that have been identified as targets for the action of anesthetic agents. Prolonging the decay of the unitary IPSP and reducing subcortical input are both found to significantly modulate modeled EEG bursting. In particular it is found that reducing excitatory extra-cortical input (p_{ee}), which



presumably dominantly arises from thalamus, leads to very long quiescent (isoelectric) periods.

One of the limitations in using SS1 is that the PSPs are, in principle, free to take any value whereas physiology would dictate that they should remain bounded. To explore the effects of this restriction we chose to define an alternative slow modulating system [SS2, equation (12)]. **Figure 7** shows the effect of utilizing this system to provide a slow activity dependent modulation of excitatory and inhibitory synaptic efficacy of our “fast” EEG system. In the left panel of this figure we have plotted $c_e \equiv \Gamma_{el}/\Gamma_{el}^0$ as a function of time. The interesting thing to note is that in contrast to **Figure 3** excitatory synaptic efficacy decreases during the quiescent inter-burst period and increases during the burst. The left hand panel however shows that there is a significant phase difference between the normalized excitatory synaptic efficacy $c_e \equiv \Gamma_{el}/\Gamma_{el}^0$ and the



normalized inhibitory synaptic efficacy $c_i \equiv \Gamma_{il}/\Gamma_{il}^0$. This suggests that there is an important dynamical interplay between excitatory and inhibitory synaptic efficacy to regulate neuronal population excitability such that bursting occurs. This may explain why there is confusion in the empirical literature regarding the role alterations in synaptic efficiency have in the genesis of BS.

An important difference between SS2 and SS1 is that parameters hypothesized to be important targets of anesthetic action p_{ee} and ε_{il} are able not only to parametrically regulate bursting but appear also able to switch bursting on (presumably through a bifurcation from a stable fixed point). **Figure 8** illustrates this. If the inhibitory neuronal IPSP decay time is not long enough then a single fixed point dominates which has an associated white noise fluctuation spectrum. But as the IPSP decay time increases (beyond $\varepsilon_{il} > 1.8$ for the parameter set chosen) then bursting emerges. However if p_{ee} is decreased, as we would expect during anesthesia, then the isoelectric period is prolonged until at some critical value of p_{ee} bursting is extinguished to be replaced by an infinitely long quiescent/isoelectric period.

4. DISCUSSION

We have described here how a well-known model of the “fast” dynamics of the EEG can be modulated by a number of slow systems to produce bursting activity that bears some resemblance to BS seen clinically. The slow systems that we used were all based on some form of activity dependent short term plasticity that has been empirically observed, and used successfully in other models of macroscopic level bursting (Tabak and Rinzel, 2005). While we were able to clearly show the existence of bursting, because we did not include any additive or multiplicative noise sources, we were unable to account for the quasi-periodicity of BS. Thus all our bursting arises from purely deterministic processes, presumably involving a range of well described bifurcations (Izhikevich,

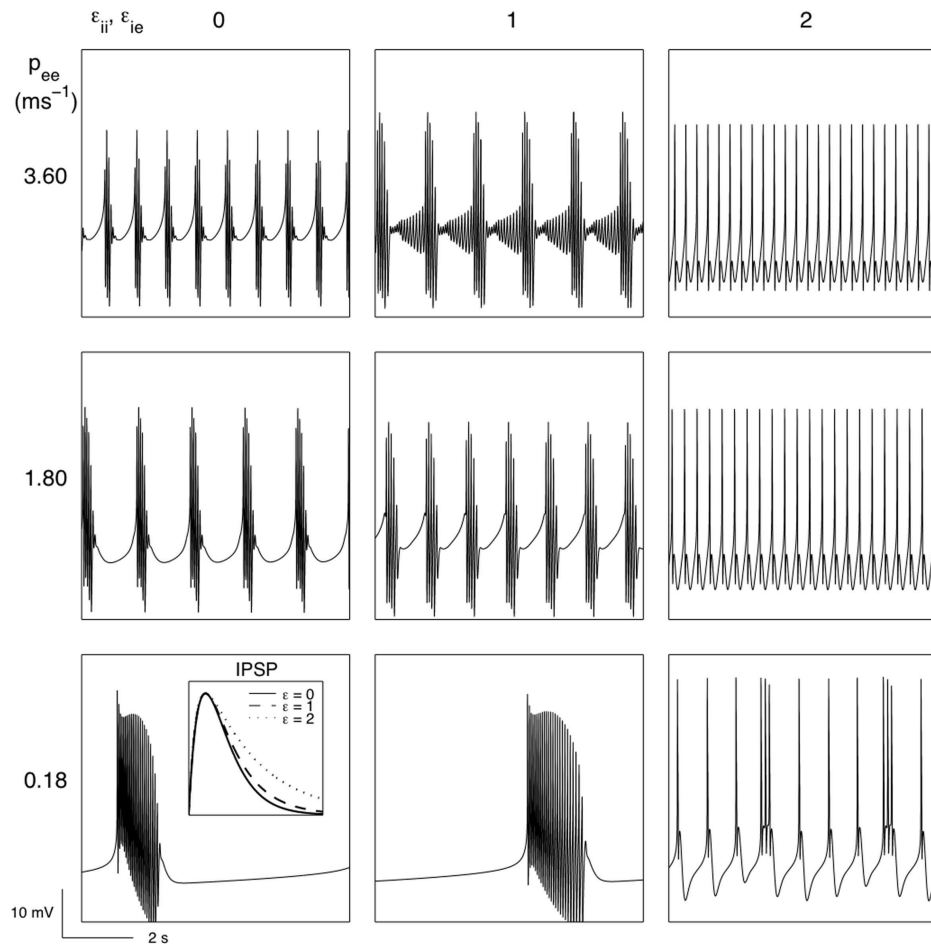


FIGURE 6 | The effects of short term synaptic depression of both excitatory and inhibitory cortical synapses in the genesis of burst suppression and its modulation by variations in extra-cortical input (p_{ee})

and the IPSP decay time (ε_{irr} , ε_{ie}) (see section 2.3 for further details).

Parameters: $\theta_e = 0.1818$ mV, $\theta_i = 0.07$ mV, $k_e = 14$ mV s. All other parameters as for **Figure 3**.

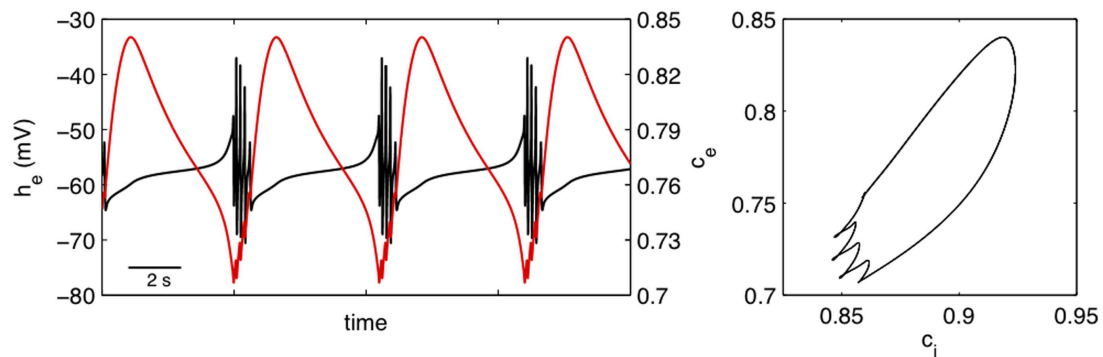
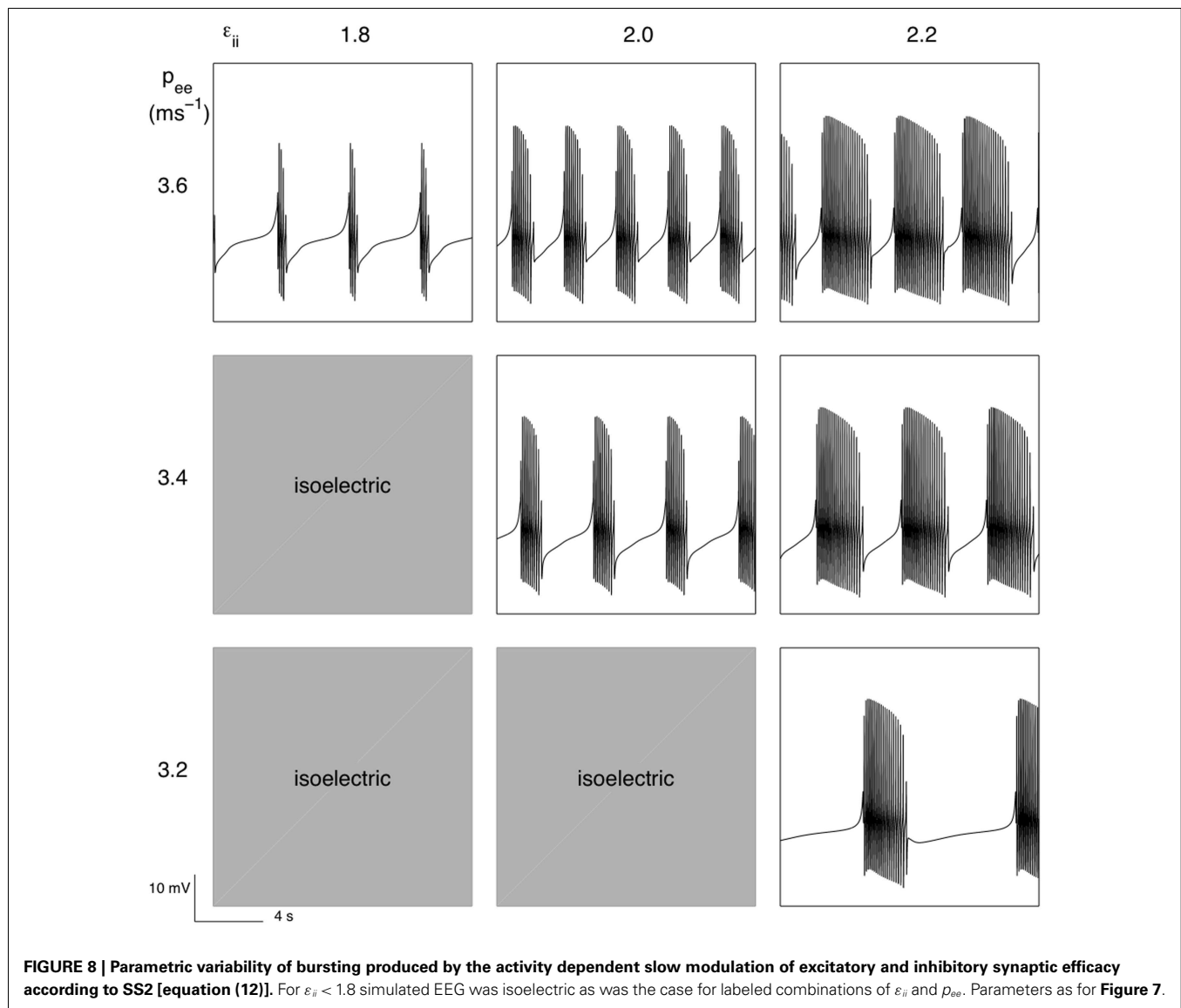


FIGURE 7 | Bursting produced by the activity dependent slow modulation of excitatory and inhibitory synaptic efficacy according to SS2 [equation (12)]. Left panel shows h_o (solid black line) and normalized excitatory synaptic efficacy $c_o \equiv \Gamma_{o\omega}/\Gamma_{o\omega}^0$ (solid red

line) as a function of time. The left hand panel shows the phase relationship between normalized excitatory and inhibitory ($c_i \equiv \Gamma_{ii}^0/\Gamma_{ii}^0$) synaptic efficacies. Parameters: $\kappa_e = 0.2 \text{ mV}^{-1}$, $\kappa_i = 0.1 \text{ mV}^{-1}$, $\varepsilon_{ii} = 1.8$, $\varepsilon_{ie} = 1.5$.

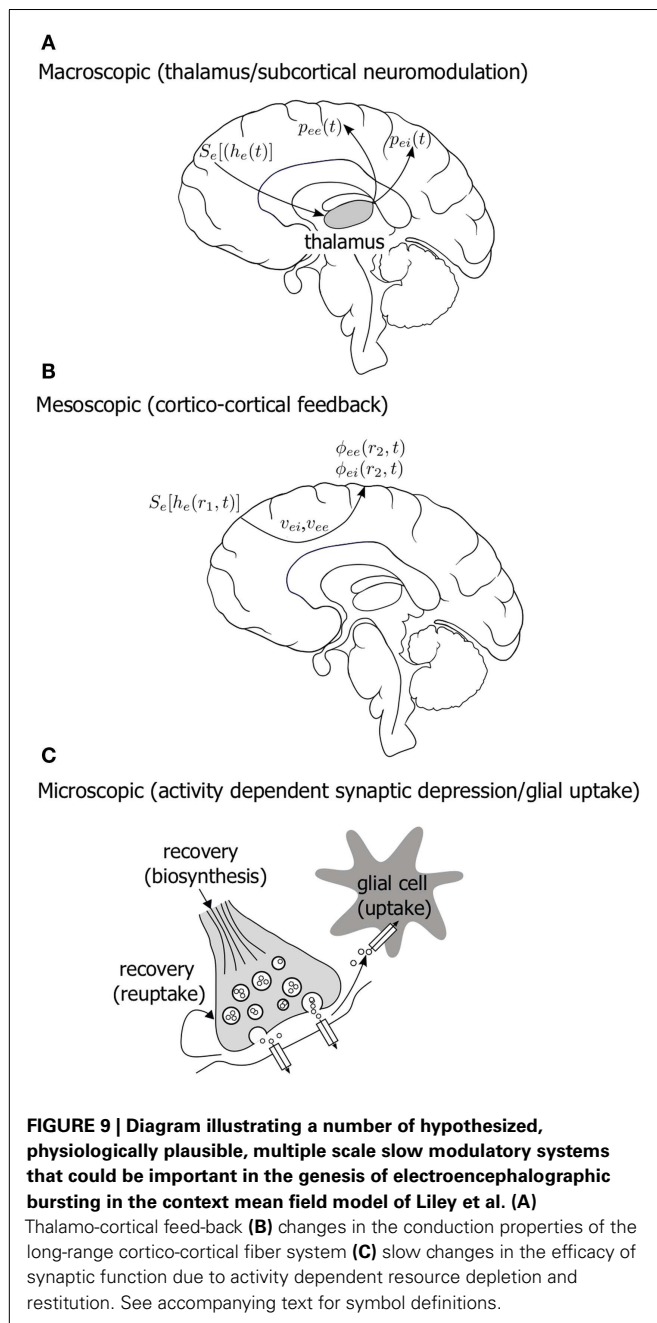


2007). However because our system clearly exhibited bistability (see Figure 4) it is almost certain that our system will be able to exhibit some form of burst excitability in response to stochastic forcing. Such burst excitability has been described experimentally. For example during BS induced by various halogenated anesthetic agents, bursts can be readily evoked by auditory, visual, or somatosensory stimuli (Hartikainen et al., 1995). Nevertheless while burst onset and duration may be random variables it would seem that the bursts themselves should reveal a high degree of determinism (weak non-linearity) when compared to EEG in which bursting or epileptiform activity is not evident.

Because the parameter space of the underlying “fast” EEG model is potentially extremely large it is not possible to systematically explore its dynamical repertoire and it may be possible that this system, not augmented with one of the slow systems described, is able to burst. Nevertheless, on the basis of our results, and what is known regarding the dynamical mechanisms

of bursting, it would seem likely that multiple pathways to BS exist through a variety of activity dependent slow modulatory systems.

Further we might hypothesize that such slow modulatory systems might span a number of functional scales in the brain. Figure 9 diagrammatically illustrates some possible candidate systems. An obvious activity driven slow modulatory system would be that associated with thalamus and the corresponding thalamo-cortical feed-back. Mean field models of the EEG that incorporate thalamo-cortical feed-back have been developed (Rennie et al., 2002) and it will be interesting to see if they are structurally configured to support BS. In addition to synaptic fatigue another obvious cortical system that might be marshaled to provide slow activity dependent modulation is the cortico-cortical conduction system. Although at this stage there is currently little evidence to suggest anesthetics slow conduction velocities, it is widely documented that axonal conduction velocities are significantly decreased in



hypothermia. Decreases of up to 5% per °C for conduction velocity have been reported (Waxman, 1980). As BS has been observed to occur in hypothermia (Schwartz et al., 1989; Akrawi et al., 1996) we can conclude that a slow system emerging in the long-range

coupling via a slowing of axonal conduction velocity is a possible route to BS.

Our attempts to account for the dynamical pathogenesis of burst suppression differ from other approaches, most notably Ching et al. (2012). Ching et al. described the scalp EEG in terms of the activity of a small scale, biophysically detailed, computational model of interacting populations of cortical and thalamic neurons. Burst suppression was modeled as arising from metabolically induced alterations in an ATP-gated slow neuronal membrane potassium current ($I_{K_{ATP}}$) on the basis that the reduction in cerebral metabolic rate (CMRO2) induced by anesthetic agents and hypoxia was associated with the depletion of ATP, and hence membrane hyperpolarization. While on this basis they claim to have accounted for a number of defining features of BS that included (i) the spatial synchrony of burst onset (ii) the parametric variability of burst duration/iselectricity and (iii) the characteristically long time scales associated with bursting/iselectricity compared to resting EEG, some caution needs to be exercised.

Firstly their model of resting/spontaneous EEG is constructed on the basis of the activity of no more than 20 model neurons. Because EEG is a distributed large scale phenomena such a model is unlikely to meaningfully account for resting/spontaneous activity particularly given the absence of any long-range excitatory cortico-cortical coupling. This has important implications for the propagation of burst activity particularly given that the onset of bursts, when examined at fine temporal scales, is probably not truly spatially homogeneous.

Secondly while the relationship between CMRO2 and ATP production cannot be reasonably disputed, not all anesthetic agents that produce reductions in CMRO2 produce BS. For example the noble gas xenon has been reported to reduce CMRO2 by up to 33% in human participants (Rex et al., 2006) yet is not associated with any discernable BS.

Thirdly the approach they have taken to producing BS essentially depends on the slow modulation of a faster system, the approach we have adopted here. The modeled time scales of $I_{K_{ATP}}$ variability are very long, of the order of tens of seconds.

For clarity and tractability the current investigations have focused on the spatially homogeneous case for the model of Liley et al. [i.e., $\nabla^2 = 0$ in equation (6)]. Clearly the emergence of BS in the spatially extended case will need to be investigated through the appropriate numerical solution of the defining partial differential equations. Because the cortical phase synchrony (Hartikainen et al., 1995) of burst suppression has not, as far as we are aware, been explicitly investigated it will be crucial to empirically determine the spatiotemporal emergence of bursts in order to assess the importance of excitatory cortico-cortical connectivity in the emergence and modulation of BS as implied by the model and as we have suggested. This will require recording high density EEG during anesthesia in which BS is present.

REFERENCES

- Akrawi, W. P., Drummond, J. C., Kalkman, C. J., and Patel, P. M. (1996). A comparison of the electrophysiologic characteristics of eeg burst-suppression as produced by isoflurane, thiopental, etomidate, and propofol. *J. Neurosurg. Anesthesiol.* 8, 40–46.
- Amari, S.-I. (1975). Homogeneous nets of neuron-like elements. *Biol. Cybern.* 17, 211–220.
- An, D. S., Straumann, D., and Wieser, H. G. (1996). 'One-way asynchrony' of burst-suppression activity. *Neurophysiol. Clin.* 26, 329–334.
- Banks, M. I., and Pearce, R. A. (1999). Dual actions of volatile anesthetics on GABA(A) IPSCs: dissociation of blocking and prolonging effects. *Anesthesiology* 90, 120–134.
- Bojak, I., and Liley, D. T. J. (2005). Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 71, 041902.
- Brenner, R. P. (1985). The electroencephalogram in altered states

- of consciousness. *Neurol. Clin.* 3, 615–631.
- Ching, S., Purdon, P. L., Vijayan, S., Kopell, N. J., and Brown, E. N. (2012). A neurophysiological-metabolic model for burst suppression. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3095–3100.
- Cohen, S. D., and Hindmarsh, A. C. (1994). *CVODE User Guide*. Technical Report. Livermore: Lawrence Livermore National Laboratory.
- Deco, G. R., Jirsa, V. K., Robinson, P. A., Breakspear, M., and Friston, K. J. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* 4:e1000092. doi:10.1371/journal.pcbi.1000092
- Derbyshire, A. J., Rempel, B., Forbes, A., and Lambert, E. F. (1936). The effects of anesthetics on action potentials in the cerebral cortex of the cat. *Am. J. Physiol.* 116, 577–596.
- Ermentrout, B. (2002). *Simulating, Analyzing and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students*. Philadelphia: SIAM.
- Ferron, J. F., Kroeger, D., Chever, O., and Amzica, F. (2009). Cortical inhibition during burst suppression induced with isoflurane anesthesia. *J. Neurosci.* 29, 9850–9860.
- Fischer-Williams, M., and Cooper, R. A. (1963). Depth recording from the human brain in epilepsy. *Electroencephalogr. Clin. Neurophysiol.* 15, 568–587.
- Frascoli, F., van Veen, L., Bojak, I., and Liley, D. T. J. (2011). Metabifurcation analysis of a mean field model of the cortex. *Physica D* 240, 949–962.
- Freeman, W. J. (1975). *Mass Action in the Nervous System: Examination of the Neurophysiological Basis of Adaptive Behavior Through the EEG*, 1st Edn. New York: Academic Press.
- Geocadin, R. G., Sherman, D. L., Hansen, H. C., Kimura, T., Niedermeyer, E., Thakor, N. V., et al. (2002). Neurological recovery by eeg bursting after resuscitation from cardiac arrest in rats. *Resuscitation* 55, 193–200.
- Grigg-Damberger, M. M., Coker, S. B., Halsey, C. L., and Anderson, C. L. (1989). Neonatal burst suppression: its developmental significance. *Pediatr. Neurol.* 5, 84–92.
- Hartikainen, K. M., Rorarius, M., Perakyla, J. J., Laippala, P. J., and Jantti, V. (1995). Cortical reactivity during isoflurane burst-suppression anesthesia. *Anesth. Analg.* 81, 1223–1228.
- Henry, C. E., and Scoville, W. B. (1952). Suppression-burst activity from isolated cerebral cortex in man. *Electroencephalogr. Clin. Neurophysiol.* 4, 1–22.
- Hockaday, J. M., Potts, F., Epstein, E., Bonazzi, A., and Schwab, R. S. (1965). Electroencephalographic changes in acute cerebral anoxia from cardiac or respiratory arrest. *Electroencephalogr. Clin. Neurophysiol.* 18, 575–586.
- Izhikevich, E. M. (2000). Neural excitability, spiking, and bursting. *Int. J. Bifurcat. Chaos* 10, 1171–1266.
- Izhikevich, E. M. (2007). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. Cambridge, MA: The MIT Press.
- Kaisti, K. K., Langsjo, J. W., Aalto, S., Oikonen, V., Sipila, H., Teras, M., et al. (2003). Effects of sevoflurane, propofol, and adjunct nitrous oxide on regional cerebral blood flow, oxygen consumption, and blood volume in humans. *Anesthesiology* 99, 603–613.
- Kroeger, D., and Amzica, F. (2007). Hypersensitivity of the anesthesia-induced comatose brain. *J. Neurosci.* 27, 10597–10607.
- Liley, D. T. J., and Bojak, I. (2005). Understanding the transition to seizure by modeling the epileptiform activity of general anesthetic agents. *J. Clin. Neurophysiol.* 22, 300–313.
- Liley, D. T. J., Cadusch, P. J., and Dafilis, M. P. (2002). A spatially continuous mean field theory of electrocortical activity. *Network* 13, 67–113.
- Liley, D. T. J., Cadusch, P. J., and Wright, J. J. (1999). A continuum theory of electro-cortical activity. *Neurocomputing* 26–27, 795–800.
- Liley, D. T. J., Foster, B. L., and Bojak, I. (2011). “A mesoscopic modelling approach to characterising anaesthetic action on brain electrical activity,” in *Sleep and Anesthesia: Neural Correlates in Theory and Experiment*, Springer Series in Computational Neuroscience, ed. A. Hutt (New York: Springer), 139–166.
- Lukatch, H. S., and MacIver, M. B. (1996). Synaptic mechanisms of thiopental-induced alterations in synchronized cortical activity. *Anesthesiology* 84, 1425–1434.
- Niedermeyer, E. (2005). “Epileptic seizure disorders,” in *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 5th Edn, eds E. Niedermeyer and F. H. Lopes da Silva (Philadelphia: Lippincott), 505–619.
- Niedermeyer, E. (2009). The burst-suppression electroencephalogram. *Am. J. Electroneurodiagnostic Technol.* 49, 333–341.
- Nunez, P. L. (1974). The brain wave equation: a model for the EEG. *Math. Biosci.* 21, 279–297.
- Nunez, P. L., and Srinivasan, R. (2005). *Electric Fields of the Brain: The Neurophysics of EEG*, 2nd Edn. New York: Oxford University Press.
- Rennie, C. J., Robinson, P. A., and Wright, J. J. (2002). Unified neurophysiological model of EEG spectra and evoked potentials. *Biol. Cybern.* 86, 457–471.
- Rex, S., Schaefer, W., Meyer, P. H., Ros-saint, R., Boy, C., Setani, K., et al. (2006). Positron emission tomography study of regional cerebral metabolism during general anesthesia with xenon in humans. *Anesthesiology* 105, 936–943.
- Rinzel, J. (1985). “Bursting oscillations in an excitable membrane model,” in *Ordinary and Partial Differential Equations: Proceedings of the 8th Dundee Conference, Number 1151 in Lecture Notes in Mathematics*, eds B. D. Sleeman and R. J. Jarvis (Berlin: Springer), 304–316.
- Schwartz, A. E., Tuttle, R. H., and Poppers, P. J. (1989). Electroencephalographic burst suppression in elderly and young patients anesthetized with isoflurane. *Anesth. Analg.* 68, 9–12.
- Steriade, M., Amzica, F., and Contreras, D. (1994). Cortical and thalamic cellular correlates of electroencephalographic burst-suppression. *Electroencephalogr. Clin. Neurophysiol.* 90, 1–16.
- Swank, R. L., and Watson, C. W. (1949). Effects of barbiturates and ether on spontaneous electrical activity of dog brain. *J. Neurophysiol.* 12, 137–160.
- Swindale, N. V. (2003). Neural synchrony, axonal path lengths and general anesthesia: a hypothesis. *Neuroscientist* 9, 440–445.
- Tabak, J., and Rinzel, J. (2005). “Bursting in excitatory neural networks,” in *Bursting: The Genesis of Rhythm in the Nervous System*, eds S. Coombes and P. C. Bressloff (Singapore: World Scientific), 273–301.
- Treiman, D. M., Walton, N. Y., and Kendrick, C. (1990). A progressive sequence of electroencephalographic changes during generalized convulsive status epilepticus. *Epilepsy Res.* 5, 49–60.
- Tsodyks, M. V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. U.S.A.* 94, 719–723.
- Voss, L. J., Ludbrook, G., Grant, C., Sleight, J. W., and Barnard, J. P. (2006). Cerebral cortical effects of desflurane in sheep: comparison with isoflurane, sevoflurane and enflurane. *Acta Anaesthesiol. Scand.* 50, 313–319.
- Waxman, S. G. (1980). Determinants of conduction velocity in myelinated nerve fibers. *Muscle Nerve* 3, 141–150.
- Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neuron. *Biophys. J.* 12, 1–24.
- Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 December 2012; accepted: 05 April 2013; published online: 30 April 2013.

Citation: Liley DTJ and Walsh M (2013) The mesoscopic modeling of burst suppression during anesthesia. *Front. Comput. Neurosci.* 7:46. doi: 10.3389/fncom.2013.00046

Copyright © 2013 Liley and Walsh. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Neural masses and fields in dynamic causal modeling

Rosalyn Moran^{1,2,3*}, Dimitris A. Pinotsis^{1†} and Karl Friston¹

¹ Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London, UK

² Virginia Tech Carilion Research Institute, Virginia Tech, Roanoke, VA, USA

³ Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA

Edited by:

Peter Robinson, University of Sydney, Australia

Reviewed by:

Peter Robinson, University of Sydney, Australia

James Roberts, Queensland Institute of Medical Research, Australia

*Correspondence:

Rosalyn Moran, Virginia Tech Carilion Research Institute, Virginia Tech, 2 Riverside Drive, Roanoke, 24016 VA, USA
e-mail: rosalynj@vtc.vt.edu

[†] These authors have contributed equally to this work.

Dynamic causal modeling (DCM) provides a framework for the analysis of effective connectivity among neuronal subpopulations that subtend invasive (electrocorticograms and local field potentials) and non-invasive (electroencephalography and magnetoencephalography) electrophysiological responses. This paper reviews the suite of neuronal population models including neural masses, fields and conductance-based models that are used in DCM. These models are expressed in terms of sets of differential equations that allow one to model the synaptic underpinnings of connectivity. We describe early developments using neural mass models, where convolution-based dynamics are used to generate responses in laminar-specific populations of excitatory and inhibitory cells. We show that these models, though resting on only two simple transforms, can recapitulate the characteristics of both evoked and spectral responses observed empirically. Using an identical neuronal architecture, we show that a set of conductance based models—that consider the dynamics of specific ion-channels—present a richer space of responses; owing to non-linear interactions between conductances and membrane potentials. We propose that conductance-based models may be more appropriate when spectra present with multiple resonances. Finally, we outline a third class of models, where each neuronal subpopulation is treated as a field; in other words, as a manifold on the cortical surface. By explicitly accounting for the spatial propagation of cortical activity through partial differential equations (PDEs), we show that the topology of connectivity—through local lateral interactions among cortical layers—may be inferred, even in the absence of spatially resolved data. We also show that these models allow for a detailed analysis of structure–function relationships in the cortex. Our review highlights the relationship among these models and how the hypothesis asked of empirical data suggests an appropriate model class.

Keywords: dynamic causal modeling, electroencephalography, magnetoencephalography (MEG), local field potential (LFP), neural mass models

INTRODUCTION

Over the past two decades, BOLD neuroimaging techniques have been successfully applied in human studies to identify regions of functional specialization, to within a scale of a few millimeters (Ashburner, 2012). Electrophysiological recordings including magneto- and electro-encephalography (M/EEG) offer an aggregate measure of neuronal activity (in the order of hundreds of thousands of neurons) but at a millisecond timescale (Baillet et al., 2001; Nunez and Srinivasan, 2006). Though localizing activity is mathematically ill-posed in these electromagnetic modalities, the wealth of spatial information from fMRI studies can now support M/EEG as a powerful modality for the analysis of functional integration in the human brain.

Dynamic causal modeling (DCM) is designed to probe the mechanisms of effective connectivity (the influence of one brain region on another) that underlie multi-region network responses in neuroimaging (fMRI, M/EEG) data. The approach uses a neurobiologically motivated model which is inverted or fitted to empirical observations using Bayesian techniques (Daunizeau et al., 2011). These comprise separate generative processes at the

neuronal level and at the observation level. For M/EEG, neural mass and neural field models in particular are used, to support this analysis by quantifying the temporal and spatiotemporal evolution of macroscopic brain activity, using physiologically plausible dynamics. In DCM; as implemented in the SPM software (Litvak et al., 2011), neural mass and field models are used as generative models to infer the synaptic parameters and effective connectivity that constitute active brain networks. More recently, these models have been applied in (single-region) DCMs as a “mathematical microscope”—to test synaptic hypotheses at the level of specific laminae and receptors (Moran et al., 2011b), and disambiguate between structural and functional hypotheses; for example, explaining intersubject variations in gamma oscillations (Pinotsis et al., 2013b). In this review, we summarize the state-of-the-art in modeling such population-based activity and demonstrate their use in the context of DCM. The development of these modeling approaches has been underpinned by pioneering developments several decades ago that produced generative models of EEG data characteristics based on neural masses (Wilson and Cowan, 1972; Nunez, 1974; Freeman, 1975, 1987;

Jansen and Rit, 1995; Valdes et al., 1999; Wendling et al., 2000). These models have since been refined and extended (Wright and Liley, 1996; Rennie et al., 2000) to examine a myriad of neurobiological processes including anesthesia (Steyn-Ross et al., 1999), epilepsy (Breakspear et al., 2006; Marten et al., 2009; Nevado-Holgado et al., 2012), “resting state” brain dynamics (Deco and Jirsa, 2012), and so on. The types of models we will review here have played a direct implementational role in DCM and do not represent an exhaustive overview—a non DCM-centric treatment is provided in Deco et al. (2008). Here we consider, in particular, how DCM treats ensemble neuronal activity as a point process (neural mass models) or explicitly incorporate a spatial dimension (neural field models). Both types describe so-called “mesoscopic” properties of neural activity, employing statistical mechanics to transform single unit activity into population activity—where appropriate composites can be used to generate macroscopic data. Since this mesoscale is hidden from direct observation, we demonstrate how these models rely upon and exploit knowledge about synaptic and cell physiology, as well as neuroanatomy.

Two distinct biological perspectives have informed the development of neural models, leading to a taxonomy of “convolution” or “conductance-based” models. These distinctions arose from the consideration of cortical mesocolumns—convolution models (Freeman, 1975); and separately from the consideration of a single cell’s electrophysiological properties—conductance models (Hodgkin and Huxley, 1952). Early work by Wilson and Cowan (1973) derived a sigmoidal relationship for transforming population membrane potential to an average population firing rate. These models consisted of sigmoidal and convolution-based operators and were refined on the basis of empirical observations by Freeman, Wendling and others (Freeman, 1987; Jansen and Rit, 1995; Wendling et al., 2000). In contrast, conductance based models were formulated as an equivalent circuit model of an excitable cell membrane: Hodgkin–Huxley’s original description of the giant squid axon is the classical example of this sort of model—and was reduced to a two dimensional form by Morris and Lecar (1981). Their reduced circuit has been scaled up for M/EEG analysis in DCM using the Fokker-Planck formalism (Breakspear et al., 2010) to describe the evolution of population densities (Marreiros et al., 2009). In this setting, when only the first order statistics (e.g., mean) of the population density are considered, the model describes a neural mass (where the population density can be regarded as a point of mass). When higher order statistics are considered, we obtain a “mean-field” model (where the full density of one population depends on the mean of another). We note that the terminology here is rather specific to DCM; the focus of this review. In other settings, and in other treatments of neuronal activity (Coombes, 2010; Buice and Chow, 2013), mean-field models often refer to population dynamics with interacting means only. However, here we use the term *neural-mass* to refer to an interaction in population means and *mean-field* to higher-order interactions to remain consistent with the DCM literature and to acknowledge the early neural-mass nomenclature developed by Valdez-Sosa and other pioneering work in this field (Valdes et al., 1999). Both neural mass and mean field

formulations can be applied to convolution and conductance based models: The choice of either convolution or conductance based model depends on the type of inference required (when applying the model to real data), with the latter offering a richer and biologically more realistic parameterization of synaptic currents.

The deployment of neural mass (or mean field) models of populations in DCM entails further neurobiological plausibility, through a laminar specification of cell types and their interconnectivity. For neocortical studies, a laminar architecture is populated with neuronal ensembles, so that forward (e.g., thalamo-cortical), backward or lateral (e.g., inter-hemispheric) extrinsic connections impinge upon pyramidal, spiny stellate or inhibitory interneurons (David et al., 2006). This construction is motivated by tracing studies in the macaque (Felleman and Van Essen, 1991) and demonstrates the first constraint under which these models were developed for DCM. Namely; that they conform to known physiological and anatomical principles. A second constraint is that they must be able to generate stereotypical features of empirical macroscopic measurements; for example, dominant alpha rhythms (David and Friston, 2003) or late potentials in evoked transients (Garrido et al., 2007a). In this sense, none of the models are “right” or “wrong”—but can be usefully compared to test a particular hypothesis (Box, 1976).

In addition to the distinction between neural mass and mean field formulations of either convolution or conductance based models, we also have to consider the distinction between models based upon ordinary differential equations and partial differential equations (PDEs) that endow neuronal populations with spatial attributes: incorporating the spatial domain into DCM was motivated by the advent of spatially resolved population recording modalities (Pinotsis et al., 2012). This use of neural fields, was proposed as a semi-quantitative treatment of electromagnetic brain activity by Jirsa and Haken (1996, 1997) and Robinson (2006). Crucially neural fields enable local axonal arborization to be modeled directly and can generate topological data features. These may be particularly resolved in high-density subdural grid electrodes (electrocorticography) and optical imaging techniques and also contribute to the topographical distribution of sensor/scalp space measurements in M/EEG.

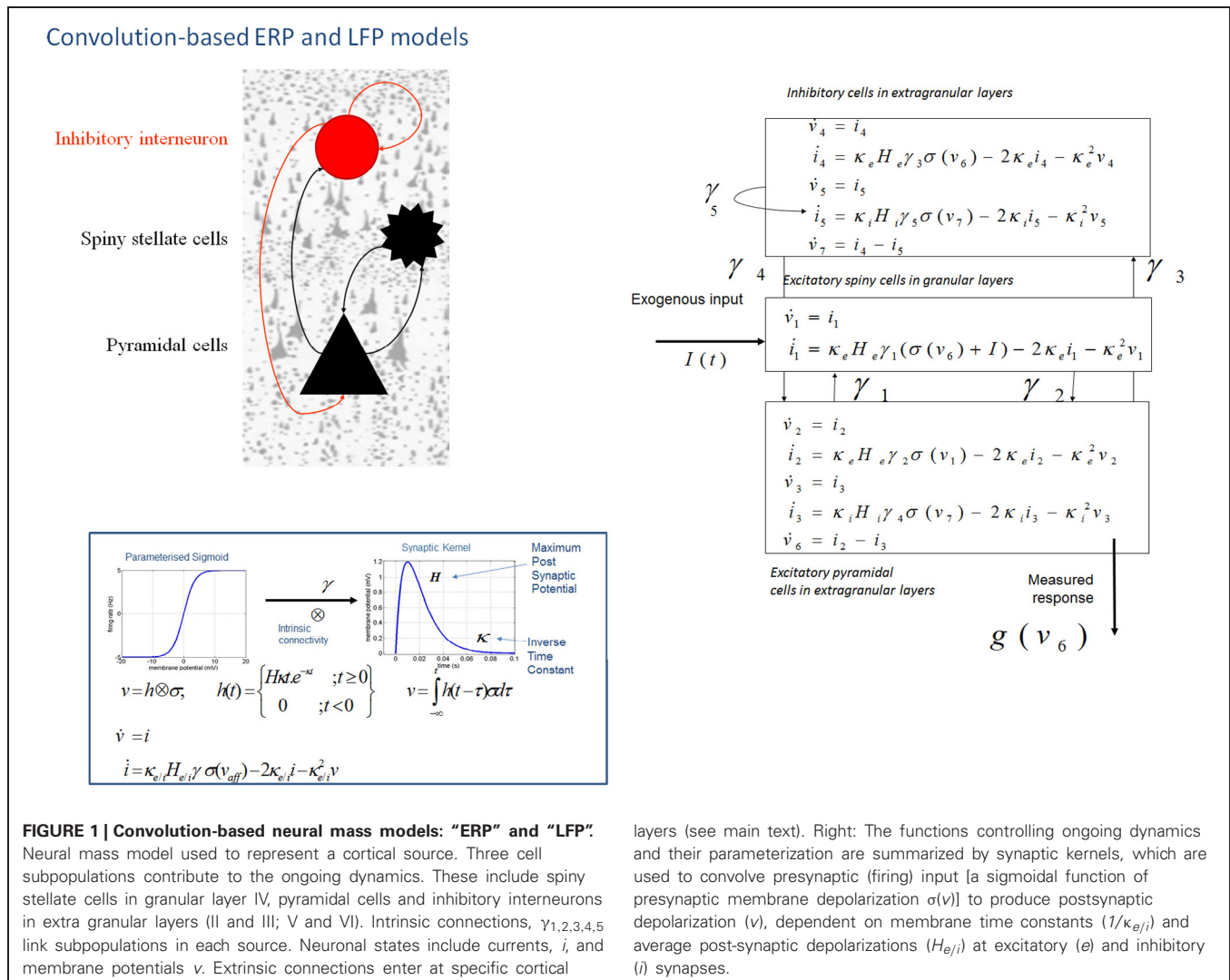
In this review, we hope to provide a didactic treatment of the neural mass and neural field models available in DCM and highlight application studies that exemplify their use. This complements more general treatments of neural population modeling (Deco et al., 2008). The first section considers convolution-based neural mass models. We will demonstrate their use in inferring causal interactions among multiple brain regions and highlight the minimal assumptions needed to form—and test—competing hypotheses. In this section, we will also introduce the important distinction between different models and different data features; noting that the same models can be used for (and indeed should be capable of generating) different data features. We will focus on the distinction between time and frequency domain responses—highlighting the use of identical neural mass models when modeling evoked and steady state responses. In the second section, we examine conductance-based models and how new currents can be added to enhance physiological detail at

the synaptic level. We also examine the impact of second-order interactions among neuronal ensembles in mean field models, particularly in the spectral domain. The third section introduces the spatial parameterization in the form of partial differential or neural field equations—and how these have been applied to test alternative explanations for gamma activity in the visual cortex. In this example, we reconsider lateral connections and the role of distinct pyramidal cell populations. This final model recapitulates “a canonical microcircuit” and provides a framework for investigating differences in directed oscillations. This development was motivated by theoretical considerations about message passing in the brain; namely predictive coding and implications for spectral asymmetries in laminar specific communication (Friston, 2005, 2009). Important asymmetries of the sort are evident in several recent empirical observations (Maier et al., 2010; Buffalo et al., 2011; Bastos et al., 2012) on the laminar specificity of oscillation frequencies. In principle, these sorts of observations can be used with DCM, to address key questions about reciprocal message passing in the brain and its hierarchical architecture.

CONVOLUTION BASED NEURAL MASS MODEL

GENERATIVE MODELS OF EVOKED RESPONSES: THE ERP MODEL

In DCM, event related potentials are modeled as the response of a dynamic input–output system to exogenous (experimental) inputs (David et al., 2006; Kiebel et al., 2006; Garrido et al., 2007b). The DCM generates a predicted ERP as the response of a network of coupled sources to sensory (thalamic) input—with the form of a narrow (16 ms) Gaussian impulse function that accounts for some temporal smoothing in thalamic volleys. Each source is modeled as a point source (c.f., equivalent current dipole) comprising three subpopulations, each assigned to a particular cortical layer. For simplicity, we place an inhibitory subpopulation in the supragranular layer. This receives inputs from excitatory deep pyramidal cells in an infra-granular layer which are, in turn, driven by excitatory spiny cells in the granular layer; layer IV. These three subpopulations are connected with intrinsic coupling parameters as shown in **Figure 1**. Though these models operate as a point process, by specifying different layers we can call on anatomical rules of extrinsic (region to region) connectivity (Felleman and Van Essen, 1991). Specifically,



layers (see main text). Right: The functions controlling ongoing dynamics and their parameterization are summarized by synaptic kernels, which are used to convolve presynaptic (firing) input [a sigmoidal function of presynaptic membrane depolarization $\sigma(v)$] to produce postsynaptic depolarization (v), dependent on membrane time constants ($1/\kappa_{e/i}$) and average post-synaptic depolarizations ($H_{e/i}$) at excitatory (e) and inhibitory (i) synapses.

the “ERP” and “LFP” convolution based models can be assembled within a generative network using three distinct types of connections. Forward connections correspond to afferent pyramidal axons and synapse on layer IV stellate cells, while backward afferents impinge upon pyramidal and inhibitory interneurons outside of layer IV. Lateral, inter-hemispheric connections are modeled with a postsynaptic response that is elicited in all layers.

The evolution of neuronal activity in this anatomical architecture is controlled by two simple operations following Jansen and Rit (1995). The first is a convolution operation (**Figure 1**) that lends the model its name and models the average membrane depolarization response as a low-pass impulse response. This transforms the average density of pre-synaptic firing arriving at the population into the average postsynaptic membrane potential (PSP). This response is specified by two biologically informed parameters; one tunes the maximum amplitude of PSPs and represents the receptor density and the second is a lumped representation of the sum of synaptic rate constants (of passive membrane and other spatially distributed delays in the dendritic tree). The output operator (**Figure 1**) then transforms this average membrane potential into the average rate of action potentials fired by the population. This transformation is assumed to be instantaneous and is described by a sigmoid function with parameters that determine its shape and location. These parameters model the voltage sensitivity or gain of the subpopulation and its average threshold. It is this function that endows the model with non-linear behaviors that are crucial for phenomena like phase-resetting of the M/EEG. The sigmoid form for these activation functions was originally motivated as arising from a unimodal distribution of threshold potentials within a population of Heaviside response units (Wilson and Cowan, 1973). More recent formulations that connect directly to full mean field (population density) treatments consider the sigmoid form to arise from the distribution of depolarizations within a population, under a fixed threshold. This output of this sigmoid function (presynaptic input) is scaled by intrinsic and extrinsic connectivity parameters from subpopulations within a source or from pyramidal cell afferents that arise from other sources in the network. Thus activity promulgates and reverberates throughout the network (**Figure 1**). Delays along these connections are also parameterized with values that correspond to the time taken for axonal propagation between layers (~ 2 ms) and regions (~ 16 ms).

A pair of ordinary differential equations completely describes the dynamics of each subpopulation within a source (**Figure 1**). These are deterministic and—for a DCM of ERPs—a thalamic impulse timed to correspond to some experimental stimulus perturbs the sources. The spatial arrangement of pyramidal cell dendrites (perpendicular to the cortical surface) renders them the prominent sources of measurable electromagnetic signals and are thus harvested from each source in the network and passed through a lead field, to produce the spatiotemporal patterns observed in M/EEG sensors.

Early applications of this model led to a series of EEG-based investigations into oddball effects by Garrido et al. (2007a, 2008, 2009). The mismatch negativity (MMN) is a negative change

in the auditory evoked potential that occurs after an unpredictable change in the acoustic stream. It is elicited, for example, when deviant frequency tones are embedded in a stream of repeated tones. The MMN has a fronto-temporal topology and is thought to reflect the updating of an internal model of the sensorium, where by a sensory prediction error is registered and a new prediction formed (Näätänen et al., 2005). Competing “ERP” neural mass models were compared (in terms of their model evidence, using standard Bayesian techniques) to probe the type of extrinsic connection changes that mediate the MMN. Indicating both a bottom-up sensory prediction error and a top-down change in predictions, a stimulus specific modulation of both forward and backward connections among hierarchically deployed sources exhibited trial specific (deviant compared to standard) modulation (Garrido et al., 2007b). This network and paradigm has since been investigated in pathological settings. A striking example by Boly et al. (2011) used the same model comparison procedure to distinguish between top-down and bottom-up extrinsic connections and their changes with levels of consciousness.

GENERATIVE MODELS OF EVOKED AND SPECTRAL RESPONSES: THE LFP MODEL

A second convolution-based model, named the “LFP” model was developed from the “ERP” model. The model was augmented to address the neurotransmitter basis of changes in intracortical local field potentials from rat prefrontal cortex (Moran et al., 2007)—and now also serves as a generative model for non-invasive EEG and MEG studies (Boly et al., 2012). The model differed from the “ERP” model in two ways: first, based on biophysical models by Whittington et al. (1995) of gamma oscillations in the hippocampus, the role of inhibitory interneurons was augmented with recurrent self-connections (**Figure 1**). This subtle addition was important from the perspective of a new set of questions. Here our goal was to develop a generative model of spectral responses can exhibit high-frequency oscillations. A second extension refined the neurophysiological input-output transforms; whereby spike-rate adaptation was modeled at the input stellate cell population. This involves the addition of currents based on the phenomenological model described by Benda and Herz (2003) and combined several ionic currents modulating spike generation—including voltage-gated potassium currents (M-type currents), the interplay of calcium currents and intracellular calcium dynamics with calcium-gated potassium channels (AHP-type currents) and the slow recovery from inactivation of the fast sodium current.

Practically these additions lead to a larger dynamic repertoire using an identical connectivity architecture (three subpopulations within a source and extrinsic forward, backward and lateral connections with laminar specificity) as that described above. The focus of the LFP model was the reproduction of fast synchronous activity as summarized with the steady-state spectral density (Fourier transform) of time series data. In order to generate steady-state spectral responses, we linearized the model’s differential equations around an equilibrium point. This equilibrium or operating point is obtained by integrating the system over a protracted time window. The linearity assumption will

accommodate parameter spaces in the region of fixed points and local bifurcations (Friston et al., 2012), known to emerge from this sort of model (Grimbert and Faugeras, 2006). In DCM, the neural masses are treated as a system that is perturbed by white and pink noise; which is explicitly parameterized. This provides a compact summary of the system, where the system's spectral responses can be obtained from its transfer functions, which depend on the physiological parameters of the model (and neural noise) (Nunez, 1974; Steyn-Ross et al., 1999; Robinson et al., 2001; Moran et al., 2007). In other words, the transfer function links unobserved physiological processes to measured spectral responses and is an essential part of forward or generative models of spectral measures. This "LFP" neural mass model was first used in a single-region DCM analysis to demonstrate how one can make inferences about synaptic function at the neuronal level, using macroscopic electrophysiological measurements. This proof of principle used microdialysis measures of extracellular neurotransmitter for validation (Moran et al., 2008) and extended the breadth of applications of DCM—in this case by inferring condition specific modulations of synaptic parameters.

CONDUCTANCE BASED NEURAL MASS MODELS

THE NMM AND MFM

In DCM, the first biophysical model of ensemble activity to receive the three-letter acronym "NMM" was described in Marreiros et al. (2008, 2009). These models parameterize neuronal dynamics at the level of a single neuron and employ density-based flow (the time derivative of neural activity) mechanics to represent the dynamics of a population of neurons. In coupling multiple subpopulations with different characteristics within and between regions, the framework used a mean-field reduction (assumption); where each type of population comprises a probability density and is only affected by the main activity of other neuronal populations or ensembles in the model. This formal population treatment was applied in the context of a conductance-based model predicated on the Morris–Lecar equivalent RC-circuit description of oscillatory membrane properties in barnacle muscle fiber (Morris and Lecar, 1981). These models equate capacitive current (according to Kirchhoff's current law) with the summed active and passive currents across the membrane. Morris–Lecar models can be thought of as reductions of Hodgkin and Huxley's model of the squid axon. They include active currents that describe ligand-gated excitatory (Na^+) and inhibitory (Cl^-) ion flow, mediated through fast glutamatergic and GABAergic receptors, with a potassium leak current used to account for all passive ionic currents (Gutkin et al., 2003); where the conductance of the active channels display first order dynamics that depend on the time constant of the channel and their current state (Figure 2).

It may seem that we are conflating the introduction of mean field (versus neural mass) formulations with the introduction of conductance-based models. However, there is a fundamental reason for doing this: in full mean field treatments, the ordinary differential equations describe the dynamics of first and higher-order statistics of population densities—such as the covariance among neuronal states within a population. Crucially,

the covariance depends upon the mean of the neuronal states when, and only when, the equations are non-linear in the states (i.e., where the states interact multiplicatively). In other words, the weakly non-linear (sigmoidal) equations of motion of convolution-based models mean that the covariances are not functions of the population averages and therefore do not change with time. This is why one only has to consider first-order statistics in neural mass models based upon (linear) synaptic convolution operators. However, when we move to conductance-based models, there is a necessary interaction between conductance and depolarization, which renders the models intrinsically non-linear. This means that the covariances depend upon the means (and vice versa). It therefore only makes sense to consider mean field treatments of conductance-based models.

Mean field treatments start with a description of a single neuronal response in terms of stochastic differential equations (presented in Figure 2) that accommodate noise or fluctuations in neuronal states (Note: In physics generally a mean-field reduction refers to any statistical summary i.e., to first or higher order, to describe population responses. In the context of neuronal population models of brain function, the literature has adopted a standard where a neural-mass refers to first order and a mean-field; first and higher order interactions (Deco et al., 2008). Marreiros et al. (2009) used these stochastic differential equations to form a set of non-linear ordinary differential equations by applying the Fokker-Planck formalism using the Laplace assumption (or method of moments). This meant that ensemble activity could be modeled without the need to simulate individual unit activity with noisy fluctuations—the neuronal fluctuations are implicit in the population density dynamics. Heuristically, the population density approach is important because it provides a unique prediction or generative model of an empirical response. Mean field formulations of conductance-based models are therefore invertible (can be fitted to data) in the setting of DCM. The resulting expression for the population dynamics; the evolution of the population's mean and variance, decomposes into deterministic flow and diffusion. In turn, these reduce to simple forms under Gaussian (Laplace) assumptions about the population density—where first order population dynamics are a function of flow and the curvature of the flow, and the second order statistics a function of the gradients of flow. However, a DCM of a single source typically comprises three coupled populations (David and Friston, 2003).

For conductance based models in DCM, we employed a similar structure to the "ERP" and "LFP" mass models—with interacting populations of excitatory spiny stellate cells, pyramidal cells and inhibitory interneurons, coupled through intrinsic connections (Figure 2). The mean field partition means that these operations can be applied to each population in turn, and lead to simple expressions; given that the gradient and curvature of the equations of motion are only non-zero within a particular ensemble. These equations describe a mean-field model; the "MFM," in which the first and second order sufficient statistics interact, influencing each other when the curvature (derivative of the flow) is non-zero. The formally reduced model, where only first order interactions are considered is termed the "NMM". This point mass interaction is identical to the ERP and LFP

Conductance-based NMM and MFM

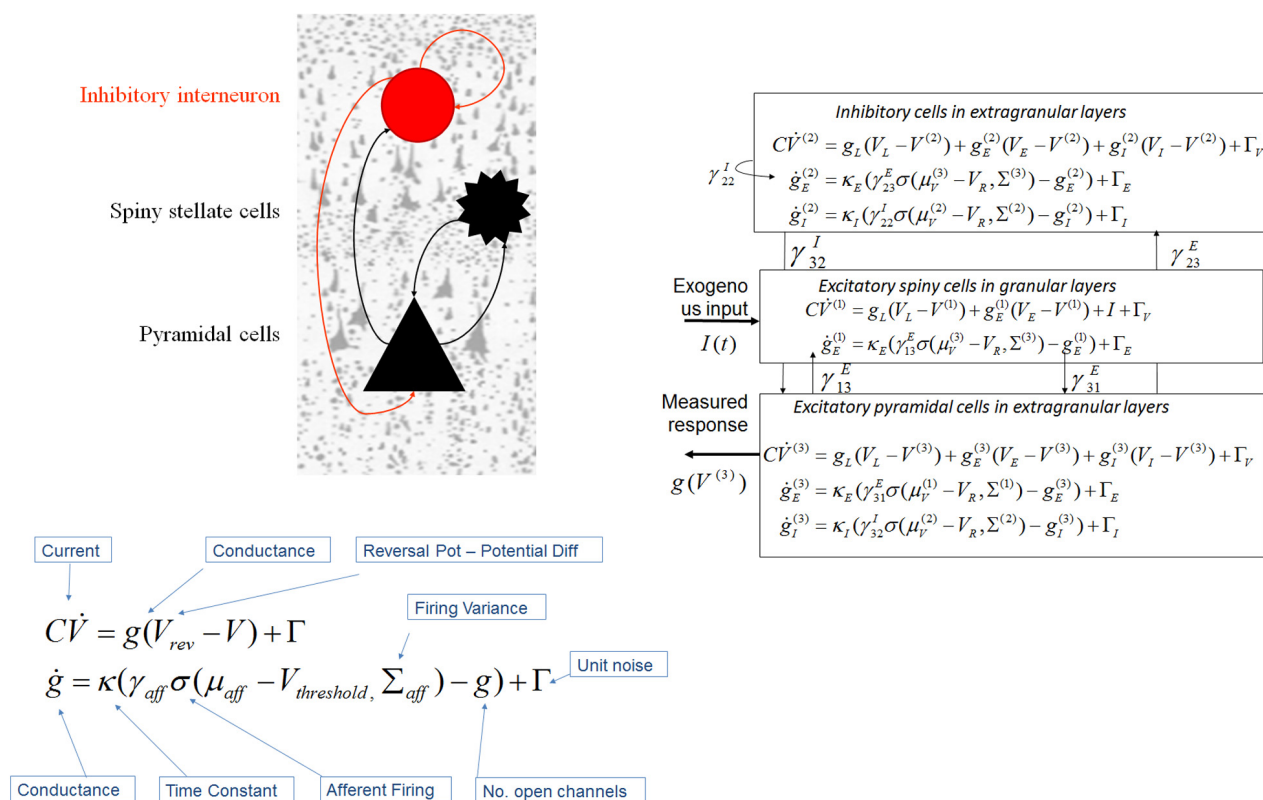


FIGURE 2 | Conductance-based neural mass models: “NMM” and “MFM”. This figure shows Morris–Lecar-type differential equations describing the time evolution of a single cell current (capacitance \times change of membrane potential: $C\dot{V}$) and conductance (g) at inhibitory interneurons (extra granular layers), spiny stellate cells (granular layers) and pyramidal cells (extra granular layers). In this model, all cell types possess AMPA receptors, GABA_A; with ion-channel time constants ($1/\kappa_{E/I}$). Layers

are connected with strengths parameterized by γ . V_L , V_E , and V_I are reversal potentials for leak potassium channels, sodium, and chloride channels, respectively, at V_T is the threshold potential. NMDA receptors at pyramidal cells and inhibitory interneurons can be added using a conductance equation of similar form, weighted by a voltage dependent switch (Moran et al., 2011a,b). For a full population Fokker-Planck characterization see Marreiros et al. (2008).

models in terms of the effects different subpopulations exert on others, but have a different dynamic form through the equivalent RC circuit description. Unlike the NMM, the MFM allows for expansion (dynamic increase in the variance of the state’s associated probability distribution) and contraction (dynamic reduction in the variance of the state’s associated probability distribution) of dispersed neuronal states to influence the average flow.

Marreiros et al. (2010) compared DCM using mean-field models (MFM) with dynamically coupled means and variances to a model where the variance was fixed (NMM). In the time domain, a simulated evoked potential elegantly demonstrated the effect of this coupling with the variance of the pyramidal cell population’s depolarization contracting to close to zero when the mean approaches its maximum. Within the spectral domain, they similarly showed qualitative differences in the dynamic repertoire—with the MFM displaying limit-cycle attractors after bifurcation from a fixed-point (Marreiros et al.,

2009). In this setting, the mean-field model is inherently more non-linear, because it entails non-linear interactions between the first and second order statistics of the hidden states (i.e., dynamic processes that affect the observations, but cannot be directly measured. the activity in interneurons affect signal propagation within regions but due to the random pattern of their dendrites do not directly contribute to the measured field). This speaks to many similar investigations of non-linearity in neural systems (Lopes Da Silva et al., 1989; Destexhe and Babloyantz, 1991; Daffertshofer et al., 2000; Breakspear, 2002; Breakspear and Terry, 2002; Stam, 2005), where the emergence of quasiperiodic, chaotic and itinerant attractors belies a rich set of dynamic phenomena with physiologically plausible interpretations.

THE NMM AND MFM WITH VOLTAGE GATED NMDA RECEPTORS

An extension to the NMM and MFM was presented in Moran et al. (2011a) through the inclusion of a third ligand-gated

ion channel to model conductances controlled by the NMDA receptor. NMDA receptor controlled ion channels were considered in a separate treatment since they are both ligand- and voltage-gated. For an NMDA channel to open, following the binding of glutamate, there must first be a large transmembrane potential to remove a magnesium ion blocking the channel. Hence, the dynamics for this particular current are given by an extended equation, which includes the magnesium component using a voltage gated function (**Figure 2**). By introducing NMDA ion channels to pyramidal cells and inhibitory interneurons (Brunel and Wang, 2001), we further constrained and distinguished laminar specific responses. This type of channel afforded another source of non-linearity due to its voltage dependency (Jahr and Stevens, 1990), and required an extension of the steady-state linearization, for both the NMM and MFM case.

To characterize the dynamic repertoire of these models, we examined steady-state responses in the frequency domain to identify regimes where the system settles to a fixed point (i.e., the average activity reaches steady-state) or a limit cycle. In the first regime the spectrum observed is generated by noise, where neuronal populations act as a filter—shaping the noise spectrum to produce a profile of output frequencies. In the second dynamic regime, the average neuronal states themselves may oscillate. In this situation, the system exhibits what is known as a quasiperiodic attractor and the frequency response to noise changes with different points on the attractor (Moran et al., 2011a). This means one has to take the average frequency response over the attractor manifold (i.e., over the limit cycle). Crucially, the frequencies that are preferentially passed by the system are also the frequency of the oscillation (limit cycle). This means the predicted spectral responses to noise under steady state can be seen (and treated mathematically) as a special case that obtains when the attractor collapses to a fixed point. This second regime was particularly evident in the MFM case, where attractor subspaces characteristic of heteroclinic channels were observed—and produced bimodal spectral peaks from local and global state space trajectories.

This richly parameterized neural mass model was then used to examine distinctions among the type of receptors underlying empirical neural activations in EEG and MEG. In Moran et al. (2011b), we tried to recover pharmacologically induced changes in receptor processing using MEG, during a visuo-spatial working memory task. Specifically, parameter estimates from the spectral response in superior frontal gyrus, disclosed an effect of L-Dopa on delay period activity—in terms of how L-dopa changed specific synaptic (connectivity) parameters (**Figure 2**). These effects were exactly commensurate with predictions from the animal and computational literature (Goldman-Rakic, 1996; Durstewitz et al., 2000; Gorelova and Yang, 2000; Gonzalez-Islas and Hablitz, 2003; Durstewitz and Seamans, 2008) and revealed the dual mechanisms of dopaminergic modulation of glutamatergic processing; where L-Dopa increased the non-linearity of post-synaptic responses mediated by NMDA receptors, and decreased AMPA coupling between pyramidal cells and stellate cells. In this study, we also found an L-Dopa-dependent change in exogenous input into the frontal region, which effectively

suppresses this region during delay-related reverberatory processing. Moreover, individual parameter estimates from the DCM correlated with individual performance indices (Moran et al., 2011b), a crucial finding that is often used to show that parameter estimates from an NMM reflect real neuronal processes.

SPATIAL HARMONICS AND THE NEURAL FIELD

THE NFM

Pinotsis et al. (Pinotsis and Friston, 2011; Pinotsis et al., 2012) introduced the spatial domain into DCM with neural field models. Neural fields model current fluxes as continuous processes on the cortical sheet, using PDEs. The key advance that neural field models offer, over conventional neural mass models, is that they embody spatial parameters (like the density and extent of lateral connections). This means that, in principle, one can infer the spatial parameters of cortical infrastructures generating electrophysiological signals (and infer changes in those parameters over different levels of an experimental factor) from empirical data. This rests on modeling responses not just in time but also over space. This sort of model should be ideally suited to exploit the temporal dynamics of observed cortical responses with a high spatial resolution; for example, with high-density recordings, at the epidural or intracortical level. However, as demonstrated in early DCM-NFM, the impact of spatially extensive dynamics is not restricted to expression over space but can also have profound effects on temporal (e.g., spectral) responses at one point (or averaged locally over the cortical surface) (Pinotsis et al., 2012). This means that neural field models may also play a key role in the modeling of non-invasive electrophysiological data that does not resolve spatial activity directly. Although, neural mass models can describe patterns in sensor space, the spatial attributes of these patterns result from the coupling among states at different points in source space and not from hidden states that are functions space (i.e., they are described by equations that are time dependent but not spatially dependent).

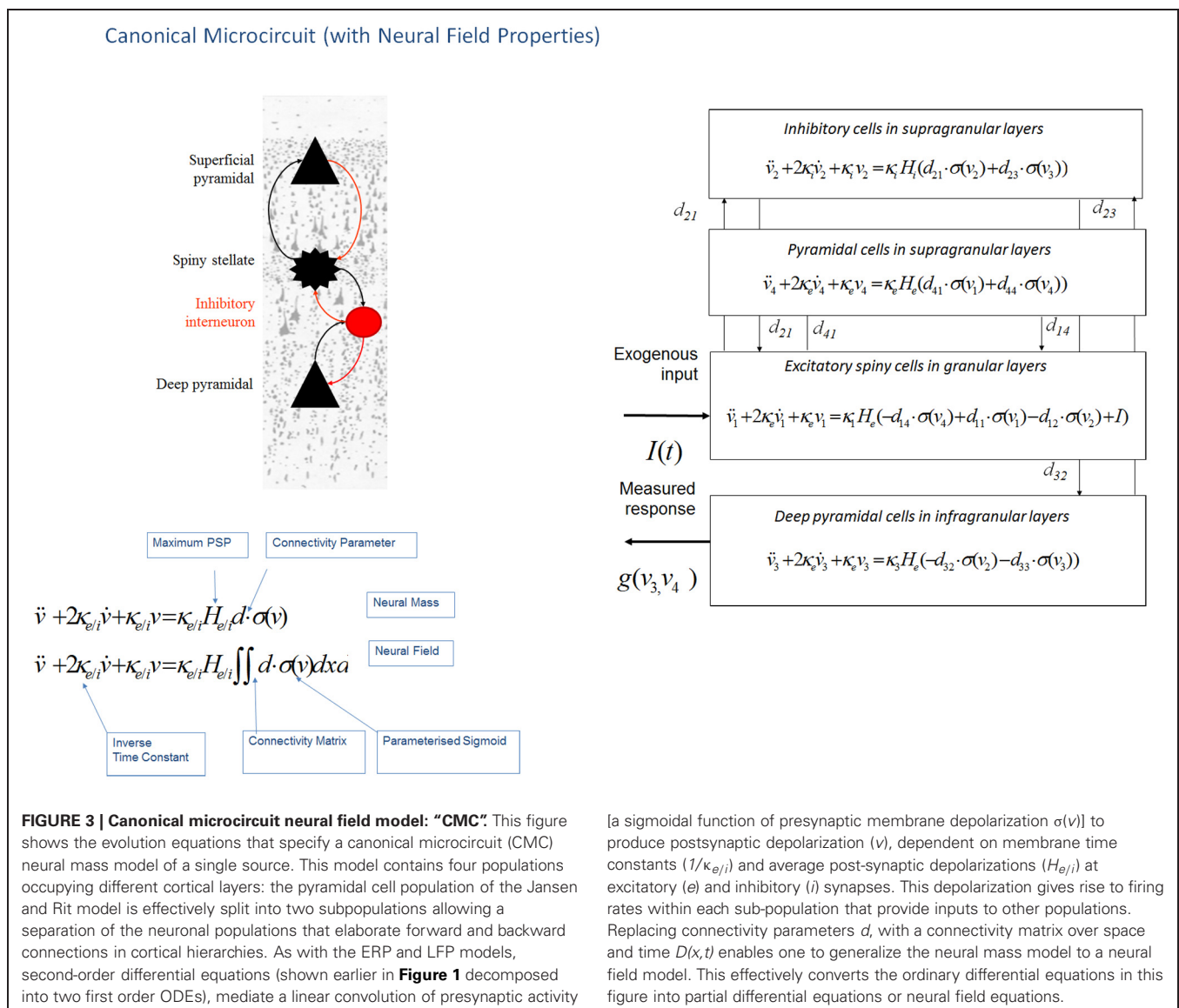
In terms of anatomical and physiological constraints, functional specialization demands that cells with common functional properties are grouped together. This architectural constraint necessitates both convergence and divergence of cortical connections (Zeki, 1990), of the sort that can be modeled with a neural field model. To model these spatial aspects of connectivity one needs partial differential or integro differential equations that accommodate lateral interactions over spatially extended cortical manifolds. In Pinotsis and Friston (2011), the repertoire of steady state regimes engendered by sparse (patches of) intrinsic connections was examined. Specifically, we considered a bimodal, non-centric distribution and showed (through a Turing instability analysis) that the dispersion relation from this particular arrangement of spatial delays leads to infinite branches of complex spectra. These branches undergo similar conformational changes, under both increased propagation velocity and decreased spatial separation (range) of lateral connections. The resulting fall in the amplitude of high frequency oscillations was also apparent in the spectral summaries of these responses, in terms of cross spectral densities. For example, as the separation of coupled neuronal populations increases, the total spectral power

decreases and falls faster for higher frequencies in a manner similar to local coherence functions based on primate recordings (see Leopold et al., 2003). In brief, both the spatial deployment and the speed of lateral connections can have a profound effect on the behavior of spatial harmonics over different scales. Interestingly, it turned out that only synaptic gain was capable of producing phase-transitions: when increasing gain, the system was driven to an unstable regime and oscillations appeared as a result of a Hopf bifurcation.

This neural field model was later extended to a three layered architecture comprising pyramidal cells, inhibitory interneurons and spiny stellate cells (Figure 3), where spatial delays result from signals propagating with finite conduction speeds along axonal arbors (Pinotsis et al., 2012). These arbors were arranged with a central distribution of synaptic densities, which decayed exponentially in space. Spatial delays operated along the same intrinsic connections used in ERP neural mass models:

the cortical micro circuitry in the “ERP” and “NFM” models. This means that the neural mass and field models are essentially the same; describing the same neurobiological dynamics over time, but where the “NFM” is equipped with spatially extended hidden states that characterize presynaptic input as a spatially-extended process that is propagated along axonal arbors. In contrast the hidden states in a neural mass model are a function of time only. As with the “ERP” and “LFP” models, this model (“NFM”) uses a convolution operator to characterize post synaptic filtering and a sigmoid function to accommodate the dispersion of the hidden states of the afferent population (Figure 3).

The advantage of neural field models is that they can accommodate spatially extended activity on cortical manifold or patches that endows the predicted responses with a complicated frequency dependency. This allows one to distinguish between spatial effects and other factors (such as intrinsic cell properties) on



the basis of observed (empirical) responses. The incorporation of neural field models in the DCM framework allowed spatial parameters of the sources—like the spatial decay rate of synaptic connections and intrinsic conduction speed—to be optimized even using spatially unresolved data, like a time series from a single LFP channel. For example, Bayesian model selection (BMS) correctly distinguished between mass and field models. This type of comparison was formalized using DCM and Bayesian model evidence in the context of invasive local field potentials from rat auditory cortex. With these invasive data, the neural-field model had a much greater evidence than the equivalent neural mass variant; this could be attributed to the increased repertoire of predictions that these models afford and indicates a key role for—and parameterization of—spatial as well as temporal dynamics on the cortical manifold.

STRUCTURE, FUNCTION, AND THE CANONICAL MICROCIRCUIT (CMC)

Pinotsis et al. (2013b) showed that DCM with neural fields can provide a detailed analysis of correlations between cortical structure and function. This analysis was motivated by previous results suggesting two hypotheses regarding the biophysical basis of inter-individual differences in peak gamma frequencies—one based on functional differences and one based on structural differences: Muthukumaraswamy et al. (2009) suggested that peak gamma frequency is determined by the level of inhibition in V1, as described by resting GABA concentration measured with MR spectroscopy. Later, Schwarzkopf et al. (2012) found a correlation between peak gamma frequency and the surface area of primary visual cortex as measured with retinotopic mapping. These authors suggested that the size of V1 and associated differences in structural microanatomy could be true determinants of peak gamma frequency. The above two hypotheses suggest that both GABA concentration and V1 size can influence gamma frequency; however, these factors may or may not be causally linked. In other words, a larger V1 may have a higher GABA concentration that may or may not be due to a higher local GABA density.

DCM with neural fields incorporate parameters pertaining to both microanatomy and the density of GABA receptors (that determine inhibitory intrinsic connection strengths) and allow one to investigate alternative explanations for differences in gamma peak frequency. These differences could be mediated by either a kinetic (functional) parameter, summarizing the level of cortical inhibition or the (structural) macrocolumn width or both of these parameters. Pinotsis et al. (2013b) looked at the correlations over subjects between peak gamma frequency, V1 surface area and the Bayesian estimates of these structural and functional parameters and found that both hypotheses were confirmed. In brief, they found correlations between columnar width, gamma peak and V1 size and also between the GABAergic parameter and the gamma peak. This correlation remained significant when controlling for V1 size and width. This suggests that the correlation between gamma peak and V1 inhibition cannot be accounted for completely by the spatial parameters (at the microscopic or macroscopic level). Structural equation modeling was used to characterize the causal dependencies among observed quantities and the model parameters: in the winning structure

equation model, peak gamma frequency was mediated proximately by excitatory drive to inhibitory (GABAergic) interneurons and the strength of this drive was determined, in part, by the size of macrocolumns. In turn, the size of the macrocolumn was constrained by the macroscopic (retinotopic) size of V1 (under the assumption that V1 size is determined genetically or epigenetically). These results suggested that both cortical microstructure and excitability may be important for visual perception and are in accord with empirical studies showing that the size of V1 is negatively correlated with the strength of visual illusions (Schwarzkopf et al., 2010)—and that GABA concentration correlates with orientation discrimination ability (Edden et al., 2009).

This work also introduced an expanded neuronal architecture based upon the canonical microcircuit. These models comprise four subpopulations (as opposed to three subpopulations in mass and field models above, see **Figure 3**). The canonical microcircuit or “CMC” models a refinement of the Jansen and Rit convolution models that explicitly accommodates the neuronal sources of forward and backward connections in cortical hierarchies (Bastos et al., 2012). These are distinct superficial and deep pyramidal cell populations, respectively that, crucially, may exhibit different spectral outputs. Specifically gamma responses have been recorded in superficial layers, while slower dynamics (in the alpha and beta range) arise concurrently in infra granular populations (Maier et al., 2010; Buffalo et al., 2011; Bastos et al., 2012). The CMC proposes an intrinsic connectivity architecture to account for this non-linear transformation through dendritic and population effects (Bastos et al., 2012). The canonical microcircuit model is based upon intracellular recordings in cat visual cortex by Douglas and Martin (1991) who investigated the laminar propagation of afferent signals and produced pathways that are thought to reflect canonical input–output processing streams for forward and backward signals throughout the cortex. These models are currently being used to test hypotheses about asymmetries in forward and backward message passing that may shed light on the distributed neural processing that underlies perceptual synthesis and inference.

SUMMARY

The basic idea behind DCM is that neural activity propagates through brain networks in a way that can be modeled as an input-state-output-system, where causal interactions within the system are mediated by unobservable (hidden) dynamics. The resulting multi-input-multi-output (MIMO) neuronal model is then supplemented with an observation model (e.g., classical electromagnetic forward model) that describes the mapping from neural activity to observed responses (Daunizeau et al., 2011). It is the dynamics of the hidden states that are prescribed by the neural mass and neural field models outlined above. The type of data and data features determine the combination of neuronal and observation models. For example, EEG and MEG data require a different observation model than LFP data, while evoked responses necessitate a parameterized Gaussian pulse input—in contrast with spectral density data features that require parameterized neuronal noise spectra. In all of these different applications, the underlying “LFP,” “ERP,”

“NMM,” “MFM,” “NFM,” and “CMC” models are, in principle, interchangeable.

The choice of the appropriate neuronal model should reflect the research question at hand: for example, whether the focus is on topographic as opposed to intrinsic neurotransmitter properties or drug effects etc. This choice may also be informed by previous applications, where a particular model has already proven useful within a DCM context (for example, fast oscillations in the gamma band for the “CMC” model). One generally designs a DCM study to assess the effects of task manipulation, group, pathology or drug on a particular parameter or set of parameters of interest. In other words, the choice of model should be evident at the outset and often conforms to the “minimal model approach” necessary to access that parameter—this is because, in general, a simpler model has more constraints and can use the degrees of freedom in the data to estimate model parameters and evidence more efficiently. For example, where differences in effective extrinsic connectivity are of interest, a convolution based model—that is agnostic to specific intrinsic ion channel mediators—will suffice to address the hypothesis (Campo et al., 2012). The direction of empirical research using DCM as a “mathematical microscope” of synaptic processes—where particular receptor and neurotransmitter changes are important—may call for the finer grained physiological details of the NMM or MFM (Moran et al., 2011b). In other cases the form of the dynamics, for example whether field (or propagation) effects should be considered, may itself embody the central hypothesis. In this case, as demonstrated in Pinotsis et al. (2013b), BMS may be sufficient to disambiguate among competing hypotheses about neuronal architecture (Penny et al., 2004). In effect, BMS usurps all other arguments as the best method to test which model should be applied to which data; though computational and time constraints, particularly as the suite of options in DCM is expanded, may determine the extensiveness and overall feasibility of such a search. In principle, researchers may employ their own neuronal model (or feature extraction process) by compositing a parameterized state space and utilizing the modularity of SPM’s source

code. Wrapper routines which specify parameter priors, integration schemes and variational expectation maximization can be applied to a function of neuronal activity that is user-specified (In SPM’s DCM toolbox: <http://www.fil.ion.ucl.ac.uk/spm/>, The current routines are described in files “spm_fx_nmm”) (Kiebel et al., 2009; Litvak et al., 2011). Construct validity would then need to be tested by simulating data from different regions of parameter space and investigating parameter identify ability i.e., whether the framework can recover the simulated parameters given different initializations. These tests are particularly important in the case of highly nonlinear state space models, given the potential of the gradient ascent to converge to local maxima (Friston et al., 2012).

The models we have reviewed in this paper may also be useful beyond the DCM inference framework. For example, in the study of large scale generative processes underlying resting-state networks in fMRI, neural mass (Deco and Jirsa, 2012) and neural field models (Pinotsis et al., 2013a) have been embedded in anatomical graphs to study emergent behaviors and dynamic properties (Gray et al., 2009; Robinson et al., 2009). This suggests that these models may have some utility as neural state equations in DCM for fMRI, though currently we deploy coarser models with far less physiological detail (Daunizeau et al., 2011). The field of inquiry using these types of models is varied and rich, with DCM applications including the locus of consciousness and unconsciousness in vegetative state patients (Boly et al., 2011), diaschisis in temporal lobe epilepsy (Campo et al., 2012) and the effects of ketamine on synaptic plasticity (Schmidt et al., 2012). The set of distinct dynamics within which these types of effects can be parameterized will no doubt grow beyond the current suite of models available. Indeed their exchangeability within the DCM framework allows researchers to define their own favorite or interesting model and proceed in the usual Bayesian way (Friston et al., 2003).

ACKNOWLEDGMENTS

This work was supported by the Wellcome Trust and Virginia Tech Carilion Research Institute.

REFERENCES

- Ashburner, J. (2012). SPM: a history. *Neuroimage* 62, 791–800.
- Baillet, S., Mosher, J. C., and Leahy, R. M. (2001). Electromagnetic brain mapping. *Signal Process. Mag. IEEE* 18, 14–30.
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711.
- Benda, J., and Herz, A. V. M. (2003). A universal model for spike-frequency adaptation. *Neural Comput.* 15, 2523–2564.
- Boly, M., Garrido, M. I., Gosses, O., Bruno, M. A., Boveroux, P., Schnakers, C., et al. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science* 332, 858–862.
- Boly, M., Moran, R., Murphy, M., Boveroux, P., Bruno, M. A., Noirhomme, Q., et al. (2012). Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. *J. Neurosci.* 32, 7082–7090.
- Box, G. E. P. (1976). Science and statistics. *J. Am. Stat. Assoc.* 71, 791–799.
- Breakspear, M. (2002). Nonlinear phase desynchronization in human electroencephalographic data. *Hum. Brain Mapp.* 15, 175–198.
- Breakspear, M., Heitmann, S., and Daffertshofer, A. (2010). Generative models of cortical oscillations: neurobiological implications of the Kuramoto model. *Front. Hum. Neurosci.* 4:190. doi: 10.3389/fnhum.2010.00190
- Breakspear, M., Roberts, J., Terry, J. R., Rodrigues, S., Mahant, N., and Robinson, P. (2006). A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cereb. Cortex* 16, 1296–1313.
- Breakspear, M., and Terry, J. (2002). Topographic organization of nonlinear interdependence in multichannel human EEG. *Neuroimage* 16, 822–835.
- Brunel, N., and Wang, X. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.* 11, 63–85.
- Buffalo, E. A., Fries, P., Landman, R., Buschman, T. J., and Desimone, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11262–11267.
- Buice, M. A., and Chow, C. C. (2013). Dynamic finite size effects in spiking neural networks. *PLoS Comput. Biol.* 9:e1002872. doi: 10.1371/journal.pcbi.1002872
- Campo, P., Garrido, M. I., Moran, R. J., Maestú, F., García-Morales, I., Gil-Nagel, A., et al. (2012). Remote effects of hippocampal sclerosis on effective connectivity during working memory encoding: a case of connectional diaschisis? *Cereb. Cortex* 22, 1225–1236.
- Coomes, S. (2010). Large-scale neural dynamics: simple and complex. *Neuroimage* 52, 731–739.
- Daffertshofer, A., Peper, C., Frank, T., and Beek, P. (2000).

- Spatio-temporal patterns of encephalographic signals during polyrhythmic tapping. *Hum. Mov. Sci.* 19, 475–498.
- Daunizeau, J., David, O., and Stephan, K. E. (2011). Dynamic causal modelling: a critical review of the biophysical and statistical foundations. *Neuroimage* 58, 312–322.
- David, O., and Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage* 20, 1743–1755.
- David, O., Kiebel, S., Harrison, L., Mattout, J., Kilner, J., and Friston, K. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. *Neuroimage* 30, 1255–1272.
- Deco, G., and Jirsa, V. K. (2012). Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *J. Neurosci.* 32, 3366–3375.
- Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., and Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* 4:e1000092. doi: 10.1371/journal.pcbi.1000092
- Destexhe, A., and Babloyantz, A. (1991). Pacemaker-induced coherence in cortical networks. *Neural Comput.* 3, 145–154.
- Douglas, R. J., and Martin, K. (1991). A functional microcircuit for cat visual cortex. *J. Physiol.* 440, 735–769.
- Durstewitz, D., Seamans, J., and Sejnowski, T. (2000). Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J. Neurophysiol.* 83, 1733.
- Durstewitz, D., and Seamans, J. K. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. *Biol. Psychiatry* 64, 739–749.
- Edden, R. A., Muthukumaraswamy, S. D., Freeman, T. C., and Singh, K. D. (2009). Orientation discrimination performance is predicted by GABA concentration and gamma oscillation frequency in human primary visual cortex. *J. Neurosci.* 29, 15721–15726.
- Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1.
- Freeman, W. J. (1975). *Mass Action in the Nervous System*. New York, NY: Academic Press.
- Freeman, W. J. (1987). Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biol. Cybern.* 56, 139–150.
- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 815–836.
- Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 293–301.
- Friston, K., Bastos, A., Litvak, V., Stephan, K., Fries, P., and Moran, R. (2012). DCM for complex-valued data: cross-spectra, coherence and phase-delays. *Neuroimage* 59, 439–455.
- Friston, K., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *Neuroimage* 19, 1273–1302.
- Garrido, M., Friston, K., Kiebel, S., Stephan, K., Baldeweg, T., and Kilner, J. (2008). The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage* 42, 936–944.
- Garrido, M., Kilner, J., Kiebel, S., and Friston, K. (2007a). Evoked brain responses are generated by feedback loops. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20961.
- Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K. E., and Friston, K. J. (2007b). Dynamic causal modelling of evoked potentials: a reproducibility study. *Neuroimage* 36, 571–580.
- Garrido, M., Kilner, J., Stephan, K., and Friston, K. (2009). The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463.
- Goldman-Rakic, P. (1996). Regional and cellular fractionation of working memory. *Proc. Natl. Acad. Sci. U.S.A.* 93, 13473.
- Gonzalez-Islas, C., and Hablitz, J. (2003). Dopamine enhances EPSCs in layer II-III pyramidal neurons in rat prefrontal cortex. *J. Neurosci.* 23, 867.
- Gorelova, N., and Yang, C. (2000). Dopamine D1/D5 receptor activation modulates a persistent sodium current in rat prefrontal cortical neurons *in vitro*. *J. Neurophysiol.* 84, 75.
- Gray, R., Fung, C., and Robinson, P. (2009). Stability of small-world networks of neural populations. *Neurocomputing* 72, 1565–1574.
- Grimbert, F., and Faugeras, O. (2006). Bifurcation analysis of Jansen's neural mass model. *Neural Comput.* 18, 3052–3068.
- Gutkin, B., Pinto, D., and Ermentrout, B. (2003). Mathematical neuroscience: from neurons to circuits to systems. *J. Physiol. Paris* 97, 209–219.
- Hodgkin, A., and Huxley, A. (1952). Propagation of electrical signals along giant nerve fibres. *Proc. R. Soc. Lond. B Biol. Sci.* 140, 177–183.
- Jahr, C. E., and Stevens, C. F. (1990). A quantitative description of NMDA receptor-channel kinetic behavior. *J. Neurosci.* 10, 1830–1837.
- Jansen, B. H., and Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.* 73, 357–366.
- Jirsa, V., and Haken, H. (1996). Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* 77, 960–963.
- Jirsa, V., and Haken, H. (1997). A derivation of a macroscopic field theory of the brain from the quasi-microscopic neural dynamics. *Physica D* 99, 503–526.
- Kiebel, S., David, O., and Friston, K. (2006). Dynamic causal modelling of evoked responses in EEG/MEG with lead field parameterization. *Neuroimage* 30, 1273–1284.
- Kiebel, S. J., Garrido, M. I., Moran, R., Chen, C. C., and Friston, K. J. (2009). Dynamic causal modelling for EEG and MEG. *Hum. Brain Mapp.* 30, 1866–1876.
- Leopold, D. A., Murayama, Y., and Logothetis, N. K. (2003). Very slow activity fluctuations in monkey visual cortex: implications for functional brain imaging. *Cereb. Cortex* 13, 422–433.
- Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., et al. (2011). EEG and MEG data analysis in SPM8. *Comput. Intell. Neurosci.* 2011:852961. doi: 10.1155/2011/852961
- Lopes Da Silva, F., Pijn, J. P., and Boeijinga, P. (1989). Interdependence of EEG signals: linear vs. nonlinear associations and the significance of time delays and phase shifts. *Brain Topogr.* 2, 9–18.
- Maier, A., Adams, G. K., Aura, C., Leopold, D. A., Maier, A., Adams, G., et al. (2010). Distinct superficial and deep laminar domains of activity in the visual cortex during rest and stimulation. *Front. Syst. Neurosci.* 4:31. doi: 10.3389/fnsys.2010.00031
- Marreiros, A., Kiebel, S., Daunizeau, J., Harrison, L., and Friston, K. (2009). Population dynamics under the Laplace assumption. *Neuroimage* 44, 701–714.
- Marreiros, A. C., Daunizeau, J., Kiebel, S. J., and Friston, K. J. (2008). Population dynamics: variance and the sigmoid activation function. *Neuroimage* 42, 147–157.
- Marreiros, A. C., Kiebel, S. J., and Friston, K. J. (2010). A dynamic causal model study of neuronal population dynamics. *Neuroimage* 51, 91–101.
- Marten, F., Rodrigues, S., Suffczynski, P., Richardson, M. P., and Terry, J. R. (2009). Derivation and analysis of an ordinary differential equation mean-field model for studying clinically recorded epilepsy dynamics. *Phys. Rev. E* 79:021911. doi: 10.1103/PhysRevE.79.021911
- Moran, R., Kiebel, S., Stephan, K., Reilly, R., Daunizeau, J., and Friston, K. (2007). A neural mass model of spectral responses in electrophysiology. *Neuroimage* 37, 706–720.
- Moran, R., Stephan, K., Kiebel, S., Rombach, N., O'Connor, W., Murphy, K., et al. (2008). Bayesian estimation of synaptic physiology from the spectral responses of neural masses. *Neuroimage* 42, 272–284.
- Moran, R. J., Stephan, K. E., Dolan, R. J., and Friston, K. J. (2011a). Consistent spectral predictors for dynamic causal models of steady state responses. *Neuroimage* 55, 1694–1708.
- Moran, R. J., Symmonds, M., Stephan, K. E., Friston, K. J., and Dolan, R. J. (2011b). An *in vivo* assay of synaptic function mediating human cognition. *Curr. Biol.* 21, 1320–1325.
- Morris, C., and Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* 35, 193–213.
- Muthukumaraswamy, S. D., Edden, R. A., Jones, D. K., Swettenham, J. B., and Singh, K. D. (2009). Resting GABA concentration predicts peak gamma frequency and fMRI amplitude in response to visual stimulation in humans. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8356–8361.
- Näätänen, R., Jacobsen, T., and Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. *Psychophysiology* 42, 25–32.
- Nevado-Holgado, A. J., Marten, F., Richardson, M. P., and Terry, J. R. (2012). Characterising the dynamics of EEG waveforms as the path through parameter space of a neural mass model: application to epilepsy seizure evolution. *Neuroimage* 59, 2374–2392.
- Nunez, P. L. (1974). The brain wave equation: a model for the EEG. *Math. Biosci.* 21, 279–297.
- Nunez, P. L., and Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. New York, NY: Oxford University Press.

- Penny, W., Stephan, K., Mechelli, A., and Friston, K. (2004). Comparing dynamic causal models. *Neuroimage* 22, 1157–1172.
- Pinotsis, D., and Friston, K. (2011). Neural fields, spectral responses and lateral connections. *Neuroimage* 55, 39–48.
- Pinotsis, D., Hansen, E., Friston, K., and Jirsa, V. (2013a). Anatomical connectivity and the resting state activity of large cortical networks. *Neuroimage* 65, 127–138.
- Pinotsis, D., Schwarzkopf, S., Litvak, V., Rees, G., Barnes, G., and Friston, K. (2013b). Dynamic causal modelling of lateral interactions in the visual cortex. *Neuroimage* 66, 563–576.
- Pinotsis, D., Moran, R., and Friston, K. (2012). Dynamic causal modeling with neural fields. *Neuroimage* 59, 1261–1274.
- Rennie, C. J., Wright, J. J., and Robinson, P. A. (2000). Mechanisms of cortical electrical activity and emergence of gamma rhythm. *J. Theor. Biol.* 205, 17–35.
- Robinson, P. (2006). Patchy propagators, brain dynamics, and the generation of spatially structured gamma oscillations. *Phys. Rev. E* 73:041904. doi: 10.1103/PhysRevE.73.041904
- Robinson, P., Henderson, J., Matar, E., Riley, P., and Gray, R. (2009). Dynamical reconnection and stability constraints on cortical network architecture. *Phys. Rev. Lett.* 103:108104. doi: 10.1103/PhysRevLett.103.108104
- Robinson, P., Rennie, C., Wright, J., Bahramali, H., Gordon, E., and Rowe, D. (2001). Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E* 63:021903. doi: 10.1103/PhysRevE.63.021903
- Schmidt, A., Diaconescu, A. O., Komater, M., Friston, K. J., Stephan, K. E., and Vollenweider, F. X. (2012). Modeling ketamine effects on synaptic plasticity during the mismatch negativity. *Cereb. Cortex*. doi: 10.1093/cercor/bhs238. [Epub ahead of print].
- Schwarzkopf, D. S., Robertson, D. J., Song, C., Barnes, G. R., and Rees, G. (2012). The frequency of visually induced gamma-band oscillations depends on the size of early human visual cortex. *J. Neurosci.* 32, 1507–1512.
- Schwarzkopf, D. S., Song, C., and Rees, G. (2010). The surface area of human V1 predicts the subjective experience of object size. *Nat. Neurosci.* 14, 28–30.
- Stam, C. (2005). Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clin. Neurophysiol.* 116, 2266–2301.
- Steyn-Ross, M. L., Steyn-Ross, D. A., Sleigh, J. W., and Liley, D. (1999). Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: evidence for a general anesthetic-induced phase transition. *Phys. Rev. E* 60, 7299.
- Valdes, P., Jimenez, J., Riera, J., Biscay, R., and Ozaki, T. (1999). Nonlinear EEG analysis based on a neural mass model. *Biol. Cybern.* 81, 415–424.
- Wendling, F., Bellanger, J., Bartolomei, E., and Chauvel, P. (2000). Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals. *Biol. Cybern.* 83, 367–378.
- Whittington, M. A., Traub, R. D., and Jefferys, J. G. R. (1995). Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature* 373, 612–615.
- Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24.
- Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biol. Cybern.* 13, 55–80.
- Wright, J., and Liley, D. (1996). Dynamics of the brain at global and microscopic scales: neural networks and the EEG. *Behav. Brain Sci.* 19, 285–294.
- Zeki, S. (1990). Parallelism and functional specialization in human visual cortex. *Cold. Spring Harb. Symp. Q. Biol.* 55, 651–661.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 January 2013; paper pending published: 21 February 2013; accepted: 21 April 2013; published online: 28 May 2013.

Citation: Moran R, Pinotsis DA and Friston K (2013) Neural masses and fields in dynamic causal modeling. *Front. Comput. Neurosci.* 7:57. doi: 10.3389/fncom.2013.00057

Copyright © 2013 Moran Pinotsis and Friston. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Implementing the cellular mechanisms of synaptic transmission in a neural mass model of the thalamo-cortical circuitry

Basabdatta S. Bhattacharya *

Engineering Hub, School of Engineering, University of Lincoln, Lincoln, UK

Edited by:

Peter Robinson, University of Sydney, Australia

Reviewed by:

Peter Robinson, University of Sydney, Australia
Pulin Gong, University of Sydney, Australia

***Correspondence:**

Basabdatta S. Bhattacharya, School of Engineering, University of Lincoln, Engineering Hub, Brayford Pool, Lincoln LN6 7TS, Lincolnshire, UK
e-mail: basab@ieee.org

A novel direction to existing neural mass modeling technique is proposed where the commonly used “alpha function” for representing synaptic transmission is replaced by a kinetic framework of neurotransmitter and receptor dynamics. The aim is to underpin neuro-transmission dynamics associated with abnormal brain rhythms commonly observed in neurological and psychiatric disorders. An existing thalamocortical neural mass model is modified by using the kinetic framework for modeling synaptic transmission mediated by glutamatergic and GABA (gamma-aminobutyric-acid)-ergic receptors. The model output is compared qualitatively with existing literature on *in vitro* experimental studies of ferret thalamic slices, as well as on single-neuron-level model based studies of neuro-receptor and transmitter dynamics in the thalamocortical tissue. The results are consistent with these studies: the activation of ligand-gated GABA receptors is essential for generation of spindle waves in the model, while blocking this pathway leads to low-frequency synchronized oscillations such as observed in slow-wave sleep; the frequency of spindle oscillations increase with increased levels of post-synaptic membrane conductance for AMPA (alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic-acid) receptors, and blocking this pathway effects a quiescent model output. In terms of computational efficiency, the simulation time is improved by a factor of 10 compared to a similar neural mass model based on alpha functions. This implies a dramatic improvement in computational resources for large-scale network simulation using this model. Thus, the model provides a platform for correlating high-level brain oscillatory activity with low-level synaptic attributes, and makes a significant contribution toward advancements in current neural mass modeling paradigm as a potential computational tool to better the understanding of brain oscillations in sickness and in health.

Keywords: neural mass model, thalamocortical circuitry, kinetic framework, brain oscillations, AMPA, GABA

1. INTRODUCTION

Neural mass computational models mimicking synchronous behavior in populations of thalamocortical neurons are often used to study brain oscillations (David and Friston, 2003; Suffczynski et al., 2004; Breakspear et al., 2006; Sotero et al., 2007; Deco et al., 2008; Izhikevich and Edelman, 2008; Pons et al., 2010; Robinson et al., 2011; de Haan et al., 2012). The term “neural mass” was coined by Freeman (1975), while the neural mass modeling paradigm is based on the mathematical framework proposed by Wilson and Cowan (1973); each cell population in a neural mass model represents a neuronal “ensemble” of mesoscopic-scale (10^4 – 10^7), which are densely packed in space and work at the same temporal-scale, so that for all practical purposes, they can be mathematically treated as a single entity (Liljenström, 2012), whence “mass”. In a seminal work, da Silva et al. (1974) used a neural mass model of a simple thalamocortical circuitry to simulate EEG (Electroencephalography) alpha rhythms (8–13 Hz).

Subsequently, this model has been the basis of several research (Zetterberg et al., 1978; Stam et al., 1999; Suffczynski, 2000; Bhattacharya et al., 2011a), albeit with modifications and enhancements; of special mention is the modification introduced by Jansen and Rit (1995) where the model is expressed as a set of ordinary differential equations (ODE). This modification, in turn, has been the basis of many significant research (Wendling et al., 2002; Grimbert and Faugeras, 2006; Ursino et al., 2010). However, the computational basis of the models remain the same—the conversion from firing rate to membrane potential by excitatory and inhibitory neurotransmitters is simulated by convolution of the input from a pre-synaptic neuronal mass with an exponential function, commonly known as the “alpha function”, proposed by Rall (1967). Although the alpha function is a fair estimate of the synaptic process (Bernard et al., 1994), it does not allow an insight into the underlying cellular mechanisms of synaptic transmission associated with abnormal brain oscillations—an aspect emphasized to be crucial as an

aid to research in brain disorders (McCormick, 1992; Basar and Guntekin, 2008). The importance of understanding the neuro-transmission mechanisms in slow wave synchronized as well as spindle oscillations is also discussed in several relevant experimental studies (Steriade et al., 1993; von Krosigk et al., 1993). Moreover, correlating synaptic kinetics with brain oscillatory activity has the potential to aid neuropharmacological advances in treating the diseased brain (Aradi and Erdi, 2006). Along these lines, Destexhe et al. (1998) argue that the alpha function is inappropriate for representing post synaptic events other than the originally proposed post-synaptic potential in spiking neural networks; they propose a kinetic framework as a more biologically plausible method of modeling synaptic transmission compared to the alpha function (Destexhe et al., 2002). The ability of such a modeling framework to capture the physiological properties of synaptic transmission was demonstrated by fitting the model outputs to experimental data from hippocampal slices. Moreover, kinetic modeling is reported to be computationally efficient (Destexhe et al., 1994), a vital prerequisite in large-scale computational models. Subsequently, the kinetic models of neurotransmission was used in several single-neuronal-level model-based studies—to investigate thalamic oscillations (Destexhe et al., 1996) and corticothalamic influence on brain oscillatory activity (Destexhe, 2008); to investigate network synchrony (Breakspear et al., 2003); to simulate synchronous behavior observed during *in vitro* experimental studies on ferret thalamic slice by Wang and Rinzel (1992), Golomb et al. (1994, 1996) and Wang et al. (1995).

A significant modification to current neural mass modeling framework was proposed by Suffczynski et al. (2004) by applying single-neuronal-level model based techniques. Toward this, they proposed an “ensemble” representation of the membrane conductance and post-synaptic current in a neuronal mass model of the thalamocortical circuitry; an integrator is used to generate the “ensemble” post-synaptic membrane potential. In the work presented here, a similar approach is adopted to implement the kinetic framework of synaptic transmission in neural mass models—each post-synaptic attribute is assumed to be an “ensemble” representation corresponding to a “neuronal mass”. For brevity, only two-state (“open” and “closed”) ion-channels (Destexhe et al., 1998) are considered, the desensitized state is ignored. While two-state models are a significant simplification of the very complex nature of ion channel dynamics in biology, they have shown a remarkable fit to biological data compared to more-than-two-state models (Destexhe et al., 1998, 2002). This work aims to interface an abstraction of the ion channel dynamics, such as the two-state ion channel kinetic models, with an abstraction of the population level neuronal behavior, such as neural mass models. The goal is to enable the correlation of higher-level brain dynamics observed in EEG with cellular-level dynamics.

The work is presented thus: first, the kinetic framework for modeling AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic-acid) and GABA (γ -amino-butyric-acid) receptor mediated synapses is introduced in an existing thalamocortical neural mass model (section 2); second, a qualitative comparison of the model behavior with experimental studies on

ferret thalamocortical tissue reported in von Krosigk et al. (1993) as well as to single-neuronal-level model based observations reported in Golomb et al. (1996; section 3) is presented; the lack of a quantitative study is mainly to avoid erratic conclusions as difference in model structure and simulation techniques are bound to induce mismatch in numerical results. The model behavior is observed to be consistent with these studies (von Krosigk et al., 1993; Golomb et al., 1996)—The post synaptic membrane conductance in both the thalamocortical relay (TCR) and thalamic reticular nucleus (TRN) cell population plays a role in effecting a bifurcation in model behavior from spindling mode [oscillations with the characteristic waxing-and-waning pattern seen in early stages of sleep (Steriade et al., 1993; Hughes et al., 2004) as well as in alpha rhythmic oscillations during resting brain state (da Silva et al., 1973)] to a limit-cycle mode (synchronized oscillations as seen in later stages of sleep or during absence seizures). The post-synaptic membrane conductance for both AMPA and GABA in the TRN cell population is responsible for sustaining and modulating spindle oscillations in the model output. Blocking the GABA-ergic synapses in the self-inhibitory loop of the TRN cell population effects a low-frequency synchronized oscillation in the model; this is aided by the secondary-messenger-gated GABA synapses in the TCR cell population. In addition, the reverse rate of transmitter binding plays a role in increasing or decreasing the frequency of synchronized oscillations, besides functioning as a bifurcation parameter, an observation that has not been reported in experimental studies. A comparison of the simulation time of the model with previous research using neural mass models based on alpha functions show a factor of 10 improvement in simulation time. This is a dramatic improvement on computational efficiency and emphasizes the appropriateness of the model proposed herein toward building large-scale software models for investigating neuronal disorders. The observations from this study as well as issues related to the modeling approach are discussed in section 4.

2. MATERIALS AND METHODS

2.1. FROM ALPHA FUNCTION TO KINETIC MODEL: A BRIEF OUTLINE

A single neuronal mass structure as used commonly in neural mass models is shown in **Figure 1** and is defined in Equations (1–5):

$$h_{\bar{w}}(t) = \frac{H_{\bar{w}}}{\tau_{\bar{w}}} \exp(-t/\tau_{\bar{w}}) \quad (1)$$

$$y_N(t) = \sum h_{\bar{w}}(t) \otimes E_{\bar{w}}^N(t) \quad (2)$$

$$\ddot{y}_N(t) = \frac{H_{\bar{w}}}{\tau_{\bar{w}}} E_{\bar{w}}^N(t) - \frac{2}{\tau_{\bar{w}}} \dot{y}_N(t) - \frac{1}{\tau_{\bar{w}}^2} y_N(t) \quad (3)$$

$$V_P(t) = \sum_{N \in \{1, 2, 3, \dots, n\}} C_N \cdot y_N(t) \quad (4)$$

$$E_{\bar{w}}^P(t) = S(V_P) = \frac{2e_0}{1 + e^{v(s_0 - V_P)}} \quad (5)$$

where $\bar{w} \in \{e, i\}$ represents pre-synaptic neuronal populations which make excitatory (e) and inhibitory (i) synapses on a post-synaptic neuronal population; $\tau_{\bar{w}}$ is the time constant and

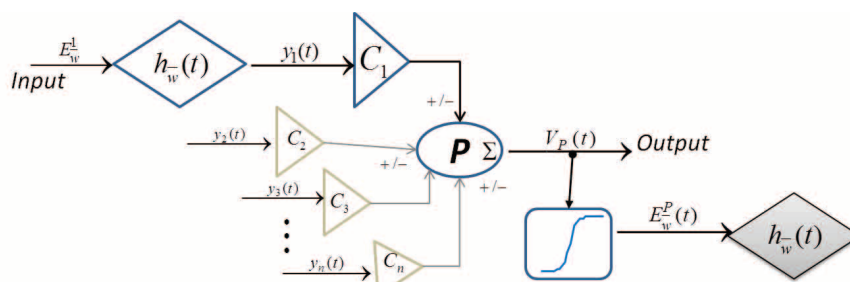


FIGURE 1 | Block diagram of a single “neuronal mass” in current state-of-the-art neural mass models.

H_w is the amplitude of the synapse; $E_w^N(t)$, $N \in \{1, 2, \dots, n\}$ is the firing frequency of an extrinsic or intrinsic cell population that is pre-synaptic to the population P ; C_N is a percentage of the total number of synapses from all afferents to P ; V_P is the “ensemble post-synaptic membrane potential”; E_w^P is the “ensemble firing rate” of P and is defined by a sigmoid function where $2e_0$ is the maximum firing rate of the population, s_0 is the threshold potential at which the neurons spike and v is the sigmoid steepness parameter.

2.1.1. A modified neural mass representation

In a recent work, Suffczynski et al. (2004) modified the neural mass representation of a cell population and introduced post-synaptic current mediated by the ligand-gated glutamatergic receptors AMPA, and the ligand- and secondary-messenger-gated GABA-ergic receptors $GABA_A$ and $GABA_B$, respectively. The input $E_{\xi}^N(t)$, $\xi \in \{\text{AMPA}, GABA_A, GABA_B\}$, is the firing rate of an excitatory (AMPA) or inhibitory ($GABA_A$ and $GABA_B$) pre-synaptic neuronal population $N \in \{1, 2, \dots, n\}$. The model (Figure 2A) is defined in Equations (6–11):

$$h_{\xi}(t) = H_{\xi} \left(\exp(-t/\tau_{\xi}^a) - \exp(-t/\tau_{\xi}^b) \right), \tau_{\xi}^b > \tau_{\xi}^a \quad (6)$$

$$g_{\xi}^N(t) = \sum h_{\xi}(t) \otimes E_{\xi}^N(t) \quad (7)$$

$$\ddot{g}_{\xi}^N(t) = \frac{1}{\tau_{\xi}^a \tau_{\xi}^b} \left[H_{\xi} \left(\tau_{\xi}^a - \tau_{\xi}^b \right) E_{\xi}^N(t) - \left(\tau_{\xi}^a + \tau_{\xi}^b \right) \dot{g}_{\xi}^N(t) - g_{\xi}^N(t) \right] \quad (8)$$

$$I_{\xi}^N(t) = g_{\xi}^N(t) (V^P(t) - V_{\xi}) \quad (9)$$

$$\kappa_m \dot{V}^P(t) = - \sum_{N \in \{1, 2, 3, \dots, n\}} C_N \cdot I_{\xi}^N(t) - I_{\lambda}(t) \quad (10)$$

$$I_{\lambda}(t) = g_{\lambda} (V^P(t) - V_{\lambda}) \quad (11)$$

where $h_{\xi}(t)$ is the synaptic transmission function with τ_{ξ}^a and τ_{ξ}^b as the rise and decay times, respectively; g_{ξ}^N denote the post-synaptic “ensemble” membrane conductance; V_{ξ} is the reversal potential for the synapse mediated by ξ ; V^P is the ensemble

post synaptic membrane potential of the population P due to PSC from all pre-synaptic cell populations $N \in \{1, 2, \dots, n\}$; κ_m is the ensemble membrane capacitance; C_N is the synaptic connectivity parameter; I_{λ} , g_{λ} and V_{λ} are the ensemble leakage current, conductance and reversal potential, respectively for P . The ensemble firing rate $E_{\xi}^P(t)$ is as defined in Equation (5) and is the pre-synaptic firing rate input to other neuronal populations.

2.1.2. Introducing kinetic model of synapses in a neural mass representation

The single neuronal mass structure presented in Figures 1, 2A is modified by replacing the alpha function with kinetic models of AMPA, $GABA_A$, and $GABA_B$ synapses; the enhanced representation (Figure 2B) is defined in Equations (12–19):

$$[T]_{\chi}(V_{\chi}) = \frac{T_{\max}}{1 + e^{-\frac{V_{\chi} - \theta_{\xi}}{\sigma_{\xi}}}} \quad (12)$$

$$\frac{dr_{\chi}^{\xi_1}(t)}{dt} = \alpha^{\xi_1} [T]_{\chi} \left(1 - r_{\chi}^{\xi_1}(t) \right) - \beta^{\xi_1} r_{\chi}^{\xi_1}(t) \quad (13)$$

$$\frac{dR_{\chi}^{\xi_2}(t)}{dt} = \alpha^{\xi_2} [T]_{\chi} \left(1 - R_{\chi}^{\xi_2}(t) \right) - \beta^{\xi_2} R_{\chi}^{\xi_2}(t) \quad (14)$$

$$\frac{d[X](t)}{dt} = \alpha^{\xi_2} R_{\chi}^{\xi_2}(t) - \beta^{\xi_2} [X](t) \quad (15)$$

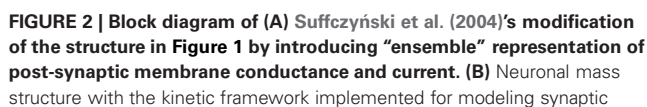
$$r_{\chi}^{\xi_2}(t) = \frac{[X]^n(t)}{[X]^n(t) + K_d} \quad (16)$$

$$I_{\chi}^{\xi}(t) = g_{\chi}^{\xi} r_{\chi}^{\xi}(t) (V_P(t) - V_{\xi}) \quad (17)$$

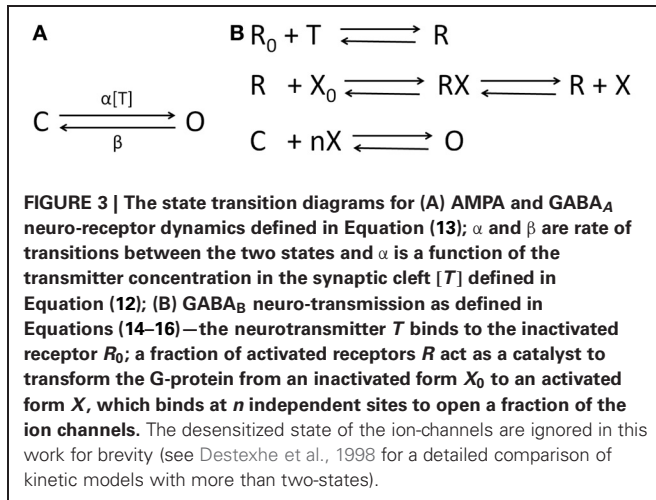
$$\kappa_m \frac{dV_P(t)}{dt} = - \sum_{\chi \in \{1, 2\}} I_{\chi}^{\xi}(t) \cdot C_{\chi} - I_P^{\lambda}(t) \quad (18)$$

$$I_P^{\lambda}(t) = g_P^{\lambda} (V_P(t) - V_P^{\lambda}) \quad (19)$$

Let V_{χ} , $\chi \in \{1, 2\}$ be the “ensemble” membrane potential of two pre-synaptic neuronal population that are afferent to the post-synaptic population P such that the synapses made by $\chi = 1$ is mediated by a ligand-gated receptor $\xi_1 \in \{\text{AMPA}, GABA_A\}$ while that made by $\chi = 2$ is mediated by a secondary-messenger-gated receptor $\xi_2 \in \{GABA_B\}$. The concentration of neurotransmitters



transmission as an alternative to the alpha function ($h_{\text{WT}}(t)$ in **Figure 1**). A diagrammatic representation of the ion-channel kinetics during synaptic transmission is presented in **Figures 3A,B**. **(C)** A thalamocortical circuitry implementing the modified neuronal mass representation in **(B)**.



$[T]_\chi$ in the synaptic cleft is defined as a function of V_χ and is expressed by a sigmoid function (Equation 12) where T_{\max} is the maximum neuronal concentration in the synaptic cleft and is well approximated by 1 mM (Destexhe et al., 1998), θ_s represents the threshold at which $[T] = 0.5T_{\max}$ and σ_s denote the steepness of the sigmoid. The proportion of open ion-channels due to the bound receptors ξ_1 on the ensemble membrane of the post-synaptic cell population corresponding to the synapse made by the population $\chi = 1$ is defined in Equation (13) where α^{ξ_1} and β^{ξ_1} are the forward and backward rate constants, respectively for transmitter binding. The transition diagram is shown in **Figure 3A**. However, GABA_B mediated synapses, unlike AMPA and GABA_A synapses, activate G-proteins which in turn act as the “secondary messengers” and initiate the opening of ion channels.

The process is defined in Equations (14–16) where $R_\chi^{\xi_2}$ is the fraction of activated ξ_2 receptors, which acts as a catalyst in activating the secondary-messenger G-protein (guanine nucleotide binding proteins); $[X]$ is the concentration of the activated G-protein; $r_\chi^{\xi_2}$ is the fraction of open ion channels caused by binding of X with independent binding sites; α^{ξ_2} and β^{ξ_2} are the binding rate constants; n is the number of bound receptor sites and K_d is the dissociation constant of binding of X with the ion channels. The transition diagram of this process is shown in **Figure 3B**. The resulting ensemble PSC mediated by the receptor $\xi \equiv \xi_1 \cup \xi_2$ due to a synapse from the pre-synaptic population χ is defined in Equation (17) where g^ξ and V^ξ are the maximum conductance and reverse potential, respectively corresponding to ξ mediated synapse. V_p (Equation 18) is the ensemble post-synaptic potential (PSP) of P , where κ_m is the ensemble membrane capacitance of P , C_χ , $\chi \in \{1, 2\}$ is the synaptic connectivity parameter. I_P^λ (Equation 19) is the ensemble leak current of the post-synaptic membrane, where g_P^λ and V_P^λ are conductance and reverse potential, respectively, corresponding to “non-specific” leak (Golomb et al., 1996; Suffczynski et al., 2004) in the ensemble membrane of the post synaptic cell population. In the following section, we implement this framework in a neural mass model of the thalamocortical circuitry.

2.2. NEURAL MASS MODEL OF A THALAMOCORTICAL CIRCUITRY WITH KINETIC SYNAPSES

The thalamocortical circuitry is shown in **Figure 2C** and consists of the two thalamic cell populations that communicate with the cortex viz. the TCR and TRN. The third group of cells viz. the Interneurons (IN) participate in intra-thalamic communications and are ignored here for brevity. The synaptic structure and connectivity are informed from experimental data based on the dorsal thalamic Lateral Geniculate Nucleus (LGNd) (Horn et al., 2000). The input to the model is assumed to be the ensemble membrane potential of pre-synaptic retinal cells (V_{ret}) in a resting state with no sensory input and is simulated using a Gaussian white noise (da Silva et al., 1973). The TCR cells make AMPA receptor mediated glutamatergic synapses on the TRN cells (other types of glutamatergic receptors are ignored in this work for brevity); the TRN cells make GABA-ergic synapses on the TCR cells mediated by both the ligand-gated GABA_A and the secondary-messenger-gated GABA_B receptors. Furthermore, the TRN cells make GABA_A receptor mediated synapses within the population. The model is defined in Equations (20–27); all variables and parameters in the model are assumed to be the ensemble representation corresponding to a neural mass:

$$[T]_{\bar{\Psi}}(V_{\bar{\Psi}}(t)) = \frac{T_{\max}}{1 + \exp\left(-\frac{V_{\bar{\Psi}}(t) - \theta_s}{\sigma_s}\right)} \quad (20)$$

$$\frac{dr_{\bar{\Psi}}^{\bar{\eta}_1}(t)}{dt} = \alpha^{\bar{\eta}_1}[T]_{\bar{\Psi}}\left(1 - r_{\bar{\Psi}}^{\bar{\eta}_1}(t)\right) - \beta^{\bar{\eta}_1}r_{\bar{\Psi}}^{\bar{\eta}_1}(t) \quad (21)$$

$$\frac{dR_{\bar{\Psi}}^{\bar{\eta}_2}(t)}{dt} = \alpha_1^{\bar{\eta}_2}[T]_{\bar{\Psi}}\left(1 - R_{\bar{\Psi}}^{\bar{\eta}_2}(t)\right) - \beta_1^{\bar{\eta}_2}R_{\bar{\Psi}}^{\bar{\eta}_2}(t) \quad (22)$$

$$\frac{d[X](t)}{dt} = \alpha_2^{\bar{\eta}_2}R_{\bar{\Psi}}^{\bar{\eta}_2}(t) - \beta_2^{\bar{\eta}_2}[X](t) \quad (23)$$

$$r_{\bar{\Psi}}^{\bar{\eta}_2}(t) = \frac{[X]^n(t)}{[X]^n(t) + K_d}, \quad (24)$$

$$I_{\bar{\Psi}}^{\bar{\eta}}(t) = g^{\bar{\eta}}r_{\bar{\Psi}}^{\bar{\eta}}(t)\left(V_{\bar{\gamma}}(t) - V^{\bar{\eta}}\right) \quad (25)$$

$$\kappa_m \frac{dV_{\bar{\gamma}}(t)}{dt} = - \sum_{\bar{\Psi} \in \{\text{ret}, \text{trn}, \text{trn}\}} I_{\bar{\Psi}}^{\bar{\eta}}(t) \cdot C_{\bar{u}\bar{v}\bar{w}} - I_{\bar{\gamma}}^{\lambda}(t), \quad (26)$$

$$I_{\bar{\gamma}}^{\lambda}(t) = g_{\bar{\gamma}}^{\lambda}\left(V_{\bar{\gamma}}(t) - V_{\bar{\gamma}}^{\lambda}\right), \quad (27)$$

where $\bar{\Psi} \in \{\text{ret}, \text{trn}, \text{trn}\}$ represent the afferent cell populations; $\bar{\gamma} = \{\text{trn}, \text{trn}\}$ represent the efferent cell populations; $\bar{\eta}_1 \in \{\text{AMPA}, \text{GABA}_A\}$, $\bar{\eta}_2 \in \{\text{GABA}_B\}$, $\bar{\eta} \equiv \bar{\eta}_1 \cup \bar{\eta}_2$; $C_{\bar{u}\bar{v}\bar{w}}$ are connectivity parameters where $\bar{u} \in \{t, n\}$ and $\bar{v} \in \{r, t, n, s\}$ denote the post-synaptic and pre-synaptic cell populations, respectively of the retina (r), TCR (t), TRN (n), while s denote an intra-population afferent; $\bar{w} \in \{e, i\}$ represent an excitatory (e) or an inhibitory (i) synapse. All other parameter nomenclatures are as defined in section 2.1. The initial parameter values are mentioned in **Table 1**.

Table 1 | Initial values of the parameters defined in Equations (21–27).

Neuroreceptors → Units↓	AMPA	GABA _A	GABA _B	
(A) NEUROTRANSMISSION PARAMETERS				
mM.msec ⁻¹	α ^ñ ₁ = 2	α ^ñ ₁ = 2	α ^ñ ₁ ² = 0.02 α ^ñ ₂ ² = 0.03	
msec ⁻¹	β ^ñ ₁ = 0.1	β ^ñ ₁ = 0.08	β ^ñ ₁ ² = 0.05 β ^ñ ₂ ² = 0.01	
mS	g ^ñ = 0.1	g ^ñ _{TRN to TCR} = 0.1 g ^ñ _{TRN to TRN} = 0.2	g ^ñ = 0.06	
mV	V ^ñ = 0	V ^ñ _{TRN to TCR} = -85 V ^ñ _{TRN to TRN} = -75	V ^ñ = -100 K _d = 100 n = 4	
(B) CELL MEMBRANE PARAMETERS				
	TCR	TRN		
g ^λ _{TR} (mS)	0.01	0.01		
V ^λ _{TR} (mV)	-55	-72.5		
V _{rest} (mV)	-61	-84		
(C) CONNECTIVITY PARAMETERS				
Efferents →		TRN		
Afferents ↓	TCR	GABA _A	GABA _B	Retinal
TCR	X	C ^a _{tni} ¾ of 30.9	C ^b _{tni} ¼ of 30.9	C _{tre} 7.1
TRN	C _{nte} 35	C _{nsi} 20	X	X

Data in (A) and (B) are as in Golomb et al. (1996) and Suffczynski et al. (2004). In Equation (20), both θ_s and σ_s act as bifurcation parameters in the model (see Bhattacharya et al., 2012). However, the emphasis here is on post-synaptic membrane attributes as in von Krosigk et al. (1993) and Golomb et al. (1996). Thus these parameters ($\theta_s = -35$ and $\sigma_s = 2$) are set by trial and error at values just before the model undergoes bifurcation from a “point-attractor” mode to a “limit-cycle” mode, based on a recent study where we observed rich model dynamics and power spectral behavior around the bifurcation point (Bhattacharya et al., 2013); $T_{\text{max}} = 1$ mM (Destexhe et al., 1994). The input noise mean $\mu = -45$ mV and standard deviation $\varphi = 20$ mV² are set by trial and error and represents the resting state membrane potential fluctuations in retinal cells. While the total number of GABA-ergic synaptic count on TCR cells is reported as 30.9%, specific data on GABA_A and GABA_B are not available in literature to the best of our knowledge. Thus, values for C^a_{tni} , C^b_{tni} and C_{nsi} in (C) are selected, within the reported biological range [see Bhattacharya et al. (2011b) for details], when the model output showed an increased frequency content within the theta (4–7 Hz) and alpha (8–13 Hz) bands. C_{tre} and C_{nte} are as in Bhattacharya et al. (2011b). All variables in the ODEs are initialized to an arbitrarily small value 0.0002.

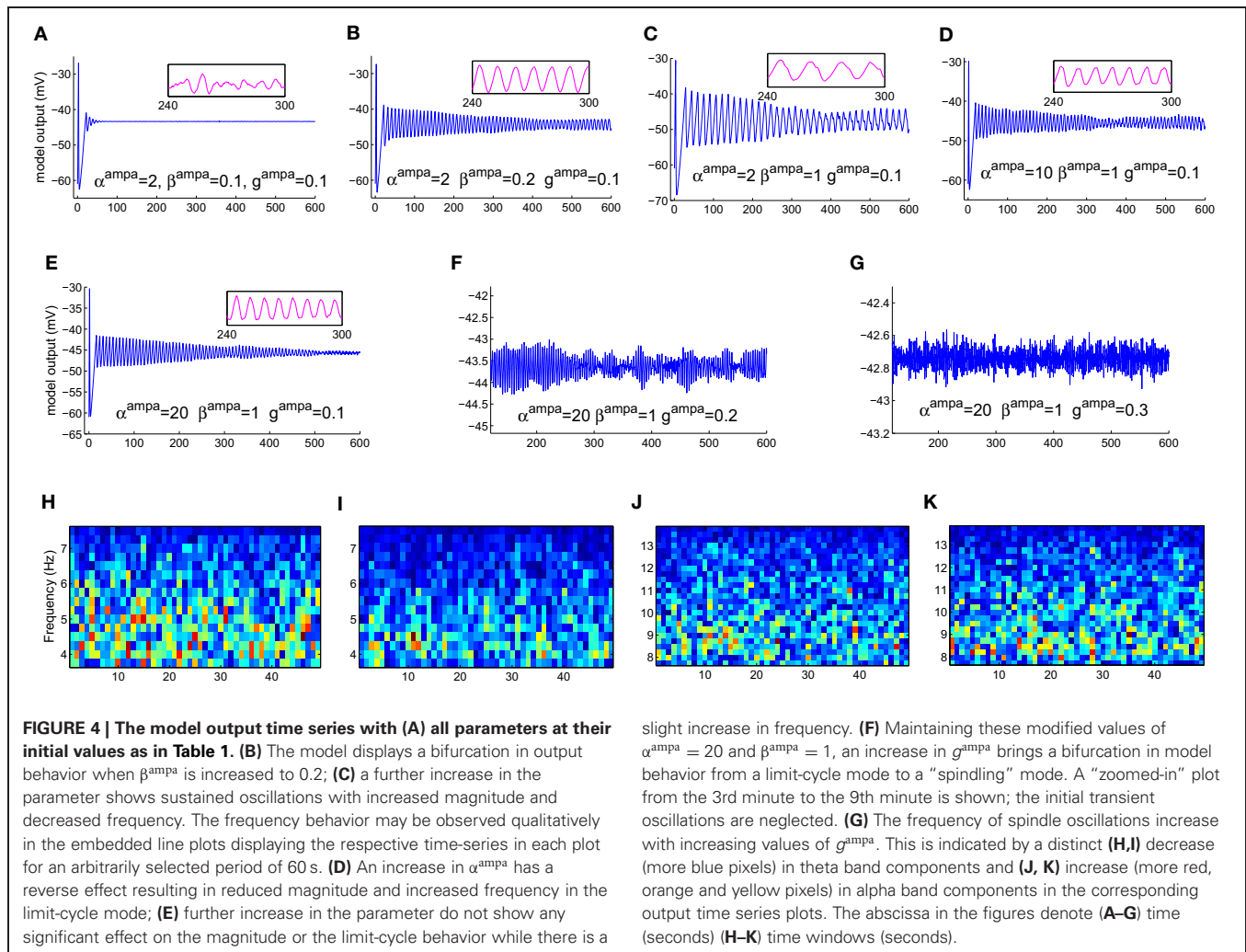
3. RESULTS

The ODEs are solved using the 4th/5th order Runge-Kutta-Fehlberg method (RK45) in Matlab for a total duration of 600 s (10 min) at a resolution of 1 ms. The output voltage time series is

averaged over 20 simulations, each simulation run with different seed for the noisy input. For frequency analysis, an epoch from 100–599 s of the output signal is sampled every 4 ms (250 Hz) and bandpass filtered between 3.5–14 Hz with a Butterworth filter of order 10. Short Time Fourier Transform (STFT) is done with a Hamming window of duration 10 s and overlap of 50%.

The model displays a point-attractor mode behavior (initial transient oscillations before settling down to a low amplitude noisy output, which reflects the noisy input of the model) corresponding to initial parameter values (Figure 4A). There is a behavioral transition in the model to a limit cycle mode with increasing values of β^{ampa} , which correlates with a decrease in the fraction of open ion channels in the post-synaptic ensemble membrane (Figures 4B,C). Varying α^{ampa} , on the other hand, does not affect the model behaviour (Figures 4D,E). A transition from the limit cycle mode to a spindling mode is effected in the model by increasing g^{ampa} , the post-synaptic membrane conductance for AMPA mediated synapses in both TCR and TRN cell population, and shown in Figures 4F,G. STFT of the output time series indicates the non-stationary behavior of the model (Figures 4H–K). A decrease and increase, respectively of the theta and alpha band components imply an overall increase in frequency with increasing values of $g^{\text{ampa}} \equiv \{g^{\text{ampa}}_{\text{TCR}}, g^{\text{ampa}}_{\text{TRN}}\}$, where $g^{\text{ampa}}_{\text{TCR}}$ and $g^{\text{ampa}}_{\text{TRN}}$ correspond to the incoming signal from the retina (to the TCR) and TCR (to the TRN), respectively in the model. These observations are consistent with similar reports of a transition in the state of the model output with increasing values of g^{ampa} in Golomb et al. (1996; pp. 756–757), accompanied by an abrupt increase in the ratio of the frequency of oscillation of the TCR and the TRN cell populations; we have not studied the latter aspect in this work. A more detailed study on the model presented herein where $g^{\text{ampa}}_{\text{TCR}}$ and $g^{\text{ampa}}_{\text{TRN}}$ are varied separately specify the $g^{\text{ampa}}_{\text{TCR}}$ as the control parameter that causes a bifurcation in the model output from a limit cycle mode to the spindling mode with an increase in its value. On the other hand, the $g^{\text{ampa}}_{\text{TRN}}$ does not effect any behavioral change in the model output, rather, it is effective in increasing the inter-spindle frequency with an increase in its value when the model is in a spindling mode. This observation implies that a change in AMPA receptor related attributes in the TRN plays a role in modulating thalamocortical spindle oscillations, which finds strong support in the experimental study by von Krosigk et al. (1993), where “activation of AMPA-kainate receptors on the PGN” (Peri-geniculate nucleus—the part of the TRN associated with the LGNd) is described as “critical to the generation of spindle waves”. Furthermore, this observation is in line with the TRN being widely implicated as being the key “ingredient” in the generation of thalamocortical spindle oscillations (McCormick, 1992; Steriade et al., 1993).

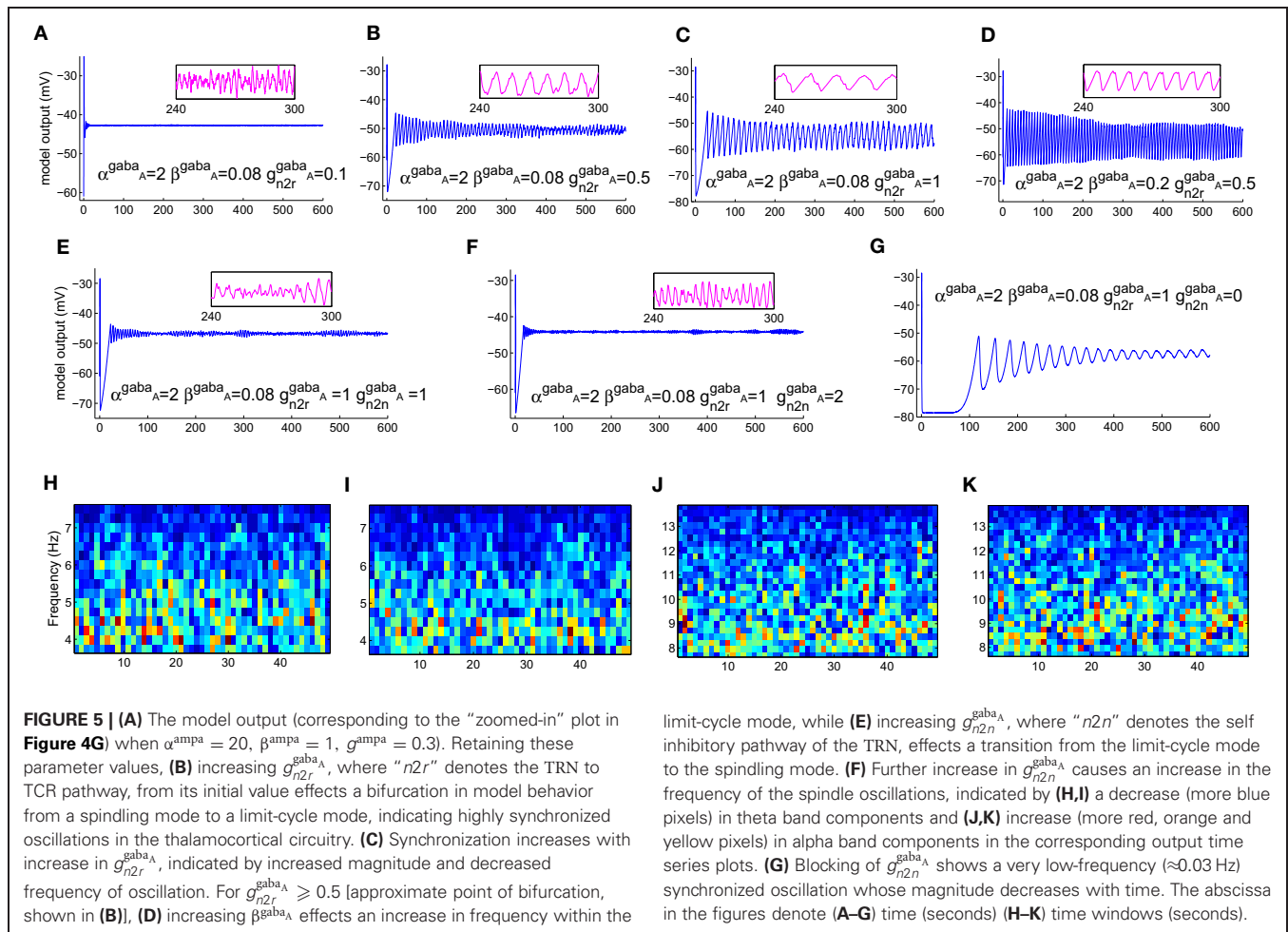
Varying the GABA-ergic synaptic attributes when the model is in a point-attractor mode does not show any change in model behavior. When the model is in a spindling mode (Figure 5A), increased synchronization within the limit cycle mode with increasing values of $g^{\text{gaba}_A}_{\text{TRN to TCR}}$ (Figures 5B,C) is observed. An increase in the parameter β^{gaba_A} affects the output only when the model is in a limit-cycle mode and counters the effect of increase in $g^{\text{gaba}_A}_{\text{TRN to TCR}}$ (Figure 5D). However, varying α^{gaba_A}



does not affect the model output. For $g_{\text{TRN to TCR}}^{\text{gaba}_A} \geq 0.5$, which is the approximate bifurcation point (Figure 5B), increasing $g_{\text{TRN to TRN}}^{\text{gaba}_A}$ causes the model to revert back to the spindling mode; the frequency of the inter-spindle oscillations increase with increasing values of the parameter (Figures 5E,F). This is also indicated by a decrease (Figures 5H,I) and increase (Figures 5J,K) of theta and alpha band components respectively in the STFT of the output time series. In other words, decreasing values of the parameter $g_{\text{TRN to TRN}}^{\text{gaba}_A}$ causes increased synchronization within the spindling mode behavior of the model along with a decrease in the inter-spindle frequency. However, blocking $g_{\text{TRN to TRN}}^{\text{gaba}_A}$ effects a switch in the model behavior to a very low-frequency oscillatory state (Figure 5G). These results are consistent with experimental findings (von Krosigk et al., 1993) where application of GABA_A inhibitor either “abolished spindle waves or decreased within-spindle frequency,” which correspond to the condition of either blocking or decreasing, respectively of $g_{\text{TRN to TRN}}^{\text{gaba}_A}$ in our model. Thus, the model implicate the intra-TRN synaptic activity to be a key factor in sustaining spindle oscillations in the thalamocortical circuitry, an observation which conforms to those made in Golomb et al.

(1996; p. 755). Furthermore, a “frequency jump” with increasing g_{gaba_A} , and associated transition in model behavior is also reported in Golomb et al. (1996; see Figure 7) as a comparative study between the TCR and TRN cells. This is similar with the increase in frequency of the spindle oscillations corresponding to increasing $g_{\text{TRN to TRN}}^{\text{gaba}_A}$ in the present model, although we have not done a comparative study with the TRN cell population behavior. However, the current study implicate the increased post-synaptic conductance for GABA_A receptors in the TCR cell population ($g_{\text{TRN to TCR}}^{\text{gaba}_A}$) to play a significant role in effecting state-transition between spindle and slow-wave oscillations, an observation that is yet to find support from experimental or model-based studies.

A quiescent state is observed corresponding to blocking either AMPA ($g^{\text{ampa}} = 0$) or both GABA_A ($g_{\text{TRN to TCR}}^{\text{gaba}_A} = 0$) and GABA_B ($g_{\text{TRN to TCR}}^{\text{gaba}_B} = 0$) mediated synapses in the TRN to TCR pathway (not shown). This is consistent with both experimental (von Krosigk et al., 1993) and model-based (Golomb et al., 1996) studies. The role of the synaptic parameters in the GABA_B pathway in our model was minimal—to sustain a non-quiescent model behavior with blockage of GABA_A; to sustain a high amplitude of limit-cycle oscillations in the model. Again, this



is in agreement with experimental studies (von Krosigk et al., 1993), where activation of GABA_B receptors are reported as “not essential” for generating synchronized oscillations, while application of GABA_B antagonist abolished “evoked or spontaneous slowed oscillations.” The model in Golomb et al. (1996; see in Discussion p. 763) is also mentioned as being consistent with these experimental results.

In a recent work (Bhattacharya et al., 2012), a simple neural mass model implementing kinetic modeling for synaptic transmission is presented; the synaptic connectivity parameters in the model correlate directly to that of an alpha function based neural mass model [modified Alpha Rhythm model (modARm) from Bhattacharya et al. (2011a)]. The model behavior is studied corresponding to changes in the synaptic connectivity parameters as well as transmitter concentration related parameters, and a relevant comparison is made with the modARm. However, the model presented in this work has a larger set of synaptic connectivity parameters; model behavior corresponding to this parameter space and its usefulness in understanding neurological disorders will be the topic of a future work.

4. DISCUSSION

The work presented here explores a novel approach toward correlating current neural mass model based studies with underlying

cellular mechanisms during synaptic transmission. The aim is to underpin the synaptic correlates of abnormal brain oscillations in neurological and psychiatric disorders such as observed in Electroencephalogram (EEG). A kinetic framework for modeling AMPA and GABA receptor mediated synapses is implemented in an existing thalamocortical neural mass model consisting of an excitatory and an inhibitory neural mass, representing cell populations of the thalamocortical relay (TCR) and the thalamic reticular nucleus (TRN), respectively. Parameters in the model are assumed to be “ensemble” representations of the corresponding attributes in a single neuron. A preliminary observation is made on the model behavior by varying the parameters corresponding to the post-synaptic membrane conductance of the cell populations as well as the forward and reverse rates of synaptic reaction; of specific interest is the transition of the model behavior between the spindle oscillatory mode and the limit-cycle mode, the latter resembling the slow-wave (high-amplitude, low-frequency) synchronized oscillations that are signatures of absence seizures as well as slow-wave sleep. Furthermore, only the alpha (8–13 Hz) and theta (4–7 Hz) frequency bands of the output power spectra are studied here, as EEG alpha and theta bands are believed to have a strong correlation with thalamocortical oscillations (Hughes et al., 2004).

The results indicate that: (1) The post synaptic membrane conductance for both AMPA and GABA_A receptors in the TRN cell population play a role in sustaining spindle oscillations of the TCR cell population (the model output). (2) Blocking the GABA_A mediated synapses in the self-inhibitory feedback pathway of the TRN cell population effects synchronized oscillations with high amplitude and increased time-period of oscillation (≈ 0.03 Hz). (3) The post-synaptic membrane conductance for GABA_B in the TCR cell population does not play any role in generating or sustaining spindle oscillations, but is responsible for sustaining the slow-wave oscillations in the model associated with blocking of the intra-TRN GABA_A synapses. (4) Blocking both GABA_A and GABA_B or only the AMPA mediated synapses in the TCR cell population results in a quiescent model output. These findings are consistent with *in vitro* studies based on multiple unit recordings from ferret thalamic slices (von Krosigk et al., 1993) as well as single-neuron-level model based studies (Golomb et al., 1996). In addition, this study identifies—(a) the reverse rate of transmitter binding as an important attribute in effecting thalamocortical synchronized oscillations that can be induced in the model by increasing (decreasing) the fraction of open channels due to GABA_A (AMPA) mediated synapses in the TCR (TRN) cell populations; (b) the post-synaptic membrane conductance for GABA_A in the TCR cell population as a control parameter for effecting a behavioral transition in the model.

It may be noted that the above-mentioned observations are only a qualitative comparison with single-neuron-level model-based (Golomb et al., 1996) and experimental (von Krosigk et al., 1993) studies; a drawback of the current work is a lack of quantitative comparison with these studies. The neural mass model presented in this work is at a mesoscopic scale, representation of a population of $\approx 10^4$ – 10^7 neurons, unlike that in Golomb et al. (1996), which is at single-neuronal-level. Similarly, the multiple unit recording based study in von Krosigk et al. (1993) observes neuronal behavior of either a single neuron or a population of $< 10^2$ neurons. In addition, the modeling and simulation methods in the current work and that in Golomb et al. (1996) are not similar. Thus, a quantitative comparison of the current work with these studies may lead to erroneous conclusions. However, model validation with experimental data is a crucial criteria when investigating brain disorders. Along these lines, an ongoing work is investigating ways to validate the model presented herein with EEG data, and will be the topic of a future study.

The model structure in the current work is a considerably simplified representation of the thalamocortical circuitry. The role of the thalamocortical circuitry in generating slow wave brain oscillations is discussed at length in Steriade et al. (1993), based on *in vivo* and *in vitro* studies. More recently, three parameters in the thalamo-cortico-thalamic loop viz. the cortico-thalamic, thalamo-cortical and intra-thalamic pathways are specified in Breakspear et al. (2006) for generating instabilities in the thalamocortical circuitry, leading to synchronized oscillations such as seen during absence seizures. Furthermore, a non-linear dynamical analysis of the model is shown to predict seizure onset by validating with patient EEG data. In a previous research (Bhattacharya et al., 2011b), we have proposed a

more elaborate alpha-function based neural mass model that have considered these vital pathways in the thalamocorticothalamic loop. Also, we have performed a non-linear dynamical analysis of a simple thalamocortical model based on alpha functions in Bhattacharya et al. (2013) to understand EEG power spectra abnormalities associated with several neurological disorders. Such research directions will be considered as an extended work based on the model presented herein.

It is worth mentioning here that biologically plausible parameterizations has been a major constraint in neural mass modeling of brain dynamics. This is largely due to insufficient experimental data, published or otherwise, as well as to a lack of “homogeneity” of published data from different experimental laboratories. The trend thus far has been to use biologically plausible data if and when available; otherwise, i.e., for parameter values that cannot be availed from experimental data, the models are tuned to estimated parameter values which provide a desirable output in context to the objectives of the research [the reader may refer to Robinson et al. (2004) for a model parameterizations related work and discussion]. Thus, the model in Breakspear et al. (2006) was based on neurophysiological parameters obtained from Robinson et al. (2002), which in turn are based on inverse parameterizations during model validation with EEG data from patients of epileptic seizures. The parameterizations of the model presented in this work is largely based on neurophysiological parameters obtained from experimental studies: the cellular-level parameters, including those of the synaptic kinetics, are based on *in vitro* studies and model-based studies of thalamocortical tissue by von Krosigk et al. (1993) and Golomb et al. (1996), respectively; the model connectivity parameters are based on experimental studies of the cat and rat thalamus obtained from Horn et al. (2000); Sherman and Guillery (2001). On the other hand, the extrinsic (retinal) input and neuro-transmitter concentration parameter values are adjusted to maintain a “dynamically active” model behavior (this is as opposed to a continuous “quiescent” state of the model corresponding to certain parameter values, and does not conform to biology). However, technological advances in the field of neuro-imaging during the last decade such as functional Magnetic Resonance Imaging (fMRI), Diffusion Tensor Imaging (DTI) and Transcranial Magnetic Stimulation (TMS) are paving the way for biologically-realistic mapping of parameter values in computational models; for example as in Izhikevich and Edelman (2008).

The observations made herein support the motivation toward this preliminary work, which is to correlate higher-level brain dynamics with underlying cellular-level synaptic mechanisms. It may be noted that in all our previous works using alpha function based neural mass models, the emphasis has been on studying the model behavior with varying values of synaptic connectivity parameters toward a meaningful mapping to Alzheimer disease-related EEG anomalies. However, such “synaptic parameter variation only” studies are highly constrained and do not make much sense when trying to understand generic brain-state conditions e.g., the sleep-awake cycle, or several other neurological and psychiatric disorders e.g., absence seizures, which rely heavily on various aspects of cellular dynamics in the thalamocortical circuitry. Rather, the emphasis of this work is on laying the

ground-work for a more elaborate, and yet computationally efficient scheme, whereby large-scale computational models may be simulated to mimic brain rhythms, which can then be correlated to model parameters emulating cellular dynamics. The synaptic transmission kinetics and subsequent post-synaptic membrane parameters are some of the key constituents of brain signaling, and are affected significantly in various brain diseases. Clearly, the alpha-function based neural mass models are inadequate in dealing with research directions where model parameters can be mapped in a biologically plausible manner to synaptic attributes. In terms of computational efficiency, the time for simulating

20 trials with the model presented in this work takes 60 s; this may be contrasted with 600 s for simulating a similar model [the modified Alpha Rhythm model in Bhattacharya et al. (2011a)] based on alpha functions. This is a dramatic improvement in computational efficiency and highlight the plausibility of using the kinetic-model based neural mass modeling framework in simulating large-scale computational models toward mimicking real-time EEG signals. This in turn will provide a powerful tool for specifying cellular pathways that need be targeted for symptomatic alleviation of anomalous brain rhythms as well as to inform effective neuropharmacological research directions.

REFERENCES

- Aradi, I., and Erdi, P. (2006). Computational neuropharmacology: dynamical approaches in drug discovery. *Trends Pharmacol. Sci.* 27, 240–243. doi: 10.1016/j.tips.2006.03.004
- Basar, E., and Guntekin, B. (2008). A review of brain oscillations in cognitive disorders and the role of neurotransmitters. *Brain Res. Rev.* 1235, 172–193. doi: 10.1016/j.brainres.2008.06.103
- Bernard, C., Ge, Y., Stockley, E., Willis, J., and Wheal, H. (1994). Synaptic integration of NMDA and non-NMDA receptors in large neuronal network models solved by means of differential equations. *Biol. Cybern.* 70, 267–273. doi: 10.1007/BF00197607
- Bhattacharya, B. S., Cakir, Y., Serap-Sengor, N., Maguire, L. P., and Coyle, D. (2013). Model-based bifurcation and power spectral analyses of thalamocortical alpha rhythm slowing in Alzheimer's disease. *Neurocomputing* 115, 11–22. doi: 10.1016/j.neucom.2012.10.023
- Bhattacharya, B. S., Coyle, D., and Maguire, L. P. (2011a). "Alpha and theta rhythm abnormality in Alzheimer's disease: a study using a computational model," in *Advances in Experimental Medicine and Biology*. Vol. 718, eds C. Hernandez, J. Gomez, R. Sanz, I. Alexander, L. Smith, A. Hussain, et al. (New York, NY: Springer), 57–73.
- Bhattacharya, B. S., Coyle, D., and Maguire, L. P. (2011b). A thalamo-cortico-thalamic neural mass model to study alpha rhythms in Alzheimer's disease. *Neural Netw.* 24, 631–645. doi: 10.1016/j.neunet.2011.02.009
- Bhattacharya, B. S., Coyle, D., Maguire, L. P., and Stewart, J. (2012). "Kinetic modelling of synaptic functions in the alpha rhythm neural mass model," in *ICANN 2012 Part I, Lecture Notes in Computer Science* 7552, eds E. P. V. Alessandro, D. Wlodzislaw, E. Peter, M. Francesco, and P. Gunther (Berlin; Heidelberg: Springer Verlag), 645–652.
- Breakspear, M., Roberts, J., Terry, J., Rodrigues, S., Mahant, N., and Robinson, P. A. (2006). A unifying explanation of primary generalized seizures through nonlinear brain modelling and bifurcation analysis. *Cereb. Cortex* 16, 1296–1313. doi: 10.1093/cercor/bhj072
- Breakspear, M., Terry, J. R., and Friston, K. J. (2003). Modulation of excitatory synaptic coupling facilitates synchronization and complex dynamics in a nonlinear model of neuronal dynamics. *Neurocomputing* 52, 151–158. doi: 10.1016/S0925-2312(02)00740-3
- da Silva, F. H. L., Hoeks, A., Smits, H., and Zetterberg, L. H. (1974). Model of brain rhythmic activity. *Kybernetik* 15, 27–37. doi: 10.1007/BF00270757
- da Silva, F. H. L., van Lierop, T. H. M. T., Schrijer, C. F., and van Leeuwen, W. S. (1973). Essential differences between alpha rhythms and barbiturate spindles: spectra and thalamo-cortical coherences. *Electroencephalogr. Clin. Neurophysiol.* 35, 641–645. doi: 10.1016/0013-4694(73)90217-4
- David, O., and Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage* 20, 1743–1755. doi: 10.1016/j.neuroimage.2003.07.015
- de Haan, W., Mott, K., van Straaten, E. C. W., Scheltens, P., and Stam, C. J. (2012). Activity dependent degeneration explains hub vulnerability in Alzheimer's disease. *PLoS Comput. Biol.* 8:e1002582. doi: 10.1371/journal.pcbi.1002582
- Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., and Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* 4:e1000092. doi: 10.1371/journal.pcbi.1000092
- Destexhe, A. (2008). "Cortico-thalamic feedback: a key to explain absence seizures," in *Computational Neuroscience in Epilepsy*, eds I. Soltes and K. Staley (Amsterdam: Elsevier), 184–214. doi: 10.1016/B978-012373649-9.50016-8
- Destexhe, A., Bal, T., McCormick, D., and Sejnowski, T. J. (1996). Ionic mechanisms underlying synchronised oscillations and propagating waves in a model of ferret thalamic slices. *J. Neurophysiol.* 76, 2049–2070.
- Destexhe, A., Mainen, Z., and Sejnowski, T. (1994). An efficient method for computing synaptic conductances based on a kinetic model of receptor binding. *Neural Comput.* 6, 14–18. doi: 10.1162/neco.1994.6.1.14
- Destexhe, A., Mainen, Z., and Sejnowski, T. (1998). "Kinetic models of synaptic transmission," in *Methods in Neuronal Modelling*, eds C. Koch and I. Segev (Cambridge, MA: MIT Press), 1–25.
- Destexhe, A., Mainen, Z., and Sejnowski, T. (2002). "Kinetic models for synaptic interactions," in *The Handbook of Brain Theory and Neural Networks*, ed M. Arbib (Cambridge, MA: MIT Press), 1126–1130.
- Freeman, W. J. (1975). *Mass Action in the Nervous System*. 1st Edn. New York, NY: Academic Press.
- Golomb, D., Wang, X.-J., and Rinzel, J. (1994). Synchronization properties of spindle oscillations in a thalamic reticular nucleus model. *J. Neurophysiol.* 72, 1109–1126.
- Golomb, D., Wang, X.-J., and Rinzel, J. (1996). Propagation of spindle waves in a thalamic slice model. *J. Neurophysiol.* 75, 750–769.
- Grimbert, F., and Faugeras, O. (2006). Bifurcation analysis of Jansen's neural mass model. *Neural Comput.* 18, 3052–3068. doi: 10.1162/neco.2006.18.12.3052
- Horn, S. C. V., Erisir, A., and Sherman, S. M. (2000). Relative distribution of synapses in the A-laminae of the lateral geniculate nucleus of the cat. *J. Compar. Neurol.* 416, 509–520. doi: 10.1002/(SICI)1096-9861(200012)416:4<509::AID-CNE7>3.0.CO;2-H
- Hughes, S. W., Lorincz, M., Cope, D. W., Blethyn, K. L., Kekesi, K. A., Parri, H. R., et al. (2004). Synchronised oscillations at α and θ frequencies in the lateral geniculate nucleus. *Neuron* 42, 253–268. doi: 10.1016/S0896-6273(04)00191-6
- Izhikevich, E. M., and Edelman, G. M. (2008). Large-scale model of mammalian thalamocortical systems. *Proc. Natl. Acad. Sci. U.S.A.* 105, 3593–3598. doi: 10.1073/pnas.0712231105
- Jansen, B. H., and Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.* 73, 357–366. doi: 10.1007/BF00199471
- Liljenström, H. (2012). Mesoscopic brain dynamics. *Scholarpedia* 7:4601. doi: 10.4249/scholarpedia.4601
- McCormick, D. A. (1992). Neurotransmitter actions in the thalamus and cerebral cortex and their role in neuromodulation of thalamocortical activity. *Progr. Neurobiol.* 39, 337–388. doi: 10.1016/0304-0082(92)90012-4
- Pons, A. J., Cantero, J. L., Atienza, M., and Garcia-Ojalvo, J. (2010). Relating structural and functional anomalous connectivity in the ageing brain via neural mass modelling. *Neuroimage* 52, 848–861. doi: 10.1016/j.neuroimage.2009.12.105
- Rall, W. (1967). Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic inputs. *J. Neurophysiol.* 30, 1138–1168.
- Robinson, P., Rennie, C., and Rowe, D. (2002). Dynamics of large-scale brain activity in normal arousal states and epileptic seizures.

- Phys. Rev. E* 65:041924. doi: 10.1103/PhysRevE.65.041924
- Robinson, P., Rennie, C., Rowe, D., and Connor, S. O. (2004). Estimation of multiscale neurophysiologic parameters by electroencephalographic means. *Hum. Brain Mapp.* 23, 53–72. doi: 10.1002/hbm.20032
- Robinson, P. A., Phillips, A. J. K., Fulcher, B. D., Puckeridge, M., and Roberts, J. A. (2011). Quantitative modelling of sleep dynamics. *Philos. Trans. R. Soc. A* 369, 3840–3854. doi: 10.1098/rsta.2011.0120
- Sherman, S. M., and Guillery, R. W. (2001). *Exploring the Thalamus*. 1st Edn. New York, NY: Academic Press.
- Sotero, R. C., Tujillo-Barreto, N. J., and Iturria-Medina, Y. (2007). Realistically coupled neural mass models can generate EEG rhythms. *Neural Comput.* 19, 479–512. doi: 10.1162/neco.2007.19.2.478
- Stam, C. J., Pijn, J., Suffczynski, P., and da Silva, F. (1999). Dynamics of the human alpha rhythm: evidence for non-linearity? *Clin. Neurophysiol.* 110, 1801–1813. doi: 10.1016/S1388-2457(99)00099-1
- Steriade, M., McCormick, D. A., and Sejnowski, T. J. (1993). Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262, 679–685. doi: 10.1126/science.8235588
- Suffczynski, P. (2000). *Neural Dynamics Underlying Brain Thalamic Oscillations Investigated with Computational Models*. Ph.D. Thesis, Institute of experimental physics, University of Warsaw.
- Suffczynski, P., Kalitzin, S., and Silva, F. H. L. D. (2004). Dynamics of non-convulsive epileptic phenomena modelled by a bistable neuronal network. *Neuroscience* 126, 467–484. doi: 10.1016/j.neuroscience.2004.03.014
- Ursino, M., Cona, F., and Zavaglia, M. (2010). The generation of rhythms within a cortical region: analysis of a neural mass model. *Neuroimage* 52, 1080–1094. doi: 10.1016/j.neuroimage.2009.12.084
- von Krosigk, M., Bal, T., and McCormick, D. A. (1993). Cellular mechanisms of a synchronised oscillation in the thalamus. *Science* 261, 361–364. doi: 10.1126/science.8392750
- Wang, X.-J., Golomb, D., and Rinzel, J. (1995). Emergent spindle oscillations and intermittent burst firing in a thalamic model: specific neuronal mechanisms. *Proc. Natl. Acad. Sci.* 92, 5577–5581. doi: 10.1073/pnas.92.12.5577
- Wang, X.-J., and Rinzel, J. (1992). Alternating and synchronous rhythms in reciprocally inhibitory model neurons. *Neural Comput.* 4, 84–97. doi: 10.1162/neco.1992.4.1.84
- Wendling, F., Bartolomei, F., Bellanger, J. J., and Chauvel, P. (2002). Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *Eur. J. Neurosci.* 15, 1499–1508. doi: 10.1046/j.1460-9568.2002.01985.x
- Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80. doi: 10.1007/BF00288786
- Zetterberg, L. H., Kristiansson, L., and Mossberg, K. (1978). Performance of a model for a local neuron population. *Biol. Cybern.* 31, 15–26. doi: 10.1007/BF00337367

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 November 2012; paper pending published: 03 February 2013; accepted: 06 June 2013; published online: 04 July 2013.

Citation: Bhattacharya BS (2013) Implementing the cellular mechanisms of synaptic transmission in a neural mass model of the thalamo-cortical circuitry. *Front. Comput. Neurosci.* 7:81. doi: 10.3389/fncom.2013.00081

Copyright © 2013 Bhattacharya. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Modulation of epileptic activity by deep brain stimulation: a model-based study of frequency-dependent effects

Faten Mina^{1,2}, Pascal Benquet^{1,2}, Anca Pasnicu³, Arnaud Biraben^{1,2,3} and Fabrice Wendling^{1,2*}

¹ INSERM, U1099, Université de Rennes 1, Rennes, France

² Laboratoire Traitement du Signal et de L'Image, Université de Rennes 1, Rennes, France

³ Unité d'Épileptologie, Service de Neurologie, CHU, Rennes, France

Edited by:

Peter Robinson, University of Sydney, Australia

Reviewed by:

Peter Robinson, University of Sydney, Australia

Anthony Burkitt, University of Melbourne, Australia

*Correspondence:

Fabrice Wendling, Laboratoire Traitement du Signal et de L'Image, INSERM, Université de Rennes 1, Campus de Beaulieu, Bat. 22 - 35042 Rennes Cedex, France
e-mail: fabrice.wendling@univ-rennes1.fr

A number of studies showed that deep brain stimulation (DBS) can modulate the activity in the epileptic brain and that a decrease of seizures can be achieved in “responding” patients. In most of these studies, the choice of stimulation parameters is critical to obtain desired clinical effects. In particular, the stimulation frequency is a key parameter that is difficult to tune. A reason is that our knowledge about the frequency-dependant mechanisms according to which DBS indirectly impacts the dynamics of pathological neuronal systems located in the neocortex is still limited. We address this issue using both computational modeling and intracerebral EEG (iEEG) data. We developed a macroscopic (neural mass) model of the thalamocortical network. In line with already-existing models, it includes interconnected neocortical pyramidal cells and interneurons, thalamocortical cells and reticular neurons. The novelty was to introduce, in the thalamic compartment, the biophysical effects of direct stimulation. Regarding clinical data, we used a quite unique data set recorded in a patient (drug-resistant epilepsy) with a focal cortical dysplasia (FCD). In this patient, DBS strongly reduced the sustained epileptic activity of the FCD for low-frequency (LFS, < 2 Hz) and high-frequency stimulation (HFS, > 70 Hz) while intermediate-frequency stimulation (IFS, around 50 Hz) had no effect. Signal processing, clustering, and optimization techniques allowed us to identify the necessary conditions for reproducing, in the model, the observed frequency-dependent stimulation effects. Key elements which explain the suppression of epileptic activity in the FCD include: (a) feed-forward inhibition and synaptic short-term depression of thalamocortical connections at LFS, and (b) inhibition of the thalamic output at HFS. Conversely, modeling results indicate that IFS favors thalamic oscillations and entrains epileptic dynamics.

Keywords: DBS, thalamocortical model, computational, centromedian nucleus, FCD, premotor cortex, epilepsy

INTRODUCTION

Deep brain stimulation (DBS) for Parkinson's disease (PD) and other movement and psychiatric disorders—including dystonia, tremor, and depression—is clinically used today as a conventional therapeutic procedure for the alleviation of symptoms (Sillay and Starr, 2009). Since the early 90s, neurologists also attempted to apply DBS to other neurological disorders, typically to intractable epilepsies in order to suppress—or at least dramatically reduce—the occurrence of seizures [see recent review in Boon et al. (2009)]. These studies followed early scientific evidence showing potentially beneficial effects of DBS on epileptic neural dynamics in animal models (Reimer et al., 1967; Hablitz, 1976) as well as in patients (Cooper et al., 1973; Davis et al., 1982; Wright and

Weller, 1983). However, contrary to PD, the optimal “antiepileptic parameters” of DBS for reducing the frequency of seizures are much more variable among patients and the number of non-responders to stimulation still perplexes scientists. Moreover, in responding patients, the fine tuning of stimulation parameters in a patient-specific manner remains indispensable for maximizing antiepileptic effects. On that account, many fundamental questions are frequently raised: where and when to stimulate, at which frequency, at which current intensity, and with which current waveform?

The answers to these questions remain bound to our current, and still limited, understanding of the mechanisms by which DBS modulates neuronal dynamics, whether normal or pathological. Today, the precise mechanisms of neuronal modulation by DBS remain elusive. In addition, these mechanisms are controversial as observed effects are sometimes opposite (McIntyre et al., 2004b). Among the many studies reported over the last decade, identified mechanisms regarding HFS include: local depolarization blockade by HFS (Beurrier et al., 2001), synaptic depression due to neurotransmitter depletion (Shen et al., 2003; Kim et al., 2012), synaptic inhibition (Filali et al., 2004), disruption of

Abbreviations: CMN, Centromedian Nucleus; DBS, Deep Brain Stimulation; EPSP, Excitatory Postsynaptic Potentials; FCD, Focal Cortical Dysplasia; FFI, Feed-Forward Inhibition; GPi, Globus Pallidus; HFS, High Frequency Stimulation; iEEG, Intracerebral EEG (depth electrodes); IFS, Intermediate Frequency Stimulation; IPSP, Inhibitory Postsynaptic Potentials; LFP, Local Field Potential; LFPs_{FCD}, Local Field Potentials recorded in the FCD; LFS, Low Frequency Stimulation; NS, No Stimulation; PMC, Premotor cortex; RtN, Reticular thalamic Nucleus, STD, Short-Term Depression; STN, Subthalamic Nucleus.

the thalamocortical network's dysrhythmia (McIntyre and Hahn, 2010; Kendall et al., 2011). As far as LFS is concerned, some studies described a transient synaptic depression that alters synaptic transmission (Jiang et al., 2003; Speechley et al., 2007). Finally, IFS is routinely used in the context of presurgical evaluation of patients with drug resistant epilepsy to map epileptogenic and functional brain areas. It has long been observed that this type of stimulation is prone to trigger epileptic afterdischarges (Goddard, 1967). This brief overview shows that the spectrum of involved mechanisms is very large and that distinct stimulation frequencies trigger distinct cellular/network processes. More precise insights into these processes will come with increased knowledge about both biophysical and neurophysiological effects of stimulation currents on underlying neuronal systems.

However, the access to cellular and network mechanisms induced by DBS is rather difficult in animal models of epilepsy and (almost) impossible in patients especially in large-scale systems like the thalamocortical loop. An alternative approach is the use of computational models based on physiological data to first reproduce and then explain changes in cerebral activity as a function of stimulation conditions (stimulation site, intensity, and frequency). This is precisely the objective of this study, with a special focus on the distinct effects of DBS frequency on cortical epileptic dynamics.

Our investigation combines computational modeling and clinical data. We explored stimulation effects in a lumped-parameter mesoscopic neural mass model of the thalamocortical loop, inspired from previously published models (Suffczynski et al., 2004; Lopes Da Silva, 2006; Roberts and Robinson, 2008; Crunelli et al., 2011).

Although these models are lumped representations of underlying neuronal systems, they offer a number of advantages in the context of this study. First, neural mass models include subpopulations of principal excitatory cells and inhibitory interneurons. Second, these models were shown to produce realistic activity as observed in LFPs or EEG under normal (Freeman, 1973; Lopes Da Silva et al., 1974) or epileptic conditions [review in Lytton (2008); Wendling (2008)]. Third, main parameters (mean membrane potential and firing rate) provide access to the investigation of several stimulation-induced (patho)physiological mechanisms. For instance, a neural mass model was successfully used in the context of direct low-intensity pulse stimulation in the hippocampus to explain the behavior of evoked responses during the transition to seizures (Suffczynski et al., 2008).

In particular, using this model, we analyzed the neurophysiological effects induced by direct thalamic stimulation on epileptic cortical dynamics at low frequency (LF, < 20 Hz), intermediate frequency (IF, 20–70 Hz) and high frequency (HF, 70–130 Hz). Model parameters were tuned to reproduce a typical pathological oscillatory activity observed in a neocortical lesion (focal cortical dysplasia, or FCD) in a patient with drug-resistant epilepsy. Intracerebral EEG (iEEG) signals observed during thalamic stimulation (centromedian nucleus) of this patient revealed particularly pronounced frequency-dependent modulation of the FCD pathological activity. Therefore, this data set offered the unique opportunity to identify key model parameters for which such a frequency-dependent modulation could be reproduced and,

subsequently to get insights regarding the mechanisms underlying the modulatory effects, in the FCD, of thalamic stimulation. Results revealed that LFS favors feed-forward inhibition and short-term depression at the cortical level and that HFS inhibits the thalamic activity, while IFS reinforces reticulothalamic oscillations thus entraining cortical pathological epileptic dynamics.

MATERIALS AND METHODS

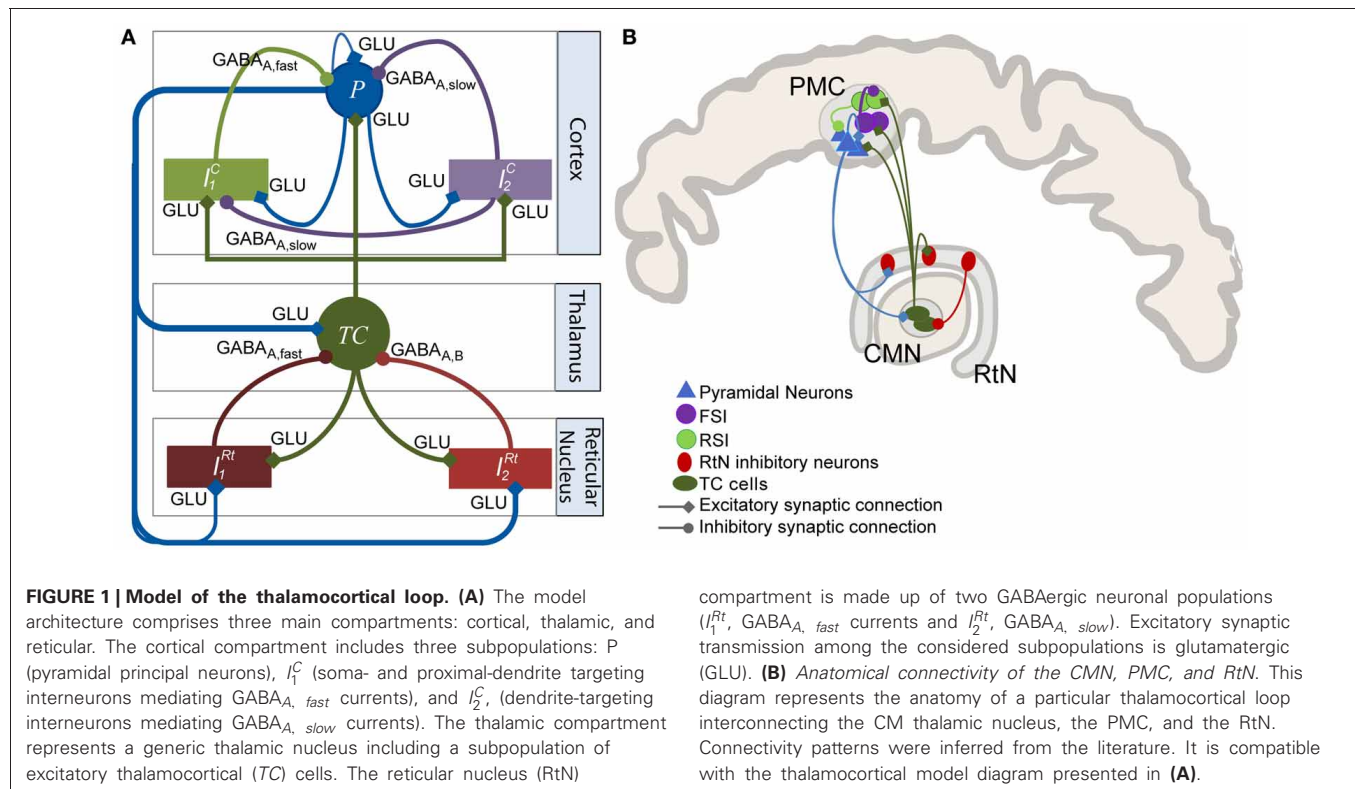
In this section, we present (1) the neurophysiologically-relevant computational model that we developed to study thalamic DBS, (2) the real depth-EEG dataset used for model tuning and, (3) the signal processing methods used for characterizing real and simulated EEG signals.

MODEL OF THE THALAMOCORTICAL LOOP

In order to study the effects of thalamic DBS on cortical dynamics, we implemented a physiologically-plausible mesoscopic model of the thalamocortical loop. This model accounts for the average activity of both cortical and thalamic compartments which include various types of neuronal populations interacting via synaptic transmission. This modeling approach was first proposed in the early 70s (Wilson and Cowan, 1972) and further enriched in order to interpret electrophysiological recordings and study brain dynamics, in the olfactory (Freeman, 1973) and the thalamocortical (Lopes Da Silva et al., 1974) system, for instance, as well as the dynamics of cortical oscillations (Nunez, 1974). This approach was then developed by other research groups in the context of state changes in brain dynamics (Wright et al., 1985), visual evoked potentials (Jansen et al., 1993), dynamics of the human alpha rhythm (Stam et al., 1999) or pathophysiological mechanisms of ictal transitions in epilepsy (Wendling et al., 2000, 2002; Suffczynski et al., 2001; Robinson et al., 2002; Liley and Bojak, 2005; Breakspear et al., 2006). Later, neural mass models were also used in studies dealing with the connectivity among cortical regions and the impact of model parameters on the power spectrum of EEG or MEG signals (Robinson et al., 1997; David and Friston, 2003; Zavaglia et al., 2006).

Model architecture

The model architecture was inspired from previously published models of the thalamocortical loop (Suffczynski et al., 2004; Lopes Da Silva, 2006; Roberts and Robinson, 2008; Crunelli et al., 2011). In a global view, the model was built of three interconnected compartments: a cortical compartment, a thalamic compartment, and a reticular compartment, in accordance with previously published models (Figure 1A) and with anatomical data (Figure 1B). Each compartment includes one or several subpopulation(s) of neurons, either excitatory or inhibitory. Generally speaking, the input/output functions of a considered subpopulation are represented by two mathematical equations that were respectively named “pulse-to-wave” (input) and “wave-to-pulse” (output) by Walter Freeman (Freeman, 1992). The former is a linear transfer function that converts the presynaptic average density of afferent action potentials into an average postsynaptic membrane potential (PSP), either excitatory (EPSP) or inhibitory (IPSP). The output function is a static nonlinear function (sigmoid) that



provides the average pulse density of action potentials fired by neurons depending on the sum of EPSPs and IPSPs at the input. This non-linear function accounts for threshold and saturation effects that take place at the somas and initial axonal segments of considered cells.

Formally, the input function is represented by a second order low-pass filter $H(s) = W/(s+1/\tau_w)^2$ (where s is the Laplace variable). The impulse response of this filter is given by

$$h(t) = \frac{W}{\tau_w} \cdot t \cdot e^{-t/\tau_w} \quad (1)$$

Parameters W and τ_w are tuned such that $h(t)$ approximates the shape of real excitatory (glutamatergic) or inhibitory (GABAergic) postsynaptic potentials (Lopes Da Silva et al., 1976). The quantity $W \cdot \tau_w^2$ is the static gain of filter h . Lumped parameter τ_w (expressed in s) is linked to the kinetics of synaptic currents. It determines both the rise time ($t_{rise} = \tau_w$) and the decay time ($t_{decay} = 3.146\tau_w$) of the second order filter impulse response h and it is usually adjusted with respect to the physiological rise and decay times of actual PSPs (Molae-Ardekani et al., 2010). Given the time constant τ_w , parameter W can be used to adjust the sensitivity of synapses (the maximal PSP amplitude is $W \cdot e^{-1}$). An alternative implementation of the h function was introduced in Bojak and Liley (2005) and is described in detail in Molae-Ardekani et al. (2013). It is based on a bi-exponential pulse-to-wave function with two time constant parameters. This implementation allows for the separate adjustment of the rise and decay times of PSPs, and therefore a better approximation of actual PSPs in some circumstances. Besides, the output function

is represented by $S(v) = \frac{2e_0}{1+e^{r(v_0-v)}}$, where $2e_0$ is the maximum firing rate, v_0 is the postsynaptic potential corresponding to a firing rate of e_0 and r is the steepness of the sigmoid.

The cortical compartment

The cortical compartment was inspired from an existing model of the neocortex which proved its capability of generating both normal and epileptiform activity. Readers may refer to Molae-Ardekani et al. (2010) for details. In brief, the cortical compartment integrates a subpopulation of pyramidal cells (P , $W = A_C$, $\tau_w = \tau_{ac}$ in Equation 1) and two inhibitory neuronal populations (I_1^C and I_2^C , **Figure 1A**) representing soma- and proximal-dendrite targeting interneurons ($GABA_{A, fast}$ currents, $W = G_C$, $\tau_w = \tau_{gc}$ in Equation 1) and dendrite-targeting interneurons ($GABA_{A, slow}$ currents, $W = B_C$, $\tau_w = \tau_{bc}$ in Equation 1), respectively. Pyramidal collateral excitation was implemented as in Jansen et al. (1993).

In addition, these three cortical subpopulations receive excitatory input from the thalamic compartment. Therefore, feed-forward inhibition (FFI) is represented in the model as the two subpopulations of interneurons project to the pyramidal subpopulation (see *The Thalamic and Reticular Compartments* paragraph below).

Short-term synaptic depression (STD)

STD is present in the neocortex (Boudreau and Ferster, 2005). It can be potentially involved in the context of direct stimulation of the thalamus as TC cells directly project to cortical pyramidal cells. Consequently, this mechanism was implemented at the interface of thalamic/cortical compartments. To our knowledge,

an implementation of STD in neural mass models has not been proposed before.

In our model, we represented a modulatory effect of the amplitude of the average EPSP (parameter A_C') at the level of subpopulation P depending on the density of action potentials [$d_{AP}(t)$] coming from the thalamic compartment. This modulatory effect was obtained by multiplying A_C' by a time-varying coefficient $\kappa(t) \in [0.6, 1]$ where the function describing the evolution of $\kappa(t)$ was derived from Chung et al. (2002). This study shows that: (i) cortical EPSPs drop by 40% under periodic low-frequency intense thalamocortical (TC) cell firing and, (ii) this drop in cortical EPSP is directly linked to transient depression of thalamocortical monosynaptic projections to pyramidal neurons.

In line with these observations, STD was implemented as follows. First, signal $d_{AP}^{(t)}$ is low-pass filtered (cutoff frequency = 10 Hz) to restrict the STD effect to LFS. Then, from each time t_η at which the filtered signal $d_{AP}^{(t)}$ exceeds a firing rate equal to η , the $\kappa(t)$ coefficient undertakes an exponential decay given by $\kappa(t) = \kappa_\eta \cdot e^{-t/\tau}$ where $\kappa_\eta = \kappa(t_\eta^-)$ and where t_η^- is the time instant that just precedes t_η . The decrease of $\kappa(t)$ is limited to the time interval [$t_\eta + 0.45$ s] and cannot exceed 40%, total. Parameters η and τ were set to 0.8 and 8 s, respectively.

The thalamic and reticular compartments

The thalamic compartment was limited to one population of excitatory neurons (known as glutamatergic thalamocortical - TC - cells) receiving glutamatergic EPSPs ($W = A_{Th}$, $\tau_w = \tau_{aTh}$ in Equation 1) from cortical pyramidal cells (P) and GABAergic IPSPs with slow ($W = B_{Th}$, $\tau_w = \tau_{bTh}$ in Equation 1) and fast ($W = G_{Th}$, $\tau_w = \tau_{gTh}$ in Equation 1) kinetics from the reticular compartment (RtN). Here, we increased the time constant (τ_{bTh}) with respect to τ_{bc} to account for both GABA_A, *slow*- and GABA_B-receptor mediated currents in a single variable. TC cells directly target both cortical pyramidal cells and interneurons. The activation of these GABAergic interneurons subsequently promotes inhibition of pyramidal cells after a di-synaptic delay. Therefore, TC cells activation induces first an EPSP followed later on by an IPSP on cortical pyramidal cells, resulting in feed-forward inhibition (FFI). The RtN compartment comprised two inhibitory subpopulations, namely I_1^{RT} and I_2^{RT} which both receive excitatory input from the cortical ($W = A_{Rt}$, $\tau_w = \tau_{aRt}$ in Equation 1) and the thalamic ($W = A_{Rt}$, $\tau_w = \tau_{aRt}$ in Equation 1) compartments.

Simulation of stimulation effects

Stimulation currents induce a perturbation of the membrane potential of neurons. At cellular level, this effect can be accounted for by the “ λE model”, which is well grounded in the biophysics of compartment models (Rattay, 1998; McIntyre et al., 2004a; Manola et al., 2005, 2007) (see Miranda et al., 2009 for a review) and supported by *in vitro* experiments (Bikson et al., 2004; Frohlich and McCormick, 2010). This model $\Delta V \approx \tilde{\lambda} \cdot \tilde{E}$ approximates the membrane potential variation ΔV as a linear function of the electrical field \tilde{E} induced by stimulation ($\tilde{\lambda}$ representing the membrane space constant). In our neural mass model, the situation is less straightforward as space is not explicitly represented, conversely to detailed or mean-field models. However, within a certain range of intensity values, it has been shown that

the membrane potential variation ΔV is modified in a linear way with respect to the electrical field which is itself proportional to the stimulation intensity (Bikson et al., 2004). These considerations led us to also assume a linear variation for the mean membrane potential as a function of stimulation intensity, in stimulated sub-populations of neurons. In addition, stimulation was represented by a train of periodic monophasic depolarizing pulses. The pulse width was fixed to 1 ms (as in clinics). Pulses were low-pass filtered to account for the average time of repolarization (set to 4.8 ms) in stimulated sub-populations of cells. The resulting stimulation signal was added to the mean membrane potential of neuronal sub-populations included in the thalamic (TC) and reticular (I_1^{RT} and I_2^{RT}) compartments of the proposed model. The depolarizing effect was weighted by three coefficients S_{TC} , S_{Rt1} and S_{Rt2} (Table 1) accounting for the possibly different stimulation impact at the thalamic and reticular level.

Model parameters, outputs, and implementation

Parameter values as well as physiological interpretation are provided in Table 1. Note that each synaptic connection in the model is weighted by a connectivity constant denoted by $C_{SP1-SP2}$ where $SP1$ and $SP2$, respectively, denote the source and target subpopulations. In addition, two Gaussian noise inputs $p_P(t) \sim N(\mu_P, \sigma_P)$ and $p_{TC}(t) \sim N(\mu_{TC}, \sigma_{TC})$ were used to represent nonspecific inputs on pyramidal and thalamocortical cell subpopulations. Finally, signals simulated at the level of pyramidal cells in the cortical compartment and at the level of TC cells in the thalamic compartment were chosen as model outputs. They correspond to the sum of PSPs at each compartment respectively. The temporal dynamics of these signals provide a good approximation of actual LFPs. The model was implemented in Simulink®, and all other complementary scripts were implemented in MATLAB®.

REAL DATA FOR MODEL TUNING

We used real clinical data to tune the model into a functioning mode which simulates pathological activity. The clinical data set was limited to a unique patient who underwent thalamic DBS during the presurgical intracerebral EEG exploration (iEEG performed with depth electrodes implanted under stereotaxic conditions) at the Epilepsy Surgery Unit, Rennes University Hospital. This particular patient was chosen for two main reasons: (1) the pronounced frequency-dependent stimulation effects observed during his preoperative diagnostic iEEG exploration at LF, IF and HF in addition to (2) the existence of an epileptogenic zone in a limited area of the premotor cortex (PMC).

In brief, this patient suffered from partial drug-resistant epilepsy since the age of two. MRI scans and EEG recordings pointed out the existence of a neuronal malformation known as FCD in the PMC at the origin of seizures. This type of cortical malformation is known for its epileptogenic features like neuronal hyperexcitability and hypersynchronization and its characteristic epileptiform discharges (continuous, rhythmic or semirhythmic spikes, and polyspikes) (Avoli et al., 2003; Palmieri, 2010) as shown in Figure 2C. Based on various clinical studies reporting the modulation of epileptic cortical activity by the

Table 1 | Model parameters, values and interpretation.

Parameter	Value	Interpretation
A_C	6 (optimized, pathological) 3 (normal) mV	Amplitude of the cortical average EPSP
A_C'	$\kappa(t).A_C$ mV	Amplitude of the cortical average EPSP in response to thalamic input (only on subpopulation P)
B_C	14 (optimized, pathological) 50 (normal) mV	Amplitude of the cortical average IPSP (GABA _{A,slow} mediated currents)
G_C	16.5 (optimized, pathological) 22 (normal) mV	Amplitude of the cortical average IPSP (GABA _{A,fast} mediated currents)
A_{Th}	3.5 mV	Amplitude of the thalamic average EPSP
B_{Th}	30 mV	Amplitude of the thalamic average IPSP (GABA _{A,slow} and GABA _B receptors)
G_{Th}	22 mV	Amplitude of the thalamic average IPSP (GABA _{A,fast} receptors)
A_{Rt}	3.5 mV	Amplitude of the reticular average EPSP
τ_{ac}	1/80 s	Time constant of cortical glutamate-mediated synaptic transmission.
τ_{bc}	1/35 s	Time constant of cortical GABA-mediated synaptic transmission (GABA _{A,slow} receptors)
τ_{gc}	1/180 s	Time constant of cortical GABA-mediated synaptic transmission (GABA _{A,fast} receptors)
τ_{aTh}	1/100 s	Time constant of thalamic glutamate-mediated synaptic transmission
τ_{bTh}	1/20 s	Time constant of thalamic GABA-mediated synaptic transmission (GABA _{A,slow} and GABA _B receptors)
τ_{gTh}	1/150 s	Time constant of thalamic GABA-mediated synaptic transmission (GABA _{A,fast} receptors)
τ_{aRt}	1/100 s	Time constant of reticular glutamate-mediated synaptic transmission
v_0, e_0, r	$v_0 = 6\text{mV}$, $e_0 = 2.5\text{ s}^{-1}$ $r = 0.56\text{mV}^{-1}$	Parameters of the nonlinear sigmoid function (transforming the average membrane potential to an average density of action potentials)
$C_{P-P'}$	135	Collateral excitation connectivity constant
$C_{P'-P}$	108	Collateral excitation connectivity constant
$C_{P-I_2^C}$	33.75	P to I_2^C connectivity constant
$C_{I_2^C-P}$	33.75	I_2^C to P connectivity constant
$C_{P-I_1^C}$	40.5	P to I_1^C connectivity constant
$C_{I_2^C-I_1^C}$	13.5	I_1^C to I_2^C connectivity constant
$C_{I_1^C-P}$	91.125	I_1^C to P connectivity constant
C_{TC-P}	120	TC to P connectivity constant
$C_{TC-I_1^C}$	30	TC to I_1^C connectivity constant
$C_{TC-I_2^C}$	45	TC to I_2^C connectivity constant
$C_{TC-I_1^{Rt}}$	20	TC to I_1^{Rt} connectivity constant
$C_{TC-I_2^{Rt}}$	20	TC to I_2^{Rt} connectivity constant
$C_{P-I_1^{Rt}}$	30	P to I_1^{Rt} connectivity constant
$C_{P-I_2^{Rt}}$	30	P to I_2^{Rt} connectivity constant
C_{P-TC}	20	P to TC connectivity constant
$C_{I_1^{Rt}-TC}$	35	I_1^{Rt} to TC connectivity constant
$C_{I_2^{Rt}-TC}$	5	I_2^{Rt} to TC connectivity constant
μ_{P1}	0	Mean of nonspecific cortical input
μ_{P2}	70	Mean of nonspecific subcortical input
σ_{P1}	20.√6	Standard deviation of nonspecific cortical input
σ_{P2}	35.√6	Standard deviation of nonspecific subcortical input
S_{TC}	5	Stimulation impact on subpopulation TC
S_{Rt1}	4	Stimulation impact on subpopulation I_1^{Rt}
S_{Rt2}	4	Stimulation impact on subpopulation I_2^{Rt}
f_s	1Hz – 150Hz	Frequency of the stimulation signal (pulse train)
A_{fs}	1	Stimulation signal amplitude

Model parameters used to reproduce LFPs_{FCD}. Stimulation impact parameters S_{TC} , S_{Rt1} and S_{Rt2} are set to zero during the simulation of the NS scenario. These parameters are held constant for all other stimulation scenarios.

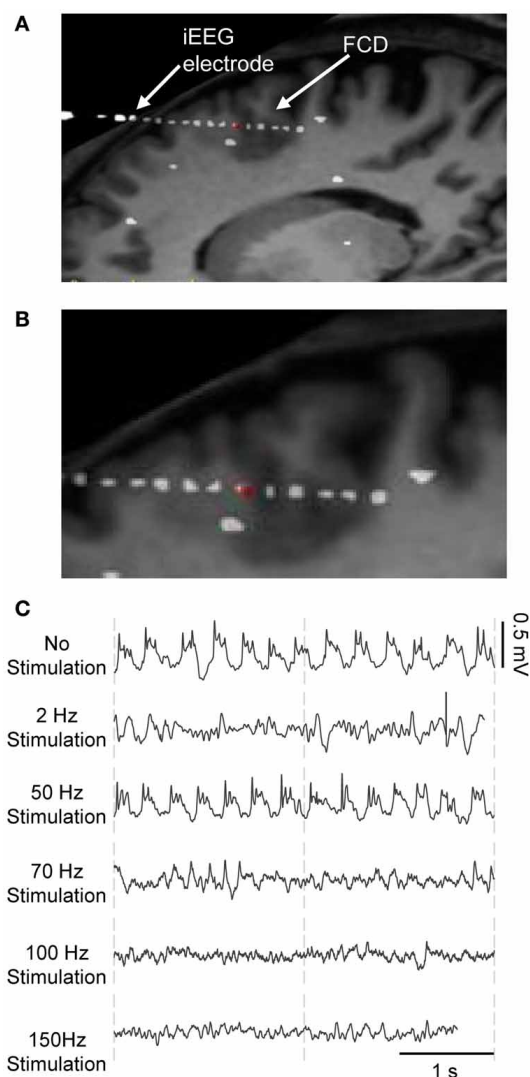


FIGURE 2 | Frequency-dependent stimulation effects: real data. iEEG signals recorded during presurgical depth-EEG exploration in a patient with drug-resistant epilepsy. **(A)** MRI data showing the FCD (focal cortical dysplasia) in the PMC and the electrode trajectory. The red dot marks the position of the depth electrode in the FCD. **(B)** Zoom on the FCD. **(C)** DBS of the CMN modulated the pathological activity of the FCD in a frequency-dependent manner. LFS (2 Hz) and HFS (≥ 70 Hz) suppressed pathological oscillations. IFS (50 Hz) had no effects.

stimulation of the CM nucleus (Velasco et al., 1995, 1997, 2000, 2001, 2007), it was decided by neurologists and neurosurgeons to implant a depth electrode in this nucleus, as potentially beneficial for the patient who gave his informed consent.

During the presurgical exploration, the stimulation of the thalamic CM nucleus (CMN) induced frequency-dependent modulation of the pathologic activity of the FCD (**Figure 2**). Readers may refer to (Pasnicu et al., 2013) for detailed information. Interestingly, LFS (2 Hz, 4 mA) and HFS (70, 100, and 150 Hz, 0.8 mA) desynchronized the pathological activity of the FCD, while IFS (50 Hz, 0.8 mA) barely affected it. These segments of

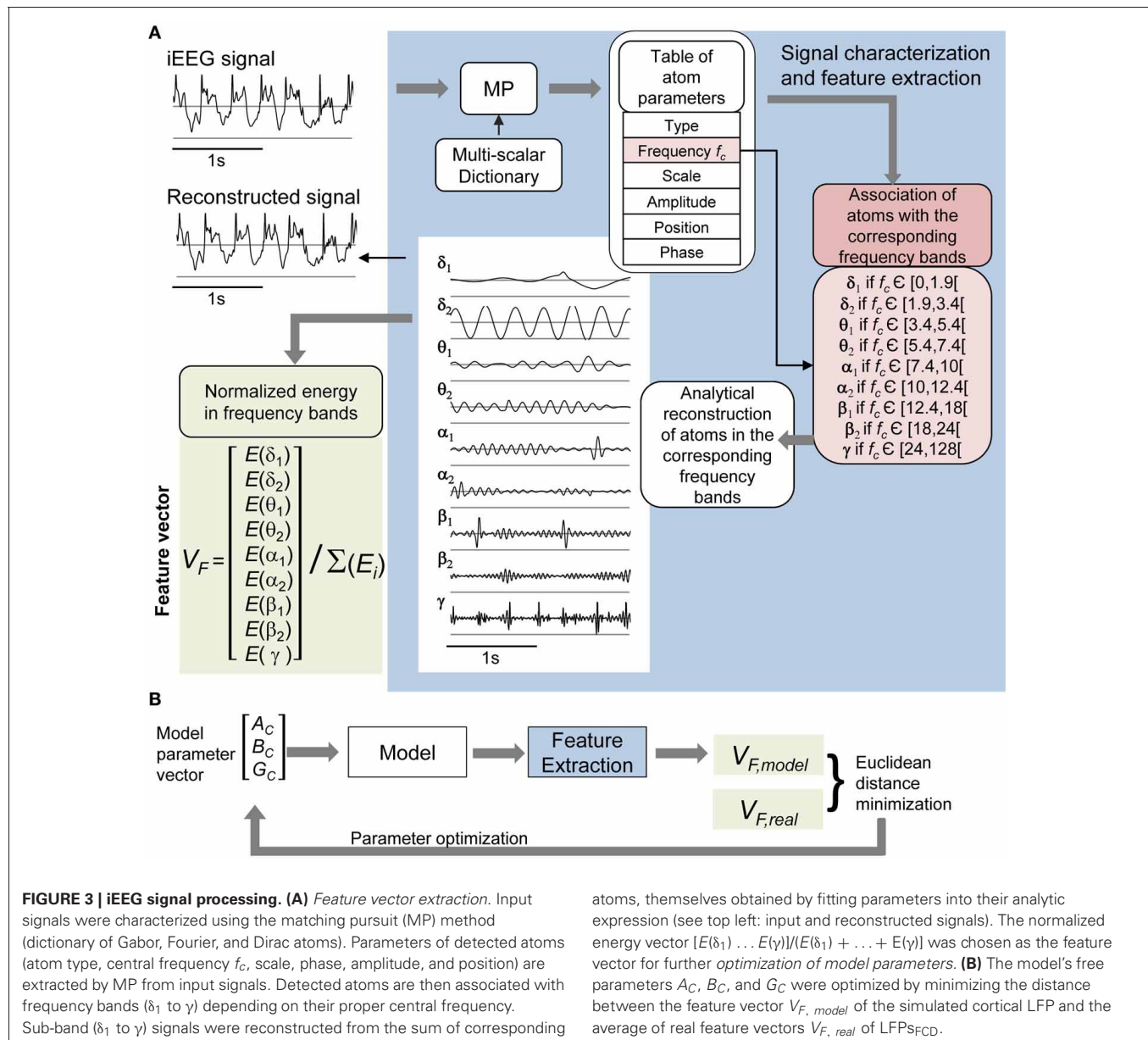
signals corresponding to either typical pathological activity or modulated activity (depending on stimulation conditions) were used to optimize the model parameters.

PROCESSING OF REAL AND SIMULATED SIGNALS

The use of signal processing techniques was necessary (i) to quantify the above-described effects of stimulation in real iEEG signals, and (ii) to define a feature-vector-based cost function for model parameter optimization. **Figure 3A** illustrates the feature extraction methodology. iEEG signals recorded in the FCD in absence of stimulation (LFP_{FCD}) and under different stimulation conditions were decomposed using an orthogonal matching pursuit algorithm [matching pursuit toolkit—MPTK—(Krstulovic and Gribonval, 2006)]. First introduced in 1993 (Mallat and Zhifeng, 1993), matching pursuit is signal processing algorithm used to decompose any time series into a linear sum of waveforms selected from a predefined dictionary based on a mother wavelet. To proceed, a proper multi-scalar dictionary of Gabor, Fourier, and Dirac atoms was first defined to account for real iEEG signal components (time-frequency atoms are waveforms well localized in both the time and the frequency domains). In line with (Krstulovic and Gribonval, 2006), the multi-scalar dictionary was formed by translation in time and amplitude/frequency modulation of atoms (defined as Gabor and Fourier functions in our case), over ten different user-defined time scales (i.e. the atom durations, ranging from 0.125 to 5 s). Then, the algorithm provided a table of time-frequency parameters associated to the detected atoms (atom type, central frequency, phase, scale, amplitude, position). Identified atoms were reconstructed using the extracted parameter table and their analytical expression. They were then associated to a given frequency band depending on their central frequency. These frequency bands corresponded to the classical EEG bands as defined in normal adults (δ_1 [0 – 1.9 Hz], δ_2 [1.9 – 3.4 Hz], θ_1 [3.4 – 5.4 Hz], θ_2 [5.4 – 7.4 Hz], α_1 [7.4 – 10 Hz], α_2 [10 – 12 Hz], β_1 [12 – 18 Hz], β_2 [18 – 24 Hz], γ [24 – 128 Hz]) (**Figure 3A**, blue). Finally, a 9D feature vector V_F was defined from the normalized energy distribution in these frequency bands, itself computed as the sum of averaged (over time) atom energies relative to the total signal energy (**Figure 3A**, green).

MODEL OPTIMIZATION UNDER THE “NO STIMULATION” CONDITION

In order to simulate LFP_{FCD} , we optimized the excitation/inhibition ratio of the cortical compartment. Thus, the average EPSP/IPSP amplitude parameters of the cortical compartment $\{A_C, B_C, G_C\}$ were considered as free parameters while all other model parameters were set to fixed values (**Table 1**). The optimization method is illustrated in **Figure 3B**. For each triplet $\{A_C, B_C, G_C\}$, the feature vector $V_{F,model}$ of the model's output signal (cortical compartment's LFP) was calculated and compared to $V_{F,real}$, i.e., the feature vector computed from the average of the 20 feature vectors, each computed on a 5 s signal segment of real LFP_{FCD} . Feature vectors $V_{F,model}$ and $V_{F,real}$ were computed as described in section Processing of Real and Simulated Signals. The optimization procedure aimed at finding the triplet $\{\hat{A}_C, \hat{B}_C, \hat{G}_C\}$ that minimizes a cost function simply corresponding to the Euclidean



distance $d(V_{F,real}, V_{F,model})$ when parameters A_C , B_C , and G_C span pre-defined ranges of values according to a Brute-Force procedure.

RESULTS

In this section, results regarding the identification of cellular mechanisms underlying the modulation of cortical activity by thalamic DBS are reported. First, the model capability to reproduce signals similar to those recorded from the FCD in the patient was assessed, under two conditions (no stimulation and during stimulation). Three mechanisms contributing to frequency-dependant stimulation effects could be identified. Then, simulations were performed to analyze the marginal or joint contribution of these mechanisms at low, intermediate or high frequency stimulation.

SIMULATION OF LFPs_{FCD} UNDER NO STIMULATION CONDITION

As a first step, we verified the ability of the model to generate signals that resemble those recorded from the FCD in the considered patient (LFPs_{FCD}). This procedure, described in sections Processing of Real and Simulated Signals and Model Optimization Under the “No Stimulation” Condition, led us to identify a minimal distance (**Figures 4A–C**) and thus an optimal parameter vector $\{\hat{A}_C, \hat{B}_C, \hat{G}_C\} = \{6, 14, 16.5\}$ for which simulated signals under the no stimulation condition have similar features as compared with those of real signals (**Figure 4D**).

SIMULATION OF LFPs_{FCD} UNDER STIMULATION CONDITIONS

Actual LFPs_{FCD} recorded at various stimulation frequencies (2, 50, 70, 100, and 150 Hz) were first characterized using the matching pursuit method described in section Processing of

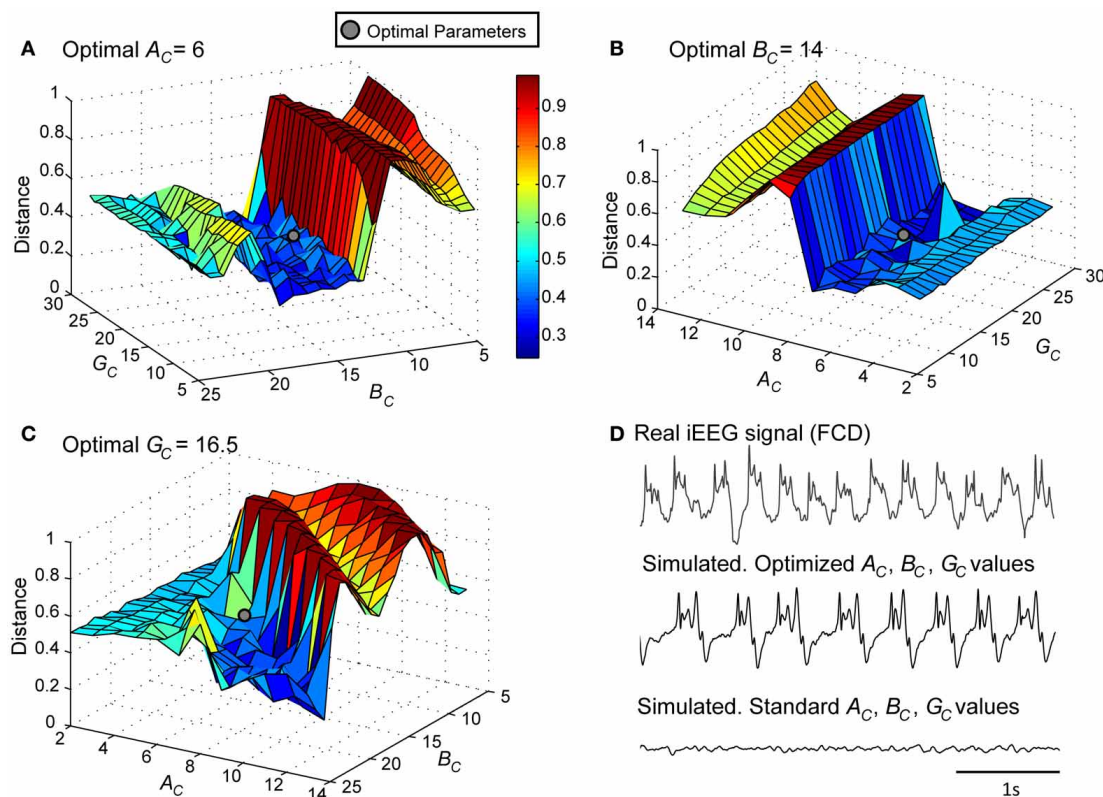


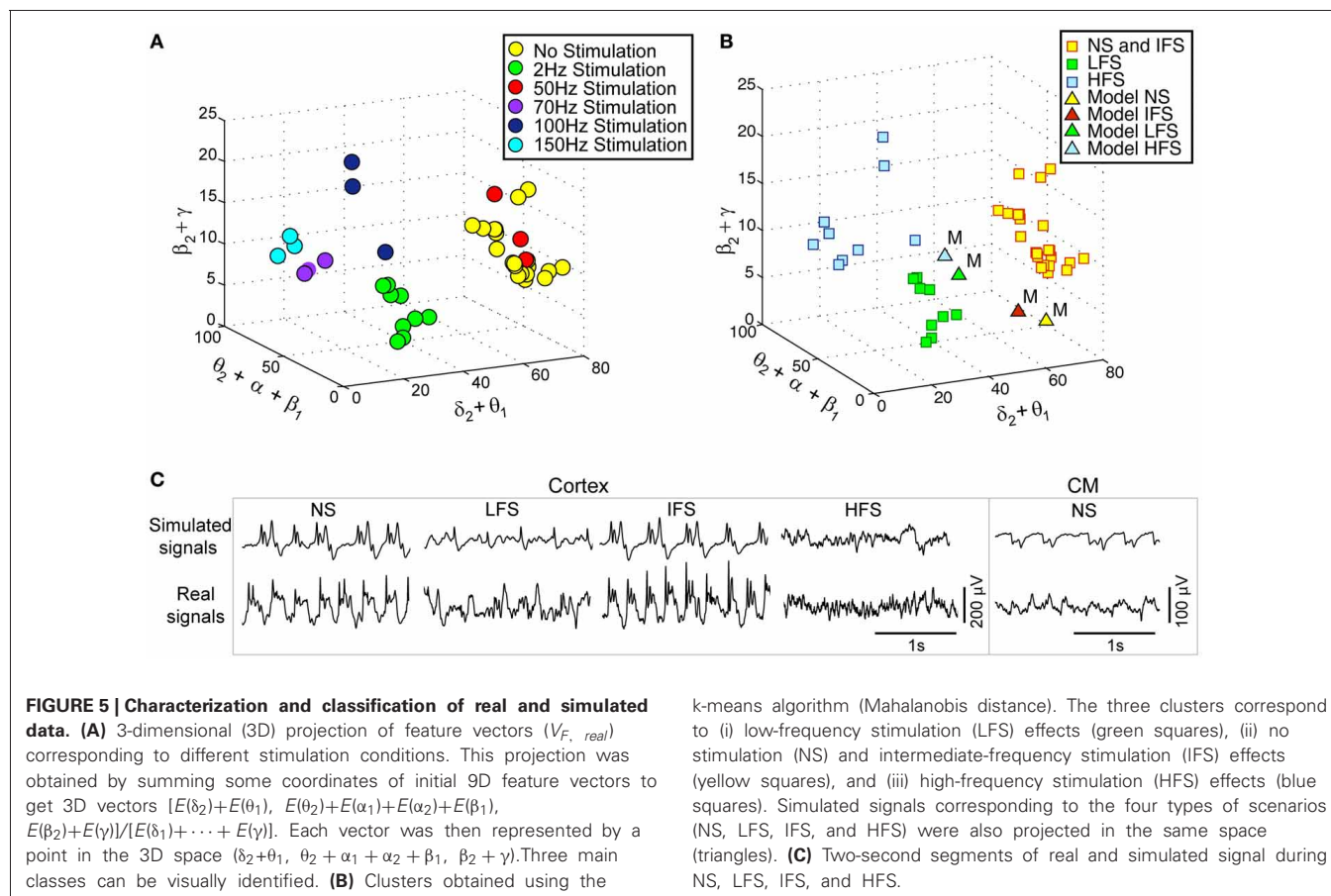
FIGURE 4 | Model parameter optimization. Normalized Euclidean distance between $V_{E, real}$ and $V_{E, model}$. Best fit (gray disk) between simulated and real $LFPs_{FCD}$ was obtained for **(A)** $A_C = 6$, **(B)** $B_C = 14$, and **(C)** $G_C = 16.5$. **(D)** For these modified values of excitation and

inhibition, the simulated signal exhibits similar characteristics as the iEEG signal recorded in the FCD. For standard values of excitation and inhibition ($A_C = 3$, $B_C = 50$, $G_C = 22$), the model generates background EEG activity.

Real and Simulated Signals (**Figure 3A**). Results are shown in **Figure 5A** where feature vectors of segments of $LFPs_{FCD}$ are represented in a 3D space where axes correspond to merged typical EEG frequency bands (δ_2 to θ_1 , θ_2 to β_1 , β_2 to γ). Results show that the distribution of points in the 3D frequency space is not random but clustered, indicating that the frequency content of $LFPs_{FCD}$ segments depends on the stimulation frequency. In addition, some clusters are very close. This is typically the case for i) the no stimulation (yellow) and the 50 Hz stimulation conditions (red) on the one hand, and ii) the 70 Hz (violet) and 150 Hz (cyan) stimulation conditions on the other hand. To go beyond the qualitative clustering performed by visual inspection of 3D plots, a K-means clustering algorithm implemented in MATLAB and using a Mahalanobis distance was used to automatically detect the three types of stimulation effects. Initial centroids were randomly chosen. The optimal clustering that globally minimizes intra-cluster inertia is presented in **Figure 5B**. $LFPs_{FCD}$ segments were automatically classified into three subgroups. The first subgroup contains $LFPs_{FCD}$ segments obtained for low-frequency stimulation (LFS). The second subgroup gathers all segments recorded for high frequency stimulations (HFS, > 70 Hz). And finally, in the third subgroup, segments obtained under the no stimulation and the intermediate stimulation frequency (IFS, 50 Hz) conditions are merged

together, suggesting that this stimulation frequency does not reduce the “epileptiform aspect” of the activity reflected in the LFP.

Based on this characterization of local field potentials recorded in the FCD ($LFPs_{FCD}$), parameters S_{TC} , S_{R1} and S_{R2} were manually tuned to lead the model to generate simulated signals which have spectral characteristics similar to those of actual $LFPs_{FCD}$. Such a manual procedure was sufficient to reproduce stimulation effects observed in one patient. However, extending the study to a larger group of patients would have made imperative an automated parameter fitting procedure based on the spectral characteristics of real EEG signals as in Rowe et al. (2004). **Figure 5B** shows the projection of representative simulated $LFPs_{FCD}$ in the 3D frequency space (“M” triangles). As depicted, simulated signals obtained for LFS, IFS and HFS were close to corresponding clusters obtained from real signals for the exact same computation of feature vectors. Shown in **Figure 5C**, these representatives simulated $LFPs_{FCD}$ do not perfectly match actual signals. However, qualitatively similar bifurcations were observed in the model when the stimulation conditions are changed. Indeed, under the no stimulation (NS) and the IFS condition the model generates rhythmic slow oscillations (δ) with superimposed faster activity (β , γ), as observed in real data. For LFS and HFS conditions, strong modulation of this activity was also obtained in the model.



At LFS, in the model, the slow wave activity was strongly reduced but spike events occurred in the signals at the instant times of stimulation, mimicking, to some extent, comparable events also present in actual LFPs_{FCD}. Finally, at HFS, slow oscillations (δ) were abolished in the model which generates quasi-normal background activity. This simulated activity was also comparable to real activity observed for HFS stimulation but disclosed less γ activity. Note that these are the effects which were quantified in **Figure 5B**. The qualitative optimization procedure of parameters S_{TC} , S_{Rt1} , and S_{Rt2} was then complemented by an evaluation of parameter sensitivity aimed at studying the impact of random changes affecting the parameter vector $\Theta = \{A_C, B_C, G_C, A_{Th}, B_{Th}, G_{Th}, A_{Rt}\}$ on simulated signals. Parameter vector Θ determines the excitability properties in the three model compartments. As shown in **Figure 6**, results show that the simulated signals obtained under the four stimulation conditions (NS, LFS, IFS, HFS) stay “quite robust” (in the sense that waveforms are conserved) when parameters stay in the range $[\Theta \pm \zeta, \Theta]$ with $0 \leq \zeta \leq 0.2$.

MECHANISMS UNDERLYING FREQUENCY-DEPENDANT STIMULATION EFFECTS

Three main mechanisms implemented in the model are required to mimic actually observed effects of the CM nucleus stimulation. These mechanisms are the following: (i) the presence of feed-forward inhibition (FFI) at the level of thalamic projections

to the FCD, (ii) the presence of short-term depression (STD) at the level of the thalamocortical glutamatergic synapses and, (iii) the depolarization of RtN inhibitory interneurons targeting TC cells.

This result raises an additional question: to what extent the joint effect of these mechanisms is necessary to reproduce frequency-dependant stimulation effects (LFS, IFS, and HFS). In order to assess their individual contribution, we performed simulations where each mechanism was either present in—or removed from—the model (the model parameters remaining unchanged). Results are displayed in **Figure 7**. First, they confirmed that both FFI and STD mechanisms are jointly necessary in the model to suppress the epileptic activity in the FCD when LFS is being used since the withdrawal of either STD or FFI leads the model to generate epileptic activity at LFS. Second, results indicated that the RtN inhibitory interneurons targeting TC cells (both I_1^{RT} and I_2^{RT} subpopulations) must be affected (i.e., depolarized) by the stimulation to obtain a suppression of epileptic activity when HFS is being used, as observed in the patient. Third, and interestingly, an unexpected effect was observed at IFS when the depolarization of I_2^{RT} interneurons was removed from the model. Indeed, epileptic activity was abolished in this case, which is really unlikely to occur during actual stimulation as both subtypes of neurons are expected to be affected by the direct stimulation of the CM nucleus.

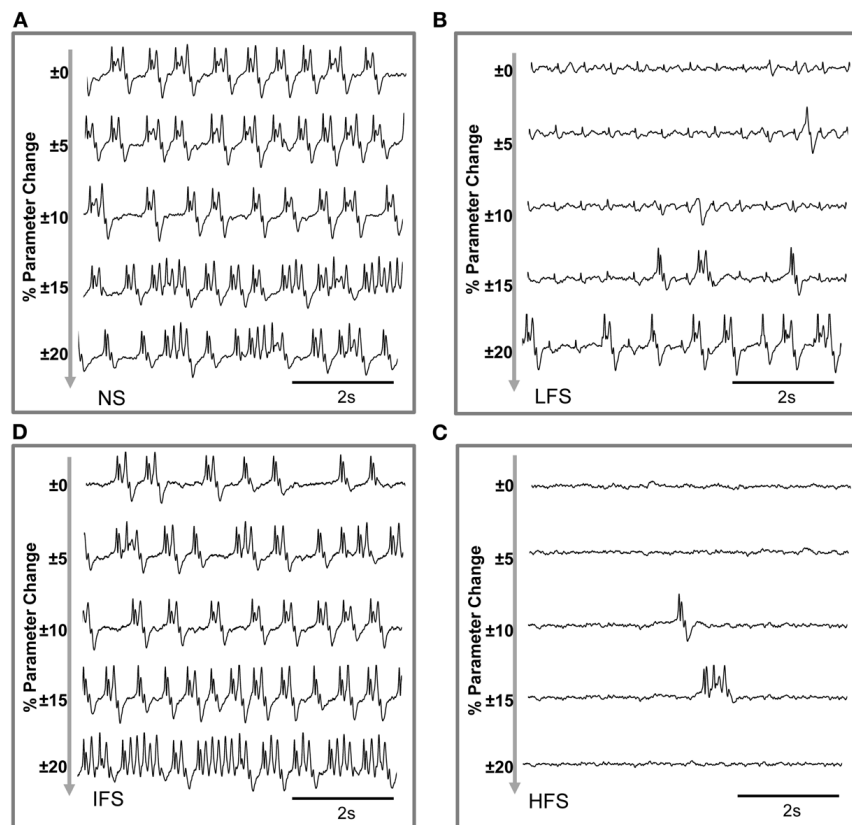


FIGURE 6 | Evaluation of parameter sensitivity. Model output sensitivity to variations of excitatory and inhibitory key parameters. Realizations of parameter vector $\Theta = \{A_C, B_C, G_C, A_{Th}, B_{Th}, G_{Th}, A_{Rt}\}$ were randomly (uniform law) generated around the optimal parameter vector Θ_0 over a

variation domain defined by $(1 \pm \zeta) \cdot \Theta_0$. For $\zeta \leq 0.2$ ($\pm 20\%$ variation), stimulation effects are preserved in the model for **(A)** no stimulation, **(B)** low-frequency stimulation, **(C)** intermediate-frequency stimulation, and **(D)** high-frequency stimulation.

These results were complemented by a deeper analysis of the thalamic output (i.e., the firing rate of *TC* cells) in response to stimulation at low, intermediate and high frequency. Results are provided in **Figure 8**. First, they showed that the thalamic output dramatically differs depending on the stimulation frequency (**Figure 8A**). Under the no stimulation condition, the firing rate continuously oscillates around a certain value (referred to as Λ , **Figure 8A**). At LFS, the firing rate was found to be lower, except at the stimulation times where it abruptly and transiently increased. At IFS, a balance was observed between time intervals for which the *TC* firing is above and below Λ . Finally, at HFS, the output of *TC* cells was found to be very low, i.e., systematically under the threshold Λ . From these observations, we could define (i) two time intervals, $\Delta 1$ and $\Delta 2$, for which the *TC* cells firing rate is either below Λ ($\Delta 1$) or above Λ ($\Delta 2$) and (ii) a “high to low firing” ratio (*HtoLR*) which provides an indication on the amount of time the *TC* cells spend firing (up state) relatively to the amount of time they do not fire (down state). **Figure 8B** provides the evolution of the *HtoLR* when the stimulation frequency is progressively changing from 0 to 150 Hz in the model. As depicted, these simulations indicated that three stimulation frequency ranges have dramatic effects on the firing of *TC* cells. First, from 0 to 20 Hz, the down state

is predominant. Then, an abrupt jump was observed around 22 Hz indicating that beyond this value, the firing rate dramatically increased. Interestingly, from 55 to 65 Hz, a progressive decrease of the *HtoLR* was observed. Then, after 70 Hz, the ratio is equal to zero indicating that *TC* cells did not fire anymore. Finally, in order to relate the thalamic activity with the cortical activity, we plotted the phase portraits (*TC* cell firing vs. cortical LFP) as illustrated in **Figure 8C**. Results confirmed the visual inspection of signals simulated at the two sites. For the no stimulation (NS) and for the intermediate frequency stimulation (IFS) conditions, phase portraits were found to be quite similar. They indicated the presence of mixed slow/fast oscillations in both signals. For the low frequency stimulation (LFS) condition, oscillations in the simulated LFP in the FCD were reduced. They came along with short-duration, abrupt and rhythmic augmentations of the *TC* firing corresponding to stimulation pulses. Finally, for the high frequency stimulation (HFS) condition, oscillations in both types of activity stayed confined to small amplitude values.

DISCUSSION

We modeled the thalamocortical loop in order to investigate frequency-dependent effects of electrical stimulation performed

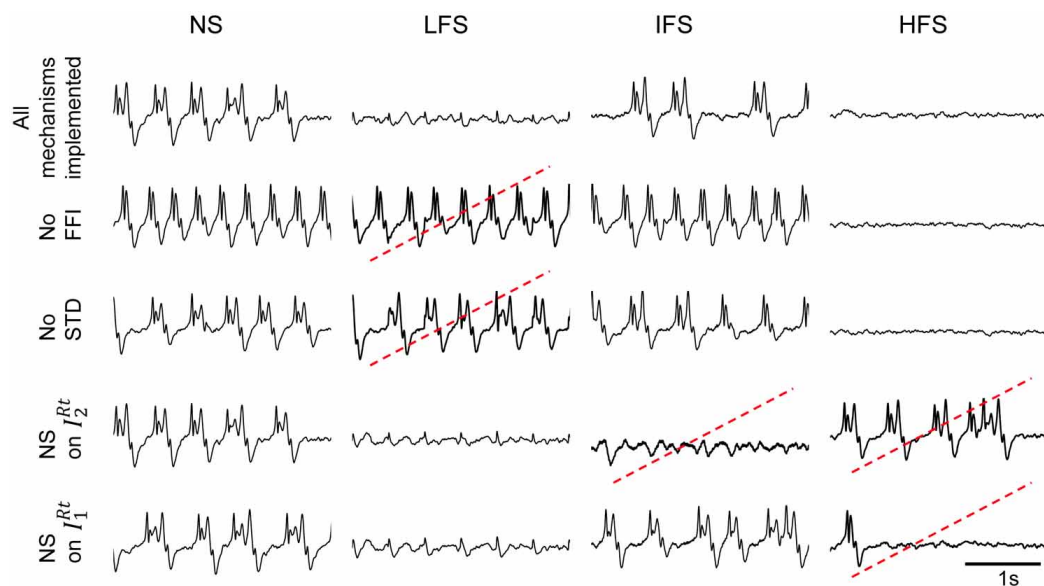


FIGURE 7 | Conditions to reproduce frequency-dependent stimulation effects. Model output in the case where one of the implemented mechanisms (FFI, STD, depolarization of I_2^{Rt} , and I_1^{Rt} , respectively) is removed at a time. LFS effects are not reproduced when the model does not account for FFI and STD. HFS effects

require the depolarization of both reticular populations I_2^{Rt} and I_1^{Rt} . Suppression of epileptic activity is observed at IFS when I_2^{Rt} interneurons are removed. Red dotted lines indicate situations where simulated signals do not match real ones for given stimulation condition.

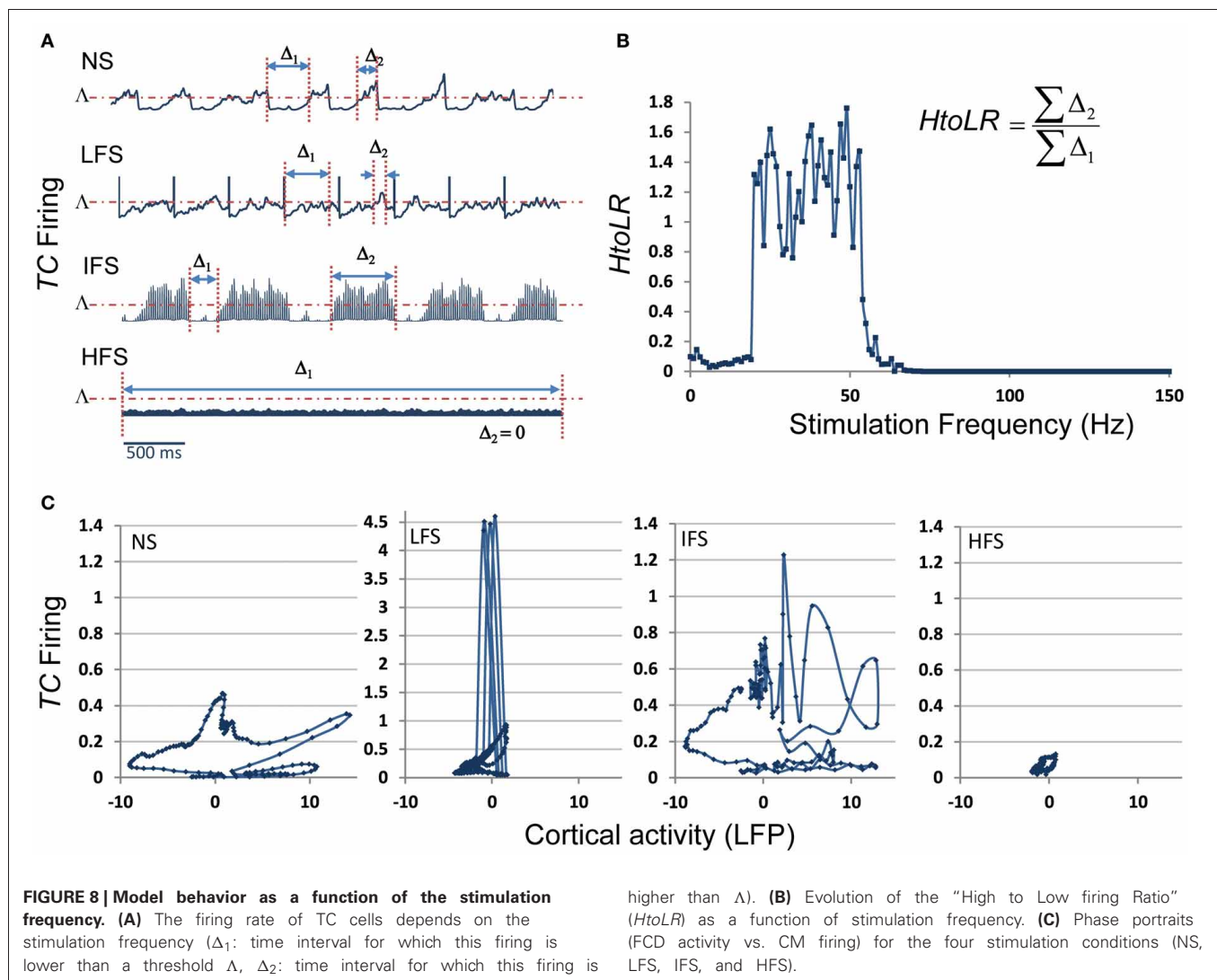
in the thalamus and aimed at modulating the neocortical activity. We chose to elaborate our model at a mesoscopic level, i.e., intermediate between microscopical and macroscopic.

Regarding the model architecture, we followed a similar approach to that used in previously proposed models of the thalamocortical loop (Robinson et al., 2002; Suffczynski et al., 2004; Breakspear et al., 2006; Roberts and Robinson, 2008; Marten et al., 2009; Crunelli et al., 2011). Our model includes three main compartments: cerebral cortex, reticular nucleus and thalamic relay. Subpopulations of neurons and interneurons located in these three structures interact via excitatory and/or inhibitory synaptic connections. The novelty with respect to aforementioned studies is threefold. First, we modified the cortical compartment in order to better approximate the temporal dynamics of epileptic signals recorded in the FCD. This modification consisted in the use of two types of interneurons (mediating GABAergic IPSPs with slow and fast kinetics on cortical principal cells), as reported in a previous study (Molae-Ardekani et al., 2010). Second, our model accounts for the direct effects of electrical stimulation. At this stage, we used the $\Delta V \approx \lambda \cdot \vec{E}$ assumption according to which the perturbation of the mean membrane potential of neurons is a linear function of the electrical field magnitude induced by bipolar stimulation. This “ λE ” assumption was already used in neural mass models in the context of low-intensity direct hippocampal stimulation to anticipate seizures (Suffczynski et al., 2008) as well as in the analysis of the stimulus-response relationship of DBS in healthy animals (Adhikari et al., 2009). However, it is worth mentioning that in our model, the three subtypes of neurons (TC cells and both subpopulations of inhibitory neurons in the RtN) are depolarized by the stimulation, as suggested in

(Molae-Ardekani et al., 2013) and conversely to (Adhikari et al., 2009) where only principal cells are impacted. And third, our model includes two well-known mechanisms at the cortical level: feed-forward inhibition (FFI) and short-term depression (STD).

As in any modeling approach, our approach has some limitations. First, the chosen modeling level does not allow for analyzing sub-cellular mechanisms involved in stimulation-evoked changes. Similarly, it does not account for direct activation of axons by stimulation versus somatic inhibition (McIntyre et al., 2004b), nor for the mechanisms of orthodromic/antidromic propagation of action potentials due to stimulation (Degos et al., 2005; Hammond et al., 2007; Dorval et al., 2008). Second, a strong assumption in the type of model we used (neural mass) is related to the intrinsic synchronization among neurons included in a given sub-population. This assumption does not allow for representing either de- or weakly-synchronized firing patterns that may be observed during epileptic activity, in particular during high frequency oscillations that can be encountered in FCDs (Brázdil et al., 2010). Nevertheless, we could accurately reproduce the abnormal rhythms generated in the FCD suggesting that main pyramidal cells have a relatively synchronized activity in this epileptogenic tissue. Third, regarding plasticity-related mechanisms, we only implemented short-term effects (i.e., STD) and neglected long-term plastic changes that may be induced by DBS (Shukla et al., 2013).

Despite these limitations, we could identify a number of mesoscopic factors which could explain the frequency-dependent mechanisms of thalamic stimulation. The model was tuned using electrophysiological data recorded in a patient in whom the centromedian nucleus (CMN) stimulation was particularly efficient



to reduce the epileptic activity of a FCD located in the premotor cortex, in a frequency-specific manner. The main findings are summarized in **Figure 9**.

“NO STIMULATION” (NS) CONDITION

In the model, under the NS condition, excitation among pyramidal cells had to be increased and inhibition had to be reduced in the cortical compartment for producing “pathological” oscillatory rhythms, as observed in the FCD. The thalamocortical loop was found to be responsible for these pathological dynamics, characteristic of FCDs. These findings are in line with histological studies showing that these typical oscillations are generated in altered brain tissue, where inhibition is partially deteriorated or dysfunctioning (Calcagnotto et al., 2005), and where excitation is heavily increased (Avoli et al., 2003). In addition to neuron alterations in the dysplastic tissue (Sisodiya et al., 2009), FCD keeps sufficient projections to—and input from—other brain structures to propagate pathological dynamics (Avoli et al., 2003). As mentioned, the presence of connections with subcortical structures was a necessary condition in the model for producing

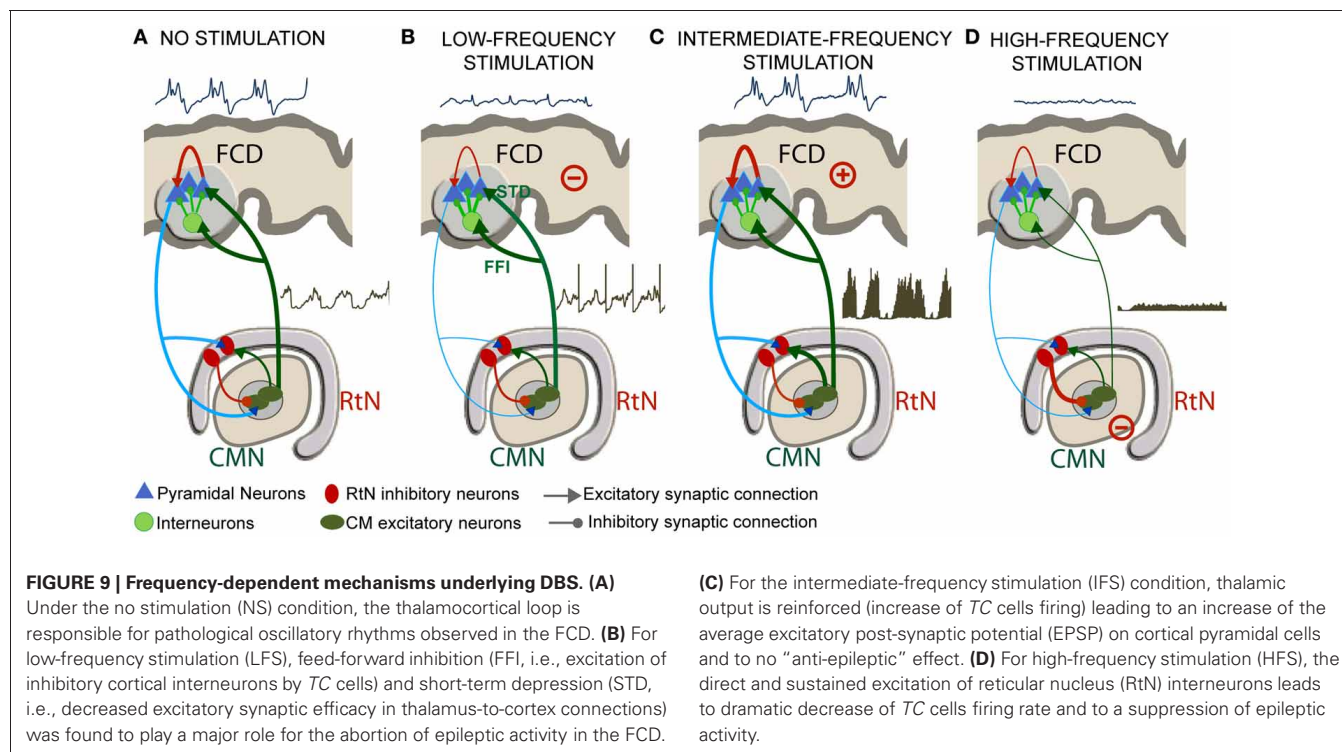
pathological oscillations resembling those actually recorded in the FCD (**Figure 9A**).

LOW-FREQUENCY STIMULATION (LFS) CONDITION

For the low-frequency stimulation (LFS, $f < 20$ Hz) condition, two mechanisms were found to play a major role for the abortion of epileptic activity in the FCD: short-term depression (STD, i.e., decreased excitatory synaptic efficacy in thalamus-to-cortex connections) and feed-forward inhibition (FFI, i.e., excitation of inhibitory cortical interneurons by TC cells) (**Figure 9B**).

STD was reported in previous studies concerning cortical adaptation to thalamic stimulation, and suggesting that electrical LFS of TC cell axons *in vivo* resulted in a 40% reduction in cortical EPSPs (Chung et al., 2002). In the same context, LFS trains in adult anaesthetized rats provoked transient long-term depression of thalamocortical synapses; this was measured by up to 40% drop in cortical EPSPs after LFS trains and under the effect of GABA antagonist (Speechley et al., 2007).

As mentioned above, the LFS effects could not be reproduced by the model without incorporating also FFI. Actually,



thalamocortical ascending fibers directly target pyramidal neurons as well as cortical GABAergic interneurons inducing EPSPs in both cell types (Pouille and Scanziani, 2001). In the model, while less efficient (STD) thalamic EPSPs arrive directly onto pyramidal neurons, IPSPs induced by thalamic stimulation also arrive on pyramidal neurons (FFI) lagging by 1–2 ms. This short latency between the onset of thalamocortical excitation and the onset of feed-forward inhibition presents a temporal “window of opportunity” for pyramidal cells to integrate excitatory and inhibitory inputs, thus keeping the transmembrane potential below firing threshold. In the literature, neuroanatomical and neurophysiological studies (Isaacson and Scanziani, 2011) showed the functional importance of FFI in regulating cortical dynamics by controlling cortical excitability (Gabernet et al., 2005). Our study suggests that LFS regulates cortical excitability by a dual mechanism of FFI and STD (Figure 9B).

INTERMEDIATE-FREQUENCY STIMULATION (IFS) CONDITION

For the intermediate-frequency stimulation (IFS, $20 < f < 70$ Hz) condition, results indicated that the thalamic output is reinforced (increase of TC cells firing) and leads to an increase of the average excitatory post-synaptic potential (EPSP) on cortical pyramidal cells (Figure 9C). This effect corresponds to an increase of the spatiotemporal summation of unitary EPSPs. In this case, both the cortical excitability and the gain in the excitatory thalamocortical loop is increased, leading to “no anti-epileptic” effect. We did not find much studies using DBS stimulation in the intermediate frequency range of (20–60 Hz) in the context of epilepsy. Nevertheless, it is noteworthy that 50 Hz stimulation frequency is classically used during the presurgical evaluation of patient with intractable partial epilepsy in order to

trigger seizures and delineate the epileptogenic zone (Talairach et al., 1974; Jayakar et al., 1992). The same frequency range is also known to provoke afterdischarges and was actually used in the kindling model of epilepsy (Goddard, 1967; Racine, 1972).

HIGH-FREQUENCY STIMULATION (HFS) CONDITION

Finally, for the high-frequency stimulation (HFS, $f > 70$ Hz) condition, the direct and sustained excitation of reticular nucleus (RtN) interneurons leads to strong inhibition of TC cells and thus to dramatic decrease of their firing rate. Despite the fact that TC neurons are also affected by stimulation, the response of reticular GABAergic neurons to stimulation and the higher efficiency of GABA-mediated currents ensure that IPSPs override EPSPs on TC cells. In this case, the reduced excitatory input to cortical pyramidal cells also leads to a suppression of epileptic activity (Figure 9D). This result corroborates reported stimulation studies where HFS (>100 Hz) was associated with significant decrease in epileptiform discharges *in vitro*, and reduction in seizure frequency in responding patients (Velasco et al., 2006; Fisher et al., 2010). This hypothesis is in line with recent findings suggesting that HFS of the globus pallidus (GPi) in dystonia patients decreased its firing by stimulation-evoked GABA release from afferent fibers and thereby the enhancement of inhibitory synaptic transmission by HFS (Liu et al., 2012). Similarly, HFS (100 Hz–130 Hz) of the STN neurons *in vitro* showed a suppression of the activity of the majority of neurons by the reinforcement of inhibitory responses (Filali et al., 2004). Other HFS studies also provided evidence on the inhibition of GPi output during HFS in human patients (Dostrovsky et al., 2000) as well as the disruption thalamocortical network's dysrhythmia (McIntyre and Hahn, 2010; Kendall et al., 2011).

CONCLUSION

In epilepsy research, it is well-admitted that there is, unfortunately, a lack of tangible results regarding the effects of electrical stimulation in the brain. Therefore, the very crucial issue of choosing the “optimal” stimulation parameters remains unsolved, whatever the stimulation procedure. Although computational models are always based on a number of simplifying assumptions, we think that they provide an efficient framework to (i) account for the many and essential factors that may intervene during stimulation procedures and (ii) analyze the links between these factors in a formal manner. This approach is particularly fruitful when models are well grounded in experimental/clinical data (Wendling et al., 2012). This is somehow a weak point of this study since we could make use of data sets recorded in one patient only. However, it should be mentioned that these very informative data sets stay relatively rare since many conditions have to be met (patient candidate to surgery, FCD, electrodes positioned in appropriate structures).

At this stage, the face value of the model is satisfactory. The next step is obviously to test the model predictions using animal

models. Experiments can be undertaken in rodents with electrodes implanted in the cerebral cortex and in the thalamus. First, we could start with control animals to assess the modulation of cortical rhythms during/after direct thalamic stimulation at various frequencies and for controlled vigilance states (sleep, awake, resting, exploratory). In these controls, some drugs can be used to alter some parameters related to synaptic transmission (in a more or less specific manner) which have a correspondence in the model, on the other hand. Then, refined experimental models could be introduced to get closer to the epilepsy context including models of developmental dysplastic lesions [see review in Schwartzkroin and Wenzel (2012)]. Hopefully, this combined computational/experimental approach will help us to disclose some of the highly intricate effects of DBS either at local or at network level.

ACKNOWLEDGMENTS

This study was supported by Region Bretagne, France. The first author benefits from a research allocation (ARED) in the context of the project entitled “STIMULE.”

REFERENCES

- Adhikari, M. H., Heeroma, J. H., Di Bernardo, M., Krauskopf, B., Richardson, M. P., Walker, M. C., et al. (2009). Characterisation of cortical activity in response to deep brain stimulation of ventral-lateral nucleus: modelling and experiment. *J. Neurosci. Methods* 183, 77–85. doi: 10.1016/j.jneumeth.2009.06.044
- Avoli, M., Louvel, J., Mattia, D., Olivier, A., Esposito, V., Pumain, R., et al. (2003). Epileptiform synchronization in the human dysplastic cortex. *Epileptic Disord.* 5(Suppl. 2), S45–S50.
- Beurrier, C., Bioulac, B., Audin, J., and Hammond, C. (2001). High-frequency stimulation produces a transient blockade of voltage-gated currents in subthalamic neurons. *J. Neurophysiol.* 85, 1351–1356.
- Bikson, M., Inoue, M., Akiyama, H., Deans, J. K., Fox, J. E., Miyakawa, H., et al. (2004). Effects of uniform extracellular DC electric fields on excitability in rat hippocampal slices *in vitro*. *J. Physiol.* 557, 175–190. doi: 10.1113/jphysiol.2003.055772
- Bojak, I., and Liley, D. T. (2005). Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 71:041902. doi: 10.1103/PhysRevE.71.041902
- Boon, P., Raedt, R., De Herdt, V., Wyckhuys, T., and Vonck, K. (2009). Electrical stimulation for the treatment of epilepsy. *Neurotherapeutics* 6, 218–227. doi: 10.1016/j.nurt.2008.12.003
- Boudreau, C. E., and Ferster, D. (2005). Short-term depression in thalamocortical synapses of cat primary visual cortex. *J. Neurosci.* 25, 7179–7190. doi: 10.1523/JNEUROSCI.1445-05.2005
- Brázdil, M., Halánek, J., Jurák, P., Daniel, P., Kuba, R., Chrástina, J., et al. (2010). Interictal high-frequency oscillations indicate seizure onset zone in patients with focal cortical dysplasia. *Epilepsy Res.* 90, 28–32. doi: 10.1016/j.epilepsyres.2010.03.003
- Breakspear, M., Roberts, J. A., Terry, J. R., Rodrigues, S., Mahant, N., and Robinson, P. A. (2006). A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cereb. Cortex* 16, 1296–1313. doi: 10.1093/cercor/bhj072
- Calcagnotto, M. E., Paredes, M. F., Tihan, T., Barbaro, N. M., and Baraban, S. C. (2005). Dysfunction of synaptic inhibition in epilepsy associated with focal cortical dysplasia. *J. Neurosci.* 25, 9649–9657. doi: 10.1523/JNEUROSCI.2687-05.2005
- Chung, S., Li, X., and Nelson, S. B. (2002). Short-term depression at thalamocortical synapses contributes to rapid adaptation of cortical sensory responses *in vivo*. *Neuron* 34, 437–446. doi: 10.1016/S0896-6273(02)00659-1
- Cooper, I. S., Amin, I., and Gilman, S. (1973). The effect of chronic cerebellar stimulation upon epilepsy in man. *Trans. Am. Neurol. Assoc.* 98, 192–196.
- Crunelli, V., Cope, D. W., and Terry, J. R. (2011). Transition to absence seizures and the role of GABA(A) receptors. *Epilepsy Res.* 97, 283–289.
- David, O., and Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage* 20, 1743–1755. doi: 10.1016/j.neuroimage.2003.07.015
- Davis, R., Engle, H., Kudzma, J., Gray, E., Ryan, T., and Duszak, A. (1982). Update of chronic cerebellar stimulation for spasticity and epilepsy. *Appl. Neurophysiol.* 45, 44–50.
- Degos, B., Deniau, J. M., Thierry, A. M., Glowinski, J., Pezard, L., and Maurice, N. (2005). Neuroleptic-induced catalepsy: electrophysiological mechanisms of functional recovery induced by high-frequency stimulation of the subthalamic nucleus. *J. Neurosci.* 25, 7687–7696. doi: 10.1523/JNEUROSCI.1056-05.2005
- Dorval, A. D., Russo, G. S., Hashimoto, T., Xu, W., Grill, W. M., and Vitek, J. L. (2008). Deep brain stimulation reduces neuronal entropy in the MPTP-primate model of Parkinson's disease. *J. Neurophysiol.* 100, 2807–2818. doi: 10.1152/jn.90763.2008
- Dostrovsky, J. O., Levy, R., Wu, J. P., Hutchison, W. D., Tasker, R. R., and Lozano, A. M. (2000). Microstimulation-induced inhibition of neuronal firing in human globus pallidus. *J. Neurophysiol.* 84, 570–574.
- Filali, M., Hutchison, W. D., Palter, V. N., Lozano, A. M., and Dostrovsky, J. O. (2004). Stimulation-induced inhibition of neuronal firing in human subthalamic nucleus. *Exp. Brain Res.* 156, 274–281. doi: 10.1007/s00221-003-1784-y
- Fisher, R., Salanova, V., Witt, T., Worth, R., Henry, T., Gross, R., et al. (2010). Electrical stimulation of the anterior nucleus of thalamus for treatment of refractory epilepsy. *Epilepsia* 51, 899–908. doi: 10.1111/j.1528-1167.2010.02536.x
- Freeman, W. J. (1973). “A model of the olfactory system,” in *Neural Modeling*, eds M. A. B. Brazier, D. O. Walter, and D. Schneider (Los Angeles, CA: University of California), 41–62.
- Freeman, W. J. (1992). Tutorial on neurobiology: from single neurons to brain chaos. *Int. J. Bifurcat. Chaos* 2, 451–482. doi: 10.1142/S0218127492000653
- Frohlich, E., and McCormick, D. A. (2010). Endogenous electric fields may guide neocortical network activity. *Neuron* 67, 129–143. doi: 10.1016/j.neuron.2010.06.005
- Gabernet, L., Jadhav, S. P., Feldman, D. E., Carandini, M., and Scanziani, M. (2005). Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. *Neuron* 48, 315–327. doi: 10.1016/j.neuron.2005.09.022
- Goddard, G. V. (1967). Development of epileptic seizures through brain stimulation at low intensity. *Nature* 214, 1020–1021. doi: 10.1038/2141020a0
- Hablit, J. J. (1976). Intramuscular penicillin epilepsy in the cat: effects of chronic cerebellar stimulation. *Exp. Neurol.* 50, 505–514. doi: 10.1016/0014-4886(76)90022-4
- Hammond, C., Bergman, H., and Brown, P. (2007). Pathological synchronization in Parkinson's disease:

- networks, models and treatments. *Trends Neurosci.* 30, 357–364. doi: 10.1016/j.tins.2007.05.004
- Isaacson, J. S., and Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron* 72, 231–243. doi: 10.1016/j.neuron.2011.09.027
- Jansen, B. H., Zouridakis, G., and Brandt, M. E. (1993). A neurophysiologically-based mathematical model of flash visual evoked potentials. *Biol. Cybern.* 68, 275–283. doi: 10.1007/BF00224863
- Jayakar, P., Alvarez, L. A., Duchowny, M. S., and Resnick, T. J. (1992). A safe and effective paradigm to functionally map the cortex in childhood. *J. Clin. Neurophysiol.* 9, 288–293. doi: 10.1097/00004691-199204010-00009
- Jiang, B., Akaneya, Y., Hata, Y., and Tsumoto, T. (2003). Long-term depression is not induced by low-frequency stimulation in rat visual cortex *in vivo*: a possible preventing role of endogenous brain-derived neurotrophic factor. *J. Neurosci.* 23, 3761–3770.
- Kendall, H. L., Frederick, L. H., Su-Youne, C., Dongchul, C. L., David, W. R., Cameron, C. M., et al. (2011). High frequency stimulation abolishes thalamic network oscillations: an electrophysiological and computational analysis. *J. Neural Eng.* 8:046001. doi: 10.1088/1741-2560/8/4/046001
- Kim, E., Owen, B., Holmes, W. R., and Grover, L. M. (2012). Decreased afferent excitability contributes to synaptic depression during high frequency stimulation in hippocampal area CA1. *J. Neurophysiol.* 108, 1965–1976. doi: 10.1152/jn.00276.2011
- Krstulovic, S., and Gribonval, R. (2006). “Mptk: Matching Pursuit Made Tractable,” in: *IEEE International Conference on Acoustics, Speech and Signal Processing*, (Toulouse), 496–499.
- Liley, D. T., and Bojak, I. (2005). Understanding the transition to seizure by modeling the epileptiform activity of general anesthetic agents. *J. Clin. Neurophysiol.* 22, 300–313.
- Liu, L. D., Prescott, I. A., Dostrovsky, J. O., Hodaie, M., Lozano, A. M., and Hutchison, W. D. (2012). Frequency-dependent effects of electrical stimulation in the globus pallidus of dystonia patients. *J. Neurophysiol.* 108, 5–17. doi: 10.1152/jn.00527.2011
- Lopes Da Silva, F. H. (2006). Event-related neural activities: what about phase? *Prog. Brain Res.* 159, 3–17. doi: 10.1016/S0079-6123(06)59001-6
- Lopes Da Silva, F. H., Hoeks, A., Smits, H., and Zetterberg, L. H. (1974). Model of brain rhythmic activity. The alpha-rhythm of the thalamus. *Kybernetik* 15, 27–37. doi: 10.1007/BF00270757
- Lopes Da Silva, F. H., van Rotterdam, A., Barts, P., van Heusden, E., and Burr, W. (1976). Models of neuronal populations: the basic mechanisms of rhythmicity. *Prog. Brain Res.* 45, 281–308. doi: 10.1016/S0079-6123(08)60995-4
- Lytton, W. W. (2008). Computer modelling of epilepsy. *Nat. Rev. Neurosci.* 9, 626–637. doi: 10.1038/nrn2416
- Mallat, S. G., and Zhifeng, Z. (1993). Matching pursuits with time-frequency dictionaries. *Signal Processing, IEEE Transactions* 41, 3397–3415. doi: 10.1109/78.258082
- Manola, L., Holsheimer, J., Veltink, P., and Buiteweg, J. R. (2007). Anodal vs cathodal stimulation of motor cortex: a modeling study. *Clin. Neurophysiol.* 118, 464–474. doi: 10.1016/j.clinph.2006.09.012
- Manola, L., Roelofsen, B. H., Holsheimer, J., Marani, E., and Geelen, J. (2005). Modelling motor cortex stimulation for chronic pain control: electrical potential field, activating functions and responses of simple nerve fibre models. *Med. Biol. Eng. Comput.* 43, 335–343. doi: 10.1007/BF02345810
- Marten, F., Rodrigues, S., Benjamin, O., Richardson, M. P., and Terry, J. R. (2009). Onset of polyspike complexes in a mean-field model of human electroencephalography and its application to absence epilepsy. *Philos. Trans. A Math. Phys. Eng. Sci.* 367, 1145–1161. doi: 10.1098/rsta.2008.0255
- McIntyre, C. C., Grill, W. M., Sherman, D. L., and Thakor, N. V. (2004a). Cellular effects of deep brain stimulation: model-based analysis of activation and inhibition. *J. Neurophysiol.* 91, 1457–1469.
- McIntyre, C. C., and Hahn, P. J. (2010). Network perspectives on the mechanisms of deep brain stimulation. *Neurobiol. Dis.* 38, 329–337. doi: 10.1016/j.nbd.2009.09.022
- McIntyre, C. C., Savasta, M., Kerkerian-Le Goff, L., and Vitek, J. L. (2004b). Uncovering the mechanism(s) of action of deep brain stimulation: activation, inhibition, or both. *J. Clin. Neurophysiol.* 115, 1239–1248.
- Miranda, P. C., Wendling, F., Ruffini, G., Merlet, I., Molaee-Ardekani, B., Dunne, S., et al. (2009). *Brain Stimulation: Models, Experiments and Open Questions, Deliverable D1.1: Review of the State of the Art in Currents Distribution and Effects HIVE* (FET Open FP7 EU project), Available online at: <http://hive-eu.org>
- Molaee-Ardekani, B., Benquet, P., Bartolomei, F., and Wendling, F. (2010). Computational modeling of high-frequency oscillations at the onset of neocortical partial seizures: from ‘altered structure’ to ‘dysfunction’. *Neuroimage* 52, 1109–1122. doi: 10.1016/j.neuroimage.2009.12.049
- Molaee-Ardekani, B., Marquez-Ruiz, J., Merlet, I., Leal-Campanario, R., Gruart, A., Sanchez-Campusano, R., et al. (2013). Effects of transcranial Direct Current Stimulation (tDCS) on cortical activity: a computational modeling study. *Brain Stimul.* 6, 25–39. doi: 10.1016/j.brs.2011.12.006
- Nunez, P. L. (1974). The brain wave equation: a model for the EEG. *Math. Biosci.* 21, 279–297. doi: 10.1016/0025-5564(74)90020-0
- Palmini, A. (2010). Electrophysiology of the focal cortical dysplasias. *Epilepsia* 51(Suppl. 1), 23–26. doi: 10.1111/j.1528-1167.2009.02437.x
- Pasnicu, A., Denoyer, Y., Haegelen, C., Pasqualini, E., and Biraben, A. (2013). Modulation of paroxysmal activity in focal cortical dysplasia by centromedian thalamic nucleus stimulation. *Epilepsy Res.* 104, 264–268. doi: 10.1016/j.eplepsyres.2012.10.012
- Pouille, F., and Scanziani, M. (2001). Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science* 293, 1159–1163. doi: 10.1126/science.1060342
- Racine, R. J. (1972). Modification of seizure activity by electrical stimulation. II. Motor seizure. *Electroencephalogr. Clin. Neurophysiol.* 32, 281–294. doi: 10.1016/0013-4694(72)90177-0
- Rattay, F. (1998). Analysis of the electrical excitation of CNS neurons. *IEEE Trans. Biomed. Eng.* 45, 766–772. doi: 10.1109/10.678611
- Reimer, G. R., Grimm, R. J., and Dow, R. S. (1967). Effects of cerebellar stimulation on cobalt-induced epilepsy in the cat. *Electroencephalogr. Clin. Neurophysiol.* 23, 456–462. doi: 10.1016/0013-4694(67)90188-5
- Roberts, J. A., and Robinson, P. A. (2008). Modeling absence seizure dynamics: implications for basic mechanisms and measurement of thalamocortical and corticothalamic latencies. *J. Theor. Biol.* 253, 189–201. doi: 10.1016/j.jtbi.2008.03.005
- Robinson, P. A., Rennie, C. J., and Rowe, D. L. (2002). Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 65:041924. doi: 10.1103/PhysRevE.65.041924
- Robinson, P. A., Rennie, C. J., and Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 56, 826–840. doi: 10.1103/PhysRevE.56.826
- Rowe, D. L., Robinson, P. A., and Rennie, C. J. (2004). Estimation of neurophysiological parameters from the waking EEG using a biophysical model of brain dynamics. *J. Theor. Biol.* 231, 413–433. doi: 10.1016/j.jtbi.2004.07.004
- Schwartzkroin, P. A., and Wenzel, H. J. (2012). Are developmental dysplastic lesions epileptogenic? *Epilepsia* 53(Suppl. 1), 35–44. doi: 10.1111/j.1528-1167.2012.03473.x
- Shen, K. Z., Zhu, Z. T., Munhall, A., and Johnson, S. W. (2003). Synaptic plasticity in rat subthalamic nucleus induced by high-frequency stimulation. *Synapse* 50, 314–319. doi: 10.1002/syn.10274
- Shukla, A. W., Moro, E., Gunraj, C., Lozano, A., Hodaie, M., Lang, A., et al. (2013). Long-term subthalamic nucleus stimulation improves sensorimotor integration and proprioception. *J. Neurol. Neurosurg. Psychiatry*. doi:10.1136/jnnp-2012-304102. [Epub ahead of print].
- Sillay, K., and Starr, P. (2009). “Chapter 42 - deep brain stimulation in Parkinson’s Disease,” in *Neuromodulation*, eds E. S. Krames, P. H. Peckham, and A. R. Rezai (San Diego, CA: Academic Press), 539–548.
- Sisodiya, S. M., Fauser, S., Cross, J. H., and Thom, M. (2009). Focal cortical dysplasia type II: biological features and clinical perspectives. *Lancet Neurol.* 8, 830–843. doi: 10.1016/S1474-4422(09)70201-7
- Speechley, W. J., Hogsdon, J. L., and Dringenberg, H. C. (2007). Continuous white noise exposure during and after auditory critical period differentially alters bidirectional thalamocortical plasticity in rat auditory cortex *in vivo*. *Eur. J. Neurosci.* 26, 2576–2584. doi: 10.1111/j.1460-9568.2007.05857.x
- Stam, C. J., Pijn, J. P., Suffczynski, P., and Lopes Da Silva, F. H. (1999). Dynamics of the human alpha rhythm: evidence for non-linearity? *Clin. Neurophysiol.* 110, 1801–1813.

- Suffczynski, P., Kalitzin, S., Da Silva, F. L., Parra, J., Velis, D., and Wendling, F. (2008). Active paradigms of seizure anticipation: computer model evidence for necessity of stimulation. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 78:051917. doi: 10.1103/PhysRevE.78.051917
- Suffczynski, P., Kalitzin, S., and Lopes Da Silva, F. H. (2004). Dynamics of non-convulsive epileptic phenomena modeled by a bistable neuronal network. *Neuroscience* 126, 467–484. doi: 10.1016/j.neuroscience.2004.03.014
- Suffczynski, P., Kalitzin, S., Pfurtscheller, G., and Lopes Da Silva, F. H. (2001). Computational model of thalamo-cortical networks: dynamical control of alpha rhythms in relation to focal attention. *Int. J. Psychophysiol.* 43, 25–40. doi: 10.1016/S0167-8760(01)00177-5
- Talairach, J., Bancaud, J., Szikla, G., Bonis, A., Geier, S., and Vedrenne, C. (1974). [New approach to the neurosurgery of epilepsy. Stereotaxic methodology and therapeutic results. 1. Introduction and history]. *Neurochirurgie* 20(Suppl. 1), 1–240.
- Velasco, A. L., Velasco, F., Jiménez, F., Velasco, M., Castro, G., Carrillo-Ruiz, J. D., et al. (2006). Neuromodulation of the centromedian thalamic nuclei in the treatment of generalized seizures and the improvement of the quality of life in patients with Lennox–Gastaut syndrome. *Epilepsia* 47, 1203–1212. doi: 10.1111/j.1528-1167.2006.00593.x
- Velasco, F., Velasco, A. L., Velasco, M., Jiménez, F., Carrillo-Ruiz, J. D., and Castro, G. (2007). Deep brain stimulation for treatment of the epilepsies: the centromedian thalamic target. *Acta Neurochir. Suppl.* 97, 337–342. doi: 10.1007/978-3-211-33081-4_38
- Velasco, F., Velasco, M., Velasco, A. L., Jiménez, F., Marquez, I., and Rise, M. (1995). Electrical stimulation of the centromedian thalamic nucleus in control of seizures: long-term studies. *Epilepsia* 36, 63–71. doi: 10.1111/j.1528-1157.1995.tb01667.x
- Velasco, M., Velasco, F., and Velasco, A. L. (2001). Centromedian-thalamic and hippocampal electrical stimulation for the control of intractable epileptic seizures. *J. Clin. Neurophysiol.* 18, 495–513. doi: 10.1097/00004691-200111000-00001
- Velasco, M., Velasco, F., Velasco, A. L., Brito, F., Jiménez, F., Marquez, I., et al. (1997). Electroconvulsive and behavioral responses produced by acute electrical stimulation of the human centromedian thalamic nucleus. *Electroencephalogr. Clin. Neurophysiol.* 102, 461–471. doi: 10.1016/S0013-4694(96)95203-0
- Velasco, M., Velasco, F., Velasco, A. L., Jiménez, F., Brito, F., and Marquez, I. (2000). Acute and chronic electrical stimulation of the centromedian thalamic nucleus: modulation of reticulo-cortical systems and predictor factors for generalized seizure control. *Arch. Med. Res.* 31, 304–315. doi: 10.1016/S0188-4409(00)00085-0
- Wendling, F. (2008). Computational models of epileptic activity: a bridge between observation and pathophysiological interpretation. *Expert Rev. Neurother.* 8, 889–896. doi: 10.1586/14737175.8.6.889
- Wendling, F., Bartolomei, F., Bellanger, J. J., and Chauvel, P. (2002). Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *Eur. J. Neurosci.* 15, 1499–1508. doi: 10.1046/j.1460-9568.2002.01985.x
- Wendling, F., Bartolomei, F., Mina, F., Huneau, C., and Benquet, P. (2012). Interictal spikes, fast ripples and seizures in partial epilepsies - combining multi-level computational models with experimental data. *Eur. J. Neurosci.* 36, 2164–2177. doi: 10.1111/j.1460-9568.2012.08039.x
- Wendling, F., Bellanger, J. J., Bartolomei, F., and Chauvel, P. (2000). Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals. *Biol. Cybern.* 83, 367–378. doi: 10.1007/s004220000160
- Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24. doi: 10.1016/S0006-3495(72)86068-5
- Wright, G. D., and Weller, R. O. (1983). Biopsy and post-mortem findings in a patient receiving cerebellar stimulation for epilepsy. *J. Neurol. Neurosurg. Psychiatr.* 46, 266–273. doi: 10.1136/jnnp.46.3.266
- Wright, J. J., Kydd, R. R., and Lees, G. J. (1985). State-changes in the brain viewed as linear steady-states and non-linear transitions between steady-states. *Biol. Cybern.* 53, 11–17. doi: 10.1007/BF00355686
- Zavaglia, M., Astolfi, L., Babiloni, F., and Ursino, M. (2006). A neural mass model for the simulation of cortical activity estimated from high resolution EEG during cognitive or motor tasks. *J. Neurosci. Methods* 157, 317–329. doi: 10.1016/j.jneumeth.2006.04.022

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 December 2012; paper pending published: 23 March 2013; accepted: 23 June 2013; published online: 16 July 2013.

Citation: Mina F, Benquet P, Pasnicu A, Biraben A and Wendling F (2013) Modulation of epileptic activity by deep brain stimulation: a model-based study of frequency-dependent effects. *Front. Comput. Neurosci.* 7:94. doi: 10.3389/fncom.2013.00094

Copyright © 2013 Mina, Benquet, Pasnicu, Biraben and Wendling. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



On conductance-based neural field models

Dimitris A. Pinotsis*, Marco Leite and Karl J. Friston

The Wellcome Trust Centre for Neuroimaging, University College London, London, UK

Edited by:

Peter B. Graben,
Humboldt-Universität zu Berlin,
Germany

Reviewed by:

G. B. Ermentrout, University of
Pittsburgh, USA
Basabdatta S. Bhattacharya,
University of Lincoln, UK

*Correspondence:

Dimitris A. Pinotsis, The Wellcome
Trust Centre for Neuroimaging,
Institute of Neurology, 12 Queen
Square, WC1N 3BG, London, UK
e-mail: d.pinotsis@ucl.ac.uk

This technical note introduces a conductance-based neural field model that combines biologically realistic synaptic dynamics—based on transmembrane currents—with neural field equations, describing the propagation of spikes over the cortical surface. This model allows for fairly realistic inter- and intra-laminar intrinsic connections that underlie spatiotemporal neuronal dynamics. We focus on the response functions of expected neuronal states (such as depolarization) that generate observed electrophysiological signals (like LFP recordings and EEG). These response functions characterize the model's transfer functions and implicit spectral responses to (uncorrelated) input. Our main finding is that both the evoked responses (impulse response functions) and induced responses (transfer functions) show qualitative differences depending upon whether one uses a neural mass or field model. Furthermore, there are differences between the equivalent convolution and conductance models. Overall, all models reproduce a characteristic increase in frequency, when inhibition was increased by increasing the rate constants of inhibitory populations. However, convolution and conductance-based models showed qualitatively different changes in power, with convolution models showing decreases with increasing inhibition, while conductance models show the opposite effect. These differences suggest that conductance based field models may be important in empirical studies of cortical gain control or pharmacological manipulations.

Keywords: neural field theory, mean field modeling, electrophysiology, conductance based models, dynamic causal modeling

INTRODUCTION

This paper introduces a conductance-based neural field model that accounts for spatial variations in synaptic transmission among neural ensembles on the cortical surface. Our modeling draws from computational neuroscience, in which spiking models are described by population density dynamics. Generally, in these mean field approaches, population activity is expressed in terms of mean post-synaptic voltages and currents; however, conductance based models that consider the geometry and topography of neuronal interactions are relatively rare in the literature (Goldstein and Rall, 1974; Elias and Grossberg, 1975; Somers et al., 1995; Ermentrout, 1998); in other words, the spatiotemporal dynamics of conductance models are often simplified to neural mass approximations, such that the population density depends upon time only. In our model, we make the statistics of neuronal states a function of space, thereby characterizing mean spike rates as fluctuations propagating over horizontal cortical connections. This involves using wave equations to describe interactions between spatially extended neuronal populations, in terms of changes in the flow of post-synaptic currents, the history of pre-synaptic inputs and the action of certain neuromodulators.

Conductance-based models have a long history in mathematical neuroscience; for a detailed review, see (Tuckwell, 2005). Within the setting of dynamic causal modeling, a treatment of conductance-based models (that ignores the spatial distribution of sources over the cortex) can be found in (Marreiros et al., 2010) that was later applied to characterize synaptic function empirically (Moran et al., 2011b). These models regard a neuron

as an electrical circuit, where the membrane response follows the inflow or outflow of current through ionic channels. These channels are associated with conductances that depend upon electrochemical gradients across the membrane and the configuration of various ion channels and receptors. The standard kinetic model for *conductance* dynamics comprises two sorts of equations: (1) an equation for the rate of change of transmembrane potential as an aggregate current flux—consisting of Ohmic components and (2) equations for the channel conductances that depend upon pre-synaptic spiking and the proportion of open channels. This form of modeling necessarily entails non-linear terms, in which changes in post-synaptic potential involve the product of synaptic conductances and potential differences associated with different channel types. In other words, the equations of motion for neuronal states are necessarily non-linear and second-order (with respect to the hidden neuronal states), in accord with electromagnetic laws. This should be contrasted with the alternative approach to neural mass and mean field modeling based upon *convolution* operators. In these models, post-synaptic depolarization is modeled as a (generally linear) convolution of pre-synaptic spiking input. Crucially, this convolution can be formulated in terms of linear differential equations.

In short, the key distinction between conductance and convolution based models is that conductance based models have non-linear dynamics and, in principle, provide a degree of biological realism that can incorporate neuromodulatory and other conductance-specific physiological effects. Here, we use this basic form of model to describe the depolarization and conductances

of neural fields on the cortical sheet—and recast pre-synaptic spike rates as fluctuations obeying a wave equation that propagates along axon collaterals. We adopt a neural mass approach, where the input to each neuron of the population is the expected firing over all neurons around a point on a local cortical surface or patch. We thus obtain a conductance-based cortical field model linking population dynamics to synaptic neurotransmission. This paper focuses on the operational aspects of this model and its ability to reproduce typical cortical responses such as event-related potentials (ERPs) and cross-spectral densities.

The use of conductance based models to simulate large networks of neurons has enjoyed recent developments, involving both direct simulations of large numbers of neurons (which can be computationally expensive); e.g., (Izhikevich, 2004) and probabilistic approaches; e.g., (De Groff et al., 1993; Nykamp and Tranchina, 2000). Probabilistic approaches model the population density directly and bypass direct simulations of individual neurons. We follow a similar approach that exploits a neural mass approximation. This effectively replaces coupled Fokker-Planck equations describing population density dynamics, with equations of motion for expected neuronal states; that is, their first moments. These equations are formulated in terms of the mean of the population density over each neuronal state, as a function of space.

Recent work has considered the link between networks of stochastic neurons and neural field theory by using convolution models (with alpha type kernels) to characterize post-synaptic filtering: some studies have focused on the role of higher order correlations, starting from neural networks and obtaining neural field equations in a rigorous manner; e.g., (Buice et al., 2010; Touboul and Ermentrout, 2011), while others have considered a chain of individual fast spiking neurons (Rose and Hindmarsh, 1989), communicating through *spike* fields (Wilson et al., 2012). These authors focused on the complementary nature of spiking and neural field models and on eliminating the need to track individual spikes (Robinson and Kim, 2012). Our focus is on the behavior of neuronal populations, where conductance dynamics replace the convolution dynamics—and the input *rate* field is a function of both time and space. This allows us to integrate field models to pre-dict responses and therefore, in principle, use these models as generative or observation models of empirical data.

When modeling pre-synaptic firing rate, we use the approximation of (Robinson et al., 1997) that yields broad temporal pulses in response to a delta input. Crucially, we characterize the neuronal input as fluctuating mean spiking activity that conforms to a wave equation. Our model is non-linear in the neuronal states, as with single unit conductance models and the model of (Liley et al., 2002). This model entails a multiplicative non-linearity, involving membrane depolarization and pre-synaptic input and has successfully reproduced the known actions of anaesthetic agents on EEG spectra, see e.g., (Steyn-Ross et al., 2001, 2011; Liley et al., 2003; Bojak and Liley, 2005; Wilson et al., 2006). Our model is distinguished by the fact that it incorporates distinct cell types with different sets of conductances and local conduction effects. More specifically, it comprises three biologically plausible populations, each endowed with excitatory and inhibitory receptors. It focuses on the propagation of spike rate fluctuations over cortical patches and the effect

this spatiotemporal dynamics has on membrane dynamics gated by ionotropic receptor proteins. We consider laminar specific connections among two-dimensional populations (layers) that conform to canonical cortical microcircuitry. The parameterization of each population or layer involves a receptor complement based on findings in cellular neuroscience. However, this model incorporates lateral propagation of neuronal spiking activity that is parameterized through an intrinsic (local) conduction velocity.

This note comprises three sections. In the first, we review the mathematical formalism that underlies conductance based neural field models and introduce a generative model that accounts for both conductance effects on membrane dynamics and propagation of activity along intrinsic connections. In the second, we compare the behavior of this model with the corresponding behavior of convolution field models and consider the effect of changing model parameters. We also compare and contrast responses obtained by the neural mass reductions of these (conductance and convolution) models; in other words, models that consider dynamics over time only. Our focus here is on the effect that propagating fluctuations along horizontal (intrinsic) connections have on spatiotemporal dynamics. We conclude with a discussion of how the neural field model based upon first-order statistics—used in this paper—relates to formal treatments of population dynamics.

A CONDUCTANCE-BASED NEURAL FIELD MODEL

We consider a group of N_R interacting neuronal populations or layers. The collective dynamics (activity) of each population evolve according to a set of coupled differential equations that depend on some scalar quantities or neuronal states $q(x, t) \in \{v(x, t), g_k(x, t), \mu_k(x, t)\}$ that are continuous functions of the location on the cortical surface $x \in X$. These neuronal states include the transmembrane potential $v(x, t)$, a set of synaptic conductances $g_k(x, t)$ modeling distinct membrane channel types and the pre-synaptic input to which they are exposed $\mu_k(x, t)$.

The resulting populations can be viewed as a set of coupled RC circuits, where channels open in proportion to pre-synaptic input and close in proportion to the number already open. Changes in conductance produce changes in depolarization in proportion to the potential difference between transmembrane potential and a reversal potential v_k that depends upon the channel type. Open channels result in hyperpolarizing or depolarizing currents depending on whether the transmembrane potential is above or below the reversal potential. These currents are supplemented with exogenous current $u(x, t)$ to produce changes in the transmembrane potential (scaled by the membrane capacitance C). The first order moments or means of neuronal states at a location x on a cortical patch evolve according to the following system of differential equations:

$$\begin{aligned} C\dot{v}(x, t) &= \sum_k g_k (v_k - v(x, t)) \\ \dot{g}_k(x, t) &= \lambda_k (\mu_k(x, t) - g_k(x, t)) \\ \mu_k(x, t) &= \iint d(x - x', t - t') \sigma_k(v(x', t')) dt' dx' + u(x, t) \end{aligned} \quad (1)$$

where pre-synaptic input to point x arises from a spatiotemporal convolution of a sigmoid activation function of depolarizations

in other locations x' (in the past at time t') and $k = E, I$ denote excitatory and inhibitory synaptic conductances or inputs. This model assumes that each neuron senses all others, so that endogenous input is the expected firing of contributing locations summarized with a sigmoid function $\sigma_k(v)$ of their transmembrane potential. It is this function that accommodates the stochastic dispersion of neuronal states: see (Marreiros et al., 2010) for a detailed discussion. The rate constants λ_k characterize the response of each channel type to afferent input. Finally, $d(x, t)$ is a connectivity kernel that accommodates axonal propagation delays. It is this connectivity kernel that specifies the spatial aspects of the ensuing spatiotemporal dynamics.

A ubiquitous choice for the connectivity kernel (Wilson and Cowan, 1973; Jirsa and Haken, 1996) is based on the assumption that the number of synaptic connections between populations decays exponentially with some characteristic spatial scale c ; namely, $d(x, t) = ae^{-c|x|}\delta(t - |x|/s)$, where a scales connection strengths and s is the speed at which neuronal spikes propagate down connections. This assumption means that we can express the dynamics of the mean firing rates as (see e.g., Pinotsis et al., 2012):

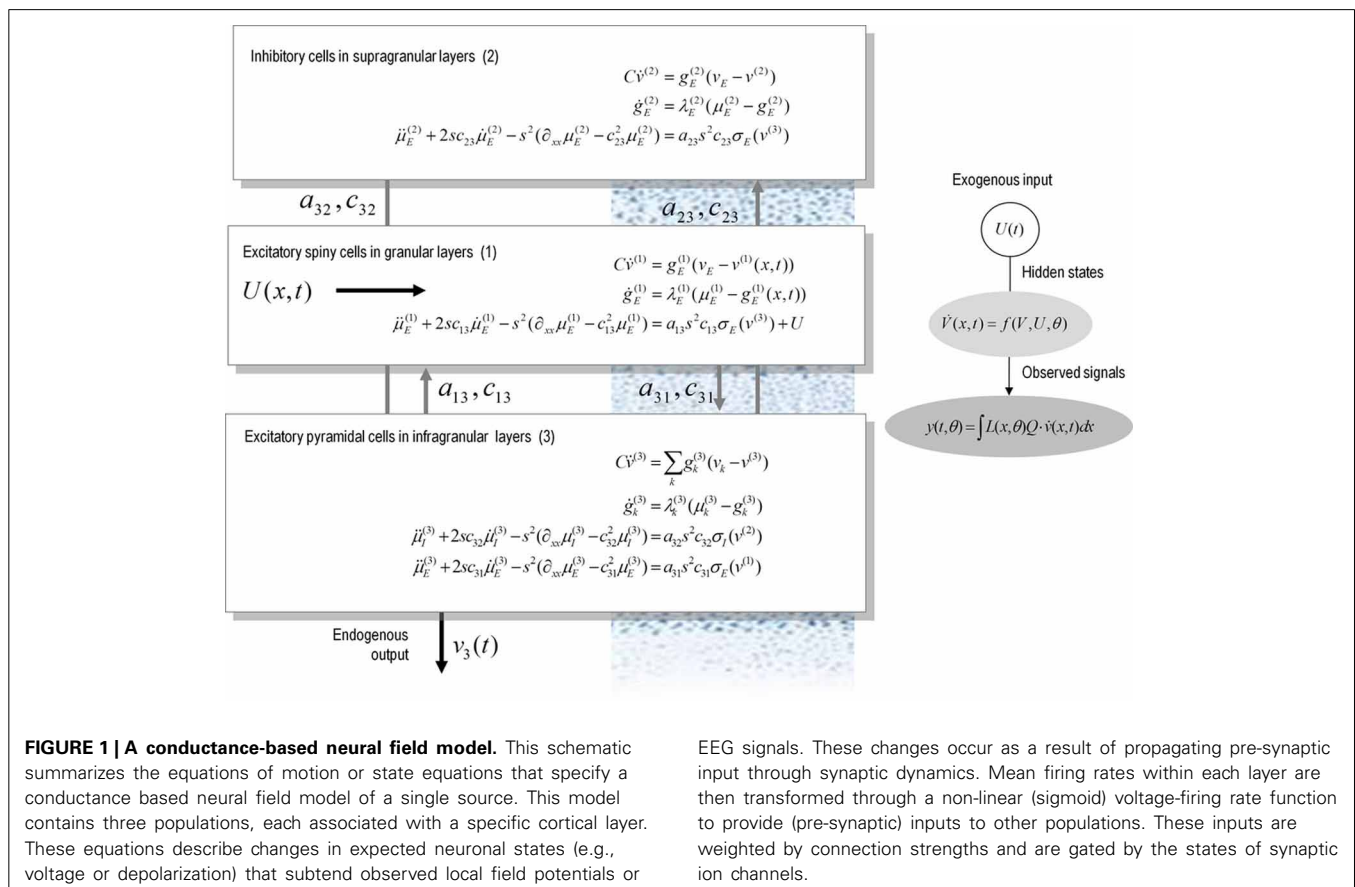
$$\begin{aligned} \ddot{\mu}_k(x, t) + 2sc\dot{\mu}_k(x, t) - s^2(\partial_{xx}\mu_k(x, t) - c^2\mu_k(x, t)) \\ = as^2c\sigma(v(x, t)) + u \end{aligned} \quad (2)$$

Combining Equations (1) and (2) gives us the equations of motion for all neuronal states:

$$\dot{q}(x, t) = \begin{bmatrix} \dot{v} \\ \dot{g}_k \\ \dot{\mu}_k \\ \dot{\mu}'_k \end{bmatrix} = \begin{bmatrix} \frac{1}{C} \sum_k g_k (v_k - v(x, t)) \\ \lambda_k (\mu_k(x, t) - g_k(x, t)) \\ \mu'_k(x, t) \\ -2sc\mu'_k(x, t) + s^2(\partial_{xx}\mu_k(x, t) - c^2\mu_k(x, t)) + as^2c\sigma(v(x, t)) + u \end{bmatrix} \quad (3)$$

for the quantitative purposes of this paper, we solve Equation (3) using a simple finite differences scheme for the second-order spatial derivatives.

Figure 1 illustrates the model for a spatially extended cortical source, which we will call a conductance-based neural field. In this model, the source consists of three layered populations; namely, spiny stellate cells, inhibitory interneurons and pyramidal cells. Each population is assigned to a cortical layer and is connected to other layers according to the principles of a typical cortical microcircuit (as described in e.g., Pinotsis et al., 2012). Each layer is equipped with neural states $q(x, t) \in \{v^{(i)}(x, t), g_k^{(i)}(x, t), \mu_k^{(i)}(x, t), \mu'_k(x, t)\}$, where the superscript i indexes different laminar populations—and the states evolve according to a system of coupled equations of the form of Equation (3). When this model is augmented with a mapping from source to sensor space, we obtain a generative model



of electrophysiological responses that can be used to infer the parameters of both synaptic kinetics—and intrinsic or lateral interactions, through the parameters of the connectivity kernel. Crucially, because of the biologically realistic construction of this model, one can examine the dependency of spatially extended dynamics of particular conductances and receptor subtypes.

RELATION TO CLASSICAL NEURAL FIELD MODELS

Equation (3) is an equation of motion, describing a neuronal field in terms of expected neuronal states. This sort of equation can accommodate both convolution and conductance based neural field models. Convolution neural field models involve kernels that are linear in the states; for example $q(x, t) \in \{v(x, t), v'(x, t), \mu(x, t), \mu'(x, t)\}$. These models can also be cast in a form similar to Equation (3):

$$\dot{q}(x, t) = \begin{bmatrix} \dot{v} \\ \dot{v}' \\ \dot{\mu} \\ \dot{\mu}' \end{bmatrix} = \begin{bmatrix} v' \\ -2\lambda v'(x, t) - \lambda^2 v(x, t) \\ + H\lambda\mu(x, t) \\ \mu'(x, t) \\ -2sc\mu'(x, t) + s^2(\partial_{xx}\mu(x, t) \\ - c^2\mu(x, t)) + as^2c\sigma(v(x, t)) + u \end{bmatrix} \quad (4a)$$

Indeed, this equation can be rewritten as

$$\begin{aligned} \ddot{v}(x, t) + 2\lambda\dot{v}'(x, t) + \lambda^2 v(x, t) &= H\lambda\mu(x, t) \\ \ddot{\mu}(x, t) + 2sc\dot{\mu}'(x, t) - s^2(\partial_{xx}\mu(x, t) - c^2\mu(x, t)) &= as^2c\sigma(v(x, t)) + u \end{aligned} \quad (4b)$$

These equations describe neural fields with constant coefficients in homogeneous media; see e.g., Pinotsis and Friston, 2011; Pinotsis et al., 2012, 2013. In a previous paper, we introduced a neural field model involving the three laminar populations depicted in **Figure 1**, which we called a Jansen and Rit neural field model. This model is similar to the classical Wilson and Cowan or Amari models (Wilson and Cowan, 1972; Amari, 1977). The model in Equation (4) assumes an alpha-type synaptic convolution kernel. This is simply the Green's function associated with a linear filtering of pre-synaptic input to produce changes in depolarization. In these mean field models, passive membrane dynamics and dendritic effects are summarized by lumped parameters (λ and H in the above equations) that model the rate that depolarization increases to a maximum and synaptic efficacy (or maximum post-synaptic potential), respectively. However, this sort of description neglects the timescales of synaptic currents that are implicit in conductance based models: in Equations (3) these timescales are characterized in terms of the rate constants λ and C ; namely, channel response and membrane capacitance.

The crucial difference between these (linear and non-linear) parameterizations is that in the conductance models, the parameters characterize the response of each population to distinct excitatory and inhibitory inputs: in other words, there is a set of synaptic rate constants (each corresponding to a distinct channel) associated with each population. The corresponding dynamics are defined over timescales that result from the parameters used

and the non-linear interaction between membrane potential and conductance. These timescales may be crucial in pharmacological manipulations that selectively affect one sort of current in a receptor specific fashion. This means that conductance-based models may be more appropriate candidates to study synaptic function at the level of specific neurotransmitter systems (Faulkner et al., 2009; Moran et al., 2011a).

SIMULATIONS

In the following, we focus on simulated responses generated by convolution and conductance variants of neural mass and field models—where these two variants incorporate fundamentally different descriptions of post-synaptic filtering. We investigate the dependence of simulated responses on model parameters with neurobiological or pharmacological significance. Specifically, we examine: (1) the effects of changing synaptic parameters and (2) the qualitative differences in the behavior of convolution and conductance based models. In this technical note, we focus only on the phenomenology of the models in the domains of the parameter space that are dynamically stable.

We generated synthetic electrophysiological responses by integrating equations (3) or (4) from their fixed points and characterized the responses to external (excitatory) impulses to spiny stellate cells, in the time and frequency domain. The spectral responses correspond to the model's transfer function. Electrophysiological signals (LFP or M/EEG data) were simulated by passing neuronal responses through a lead field that varies with location on the cortical patch. The resulting responses in sensor space (see **Figures 5–7**) are given by a mixture of currents flowing in and out of pyramidal cells in **Figure 1**:

$$y(t, \theta) = \int L(x, \theta) Q \cdot \dot{v}(x, t) dx \quad (5)$$

In this equation, $Q \subset \theta$ is a vector of coefficients that weight the relative contributions of different populations to the observed signal and $L(x, \theta)$ is the lead field. This depends upon parameters θ and we assume it is a Gaussian function of location—as in previous models of LFP or MEG recordings; see (Pinotsis et al., 2012). This equation is analogous to the usual (electromagnetic) gain matrix for equivalent current dipoles. We assume here that these dipoles are created by pyramidal cells whose current is the primary source of an LFP signal. With spatially extended sources (patches), this equation integrates out the dependence on the source locations within a patch and provides a time series for each sensor.

We modeled a cortical source (approximated with 11 grid points) and used the model equations (see **Figure 1**) to generate evoked responses (impulse response functions) and associated transfer functions (their Fourier transform). The parameters of this model are provided in **Table 1**. The results reported below were chosen to illustrate key behaviors in terms of ERP (impulse response) and frequency responses (transfer functions) in sensor space, following changes in parameter values. We compare and contrast results from the two classes of models (conductance and convolution models). We also consider the corresponding result for their mass variants, which use the same equations but assume

Table 1 | Parameters of conductance-based neural field and mass models.

Parameter	Physiological interpretation	Value
g_L	Leakage conductance	1
$\alpha_{13}, \alpha_{23}, \alpha_{31}, \alpha_{32}$	Amplitude of intrinsic connectivity kernels	(1/10, 1, 1/2, 1)*3/10 (field) 1/2, 1, 1/2, 1 (mass)
c_{ij}	Intrinsic connectivity decay constant	1 (mm ⁻¹)
v_L, v_E, v_I	Reversal potential	-70, 60, -90 (mV)
v_R	Threshold potential	-40 (mV)
C	Membrane capacitance	8 (pFnS ⁻¹)
s	Conduction speed	0.3 m/s
$\lambda, \tilde{\lambda}$	Post-synaptic rate constants	1/4, 1/16 (ms ⁻¹)
ℓ	Radius of cortical patch	7 (mm)

that all neurons of a population are located at (approximately) the same point.

The resulting mass models include the well-known Jansen and Rit mass model, see (David and Friston, 2003) and the simplified Morris-Lecar type model (that neglects fast voltage-dependent conductances) introduced in (Marreiros et al., 2010). This conductance-based model is based on the Rall and Goldstein equations (Goldstein and Rall, 1974) and is formally related to Ermentrout's (Ermentrout, 1998) reduction of the (Somers et al., 1995) model. Mass models have often been used to characterize pharmacological manipulations and the action of sedative agents (Traub et al., 1999; Liley et al., 2003; Bojak and Liley, 2005; Moran et al., 2008; Hutt and Longtin, 2010; Steyn-Ross et al., 2011). This usually entails assuming that a neurotransmitter manipulation changes a particular parameter, whose effects are quantified using a contribution or structural stability analysis, where structural stability refers to how much the system changes with perturbations to the parameters.

Our aim here was to illustrate changes in responses with changes in the parameters of the convolution and conductance field models. A range of anaesthetics has been shown to increase inhibitory neurotransmission. This effect has been attributed to allosteric activators that sensitize GABA_A receptors. In the context of our models, these effects correspond to an increase of the strength of inhibitory input to pyramidal cells α_{32} . We here focus on spectral responses in the alpha and beta range, as this is the range of interest for many applications involving drug effects.

We first consider generic differences in non-linear processes mediated by conductance and convolution based models. To do this, we integrated the corresponding equations for (impulse) inputs of different amplitudes and plotted temporal responses resulting from fixed point perturbations. Linear models are insensitive to the amplitude of the input, in the sense that the impulse responses scale linearly with amplitude. Our interest here was in departures from linearity—such as saturation—that belie the non-linear aspects of the models. **Figure 2** shows the responses of the mass models to an impulse delivered to stellate cells. Note that these responses have been renormalized with respect to the

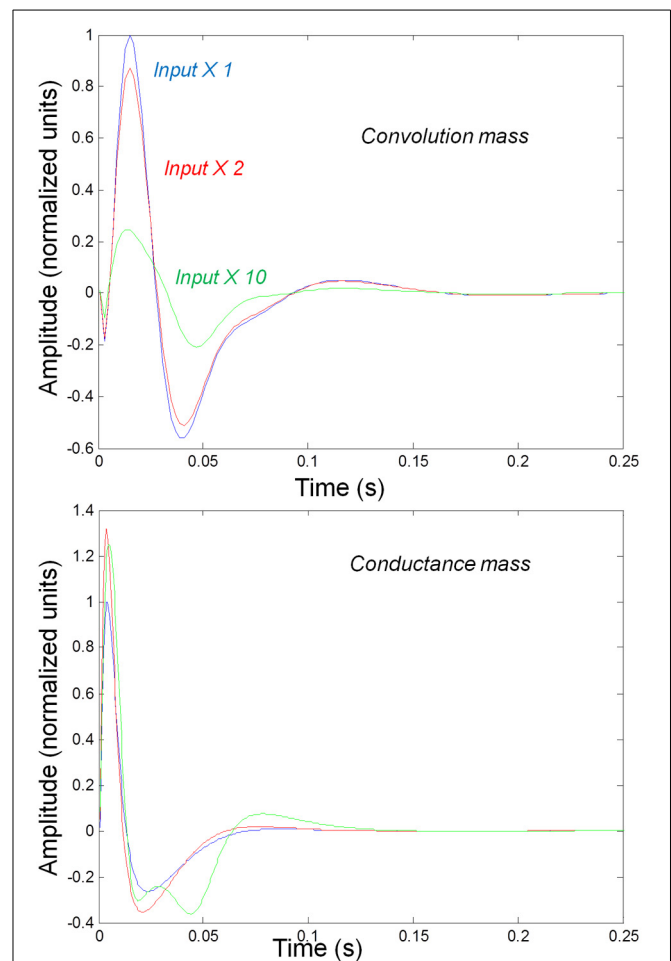


FIGURE 2 | Responses to impulses of different amplitudes for convolution (top) and conductance (bottom) based neural mass models. The responses are normalized with respect to the amplitude of each input. The blue lines illustrate responses to small perturbations. The red lines illustrate responses to intermediate sized inputs, where conductance based models show an augmented response, due to their non-linearity. The green lines show responses for larger inputs, where the saturation effects due to the sigmoid activation function are evident.

amplitude of each input. The red (green) curves depict responses to double (ten times) the input reported by the blue curves. We used the same parameters for both models: see **Table 1**; where additional parameters for the Jansen and Rit model are provided in **Table 2** below.

It can be seen that there are marked differences between the model responses. The top panel depicts the response of the convolution mass model and the lower panel shows the equivalent results for the conductance model. One can see that large inputs produce substantial sub-additive saturation effects (blue vs. green lines in **Figure 2**): for the convolution model, increasing the input amplitude produces a sub additive increase in response amplitude; whereas for the conductance model, the non-linearities produce an inverted U relationship between the amplitude of the response, relative to the input. In summary, the form of the input-output amplitude relationship differs quantitatively for the

Table 2 | Parameters of convolution-based neural field and mass models.

Parameter	Physiological interpretation	Prior mean
H_E, H_I	Maximum post-synaptic depolarizations	8 (mV)
$\alpha_{13}, \alpha_{23}, \alpha_{31}, \alpha_{32}$	Amplitude of intrinsic connectivity kernels	(1/2, 1, 1/2, 1)*3/10 (field) 1, 4/5, 1/4, 1 (mass)

Other parameters as in **Table 1**.

conductance (inverted U) and convolution (decreasing) models (see **Figure 2**).

Figure 3 shows the impulse responses of the field models described by Equations (3) and (4). Here we observe sub-additive saturation effects that are similar to the responses of the convolution mass model—with relatively stronger attenuation of the response amplitude than the mass model even for intermediate input amplitudes.

We next characterized the spectral responses of convolution and conductance-based neural fields and their mass variants. It should be noted that this analysis is purely phenomenological and a complete bifurcation analysis will be presented elsewhere. Here, we focus on transfer functions associated with the models. These are shown in subsequent figures for a range of physiological parameters. The transfer functions can be regarded as the spectral density that would be seen if the field and mass models were driven by independent fluctuations. It is interesting that—for the biologically plausible parameter values we use—both field and mass models exhibit alpha peaks (as opposed to a $1/f$ scale invariant form) that are typical of neural field models (Nunez, 1995; Robinson et al., 2001; Liley et al., 2002). Note that the transfer function characterizations used below assume a linearization around the fixed point and therefore do not capture the non-linear behavior of the models.

We varied the inhibitory intrinsic connectivity, a_{32} and excitatory time constant, $1/\lambda$, of the inhibitory populations between 10 and 36% and between 10 and 270%, respectively, of the values in **Tables 1, 2** (this corresponds to a log-scaling of between minus two and minus one and minus one and plus one, respectively). We denote these new values by \tilde{a}_{32} and $1/\tilde{\lambda}$, respectively. The transfer functions for the neural mass variants of the convolution and conductance models are shown in **Figures 4, 5**, respectively. The images in subsequent figures report the peak frequency of the spectral and response as a function of the two model parameters (the peak frequency corresponds to maximum system response). Exemplar transfer functions for selected parameter value pairs are shown as functions of frequency. We focus on spectral responses produced by fixed point perturbations; where lack of convergence to a fixed point is encoded by dark blue regions in the images.

In mass models, the peak frequencies of the spectra reflect the alpha and beta activity that these models are known to produce. It is interesting that the most parsimonious among all models considered (the convolution mass model) seems to support the widest range of simulated peak frequencies; this is, however, not a conclusive result as it is heavily dependent on the

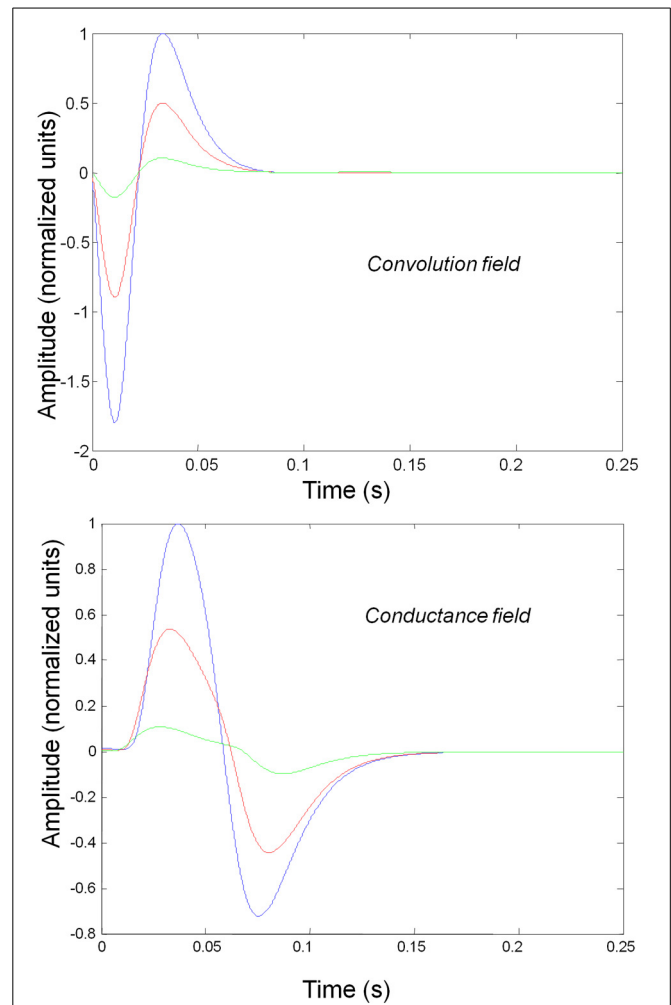


FIGURE 3 | Impulse response of conductance and convolution field models to inputs of various amplitudes distinguished by different colours as in Figure 2. The system's flow is generated by Equations (3) and (4a) and the model parameters are given in **Tables 1, 2. Non-linear effects are more pronounced—with attenuation of the response amplitude, even for intermediate input amplitudes.**

particular parameterization chosen—a fuller exploration of the parameter space will be the focus of future work. A common pattern observed in all models is an increase of peak frequencies with smaller time constants of the inhibitory populations. In other words, as the strength of inhibition increases, activity becomes progressively faster (power shifts to higher frequencies). Conversely, convolution and conductance mass models showed quantitatively different changes in power, with convolution models showing decreases with increasing inhibition, while conductance models show the opposite effect. The transfer functions for the corresponding field models are shown in **Figures 6, 7**. Here, one observes that responses of the convolution model are similar to those obtained from the mass models above—dominated by changes in the rate $\tilde{\lambda}$ parameter with less sensitivity to changes in the connectivity parameter. Again, we see a common increase in frequency as the inhibitory rate parameter is increased (or

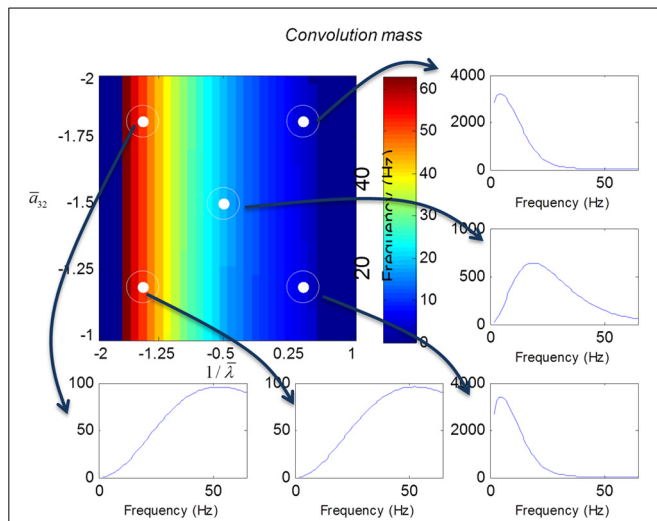


FIGURE 4 | Transfer functions associated with a convolution mass model when changing the excitatory time constant and the connection driving the pyramidal cells over a log-scaling range of $(-2, 1) \times (-2, -1)$ (from top to bottom and left to right). The image format summarizes the transfer function in terms of its peak frequency. Transfer functions can be regarded as the spectral response that would be seen if the model was driven by independent (white) fluctuations. They are also the Fourier transform of the impulse response functions of the previous figures.

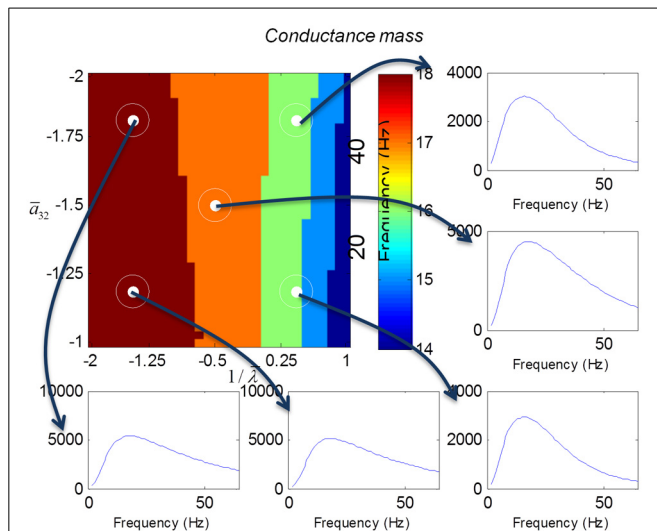


FIGURE 5 | This figure shows the transfer functions of a cortical source described by a conductance mass model. Here, the intrinsic connectivity and excitatory time constant are changed as in Figure 4. Note the alpha and beta peaks that are typical of these models.

the time constant is decreased)—and the opposite effects under convolution and conductance models, in terms of power.

The above illustrations of system's predictions assume that spectral responses result from fixed point perturbations. For conductance models, a change in the parameters changes both the expansion point and the system's flow (provided the flow is non-zero). **Figure 8** shows the dependence of the conductance model's

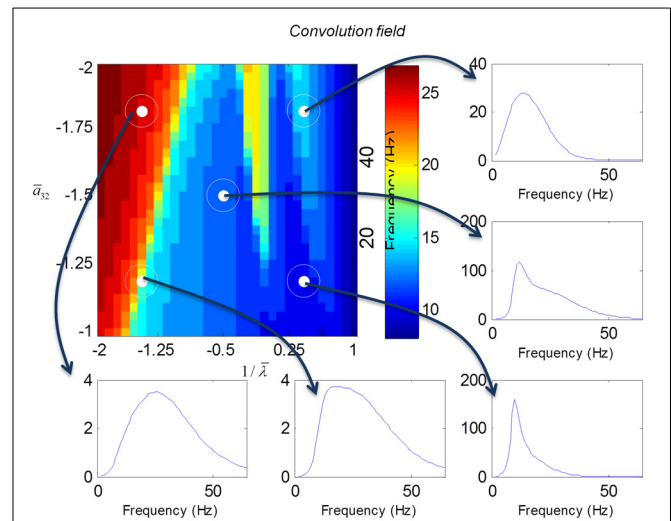


FIGURE 6 | Transfer functions associated with a convolution field model. These are equivalent to the transfer functions shown in Figure 4, where we now model spatial propagation effects with a wave equation. Here, one observes the characteristic increase in frequency when the time constants decrease.

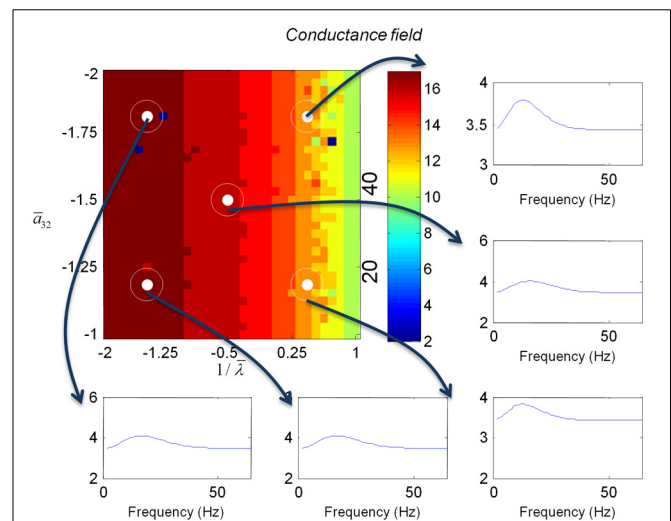
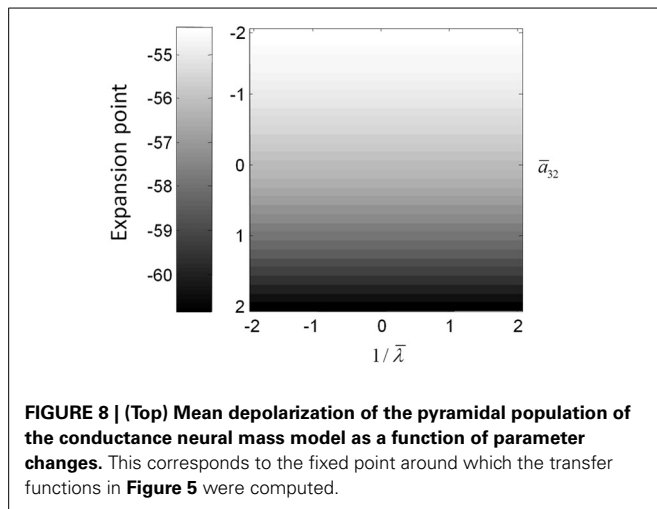


FIGURE 7 | This figure shows the changes in the transfer function of a conductance field model. This is the equivalent to the results for the mass model in Figure 5, where we now include spatial propagation effects.

fixed points on parameter perturbations. The model parameterization used here renders the expansion point relatively insensitive to changes in the synaptic time constant. **Figure 8** shows the results for the conductance mass model; results for its field variant were very similar.

DISCUSSION

In this paper, we have introduced a conductance based neural field model that combines biologically realistic synaptic dynamics—based explicitly on transmembrane currents—with neural field equations, describing the propagation of spikes over the cortical



surface. This model allows for fairly realistic inter- and intra-laminar intrinsic connections over a spatially extended cortical surface that give rise to neuronal dynamics. We have focused on the time evolution of expected neuronal states that underlie observed electrophysiological signals (such as LFP recordings and EEG). This time evolution characterizes the model's transfer functions and implicit spectral responses to uncorrelated input. Our main finding is that both the evoked responses (impulse response functions) and induced responses (transfer functions) show quantitative differences depending upon whether one uses a neural mass or field model. It is interesting that field models do not always produce a wider range of spectral responses for equivalent changes in their parameters (despite their greater degrees of freedom, compare **Figures 4, 6**). Similarly, conductance field models do not necessarily show a greater sensitivity to small parameter perturbations in comparison with their convolution counterparts (that appear more parsimonious). Although, overall, all models reproduce the characteristic increase in frequency when the rate constants of inhibitory populations increase, the precise frequency dependency depends sensitively on model type. The choice of the appropriate model might therefore depend on the particular research question at hand: for example, whether the focus is on topographic as opposed to intrinsic neurotransmitter properties or drug effects etc. This choice may also be informed by previous applications, where similar models have already proven useful along with the particular modality considered (see also the discussion in Pinotsis et al., 2013). Conductance field models may be useful in applications such as dynamic causal modeling, that try to quantify changes in gain control in cortical circuits or explain pharmacological manipulations.

The models considered in this paper deal only with the expected values (means) of neuronal states. This contrasts with higher order field treatments that would consider not just fluctuations in the means or first-order statistics of population dynamics but also higher-order statistics—such as the covariance among different neuronal states within a population or ensemble. In principle, it is relatively easy to extend the formalism described in this paper to cover the dynamics of both means and covariances using the Laplace approximation (a.k.a. the method of moments).

In these generalizations, one considers the distribution over the neuronal states of a given population to have a Gaussian form $\mathcal{N}(q(x, t), \Sigma(x, t))$. Crucially, the equations of motion now pertain to both the expectations and the covariances (Marreiros et al., 2010). The interesting challenge for the neural field variants of these Laplace models is that the covariances have a spatial dimension and, essentially, become spatial covariance functions (cf., Gaussian processes or random fields). The implicit covariance functions of space have a smoothness that is determined by the intrinsic connectivity kernels and the dynamics of the first order statistics.

These equations of motion for the means and covariances reduce to the neural fields considered in this paper when the off-diagonal terms of the covariance matrix $\Sigma(x, t)$ are zero. In this special case, the dynamics of the means and covariances are uncoupled and one can assume a fixed covariance (as in Equations 3 and 4): see Marreiros et al. (2010) for details. More generally, full mean field treatments can provide higher order corrections to *stochastic* neural field models and offer an alternative description of the motion of their sufficient statistics, cf., (Buice et al., 2010; Touboul and Ermentrout, 2011).

The conductance based model introduced in this paper describes the propagation of spikes over the cortical surface and how their effects on post-synaptic responses can be modeled in a channel-specific fashion. In principle—as illustrated in the transfer function analyses—changes in the balance of cortical excitation and inhibition may be modeled more appropriately with conductance based models, relative to classical convolution based models. In particular, these sorts of neural field models characterize the geometry and spatiotemporal dynamics that are supported by intrinsic or lateral interactions on the cortical surface and, implicitly, pharmacological effects on these interactions (such as anaesthetic administration). In the next phase of this work, we will use the conductance based field model described here as an observation or generative model of empirical electrophysiological responses to establish its validity, within the setting of dynamic causal modeling.

ACKNOWLEDGMENTS

The Wellcome trust funded this work. We thank Professor Viktor Jirsa and Dr Biswa Sengupta for useful discussions.

REFERENCES

- Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87. doi: 10.1007/BF00337259
- Bojak, I., and Liley, D. T. (2005). Modeling the effects of anesthesia on the electroencephalogram. *Phys. Rev. E* 71, 041902. doi: 10.1103/PhysRevE.71.041902
- Buice, M. A., Cowan, J. D., and Chow, C. C. (2010). Systematic fluctuation expansion for neural network activity equations. *Neural Comput.* 22, 377–426. doi: 10.1162/neco.2009.02-09-960
- David, O., and Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage* 20, 1743–1755. doi: 10.1016/j.neuroimage.2003.07.015
- De Groot, D., Neelakanta, P. S., Sudhakar, R., and Aalo, V. (1993). Stochastic aspects of neuronal dynamics: Fokker-Planck approach. *Biol. Cybern.* 69, 155–164. doi: 10.1007/BF00226199
- Ellias, S. A., and Grossberg, S. (1975). Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks. *Biol. Cybern.* 20, 69–98. doi: 10.1007/BF00327046

- Ermentrout, B. (1998). Neural networks as spatio-temporal pattern-forming systems. *Rep. Prog. Phys.* 61, 353. doi: 10.1088/0034-4885/61/4/002
- Faulkner, H. J., Traub, R. D., and Whittington, M. A. (2009). Disruption of synchronous gamma oscillations in the rat hippocampal slice: a common mechanism of anaesthetic drug action. *Br. J. Pharmacol.* 125, 483–492. doi: 10.1038/sj.bjp.0702113
- Goldstein, S. S., and Rall, W. (1974). Changes of action potential shape and velocity for changing core conductor geometry. *Biophys. J.* 14, 731–757. doi: 10.1016/S0006-3495(74)85947-3
- Hutt, A., and Longtin, A. (2010). Effects of the anesthetic agent propofol on neural populations. *Cogn. Neurodyn.* 4, 37–59. doi: 10.1007/s11571-009-9092-2
- Izhikevich, E. M. (2004). Which model to use for cortical spiking neurons? *IEEE Trans. Neural Netw.* 15, 1063–1070. doi: 10.1109/TNN.2004.832719
- Jirsa, V. K., and Haken, H. (1996). Derivation of a field equation of brain activity. *J. Biol. Phys.* 22, 101–112. doi: 10.1007/BF00954458
- Liley, D. T., Cadusch, P. J., and Dafilis, M. P. (2002). A spatially continuous mean field theory of electrocortical activity. *Network* 13, 67–113. doi: 10.1088/0954-898X/13/1/303
- Liley, D. T., Cadusch, P. J., Gray, M., and Nathan, P. J. (2003). Drug-induced modification of the system properties associated with spontaneous human electroencephalographic activity. *Phys. Rev. E* 68, 051906. doi: 10.1103/PhysRevE.68.051906
- Marreiros, A. C., Kiebel, S. J., and Friston, K. J. (2010). A dynamic causal model study of neuronal population dynamics. *Neuroimage* 51, 91–101. doi: 10.1016/j.neuroimage.2010.01.098
- Moran, R. J., Jung, F., Kumagai, T., Endepols, H., Graf, R., Dolan, R. J., et al. (2011a). Dynamic causal models and physiological inference: a validation study using isoflurane anaesthesia in rodents. *PLoS ONE* 6:e22790. doi: 10.1371/journal.pone.0022790
- Moran, R. J., Symmonds, M., Stephan, K. E., Friston, K. J., and Dolan, R. J. (2011b). An *in vivo* assay of synaptic function mediating human cognition. *Curr. Biol.* 21, 1320–1325. doi: 10.1016/j.cub.2011.06.053
- Moran, R. J., Stephan, K. E., Kiebel, S. J., Rombach, N., O'Connor, W. T., Murphy, K. J., et al. (2008). Bayesian estimation of synaptic physiology from the spectral responses of neural masses. *Neuroimage* 42, 272–284. doi: 10.1016/j.neuroimage.2008.01.025
- Nunez, P. L. (1995). *Neocortical Dynamics and Human EEG Rhythms*. Oxford: Oxford University Press.
- Nykamp, D. Q., and Tranchina, D. (2000). A population density approach that facilitates large-scale modeling of neural networks: analysis and an application to orientation tuning. *J. Comput. Neurosci.* 8, 19–50. doi: 10.1023/A:1008912914816
- Pinotsis, D. A., and Friston, K. J. (2011). Neural fields, spectral responses and lateral connections. *Neuroimage* 55, 39–48. doi: 10.1016/j.neuroimage.2010.11.081
- Pinotsis, D. A., Moran, R. J., and Friston, K. J. (2012). Dynamic causal modeling with neural fields. *Neuroimage* 59, 1261–1274. doi: 10.1016/j.neuroimage.2011.08.020
- Pinotsis, D. A., Schwarzkopf, D. S., Litvak, V., Rees, G., Barnes, G., and Friston, K. J. (2013). Dynamic causal modelling of lateral interactions in the visual cortex. *Neuroimage* 66, 563–576. doi: 10.1016/j.neuroimage.2012.10.078
- Robinson, P. A., and Kim, J. W. (2012). Spike, rate, field, and hybrid methods for treating neuronal dynamics and interactions. *J. Neurosci. Methods* 205, 283–294. doi: 10.1016/j.jneumeth.2012.01.018
- Robinson, P. A., Loxley, P. N., O'Connor, S. C., and Rennie, C. J. (2001). Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. *Phys. Rev. E* 63, 041909–1–041909-13. doi: 10.1103/PhysRevE.63.041909
- Robinson, P. A., Rennie, C. J., and Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* 56, 826–840. doi: 10.1103/PhysRevE.56.826
- Rose, R. M., and Hindmarsh, J. L. (1989). The assembly of ionic currents in a thalamic neuron I. The three-dimensional model. *Proc. R. Soc. Lond. B Biol. Sci.* 237, 267–288. doi: 10.1098/rspb.1989.0049
- Somers, D. C., Nelson, S. B., and Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.* 15, 5448–5465.
- Steyn-Ross, D. A., Steyn-Ross, M. L., Sleigh, J. W., and Wilson, M. T. (2011). Progress in modeling EEG Effects of General Anesthesia: Biphasic Response and Hysteresis. *Sleep Anesth.* 167–194. doi: 10.1007/978-1-4614-0173-5_8
- Steyn-Ross, M. L., Steyn-Ross, D. A., Sleigh, J. W., and Wilcocks, L. C. (2001). Toward a theory of the general-anesthetic-induced phase transition of the cerebral cortex. I. A thermodynamics analogy. *Phys. Rev. E* 64, 011917. doi: 10.1103/PhysRevE.64.011917
- Touboul, J. D., and Ermentrout, G. B. (2011). Finite-size and correlation-induced effects in mean-field dynamics. *J. Comput. Neurosci.* 31, 453–484. doi: 10.1007/s10827-011-0320-5
- Traub, R. D., Jefferys, J. G., and Whittington, M. A. (1999). *Fast Oscillations in Cortical Circuits*. Cambridge, MA: MIT press.
- Tuckwell, H. C. (2005). *Introduction to Theoretical Neurobiology: Volume 2, Nonlinear and Stochastic Theories*. Cambridge, MA: Cambridge University Press.
- Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24. doi: 10.1016/S0006-3495(72)86068-5
- Wilson, H. R., and Cowan, J. D. (1973). Mathematical theory of functional dynamics of cortical and thalamic nervous-tissue. *Kybernetik* 13, 55–80. doi: 10.1007/BF00288786
- Wilson, M. T., Robinson, P. A., O'Neill, B., and Steyn-Ross, D. A. (2012). Complementarity of spike-and rate-based dynamics of neural systems. *PLoS Comput. Biol.* 8:e1002560. doi: 10.1371/journal.pcbi.1002560
- Wilson, M. T., Sleigh, J. W., Steyn-Ross, D. A., and Steyn-Ross, M. L. (2006). General anesthetic-induced seizures can be explained by a mean-field model of cortical dynamics. *Anesthesiology* 104, 588–593. doi: 10.1097/00000542-200603000-00026

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 August 2013; accepted: 21 October 2013; published online: 12 November 2013.

Citation: Pinotsis DA, Leite M and Friston KJ (2013) On conductance-based neural field models. *Front. Comput. Neurosci.* 7:158. doi: 10.3389/fncom.2013.00158

This article was submitted to the journal *Frontiers in Computational Neuroscience*. Copyright © 2013 Pinotsis, Leite and Friston. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.