Theoretical predictions for spatial covariance of the electroencephalographic signal during the anesthetic-induced phase transition: Increased correlation length and emergence of spatial self-organization

Moira L. Steyn-Ross,[1,*] D. A. Steyn-Ross,[1] J. W. Sleigh,[2] and D. R. Whiting[1]
[1]*Department of Physics and Electronic Engineering, Private Bag 3105, University of Waikato, Hamilton, New Zealand*
[2]*Department of Anaesthetics, Waikato Hospital, Hamilton, New Zealand*
(Received 5 December 2002; published 7 August 2003)

In a recent series of papers, the authors have developed a stochastic theory to describe the electrical response of a spatially homogeneous cerebral cortex to infusion of a general anesthetic agent. We showed that by modeling the GABAergic (propofol-like) drug effect as a prolongation of the inhibitory postsynaptic impulse response, we obtain a prediction that there will be a hysteretically separated pair of first-order phase transitions in the population-average excitatory soma voltage, the first occurring at the point of induction of unconsciousness, and the second at the point of emergence from unconsciousness. In the present paper we generalize our earlier ''zero-dimensional'' homogeneous cortex to a one-dimensional (1D) line of cortical ''mass,'' thus allowing for the possibility of spatial inhomogeneities in neural activity. Following the spirit of our earlier adiabatic (''slow membrane'') philosophy, we impose a spatioadiabatic approximation that permits us to compute analytic expressions for changes in EEG (electroencephalographic) correlation length and EEG spatial covariance as a function of anesthetic effect. We establish that the correlation length of the EEG fluctuations is expected to *increase* at the approach to the transition points, and this finding is consistent with both the homogeneous-cortex prediction of increased correlation time (''critical slowing down'') near transition, and the recent, comprehensive anesthetic study by John *et al.* [Conscious. Cogn. **10**, 165 (2001)] reporting an increase in EEG coherence near the points of loss and recovery of consciousness. In addition, we find that if the long-range (corticocortical) excitatory-to-inhibitory connectivity in the 1D cortex is stronger than the long-range excitatory-to-excitatory connectivity, then the spatioadiabatic system can organize itself into large-amplitude spatial patterns (''dissipative structures'') consisting of giant stationary quasiperiodic voltage fluctuations distributed along the cortical rod.

DOI: 10.1103/PhysRevE.68.021902 PACS number(s): 87.19.La, 05.10.Gg, 05.70.Fh

I. INTRODUCTION

In a recent series of papers [1–3], the authors have been developing a theoretical model that explores the electrodynamic behavior of a spatially homogeneous cerebral cortex subject to the action of a GABAergic general-anesthetic agent. We model the cortex as a collection of interconnected *macrocolumns*: aggregates of cooperating neurons over whose spatial extent it is assumed possible to replace individual neuron properties with population averages. This is the mean-field philosophy. In the present paper we investigate the implications of retaining the mean-field requirement within the macrocolumn, but allowing intermacrocolumn spatial variability to develop along a one-dimensional (1D) rod of cortical mass.

Our theoretical framework is a set of coupled stochastic differential equations (SDEs) for h_e and h_i, the macrocolumn-averaged transmembrane soma voltage of the excitatory and inhibitory neurons. The h_e state variable is the key observable, as its fluctuations are assumed to be the source of the scalp-measured electroencephalographic (EEG) signal [1]. The driving force for these fluctuations is the nonspecific subcortical input p_{jk} $(j,k \in \{e,i\})$, modeled as four independent, Gaussian-distributed random processes.

As described by Franks and Lieb [4], the electrical effect of GABAergic anesthetic agents is to *prolong* the duration of the inhibitory postsynaptic potential (IPSP) event generated in response to an impulsive influx of GABA (γ-amino butyric acid) neurotransmitter arriving at the $GABA_A$ receptors of the postsynaptic neuron. We assume that this prolongation of the inhibitory impulse response depends on anesthetic concentration, and model this in terms of λ_{GABA}, a dimensionless number that scales the IPSP rate constant γ_i [see Eq. (2.9) below] appearing in the equations for inhibitory presynaptic input into the excitatory [Eq. (2.2c)] and inhibitory [Eq. (2.2d)] neural populations.

Central to our approach is the adiabatic or ''slow-membrane'' approximation. In this limit, we assume that the time constants associated with the excitatory and inhibitory membrane ''capacitors'' are very much longer than the time constants for the electrical activity generated by the various neurotransmitter input events. That is, the time course for the h_e and h_i membrane voltages is taken to be much slower than the time course of the incoming excitatory and inhibitory postsynaptic potentials. This enables a considerable simplification of the macrocolumn equations of motion, since the synaptic input voltages, having rapidly equilibrated, can be replaced by their steady-state values.

Also central to our theory is the notion that the cortex always operates close to a homogeneous equilibrium state

*Electronic address: msr@waikato.ac.nz; URL: phys.waikato.ac.nz/cortex

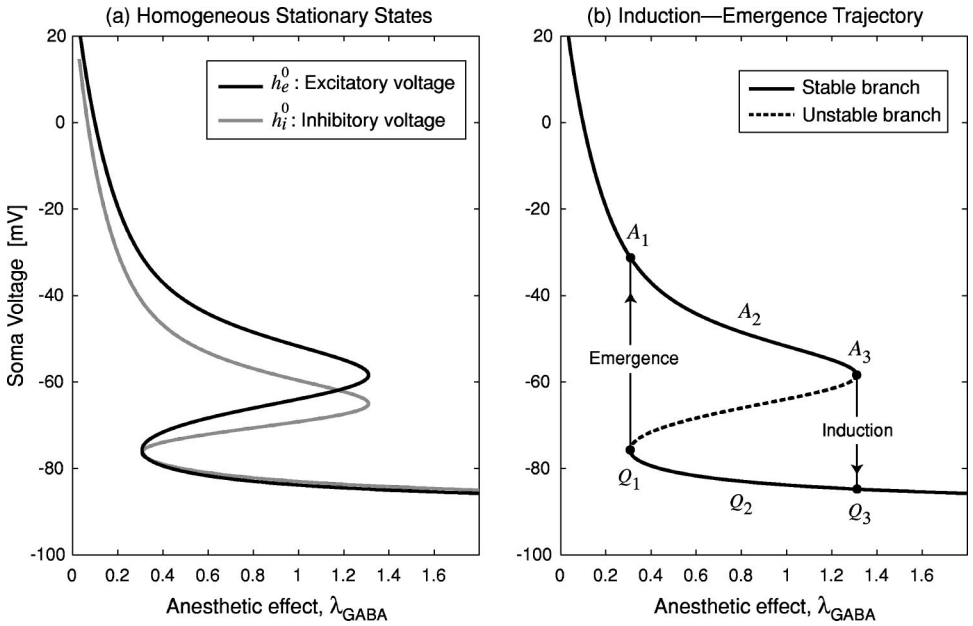

FIG. 1. (a) Predicted locus of excitatory (h_e, black curve) and inhibitory (h_i, gray curve) equilibrium voltage states for the homogeneous cortex as a function of anesthetic effect λ_{GABA}. (b) Trajectory of steady-state excitatory soma voltage for *induction* of anesthetic unconsciousness follows path $A_1 \rightarrow A_3 \rightarrow Q_3$; the return trajectory for *emergence* into consciousness follows the hysteretically distinct path $Q_3 \rightarrow Q_1 \rightarrow A_1$. (Similar to Fig. 1 of [2], but with maximum neuronal firing rate lowered from $S_{e,i}^{\max}=1000\ \mathrm{s}^{-1}$ to the physiologically more plausible rate of $100\ \mathrm{s}^{-1}$.)

determined by the stationary solution of the system SDEs. The stationary behavior of the variables h_e and h_i is shown in the inverted-S trajectory of Fig. 1.

These equilibrium curves were obtained by setting to zero all time and space derivatives, as well as zeroing the four subcortical noise terms (Γ_m), in Eqs. (2.1)–(2.3) below, and then solving for the steady-state voltages h_e^0, h_i^0 as a function of λ_{GABA}. Unfortunately, because of the strongly coupled and nonlinear nature of the system equations, it is not possible to write down an analytical expression for these equilibrium curves; instead one needs to follow an iterative numerical scheme similar in concept to Wilson and Cowan's [5] ''isocline intersection'' technique in order to map out the locus of excitatory and inhibitory steady-state soma voltages. (Full numerical details are given in Chapter 3 of [6].)

The equilibrium manifold of Fig. 1 provides the conceptual landscape on which we can identify distinct cortical states. The upper branch represents the *active* state characterized by high neuronal firing rates; the lower branch represents the low-firing *quiescent* state. Points on the active and quiescent branches are stable with respect to small perturbations; this pair of stable loci are separated by the reentrant midbranch identifying the locus of unstable equilibrium points. Although the cortex could ''visit'' a midbranch point, it cannot remain there, since the slightest fluctuation will cause it to ''fall'' onto either the upper or lower stable branch.

In our adiabatic picture, the EEG signal arises from fluctuations in h_e about its equilibrium manifold; these fluctuations originate from the white-noise stochastic p_{jk} input and enter h_e and h_i by way of nonlinear interdependencies [e.g., the $S_{e,i}$ sigmoidal voltage-to-firing-rate transfer functions Eqs. (2.5); and the ψ_{jk} synaptic-input weighting functions Eq. (2.4)] whose strengths depend directly on neuron voltage and indirectly on anesthetic concentration.

As discussed in [1,2], the main result of our white-noise-driven cortex model was the prediction of an abrupt ''anesthetodynamic'' change of state. This state change is characterized by a pair of distinct, general-anesthetic-induced, first-order phase transitions, the first transition occurring at the point of loss of consciousness (LOC), and the second at the point of recovery of consciousness (ROC).

These state-change predictions follow from the inverted-S form of Fig. 1: because the middle branch is unstable, there will be an abrupt, discontinuous change in h_e at LOC (point A_3) as the anesthetic amount λ_{GABA} is increased; and similarly at the Q_1 emergence point as the anesthetic amount is reduced.

If this cortical phase-change picture is correct, then the model predicts that there should be gross changes in the nature of the EEG signal in the vicinity of the jump points: (i) a ''biphasic'' effect (EEG power surge at LOC and ROC) [7]; (ii) a redistribution of spectral energy toward lower frequencies in the unconscious state; (iii) a reduction of spectral entropy at LOC [8]. These effects have all been observed clinically. In addition, if the unconscious state is the more ordered state, then a thermodynamic analogy argument [2] suggests that there should be a sudden release of ''latent heat'' at LOC. Stullken *et al.* [9] found that the metabolic energy requirement of the cortex (as measured by oxygen depletion in the cerebral blood flow) generally declined steadily with increasing anesthetic concentration, but declined precipitously in the vicinity of the cortical switchover detected by alteration in EEG activity. This result is consistent with the notion of a thermodynamic ''latent heat'' effect.

In this paper we extend the adiabatic approach of our earlier papers to allow for small *spatial* fluctuations about the equilibrium manifold. In the original adiabatic approximation, both the long-range inputs ϕ_k and the local inputs I_{jk} to the macrocolumn were assumed to be spatially homogeneous and to equilibrate instantaneously. Mathematically, this was accomplished by setting the space and time derivatives equal to zero in the equations governing the behavior of the ϕ_k (two equations) and I_{jk} (four equations). This is equivalent to collapsing the impulse responses for ϕ_k and I_{jk} into δ-function ''spikes.''

Here we allow for the possibility of spatial variability across a 1D cortex by developing a first-order treatment we refer to as the *spatioadiabatic approximation*. This treatment still assumes rapid equilibration in time for the ϕ_k and I_{jk}, but retains the second-order space derivative in the pair of ϕ_k wave equations. The result is a modified Green's function for the ϕ_k that is a δ function in time, scaled by an exponential decay in space.

The presumption of a spatioadiabatic limit allows us to derive analytically the spatial covariance of h_e, presumed equivalent to the spatial covariance of the scalp-measured EEG signal. The model predicts two interesting and distinct types of behavior, depending on the ''wiring'' ratio $R = \Lambda_{ee}/\Lambda_{ei}$, where the Λ_{ek} are the corticocortical inverse-length scales for the long-range $e \rightarrow e$ and $e \rightarrow i$ intermacrocolumn connections.

(1) For $R \lesssim 1.5$ (the default case: long-range $e \rightarrow e$ connectivity dominates long-range $e \rightarrow i$ connectivity), a stability analysis shows that the system relaxes over time to the homogeneous steady state defined by the Fig. 1 inverse-S curve. For this R wiring regime, the h_e covariance decays exponentially in space, with a correlation length that diverges at the induction point A_3 and again at the emergence point Q_1. We find that the correlation length for the quiescent state is longer than that in the active state. This is consistent with our phase transition picture, which views the unconscious state as being more ordered than the conscious state. There is good supporting evidence of increased correlation length at the LOC transition in the quantitative-EEG coherence experiments reported recently by John *et al.* [10]. (We briefly discuss these results in Sec. V.)

(2) In contrast, for $R \gtrsim 1.5$ (long-range $e \rightarrow i$ is now dominant), we find that the homogeneous steady state is no longer stable with respect to small perturbations. As a result, the system undergoes a transition to a *nonequilibrium* steady state, as verified by numerical simulation of the 1D cortex in the spatioadiabatic limit. Our *linearized* theory suggests that, in this far-from-equilibrium regime, the EEG distribution will organize itself to display a spatial covariance that is the sum of an exponentially decaying term plus a term that is *periodic* in space. However, simulation runs of the *nonlinearized* spatioadiabatic equations show that the outcome is actually more complicated than linear theory can predict: instead of exhibiting small-amplitude spatially periodic variations about the homogeneous steady state, we find that the cortical activity along the 1D cortical ''rod'' explodes into giant fluctuations in soma voltage that rapidly coalesce into stable, clumped regions of high-firing activity adjoining regions of low-firing activity. This emergence of a sustained mosaic of EEG activity is akin to the self-organization that can emerge in other dissipative systems in physics such as the chemical spiral waves of the Brusselator and the hexagonal convection cells in a Rayleigh-Bénard fluid (both discussed in [11]).

It is useful to place the present work in the context of other mean-field cortical models. Robinson *et al.* [12] presented a 2D continuum model that incorporates sigmoidal voltage-to-spike-rate conversion, excitatory and inhibitory synaptic connections, and the effects of dendritic and axonal time lags with axonal propagation formulated in terms of a wave equation. They investigated analytically the stability of the spatially homogeneous stationary state with respect to small perturbations and found temporally damped traveling waves for a variety of boundary conditions. They did not consider the possibility of a soft-mode transition to a spatially organized system in their stability analysis, nor did they look for such behavior in their nonlinear simulations.

The earlier cortical model of Nunez [13,14] also expresses axonal propagation in terms of a wave equation, but assumes that the cortical system is linear. Nunez solved his model for a 1D loop cortex and for a 2D cortex with periodic and spheroidal boundary conditions, obtaining global modes—weakly damped traveling waves that interfere to form standing waves—whose temporal frequencies approximately match observed cerebral rhythms. Jirsa and Haken [15,16] generalized the linear Nunez model [13] to include a sigmoid nonlinearity for the averaged effect of action potentials and derived a nonlinear wave equation for excitatory synaptic activity in a 1D neural tissue. When driven by a temporal sine-wave stimulus on spatially periodic boundaries, the field modes of synaptic activity formed nonsinusoidal standing waves. In a later paper, Jirsa and Kelso [17] demonstrated that the introduction of long-range connection inhomogeneities can cause their spatially homogeneous model to become unstable, leading to the formation of a series of distinct spatiotemporal patterns whose character is determined by the strength and location of the two-point long-range connection within the 1D loop of cortical tissue.

II. THEORY

A. One-dimensional cortical model

We model the cerebral cortex as a collection of interconnected macrocolumns (see Fig. 2 of [1]). These are neural masses containing approximately 100 000 synaptically connected neurons of which 85% are excitatory (e) and 15% are inhibitory (i). The primary variables of interest in the model are the macrocolumn-averaged *excitatory* soma voltage h_e and the macrocolumn-averaged *inhibitory* soma voltage h_i, both of which can vary in time and space.

We utilize a set of coupled stochastic differential equations for the macrocolumn developed in [1], and based on a set of eight partial differential equations (PDEs) by Liley, Cadusch, and Wright [18] for a one-dimensional cortex. These foundation PDEs can be regarded as a mean-field generalization of Tuckwell's ''subthreshold neuron'' [19]. For the Tuckwell neuron, there is only one ion species involved in excitation (e.g., sodium) and a second ion species involved in inhibition (e.g., potassium)—thus their Nernst potentials become the respective synaptic reversal potentials h_e^{rev} and h_i^{rev} (see Table I for values), and the membrane conductances of the neuron are taken as being proportional to the signed deviation of the soma voltage from the relevant reversal potential [the Tuckwell conductances map to the dimensionless ψ_{jk} weighting factors in Eq. (2.4) below].

In the mean-field picture of macrocolumn-average populations of cooperating neurons, no attempt is made to follow the detailed time dynamics for the formation of individual

TABLE I. Symbol definitions and model constants for the 1D Liley-Cadusch-Wright neural macrocolumn model. In contrast to our earlier papers [1–3], here the sigmoid maximum firing rate $S_{e,i}^{\max}$ has been lowered from 1000 to 100 s^{-1}.

Symbol	Description	Value	Unit
$e,\ i$	(As subscript) excitatory, inhibitory cell populations		
$h_{e,i}$	Population mean soma voltage		mV
$\tau_{e,i}$	Membrane time constant	0.040, 0.040	s
$h_{e,i}^{\text{rest}}$	Cell resting potential	$-70,\ -70$	mV
$h_{e,i}^{\text{rev}}$	Cell reversal potential (Nernst potential)	$45,-90$	mV
$I_{ee,ie}$	Total $e\to e, i\to e$ voltage input to excitatory synapses		mV
$I_{ei,ii}$	Total $e\to i, i\to i$ voltage input to inhibitory synapses		mV
$\psi_{jk\ (j,k\in\{e,i\})}$	Weighting factors for the I_{jk} inputs		
$p_{ee,ie}$	Exogenous (subcortical) spike input to e population	1100, 1600	s^{-1}
$p_{ei,ii}$	Exogenous (subcortical) spike input to i population	1600, 1100	s^{-1}
$\phi_{e,i}$	Long-range (corticocortical) spike input to e,i populations		s^{-1}
$\Lambda_{ee,ei}$	Characteristic corticocortical inverse-length scale	0.040, 0.065	$(\text{mm})^{-1}$
EPSP, IPSP	Excitatory, inhibitory postsynaptic potential		mV
$\gamma_{e,i}$	Neurotransmitter rate constant for EPSP, IPSP	300, 65	s^{-1}
$G_{e,i}$	Peak amplitude of EPSP, IPSP	0.18, 0.37	mV
e	[e.g., Eqs. (2.2),(2.6)] Base of natural logarithms	2.71828 . . .	
$N_{ee,ei}^{\beta}$	Total number of local $e\to e, e\to i$ synaptic connections	3034, 3034	
$N_{ie,ii}^{\beta}$	Total number of local $i\to e, i\to i$ synaptic connections	536, 536	
$N_{ee,ei}^{\alpha}$	Total number of synaptic connections from distant e populations	4000, 2000	
$\overline{v}$	Mean axonal conduction speed	7000	$\text{mm}\,s^{-1}$
$S_e(h_e), S_i(h_i)$	Sigmoid function mapping soma voltage to firing rate		s^{-1}
$S_e^{\max}, S_i^{\max}$	Maximum value for sigmoid function	100, 100	s^{-1}
$\theta_{e,i}$	Inflection-point voltage for sigmoid function	$-60,\ -60$	mV
$g_{e,i}$	Sigmoid slope at inflexion point	0.28, 0.14	$(\text{mV})^{-1}$
ℓ	Length of macrocolumn "cell"	1	mm

action potentials as described by the classical Hodgkin-Huxley [20] equations. But the concept of a *threshold voltage* is retained, and this appears within the sigmoid transfer functions that map from (average) membrane voltage $h_{e,i}$ to (average) firing rate $S(h_{e,i})$ [Eqs. (2.5) below]. Effectively, the sigmoids are "smoothed" step functions that describe the distribution of threshold voltages after averaging across the 10^5 neurons in the macrocolumn.

Each sigmoid nonlinearity enters the system equations scaled by a number representing the "strength" of the relevant *local* (symbol β) or *distant* (symbol α) synaptic connection. Thus, for example, N_{ei}^{β} is the number of $e\to i$ synaptic connections local to the macrocolumn, and appears in the term $N_{ei}^{\beta}S(h_e)$ on the right-hand side of Eq. (2.2b) for the I_{ei} excitatory-to-inhibitory synaptic input. To this locally generated flux in Eq. (2.2b) is added the incoming $e\to i$ flux ϕ_i originating from excitatory activity at distant macrocolumns [Eq. (2.3b)], plus $e\to i$ subcortical activity p_{ei} flowing into the inhibitory neural population from excitatory sources in the subcortex. This subcortical activity is assumed to be stochastic with a mean spike rate $\langle p_{ei}\rangle$ and a random component proportional to $\sqrt{\langle p_{ei}\rangle}$ as defined in Eq. (2.6b).

All synaptic inputs—local (intracortical), distant (corticocortical), and exogenous (subcortical)—into the population-average inhibitory (excitatory) neuron are scaled by the appropriate ψ_{ei} or ψ_{ii} (ψ_{ee} or ψ_{ie}) reversal-potential weighting function, then summed on the right of Eq. (2.1b) [Eq. (2.1a)] to give the deviation of h_i (h_e) from its resting voltage h_i^{rest} (h_e^{rest}). In the absence of any synaptic input, h_i (h_e) will relax exponentially, with time constant τ_i (τ_e), to its resting voltage. By assumption, the soma voltage time constants τ_e and τ_i are taken as being much larger than the relaxation times for the synaptic input events, allowing the $\partial/\partial t$ time derivatives appearing on the left of Eqs. (2.2) to be set to zero. This separation of fast (postsynaptic) and slow (soma) time scales is the *adiabatic approximation* invoked in our earlier homogeneous cortex modeling.

The wave equations (2.3) for ϕ_e and ϕ_i describe, respectively, the propagation of flux activity from distant excitatory cortical sources coupling into the excitatory and inhibitory synaptic inputs of the macrocolumn mass. The symbols ϕ_e and ϕ_i are shorthand for ϕ_{ee} and ϕ_{ei}, but the double-subscript notation can be abbreviated safely since all long-range cortical connections are exclusively from excitatory sources (therefore equations for ϕ_{ie} and ϕ_{ii} are not needed). The two Λ_{ek} constants are the inverse-length connectivity scales for the corticocortical fibers. As in our previous work, we assume the $\phi_{e,i}$ are "fast" variables (compared with soma-voltage time scales) so set their time derivatives to zero (adiabatic assumption). However, the approach used

here (described in Sec. II B) does *not* assume spatial homogeneity, so retains the $\partial^2/\partial x^2$ space derivatives, and this limit we refer to as the *spatioadiabatic* approximation.

The set of stochastic differential equations for the nonhomogeneous 1D cortex now follow:

$$\tau_e \frac{\partial h_e}{\partial t} = (h_e^{\rm rest} - h_e) + \psi_{ee}(h_e) I_{ee}(h_e) + \psi_{ie}(h_e) I_{ie}(h_i), \tag{2.1a}$$

$$\tau_i \frac{\partial h_i}{\partial t} = (h_i^{\rm rest} - h_i) + \psi_{ei}(h_i) I_{ei}(h_e) + \psi_{ii}(h_i) I_{ii}(h_i), \tag{2.1b}$$

$$\left(\frac{\partial}{\partial t} + \gamma_e\right)^2 I_{ee}(h_e) = [N_{ee}^{\beta} S_e(h_e) + \phi_e + \langle p_{ee}\rangle] G_e \gamma_e e + \Gamma_1(x,t), \tag{2.2a}$$

$$\left(\frac{\partial}{\partial t} + \gamma_e\right)^2 I_{ei}(h_e) = [N_{ei}^{\beta} S_e(h_e) + \phi_i + \langle p_{ei}\rangle] G_e \gamma_e e + \Gamma_2(x,t), \tag{2.2b}$$

$$\left(\frac{\partial}{\partial t} + \gamma_i\right)^2 I_{ie}(h_i) = [N_{ie}^{\beta} S_i(h_i) + \langle p_{ie}\rangle] G_i \gamma_i e + \Gamma_3(x,t), \tag{2.2c}$$

$$\left(\frac{\partial}{\partial t} + \gamma_i\right)^2 I_{ii}(h_i) = [N_{ii}^{\beta} S_i(h_i) + \langle p_{ii}\rangle] G_i \gamma_i e + \Gamma_4(x,t), \tag{2.2d}$$

$$\left[\left(\frac{\partial}{\partial t} + \overline{v}\Lambda_{ee}\right)^2 - \overline{v}^2 \frac{\partial^2}{\partial x^2}\right] \phi_e = \overline{v}\Lambda_{ee} N_{ee}^{\alpha} \left(\frac{\partial}{\partial t} + \overline{v}\Lambda_{ee}\right) S_e(h_e), \tag{2.3a}$$

$$\left[\left(\frac{\partial}{\partial t} + \overline{v}\Lambda_{ei}\right)^2 - \overline{v}^2 \frac{\partial^2}{\partial x^2}\right] \phi_i = \overline{v}\Lambda_{ei} N_{ei}^{\alpha} \left(\frac{\partial}{\partial t} + \overline{v}\Lambda_{ei}\right) S_e(h_e), \tag{2.3b}$$

where the four ψ_{jk} are normalized weighting functions proportional to the displacement of the h_k soma voltage from the ionic reversal potential $h_j^{\rm rev}$,

$$\psi_{jk} = \frac{h_j^{\rm rev} - h_k}{|h_j^{\rm rev} - h_k^{\rm rest}|}. \tag{2.4}$$

The $S_e(h_e), S_i(h_i)$ are sigmoidal transfer functions that map the soma voltage (in mV) to average output spike rate (in pulses per second),

$$S_e(h_e) = \frac{S_e^{\max}}{1 + \exp[-g_e(h_e - \theta_e)]}, \tag{2.5a}$$

$$S_i(h_i) = \frac{S_i^{\max}}{1 + \exp[-g_i(h_i - \theta_i)]}. \tag{2.5b}$$

The four Γ_m terms in Eqs. (2.2) represent the stochastic components of the system, and these are assumed to enter via random noise in the four p_{jk} subcortical inputs,

$$\Gamma_1(x,t) = \alpha_{ee}\sqrt{\langle p_{ee}\rangle} G_e \gamma_e e \sqrt{\ell}\, \xi_1(x,t), \tag{2.6a}$$

$$\Gamma_2(x,t) = \alpha_{ei}\sqrt{\langle p_{ei}\rangle} G_e \gamma_e e \sqrt{\ell}\, \xi_2(x,t), \tag{2.6b}$$

$$\Gamma_3(x,t) = \alpha_{ie}\sqrt{\langle p_{ie}\rangle} G_i \gamma_i e \sqrt{\ell}\, \xi_3(x,t), \tag{2.6c}$$

$$\Gamma_4(x,t) = \alpha_{ii}\sqrt{\langle p_{ii}\rangle} G_i \gamma_i e \sqrt{\ell}\, \xi_4(x,t), \tag{2.6d}$$

where the four corresponding ξ_m terms are zero-mean Gaussian-distributed δ-function-correlated spatiotemporal white-noise sources,

$$\langle \xi_m(x,t)\rangle = 0, \tag{2.7a}$$

$$\langle \xi_m(x,t)\xi_n(x',t')\rangle = \delta_{mn}\delta(x-x')\,\delta(t-t'), \tag{2.7b}$$

and the α_{jk} are dimensionless constants introduced to ensure that the fluctuations are small. The ℓ appearing in Eqs. (2.6) is the length of the macrocolumn "cell," taken as $\ell = 1$ mm, and ensures that the product $\sqrt{\ell}\,\xi(x,t)$ has units $\mathrm{s}^{-1/2}$, independent of space. In numerical simulation, each infinite-variance white-noise source $\xi(x,t)$ is approximated using a computer-generated sequence of unit-variance Gaussian-distributed random numbers $\mathcal{R}(m,n)$,

$$\xi(x,t) \to \frac{\mathcal{R}(m,n)}{\sqrt{\Delta x \Delta t}} \quad [\text{units of } (\mathrm{mm})^{-1/2}\mathrm{s}^{-1/2}] \tag{2.8}$$

at discrete position and time coordinates $x = m\Delta x$, $t = n\Delta t$, where Δx is the grid spacing (in mm) and Δt is the time increment (in s). These scaled random numbers have variance $\sigma^2 = (\Delta x \Delta t)^{-1}$, with $\sigma^2 \to \infty$ in the continuous limit $\Delta x \to 0$, $\Delta t \to 0$.

The effect of GABAergic anesthetic is included in the model as a modulation of the γ_i inhibitory neurotransmitter rate constant,

$$\gamma_i \to \overline{\gamma}_i = \frac{\gamma_i}{\lambda_{\rm GABA}}, \tag{2.9}$$

where $\lambda_{\rm GABA}$ is a dimensionless scale factor assumed to be proportional to the GABA anesthetic concentration. Thus an increase in $\lambda_{\rm GABA}$ *reduces* the IPSP rate constant and *increases* the IPSP duration.

For a complete definition of all other constants appearing in the foregoing equations, see Table I.

B. Spatioadiabatic approximation

In common with our earlier papers, we retain the adiabatic or "slow-membrane" approximation in which the I_{jk}, p_{jk}, and ϕ_k synaptic inputs are assumed to be "fast" processes that equilibrate on time scales much shorter than the time scales for h_e and h_i. However, in contrast to our earlier work, we will now explicitly allow (weak) spatial variation in the system by retaining the $\partial^2/\partial x^2$ terms in Eqs. (2.3). We

refer to this treatment as the spatioadiabatic approximation, and its details are given in the Appendix, where we find that the system equations become

$$\frac{\partial h_e}{\partial t} = F_1(h_e,h_i) + \Gamma_e(x,t), \tag{2.10a}$$

$$\frac{\partial h_i}{\partial t} = F_2(h_e,h_i) + \Gamma_i(x,t), \tag{2.10b}$$

where the F_1, F_2 are *drift* terms defined by

$$F_1(h_e,h_i) = \frac{1}{\tau_e}\Bigg\{ (h_e^{\rm rest} - h_e) + \psi_{ee}(h_e)\Bigg[(N_{ee}^{\alpha} + N_{ee}^{\beta}) S_e(h_e) + \frac{1}{\Lambda_{ee}^2}\frac{\partial^2 \phi_e}{\partial x^2} + \langle p_{ee}\rangle \Bigg] G_e e/\gamma_e + \lambda_{\rm GABA}\psi_{ie}(h_e) \times [N_{ie}^{\beta} S_i(h_i) + \langle p_{ie}\rangle] G_i e/\gamma_i \Bigg\}, \tag{2.11a}$$

$$F_2(h_e,h_i) = \frac{1}{\tau_i}\Bigg\{ (h_i^{\rm rest} - h_i) + \psi_{ei}(h_i)\Bigg[(N_{ei}^{\alpha} + N_{ei}^{\beta}) S_e(h_e) + \frac{1}{\Lambda_{ei}^2}\frac{\partial^2 \phi_i}{\partial x^2} + \langle p_{ei}\rangle \Bigg] G_e e/\gamma_e + \lambda_{\rm GABA}\psi_{ii}(h_i) \times [N_{ii}^{\beta} S_i(h_i) + \langle p_{ii}\rangle] G_i e/\gamma_i \Bigg\}, \tag{2.11b}$$

and the Γ_e, Γ_i are *diffusive noise* terms,

$$\Gamma_e(x,t) = b_{ee}\xi_1(x,t) + b_{ie}\xi_3(x,t), \tag{2.12a}$$

$$\Gamma_i(x,t) = b_{ei}\xi_2(x,t) + b_{ii}\xi_4(x,t), \tag{2.12b}$$

whose b_{jk} coefficients depend on the GABA anesthetic effect $\lambda_{\rm GABA}$ and the (h_e,h_i) soma-voltage coordinate:

$$b_{ee} = \psi_{ee}(h_e)\,\alpha_{ee}\sqrt{\langle p_{ee}\rangle} G_e e \sqrt{\ell}/\gamma_e\tau_e, \tag{2.13a}$$

$$b_{ie} = \lambda_{\rm GABA}\psi_{ie}(h_e)\,\alpha_{ie}\sqrt{\langle p_{ie}\rangle} G_i e \sqrt{\ell}/\gamma_i\tau_e, \tag{2.13b}$$

$$b_{ei} = \psi_{ei}(h_i)\,\alpha_{ei}\sqrt{\langle p_{ei}\rangle} G_e e \sqrt{\ell}/\gamma_e\tau_i, \tag{2.13c}$$

$$b_{ii} = \lambda_{\rm GABA}\psi_{ii}(h_i)\,\alpha_{ii}\sqrt{\langle p_{ii}\rangle} G_i e \sqrt{\ell}/\gamma_i\tau_i. \tag{2.13d}$$

We note that, apart from the $\partial^2\phi_{e,i}/\partial x^2$ 1D Laplacian terms in Eqs. (2.11), these spatioadiabatic equations (2.10)–(2.12) are identical in form to the homogeneous adiabatic set listed in our previous paper [see Eqs. (2.9)–(2.10) in [2]]. In the Appendix we show that the ϕ_e and ϕ_i long-range corticocortical inputs can each be expressed as a series expansion in h_e about the homogeneous steady state [see Eqs. (A11a)–(A11b)]. This leads to the following approximation for the $\partial^2\phi_{e,i}/\partial x^2$ terms, valid for small perturbations about the homogeneous steady state:

$$\frac{\partial^2 \phi_e}{\partial x^2} \approx N_{ee}^{\alpha}\left.\frac{dS_e}{dh_e}\right|_{\rm eq}\frac{\partial^2 h_e}{\partial x^2}, \tag{2.14a}$$

$$\frac{\partial^2 \phi_i}{\partial x^2} \approx N_{ei}^{\alpha}\left.\frac{dS_e}{dh_e}\right|_{\rm eq}\frac{\partial^2 h_e}{\partial x^2}. \tag{2.14b}$$

This approximation will be utilized in the linearized spatioadiabatic cortex developed in the following section.

C. Linearized stochastic equations

Our principal assumption is that the cortex always operates in a state close to its homogeneous equilibrium defined by setting $\partial h_{e,i}/\partial t = 0$ in Eq. (2.10) and $\partial^2\phi_{e,i}/\partial x^2 = 0$ in Eq. (2.11). The distribution of equilibrium states available to the cortex is depicted in Fig. 1. The inverted-S form is characteristic of a system capable of making first-order (i.e., discontinuous) jumps between stable branches at the ''knee'' turning points labeled A_3 and Q_1. Cortical phase-transition properties, such as the ''biphasic'' power surge at induction and emergence, were discussed in [1].

Suppose there is a small voltage deviation $(\hat{h}_e(x,t),\hat{h}_i(x,t))$ away from the homogeneous steady-state coordinate (h_e^0,h_i^0), and that this deviation is both time and space dependent. Then the excitatory and inhibitory soma voltages at position x and time t can be written

$$h_e(x,t) = h_e^0 + \hat{h}_e(x,t), \tag{2.15a}$$

$$h_i(x,t) = h_i^0 + \hat{h}_i(x,t). \tag{2.15b}$$

Substituting Eq. (2.15) into Eq. (2.10) and linearizing about the homogeneous stationary state gives

$$\frac{\partial}{\partial t}\begin{bmatrix}\hat{h}_e(x,t)\\ \hat{h}_i(x,t)\end{bmatrix} = \begin{bmatrix} J_{11} + \kappa_e\dfrac{\partial^2}{\partial x^2} & J_{12} \\ J_{21} + \kappa_i\dfrac{\partial^2}{\partial x^2} & J_{22}\end{bmatrix}\begin{bmatrix}\hat{h}_e(x,t)\\ \hat{h}_i(x,t)\end{bmatrix} + \begin{bmatrix}\Gamma_e(x,t)\\ \Gamma_i(x,t)\end{bmatrix}_{\rm eq}, \tag{2.16}$$

where the κ_e, κ_i are Fick's-law spatial diffusion coefficients (units $\rm mm^2/s$),

$$\kappa_e = \frac{\psi_{ee}(h_e^0)S_e^{(1)}N_{ee}^{\alpha}G_e e}{\Lambda_{ee}^2\gamma_e\tau_e}, \tag{2.17a}$$

$$\kappa_i = \frac{\psi_{ei}(h_i^0)S_e^{(1)}N_{ei}^{\alpha}G_e e}{\Lambda_{ei}^2\gamma_e\tau_i}, \tag{2.17b}$$

and the J_{mn} are the four Jacobian elements for the homogeneous system [i.e., the F_1, F_2 drifts of Eq. (2.11) have their respective $\partial^2/\partial x^2$ terms set to zero],

$$J_{11} \equiv \left.\frac{\partial F_1}{\partial \hat{h}_e}\right|_{\text{eq}} = \{-1 + \psi_{ee}^{(1)}[(N_{ee}^{\alpha} + N_{ee}^{\beta})S_e(h_e^0) + \langle p_{ee}\rangle]G_e e/\gamma_e + \psi_{ee}(h_e^0)[(N_{ee}^{\alpha} + N_{ee}^{\beta})S_e^{(1)}]G_e e/\gamma_e$$

$$+ \lambda_{\text{GABA}}\psi_{ie}^{(1)}[N_{ie}^{\beta}S_i(h_i^0) + \langle p_{ie}\rangle]G_i e/\gamma_i\}\frac{1}{\tau_e}, \tag{2.18a}$$

$$J_{12} \equiv \left.\frac{\partial F_1}{\partial \hat{h}_i}\right|_{\text{eq}} = \lambda_{\text{GABA}}\psi_{ie}(h_e^0)N_{ie}^{\beta}S_i^{(2)}G_i e/\gamma_i\tau_e, \tag{2.18b}$$

$$J_{21} \equiv \left.\frac{\partial F_2}{\partial \hat{h}_e}\right|_{\text{eq}} = \psi_{ei}(h_i^0)(N_{ei}^{\alpha} + N_{ei}^{\beta})S_e^{(1)}G_e e/\gamma_e\tau_i, \tag{2.18c}$$

$$J_{22} \equiv \left.\frac{\partial F_2}{\partial \hat{h}_i}\right|_{\text{eq}} = \{-1 + \psi_{ei}^{(2)}[(N_{ei}^{\alpha} + N_{ei}^{\beta})S_e(h_e^0) + \langle p_{ei}\rangle]G_e e/\gamma_e + \lambda_{\text{GABA}}\psi_{ii}^{(2)}[N_{ii}^{\beta}S_i(h_i^0) + \langle p_{ii}\rangle]G_i e/\gamma_i$$

$$+ \lambda_{\text{GABA}}\psi_{ii}(h_i^0)N_{ii}^{\beta}S_i^{(2)}G_i e/\gamma_i\}\frac{1}{\tau_i}, \tag{2.18d}$$

with partial derivatives of the weighting and sigmoid functions also evaluated at equilibrium,

$$\psi_{je}^{(1)} = \frac{\partial \psi_{je}}{\partial h_e} = \frac{-1}{|h_j^{\text{rev}} - h_e^{\text{rest}}|}, \tag{2.19a}$$

$$\psi_{ji}^{(2)} = \frac{\partial \psi_{ji}}{\partial h_i} = \frac{-1}{|h_j^{\text{rev}} - h_i^{\text{rest}}|}, \tag{2.19b}$$

and

$$S_e^{(1)} = \left.\frac{\partial S_e}{\partial h_e}\right|_{\text{eq}}, \quad S_i^{(2)} = \left.\frac{\partial S_i}{\partial h_i}\right|_{\text{eq}}. \tag{2.19c}$$

D. Infinite brain

We can picture the adult human cerebral cortex as having a ''length'' (in a 1D modeling sense) of $\sim$40 cm, whereas the extent of a single macrocolumn is $\sim$1 mm. Therefore, on the macrocolumn scale, it is not too unreasonable to think of the cortex as having an ''infinite'' extent. This idealization simplifies the Fourier transform mathematics to follow, and allows us to ignore completely the boundary conditions that would otherwise need to be imposed on a rod of 1D macrocolumn ''mass'' whose length was finite.

For a continuous, infinite, 1D neural system we define the spatial Fourier transforms of the $\hat{h}_e$ excitatory and $\hat{h}_i$ inhibitory soma-voltage fluctuations and of the ξ white-noise sources, respectively, as the tilde variables $\tilde{h}_e, \tilde{h}_i, \tilde{\xi}$:

$$\tilde{h}_e(q,t) = \int_{-\infty}^{+\infty} e^{-iqx}\hat{h}_e(x,t)dx, \tag{2.20a}$$

$$\tilde{h}_i(q,t) = \int_{-\infty}^{+\infty} e^{-iqx}\hat{h}_i(x,t)dx, \tag{2.20b}$$

$$\tilde{\xi}(q,t) = \int_{-\infty}^{+\infty} e^{-iqx}\xi(x,t)dx. \tag{2.20c}$$

These relations define mappings from the linear position x to the wave number q. The corresponding definitions for the inverse $q \to x$ transforms are

$$\hat{h}_e(x,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{iqx}\tilde{h}_e(q,t)dq, \tag{2.21a}$$

$$\hat{h}_i(x,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{iqx}\tilde{h}_i(q,t)dq, \tag{2.21b}$$

$$\xi(x,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{iqx}\tilde{\xi}(q,t)dq. \tag{2.21c}$$

Fourier transforming the Langevin equations (2.16), we obtain

$$\frac{\partial}{\partial t}\begin{bmatrix}\tilde{h}_e(q,t)\\ \tilde{h}_i(q,t)\end{bmatrix} = \begin{bmatrix}J_{11} - \kappa_e q^2 & J_{12}\\ J_{21} - \kappa_i q^2 & J_{22}\end{bmatrix}\begin{bmatrix}\tilde{h}_e(q,t)\\ \tilde{h}_i(q,t)\end{bmatrix} + \begin{bmatrix}\tilde{\Gamma}_e(q,t)\\ \tilde{\Gamma}_i(q,t)\end{bmatrix}_{\text{eq}}, \tag{2.22}$$

where

$$\tilde{\Gamma}_e(q,t) = b_{ee}\tilde{\xi}_1(q,t) + b_{ie}\tilde{\xi}_3(q,t), \tag{2.23a}$$

$$\tilde{\Gamma}_i(q,t) = b_{ei}\tilde{\xi}_2(q,t) + b_{ii}\tilde{\xi}_4(q,t), \tag{2.23b}$$

and where the Fourier transformed white-noise terms satisfy the correlation property

$$\langle \tilde{\xi}_m(q,t)\tilde{\xi}_n(q',t')\rangle = 2\pi\delta_{mn}\delta(q+q')\delta(t-t'). \tag{2.24}$$

In order to make Eq. (2.22) amenable to analysis via standard stochastic calculus techniques, we transform it into a two-variable Ornstein-Uhlenbeck (OU) system of equations. This is done in two steps. First, we define a diagonal 2×2 diffusion matrix **D**,

$$\mathbf{D} = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \tag{2.25}$$

whose elements are obtained by using the Eq. (2.24) correlation identity to compute the expectation values of the $\tilde{\Gamma}_e, \tilde{\Gamma}_i$ noise terms, giving

$$\langle \tilde{\Gamma}_e(q,t)\tilde{\Gamma}_e(q',t')\rangle = 2\pi D_1\delta(q+q')\delta(t-t'), \tag{2.26a}$$

$$\langle \tilde{\Gamma}_i(q,t)\tilde{\Gamma}_i(q',t')\rangle = 2\pi D_2\delta(q+q')\delta(t-t'), \tag{2.26b}$$

where

$$D_1 = (b_{ee})^2_{\rm eq} + (b_{ie})^2_{\rm eq}, \tag{2.27a}$$

$$D_2 = (b_{ei})^2_{\rm eq} + (b_{ii})^2_{\rm eq}, \tag{2.27b}$$

and the $(b_{jk})_{\rm eq}$ are the equilibrium values of the noise coefficients defined earlier in Eq. (2.13).

The second step is to redefine the drift matrix in Eq. (2.22) so that it carries an explicit negative sign (here we are following the Chaturvedi *et al.* [21] sign convention). The OU system resulting from these two steps then reads

$$\frac{\partial}{\partial t}\begin{bmatrix} \tilde{h}_e(q,t) \\ \tilde{h}_i(q,t) \end{bmatrix} = -\mathbf{A}(q)\begin{bmatrix} \tilde{h}_e(q,t) \\ \tilde{h}_i(q,t) \end{bmatrix} + \sqrt{\mathbf{D}}\begin{bmatrix} \tilde{\xi}_e(q,t) \\ \tilde{\xi}_i(q,t) \end{bmatrix}, \tag{2.28}$$

where

$$\mathbf{A}(q) = \begin{bmatrix} -J_{11}+\kappa_e q^2 & -J_{12} \\ -J_{21}+\kappa_i q^2 & -J_{22} \end{bmatrix} \tag{2.29}$$

and the J_{mn} Jacobian elements are as defined earlier in Eqs. (2.18).

III. STABILITY OF SPATIOADIABATIC SYSTEM

A. Eigenvalue analysis

The steady states of the spatioadiabatic cortex model of Eq. (2.10) can be stable only if the voltage fluctuations about the homogeneous steady state remain damped. Because of the leading minus sign in Eq. (2.28), system stability requires that both eigenvalues of the $\mathbf{A}(q)$ drift matrix have *positive* real parts. The eigenvalues λ_1, λ_2 of matrix **A** are given by the solutions of the following quadratic equation in λ:

$$\lambda^2 + (J_{11}+J_{22}-\kappa_e q^2)\lambda + J_{11}J_{22} - J_{12}J_{21} + (\kappa_i J_{12} - \kappa_e J_{22})q^2 = 0. \tag{3.1}$$

On making the identifications

$$c_1 = J_{11} + J_{22}, \tag{3.2a}$$

$$c_2 = J_{11}J_{22} - J_{12}J_{21}, \tag{3.2b}$$

$$c_3 = \kappa_i J_{12} - \kappa_e J_{22}, \tag{3.2c}$$

we can rewrite Eq. (3.1) as

$$\lambda^2 + (c_1 - \kappa_e q^2)\lambda + (c_2 + c_3 q^2) = 0 \tag{3.3}$$

with eigenvalue solutions

$$\lambda_1, \lambda_2 = -\tfrac{1}{2}(c_1 - \kappa_e q^2) \pm \tfrac{1}{2}\sqrt{(c_1 - \kappa_e q^2)^2 - 4(c_2 + c_3 q^2)}. \tag{3.4}$$

From previous analysis of the $q=0$ homogeneous system, we know that for all *stable* equilibrium points $c_1 < 0$ and $c_2 > 0$. Further, provided the soma voltage never exceeds the excitatory reversal potential $h_e^{\rm rev} = +45$ mV, then the ψ_{ee} weighting function of Eq. (2.4) will always be positive, and hence, by Eq. (2.17a), $\kappa_e > 0$ also. We find that for the default model values listed in Table I, $c_3 > 0$. This follows because $0 < \kappa_i < \kappa_e$ (i.e., the long-range corticocortical *e-i* diffusivity is *weaker* than the long-range *e-e* diffusivity), while both J_{12} and J_{22} are negative with $|J_{12}| < |J_{22}|$. Nevertheless, the possibility exists for c_3 to become negative if, for some collection of communicating macrocolumns, the *e-i* diffusivity dominates. This will be the case if

$$\kappa_i > \frac{|J_{22}|}{|J_{12}|}\kappa_e \Rightarrow c_3 < 0. \tag{3.5}$$

Taking the known signs of c_1 and κ_e into account, we can rewrite Eq. (3.4) as

$$\lambda_1, \lambda_2 = \tfrac{1}{2}(|c_1| + \kappa_e q^2) \pm \tfrac{1}{2}\sqrt{(|c_1| + \kappa_e q^2)^2 - 4(c_2 + c_3 q^2)}. \tag{3.6}$$

The homogeneous steady state becomes *unstable* when the real part of either eigenvalue becomes negative. If the real part of either eigenvalue goes to zero, then the homogeneous steady state is *marginally stable*, and there are two distinct possibilities to consider. (a) If the eigenvalues are complex and $\mathrm{Re}(\lambda_1, \lambda_2) \to 0^+$, then there is a so-called *hard-mode instability* leading to *temporal oscillations*. (b) If both eigenvalues are real and positive with the smaller eigenvalue $\lambda_2 \to 0^+$, then there is a *soft-mode instability*, meaning that the system will exhibit *spatial oscillations*. We now examine each possibility in turn.

(a) When the discriminant of Eq. (3.6) is negative, the eigenvalues will form a complex conjugate pair with real part

$$\mathrm{Re}(\lambda_1, \lambda_2) = \tfrac{1}{2}(|c_1| + \kappa_e q^2)$$

which is always positive. Therefore there is no possibility of a hard-mode instability, with its attendant temporal oscillations, developing in the present spatioadiabatic model of the 1D cortex.

(b) The eigenvalues will be real when the discriminant is positive,

$$(|c_1|+\kappa_e q^2)^2 > 4(c_2+c_3 q^2). \tag{3.7}$$

They will be real and positive if, in addition to (3.7), we require that the smaller eigenvalue λ_2 be positive, where λ_2 is given by,

$$\lambda_2 = \tfrac{1}{2}(|c_1|+\kappa_e q^2) - \tfrac{1}{2}\sqrt{(|c_1|+\kappa_e q^2)^2 - 4(c_2+c_3 q^2)}. \tag{3.8}$$

We see that $\lambda_2 \to 0^+$ along the positive real axis with λ_1 remaining real and positive if (3.7) is satisfied. In particular, at the point where $\lambda_2 = 0$ we have

$$c_2 + c_3 q^2 = 0. \tag{3.9}$$

As outlined above, the default model has $c_2 > 0$ and $c_3 > 0$, so the (3.9) condition cannot be satisfied for any real wave vector q. However, if the corticocortical connectivity alters in such a way that the long-range *e-i* diffusivity κ_i dominates the *e-e* diffusivity κ_e and condition (3.5) becomes true, then c_3 changes sign, and the system can exhibit sustained spatially periodic voltage perturbations about the homogeneous steady state. From Eq. (3.9), this marginally stable configuration will be characterized by wave vector q_s,

$$q_s = \sqrt{\frac{c_2}{|c_3|}} \quad \text{(soft-mode instability: spatial oscillations)}. \tag{3.10}$$

We note that from Eq. (3.8) any spatial mode with wave number $q < q_s$ will render a positive eigenvalue, so the mode will decay exponentially with time back to the homogeneous steady state.

In contrast, if $q > q_s$, the eigenvalue will be negative, so the amplitude of the eigenmode will grow exponentially. Therefore the homogeneous steady state cannot support spatial modes that oscillate faster than the critical q_s wave vector, and the presence of such modes will force the cortex to seek an alternative resting state. Although the form and location of this new state cannot be predicted from linearized theory, by running numerical simulations of the nonlinearized cortical equations we can gain some insights into the nature and structure of the far-from-equilibrium steady state. Nonlinearized model runs are described later in Sec. IV C. First, though, we describe linearized simulation runs designed to verify the existence of marginally stable steady states.

B. Simulation tests for soft-mode instability

To test the theoretical prediction of sustained spatial oscillations about the homogeneous steady state, we ran a series of nonstochastic numerical simulations of the linearized equations of motion (2.16) for $\hat{h}_e(x,t)$ and $\hat{h}_i(x,t)$ with noise terms $\Gamma_{e,i}(x,t)$ set to zero. We approximated the "infinite brain" of Sec. II D by a ~100-cm length of "cortical rod" on which periodic boundary conditions were imposed (i.e., the rod effectively is deformed into a 100-cm closed loop).

In order to increase the strength of the *e-i* diffusivity relative to the *e-e* diffusivity (thereby allowing the c_3 parameter to go negative) we introduce a dimensionless factor $f \geqslant 1$ to scale the inverse-length connectivity constants Λ_{ee} and Λ_{ei} in opposite directions:

$$\Lambda'_{ee} = f\Lambda_{ee}, \quad \Lambda'_{ei} = \Lambda_{ei}/f, \tag{3.11}$$

and define a dendritic "wiring" ratio

$$R = \frac{\Lambda'_{ee}}{\Lambda'_{ei}}. \tag{3.12}$$

From Eqs. (2.17), $\kappa_{e,i} \propto 1/\Lambda^2_{ee,ei}$, so the respective *e* and *i* diffusivities will scale inversely and directly as f^2:

$$\kappa'_e = \kappa_e/f^2, \quad \kappa'_i = f^2\kappa_i. \tag{3.13}$$

Thus an increase in f factor simultaneously strengthens κ_i while weakening κ_e. Setting $f=1$ gives the unadjusted model diffusivity values and default wiring ratio $R=0.615$; $f=1.58$ ($R=1.536$) is sufficient to cause c_3 to just go negative in the immediate vicinity of the A_3 conscious $\rightarrow$ unconscious transition point at $\lambda_{\mathrm{GABA}}=1.31$ (see Fig. 1); and setting $f=5.0$ causes c_3 to be negative for the entire range of (h_e^0, h_i^0) equilibrium values plotted in Fig. 1.

Shown in Fig. 2(a) is the spatioadiabatic eigenvalue distribution for $\lambda_{\mathrm{GABA}}=1.25$ on the top branch of Fig. 1 for an assigned connectivity scale factor $f=1.80$, wiring ratio $R=1.994$. The zero crossing occurs at wave number $q_s = 0.5205\ \mathrm{cm}^{-1}$, predicting a marginally stable eigenmode of wavelength $2\pi/q_s = 12.1$ cm. To demonstrate that this mode is long lived and that slower spatial modes die away, we ran a numerical simulation of the deterministic part of the linearized spatioadiabatic system defined in Eqs. (2.16). We gave the 1D system (the cortical "rod") an initial excitatory voltage perturbation away from homogeneous steady state defined by

$$\hat{h}_e(x,t=0) = \sum_{k=1}^{8} \frac{(-1)^{k+1}}{k} \sin(kq_s x/8) \tag{3.14a}$$

and with inhibitory voltage left unperturbed,

$$\hat{h}_i(x,t=0) = 0. \tag{3.14b}$$

The summation (3.14a) for $\hat{h}_e$ represents the first eight terms of a harmonic expansion of a sawtooth wave form [[22], see p. 132, Eq. (23.9)] in which the eighth and final harmonic is the long-lived q_s soft mode. The soma voltages were sampled at $N=100$ equally spaced grid points along the cortical rod, with grid spacing $\Delta x = 0.966$ cm chosen to ensure that the $k=1$ sawtooth fundamental is exactly periodic on the adjusted rod length $L = N\Delta x$. (Failure to ensure an exact spatial fit results in a step discontinuity at the "join" of the

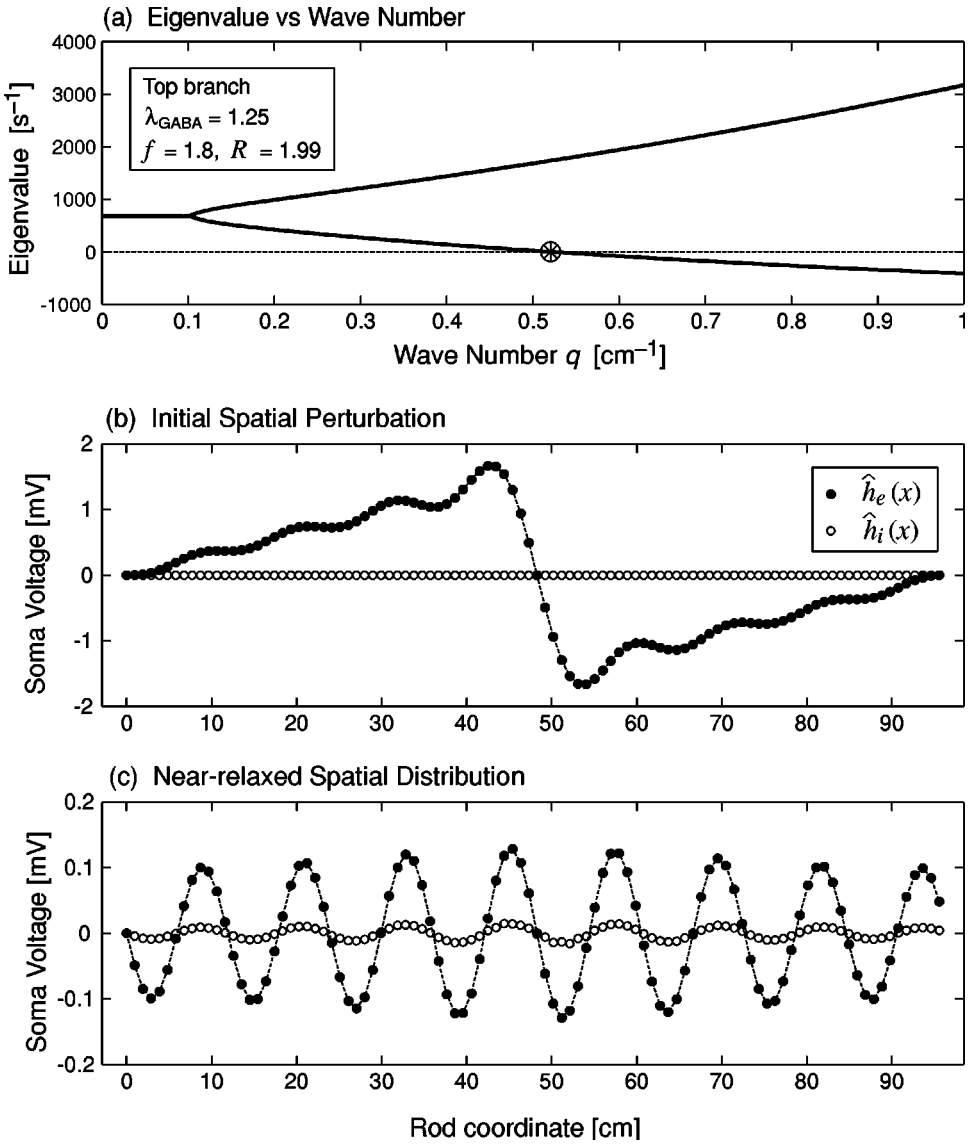

FIG. 2. (a) Eigenvalues and (b),(c) $q=0.5205\ \mathrm{cm}^{-1}$ eigenmode behavior for $f=1.8$ at $\lambda_{\mathrm{GABA}}=1.25$ on the top branch of Fig. 1. (a) plots the eigenvalue vs wave number distribution. Eigenvalue pair is complex with positive real part for $0\leqslant q \leqslant 0.101\ \mathrm{cm}^{-1}$, then becomes purely real for $q>0.101\ \mathrm{cm}^{-1}$. The smaller eigenvalue crosses the wave number axis at $q=q_s=0.5205\ \mathrm{cm}^{-1}$. For $q<q_s$, both eigenvalues are positive (or have positive real parts), so their eigenmodes will decay with time. For $q>q_s$ the smaller eigenvalue is negative and therefore the high-spatial-frequency eigenmodes will be unstable. (b) Initial voltage perturbation (in mV) at 100 points along the cortical "rod:" $\hat{h}_i(x,0)=0$, $\hat{h}_e(x,0)=\Sigma_{k=1}^{8}k^{-1}(-1)^{k+1}\sin(kq_sx/8)$ (i.e., a band-limited sawtooth). (c) Near-relaxed voltage state of the cortical rod after 8000 iterations (time step $\Delta t=4\ \mu$s) of the coupled nonstochastic equations of motion for $\hat{h}_e(x,t)$ and $\hat{h}_i(x,t)$. The low-frequency eigenmodes of (b) have decayed away, leaving only the marginally stable q_s eigenmode (of wavelength $2\pi/q_s=12.07$ cm) carried in both $\hat{h}_e(x)$ and $\hat{h}_i(x)$.

two ends of the loop; this discontinuity generates high-frequency wave numbers $q>q_s$ and these cause the simulation to explode exponentially at the join.)

The initial voltage configuration of the rod is shown in Fig. 2(b). Panel (c) illustrates the nearly relaxed voltage configuration after 8000 FTCS (forward-time, centered-space) iterations: the low-frequency modes have substantially decayed, leaving only the q_s long-lived mode evident in the in-phase voltage variations in both the excitatory and inhibitory rod voltages.

IV. SPATIAL DISTRIBUTIONS FOR EEG

A. Spatial covariance of $\hat{h}_e$

To quantify the degree to which the voltage fluctuations at separated points x and x' are correlated, we define a steady-state spatial covariance $G(x,x')$,

$$G(x,x')=\lim_{t\to\infty}\langle \hat{h}_e(x,t)\hat{h}_e(x',t)\rangle, \tag{4.1}$$

and we assume that the fluctuation covariance depends only on the separation $|x-x'|$ of the sensing electrodes (and not on their absolute positions x and x') so that the two-variable spatial covariance $G(x,x')$ can be collapsed to an explicit dependence on a single variable, the (unsigned) electrode separation,

$$G(x,x')=G(|x-x'|). \tag{4.2}$$

We will calculate the spatial covariance $G(x,x')$ by first determining the covariance in Fourier space,

$$\tilde{G}(q,q')=\langle \tilde{h}_e(q)\tilde{h}_e(q')\rangle, \tag{4.3}$$

and then Fourier inverting to give $G(x,x')$,

$$G(x,x')=\frac{1}{(2\pi)^2}\int\int_{-\infty}^{+\infty} e^{iqx}e^{iq'x'}\tilde{G}(q,q')\,dq\,dq' \tag{4.4}$$

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{iq|x-x'|}\tilde{G}(q)dq, \tag{4.5}$$

where $\tilde{G}(q)$ is defined below in Eqs. (4.7) and (4.8).

In Sec. II C we transformed the Langevin cortical equations into the two-variable Ornstein-Uhlenbeck system of Eq. (2.28). We may now apply the stochastic methods described by Chaturvedi *et al.* [21] and Gardiner [23] to compute the 2×2 steady-state covariance matrix $\tilde{\mathbf{G}}$,

$$\tilde{\mathbf{G}}(q,q')=\lim_{t\to\infty}\begin{bmatrix}\langle\tilde{h}_e(q,t)\tilde{h}_e(q',t)\rangle & \langle\tilde{h}_e(q,t)\tilde{h}_i(q',t)\rangle\\ \langle\tilde{h}_i(q,t)\tilde{h}_e(q',t)\rangle & \langle\tilde{h}_i(q,t)\tilde{h}_i(q',t)\rangle\end{bmatrix} \tag{4.6a}$$

$$=2\pi\delta(q+q')\tilde{\mathbf{G}}(q), \tag{4.6b}$$

where

$$\tilde{\mathbf{G}}(q)=\frac{\mathbf{D}}{2\,\mathrm{tr}(\mathbf{A})}+\frac{[\mathbf{A}-\mathrm{tr}(\mathbf{A})\mathbf{I}]\mathbf{D}[\mathbf{A}-\mathrm{tr}(\mathbf{A})\mathbf{I}]^{\mathrm{T}}}{2\,\mathrm{tr}(\mathbf{A})\det(\mathbf{A})}. \tag{4.7}$$

Here, $\mathbf{A}=\mathbf{A}(q)$ is the q-space drift matrix of Eq. (2.29), and $\mathbf{D}$ is the diffusion matrix of Eq. (2.25). The operators $\mathrm{tr}(\cdot)$ and $\det(\cdot)$ denote the matrix *trace* and *determinant*, respectively.

Our primary interest lies with the $\tilde{G}_{11}$ element of the $\tilde{\mathbf{G}}$ matrix, since this element gives the prediction for the Fourier space covariance for the *e-e* fluctuations. Expanding Eq. (4.7), we obtain

$$[\tilde{\mathbf{G}}(q)]_{11}=\frac{D_1}{2(A_{11}+A_{22})}+\frac{A_{22}^2D_1+A_{12}^2D_2}{2(A_{11}+A_{22})(A_{11}A_{22}-A_{12}A_{21})} \tag{4.8}$$

where all quantities are understood to be evaluated at equilibrium. Substituting Eq. (2.29) and applying some partial-fraction algebra allows us to rewrite Eq. (4.8) as

$$[\tilde{\mathbf{G}}(q)]_{11}=\frac{D_1+\kappa_e c_4}{2(\kappa_e q^2-c_1)}-\frac{c_3c_4}{2(c_3q^2+c_2)}, \tag{4.9}$$

where c_1,c_2,c_3 were defined in Eqs. (3.2), and c_4 is given by

$$c_4=\frac{J_{12}^2D_2+J_{22}^2D_1}{c_1c_3+\kappa_e c_2}. \tag{4.10}$$

Equation (4.9) is now amenable to Fourier inversion from q space back to x space, but the precise form of the result will depend on the relative signs for the coefficient pair κ_e and c_1 [first denominator of Eq. (4.9)]; and on the coefficient pair c_3 and c_2 (second denominator). For the default model values listed in Table I, we find $c_4<0$. As pointed out in Sec. III A, we have $\kappa_e>0$, $c_1<0$, $c_2>0$, and $c_3>0$, but with the possibility that c_3 could change sign if the long-range *e-i* diffusivity is allowed to dominate the long-range *e-e* diffusivity.

Taking signs into account, we can rewrite Eq. (4.9) as

$$[\tilde{\mathbf{G}}(q)]_{11}=\frac{D_1/\kappa_e+c_4}{2(q^2+|c_1/\kappa_e|)}-\frac{c_4}{2(q^2\pm|c_2/c_3|)}, \tag{4.11}$$

where, in the second denominator, the "$\pm$" operator follows the sign of c_3 (i.e., "+" if $c_3>0$; "−" if $c_3<0$).

We can calculate the inverse Fourier transforms of Eq. (4.11) by way of the following Fourier identities:

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{iqr}\frac{dq}{q^2+\alpha^2}=\frac{1}{2\alpha}\exp(-\alpha|r|), \tag{4.12}$$

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{iqr}\frac{dq}{q^2-\beta^2}=-\frac{1}{2\beta}\sin(\beta|r|). \tag{4.13}$$

Thus, depending on the sign of c_3, the form of the *e-e* spatial covariance is either the difference of *two exponential decays* whose respective $1/e$ decay lengths are L_1 and L_2,

$$G(|x-x'|)=\frac{D_1/\kappa_e+c_4}{4}L_1\exp(-|x-x'|/L_1)$$
$$-\frac{c_4}{4}L_2\exp(-|x-x'|/L_2),\quad c_3>0, \tag{4.14}$$

or the sum of an L_1-length *exponential decay plus a sine-wave variation* of wavelength $2\pi L_2$,

$$G(|x-x'|)=\frac{D_1/\kappa_e+c_4}{4}L_1\exp(-|x-x'|/L_1)$$
$$+\frac{c_4}{4}L_2\sin(|x-x'|/L_2),\quad c_3<0, \tag{4.15}$$

where the quantities

$$L_1\equiv\sqrt{|\kappa_e/c_1|}, \tag{4.16}$$

$$L_2\equiv\sqrt{|c_3/c_2|} \tag{4.17}$$

define the characteristic lengths (either a correlation length or a scaled wavelength) for the spatial covariance patterns.

B. Correlation-length predictions

Figure 3 plots the Eqs. (4.16) and (4.17) predictions for the L_1 and L_2 correlation lengths as the anesthetic effect is varied. Panel (a) shows that the domain over which EEG fluctuations are strongly correlated is expected to *increase* significantly on approach to I, the point of induction of unconsciousness (labeled A_3 on the active branch of Fig. 1),

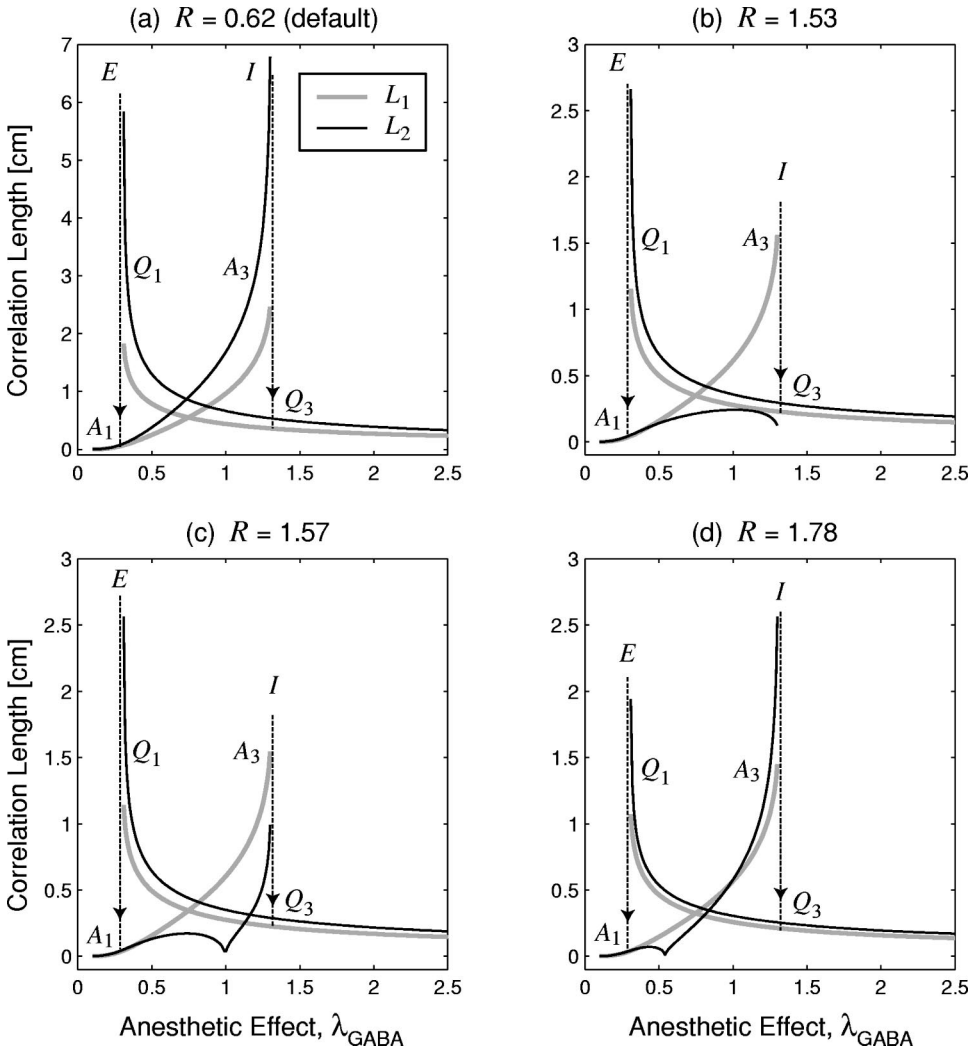

FIG. 3. Variation of L_1 and L_2 correlation lengths as a function of anesthetic effect λ_{GABA} for the spatioadiabatic 1D cortex. Four representative values for the long-range corticocortical wiring ratio $R=\Lambda_{ee}/\Lambda_{ei}$ are shown; larger R values correspond to a relatively more effective *inhibitory* connectivity. Points A_1 and A_3 lie on the active, high-firing branch; Q_1 and Q_3 are on the quiescent, low-firing branch. In all cases, fluctuations become much more correlated in space on approach to induction (I) of unconsciousness, and again for the return path on approach to emergence (E) back into consciousness. The cusp in the L_2 graphs [at $\lambda_{\mathrm{GABA}}\approx 1.0$ in (c), and at $\lambda_{\mathrm{GABA}}\approx 0.6$ in (d)] marks the point at which the c_3 parameter of Eq. (3.2c) goes negative; for points to the right of the cusp, L_2 becomes an inverse wave number for a sinusoidal component, of wavelength $2\pi L_2 = 2\pi/q_s$, in the predicted spatial covariance pattern.

and again on approach to E, the point of emergence from unconsciousness (Q_1 on the quiescent branch of Fig. 1). This increase in correlation length for the 1D spatioadiabatic cortex is consistent with the increase in correlation time (critical slowing down) reported in [3] for a homogeneous cortex near the anesthetodynamic transition.

We note, however, that there is a marked difference in the active-branch behaviors of L_1 and L_2 when R, the relative strength of the long-range inhibitory connectivity, is increased. In Fig. 3(b), whereas L_1 (thick gray curve) increases strongly on approach to the transition point A_3, we see that L_2 (thin dark curve) turns over and approaches zero at the A_3 transition. In panel (c), the turnover has become a cusp near $\lambda_{\mathrm{GABA}}=1.0$, followed by a renewed increasing trend. The cusp marks the point at which parameter c_3 becomes zero [recall Eq. (4.17): $L_2=\sqrt{|c_3/c_2|}$], and the increasing trend corresponds to c_3 going increasingly negative; therefore the interpretation of L_2 postcusp changes from being a $1/e$-decay correlation length (precusp) to being an indication that the covariance now has a spatially periodic component whose wavelength $2\pi L_2$ increases on approach to the first-order transition at A_3. Sadly, this wavelength prediction is of dubious utility, since, as will become apparent in the following section, the linearized system becomes unstable in the $c_3<0$ regime when driven by white noise.

C. Stochastic simulations of spatioadiabatic 1D cortex

1. Case $c_3>0$: Strong e→e diffusivity

In order to test the $c_3>0$ linearized spatial covariance predictions of Eq. (4.14), we ran a series of stochastic simulations of the *nonlinearized* spatioadiabatic equation set defined by Eqs. (2.1) and (A1),(A2). Illustrated in Fig. 4 are representative comparisons between the linearized theory and nonlinearized simulation for six different locations (three on the active branch and three on the quiescent branch) chosen from the Fig. 1 trajectory of homogeneous steady states.

For each simulation run, the cortex was represented by $N=100$ equally spaced points along a rod with joined ends. Each grid point was driven by the four independent white-noise sources defined in Eqs. (2.6) and (2.8), with the four noise-scale factors set to $\alpha_{jk}=0.01$. The rod was allowed to evolve from its homogeneous steady state for 1000 time steps, then $G_{\mathrm{sim}}(|x-x'|)$, the spatial covariance of the $\hat{h}_e$ excitatory-voltage fluctuations, was calculated using the MATLAB XCORR cross-correlation function, called $f_{\mathtt{xcorr}}$ in the expression

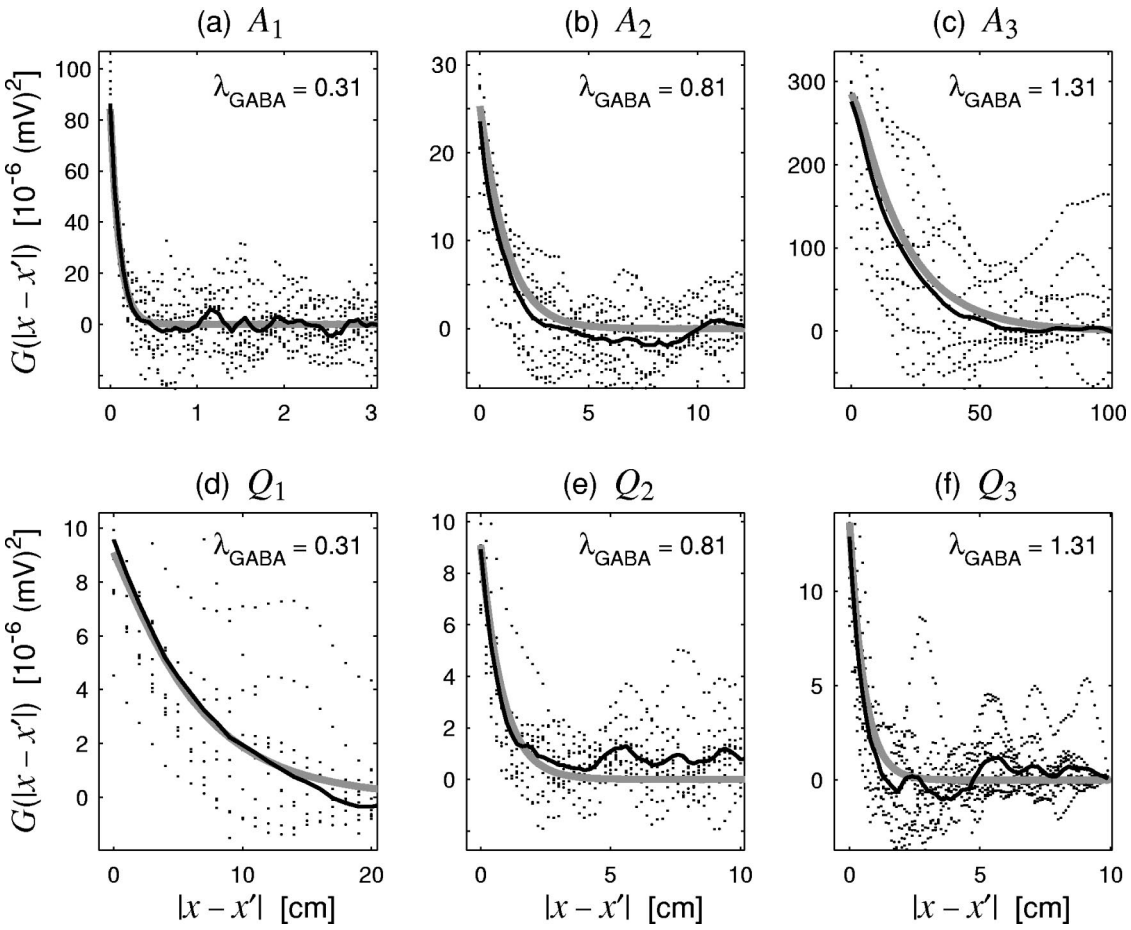

FIG. 4. Comparison of Eq. (4.14) covariance predictions (thick gray curves) and nonlinear stochastic simulation results (black dots) at discrete points along the cortical rod. The top three panels (a), (b), and (c) trace the induction trajectory $A_1 \rightarrow A_2 \rightarrow A_3$ along the active branch of Fig. 1; the bottom three panels (d), (e), and (f) trace the emergence trajectory $Q_1 \leftarrow Q_2 \leftarrow Q_3$ along the quiescent branch. For all six panels, the $\kappa_{e,i}$ diffusivities were maintained at their default values [see Fig. 3(a): $f=1$, $R=0.62$], so $c_3>0$ everywhere. For each panel, ten numerical simulations were run, the ten-run average covariance was calculated at each cortical position, and joined with a thin black pen. The spatial width of the covariance curve increases strongly on approach to induction (a) $\rightarrow$ (c) (note the change in x-axis scale), and also on approach to emergence (d) $\leftarrow$ (f). For all runs, the integration time step was set at $\Delta t=0.01$ ms, but the spatial resolution Δx and rod length $L=N\Delta x$ were adjusted to ensure numerical stability and accuracy. The Δx spatial resolutions (in cm) were (a) 0.05, (b) 0.2, (c) 2, (d) 1, (e) 0.2, and (f) 0.1.

$$G_{\mathrm{sim}}(|x-x'|)=\frac{f_{\mathtt{xcorr}}(\mathbf{h}_e)}{N} \tag{4.18}$$

where $\mathbf{h}_e$ is the 100-element vector of $\hat{h}_e$ values along the rod. The average of ten such runs was calculated (thin black line) and compared with the theoretical prediction (thick gray line). Despite the large run-to-run variation in covariance plots within a given panel [particularly evident at the near-transition points A_3 (induction) and Q_1 (emergence)], the ten-run averages of Fig. 4 generally show excellent agreement with theory. We also ran stochastic simulations of the linearized (Langevin) equations (2.16), and, unsurprisingly, these simulations also confirmed excellent agreement with the predicted fluctuation covariance curves (results not shown).

Both sets of simulations confirm the important prediction that the EEG fluctuations should become increasingly correlated over space near the transition points. This result is significant because it may provide clinicians with an early warning that a patient's state of consciousness is about to change.

2. Case $c_3<0$: Strong e→i diffusivity

We then turned our attention to the $c_3<0$ regime. We first ran numerical simulations of the linearized spatioadiabatic equations, attempting to generate the spatially periodic covariance pattern predicted by Eq. (4.15) and characterized by the soft-mode wave vector q_s defined in Eq. (3.10)—but without success. It soon became apparent that the inhibition-dominance condition $c_3<0$ is also the condition for system instability when driven by white noise. Spatial white noise contains *all* wave numbers $0 \leqslant q < \infty$; any wave numbers $q > q_s$ will have eigenmodes that grow exponentially, and therefore the homogeneous stationary state is no longer stable. See the discussion following Eq. (3.10).

Since the homogeneous state is unstable for $c_3<0$, the cortex must move to a new state of lower energy. But where

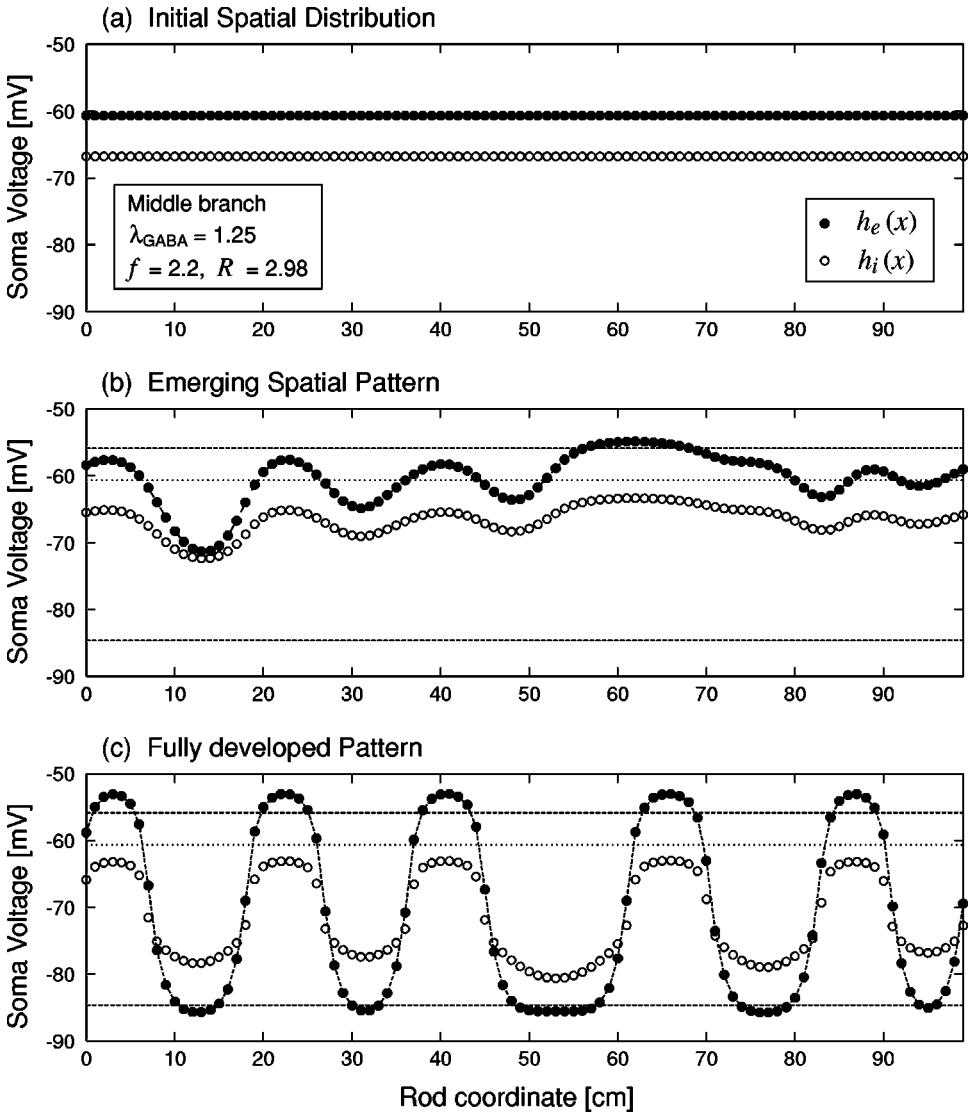

FIG. 5. Simulation run showing emergence of self-organized states in the nonlinearized spatioadiabatic 1D cortical rod. Simulation parameters: $N=100$ points with periodic boundaries, $\Delta x=1$ cm, $\Delta t=0.1$ ms, $\lambda_{\mathrm{GABA}}=1.25$, connectivity scale factor $f=2.20\Rightarrow R=2.98$, noise scale factor $\alpha_{jk}=0.01$. (a) Simulation is started as a homogeneous distribution of soma voltages set at the midbranch (i.e., unstable) steady-state values $h_{e,i}(x)=h_{e,i}^{0,\mathrm{mid}}$. (b) Each cortical point is driven by spatial and temporal white noise, so tends to ''slide off'' the potential crest represented by the midbranch steady state, and ''falls'' toward one or other of the potential valleys near the upper branch ($h_e^{0,\mathrm{top}}$) or lower branch ($h_e^{0,\mathrm{bot}}$) homogeneous steady states (drawn with a dashed pen). (b) shows the voltage pattern after 150 iterations, and (c) shows the fully developed spatial pattern after 5000 iterations.

is this alternative state? And what is its nature? Because linearized theory cannot answer these questions, we ran a series of stochastic simulations of the nonlinearized spatioadiabatic equations (2.1) and (A1),(A2).

We altered the Eq. (3.12) long-range ''wiring'' ratio R in order to make the $e\rightarrow i$ connectivity relatively stronger than the $e\rightarrow e$ connectivity, thereby causing c_3 to become negative. We found that the nonlinear simulation behaviors were dramatically different for this regime: sooner or later the stochastic fluctuations would become completely swamped by the formation of strongly growing diffusive ''fingers'' that would grow toward, and then overshoot, the second homogeneous stable state. For certain settings of λ_{GABA} and R, these diffusion finger-pattern inhomogeneities would persist indefinitely. We found that the easiest way of stimulating pattern formation is to start the cortical rod in its homogeneous configuration on the unstable midbranch. Figure 5 shows a typical $c_3<0$ nonlinearized simulation run for $\lambda_{\mathrm{GABA}}=1.25$, $R=2.98$, and Fig. 6 illustrates the range of diffusive patterns that can evolve. Although each finger has a very similar upside-down catenary- (suspended-chain-)like curvature, the location on the rod at which the fingers form is random, depending delicately on the noise sequence at the onset of pattern formation. But once the rod has formed a dissipative pattern, noise loses its dominant role.

V. DISCUSSION

The controlled induction of anesthetic unconsciousness is sudden, dramatic, and—of vital importance—reversible. The anesthetodynamic model of cortical function we have been developing suggests that this change of brain state can be viewed—at least at a gross electrical level—as a drug-induced first-order phase transition as populations of cooperating neurons switch from an active (''conscious'') state to a quiescent (''unconscious'') state.

Encouragingly, there is an accumulating body of clinical support for the major predictions of the phase-transition theory: (a) EEG power surge at the induction and emergence transition points [7,24]; (b) spectral redistribution toward lower frequencies at the transition, resulting in a reduction in EEG spectral entropy [8]; (c) nonlinear reduction in oxygen consumption (measured via changes in cerebral metabolic rate) as anesthetic concentration is increased [9]; (d) hyster-

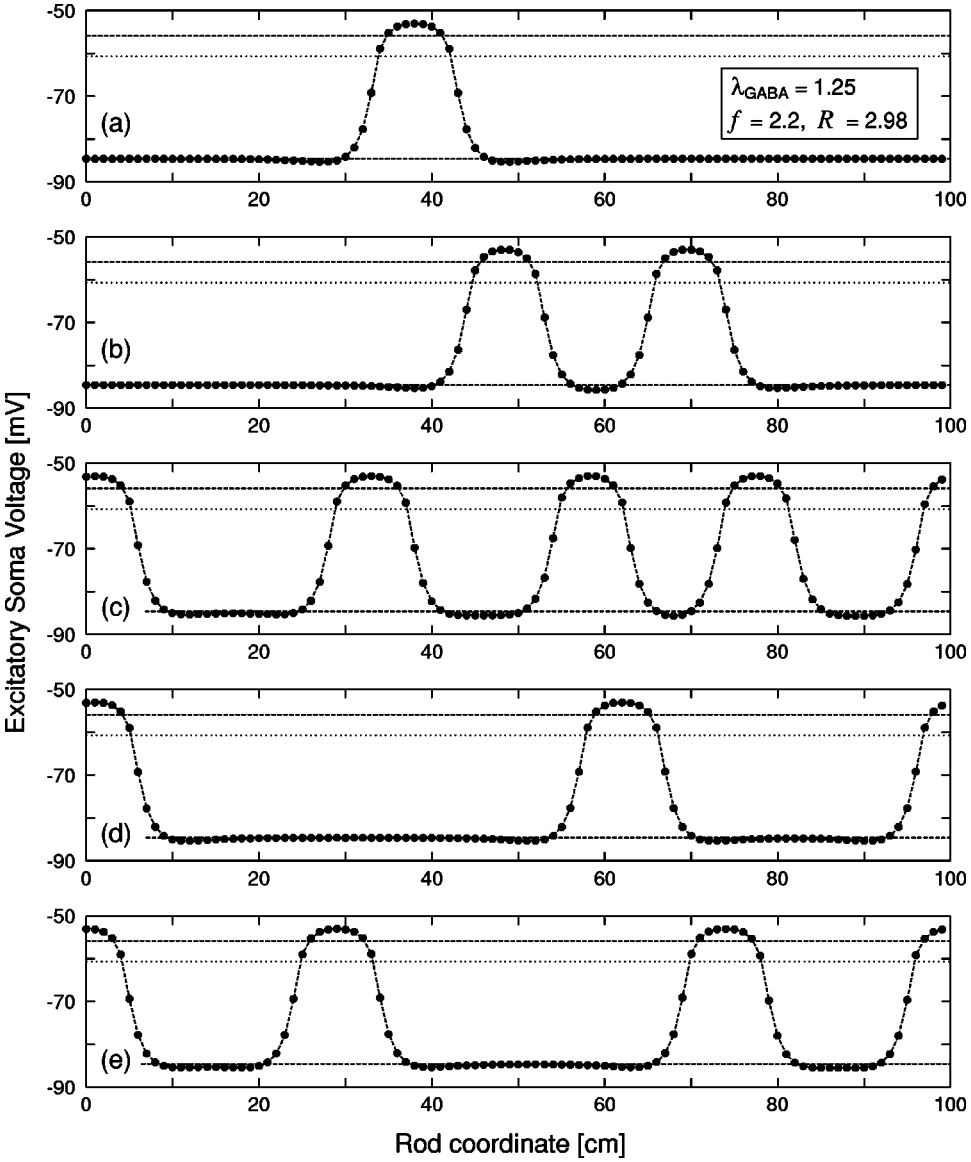

FIG. 6. Selection of excitatory voltage patterns formed spontaneously on the spatioadiabatic 1D cortical rod $\lambda_{GABA}=1.25$, $R=2.98$. Simulation settings are identical to those used for Fig. 5. In each case, the rod was started at its spatially homogeneous mid-branch value (dotted horizontal line), then driven by temporal and spatial white noise. These are the stationary dissipative patterns that have stabilized after 5000 iterations. (Dashed horizontal lines show the upper- and lower-branch homogeneous stationary states.)

esis separation, with respect to drug concentration, between the point of loss of consciousness and the point of recovery of consciousness [7,25]: the patient wakes up at a *lower* level of drug concentration than that required to put her to sleep. However, we need to point out that our interpretation of the hysteresis separation between LOC and ROC (corresponding to points A_3 and Q_1 on Fig. 1)—as a confirmation of our phase-transition theory—is probably controversial. The pharmacokinetic modelers in the anesthetics community would argue that the drug-effect hysteresis is an artifact caused by the fact that the drug *concentration* is (usually) measured in the blood, while the drug's *effect* occurs some time later at the brain, and that this displacement error causes an apparent hysteresis. Our model suggests that even after allowance for displacement error, a real hysteresis should remain, and that attempts to cancel the hysteresis loop by adjusting pharmacokinetic parameters are, in fact, an overcorrection.

Our three earlier papers [1–3] assumed a steady-state, *homogeneous* cortex in which the anesthetic-driven properties of a single macrocolumn could be taken as a proxy for the entire gray-matter cerebral cortex. This is a ''single-electrode'' theory in the sense that the overall average state of the (presumed homogeneous) cortex can, in principle, be determined with a single electrode pair: the reference wire located in the extracellular fluid (defining the zero of potential), and the sensing wire located in the ''population average'' excitatory neuron. Because there is no notion of space in a spatially homogeneous model, the precise location of the sensing electrode is, by assumption, unimportant. We note that existing commercial monitors for depth of anesthesia (e.g., Aspect Systems A-2000, which measures the EEG bispectral index; Datex-Ohmeda Entropy Module, which measures EEG spectral entropy) are *single-channel* devices consisting of an electrode pair giving a differential scalp voltage, plus a third wire for the ground attachment (the ''true'' ground reference—the extracellular fluid potential—is, of course, unavailable to a scalp-based monitor). Thus, both in theory and in clinical practice, a ''single-electrode'' picture provides a reasonable foundation for understanding the electrical properties of general anesthesia.

The present paper represents a first step toward a *multiple-electrode* theory. We accommodate the possibility of spatial inhomogeneity in electrical activity by modeling the cerebral cortex as an infinite 1D rod of macrocolumn ''mass:'' by applying a spatioadiabatic approximation, the cortical rod can develop spatial variations about the homogeneous steady state. We linearized the spatioadiabatic system equations about the homogeneous state, investigated their stability, and discovered the possibility of long-lived spatial modes (''soft-mode instability'') when the long-range (corticocortical) $e \rightarrow e$ diffusivity κ_e was dominated by the long-range $e \rightarrow i$ diffusivity κ_i such that parameter c_3 becomes negative [see Eq. (3.5)]. We verified the existence and longevity of the q_s soft mode via numerical simulation on a noise-free $\sim$100-cm rod with periodic boundary conditions.

We computed the theoretical two-electrode covariance for the EEG fluctuations, $G(|x-x'|)$, and found two different regimes of behavior, depending on the sign of c_3. For $c_3 > 0$ (which includes the default model settings), the covariance will be the difference of two exponential decays in space. The characteristic lengths (correlation lengths) for these decays are predicted to increase strongly on approach to the LOC and ROC transition points [see Figs. 3(a) and 4], and these theoretical predictions were confirmed in a series of stochastic simulations on a joined cortical rod (Fig. 4).

For clinical confirmation of increased correlation lengths near transition, we turn to the recently reported results of John *et al.* [10]. John and colleagues analyzed the quantitative-electroencephalographic changes for 176 patients undergoing general anesthesia. The EEG was recorded from 19 electrodes deployed across the scalp at positions corresponding to the international 10/20 system. A range of different anesthetic agents were used, the focus of the investigation being the identification of EEG changes that are common to all anesthetics.

There are two main findings pertinent to the present work. First, the same qualitative changes in EEG power spectrum were observed in the vicinity of LOC for every anesthetic agent—namely, a large increase in power, particularly at low frequencies, on approach to LOC, followed by a dramatic collapse in high-frequency (γ band, 25–45 Hz) power at LOC. This behavior is consistent with our anesthetodynamic prediction of a pronounced growth in fluctuation power concomitant with a spectral redistribution toward zero frequency (i.e., critical slowing down) on approach to the transition point.

The second finding was a general increase in EEG *coherence* just prior to LOC and again at ROC, particularly for the frontal (forehead) pair of electrodes. While our spatioadiabatic model cannot explain why the frontal electrodes are favored, this enhanced coherence is consistent with a model prediction of increases in correlation length and correlation time for the EEG fluctuations near the phase-change jump points. We are presently investigating the possibility of using the spatioadiabatic model to make quantitative predictions for coherence changes as a function of anesthetic effect.

For $c_3 < 0$, the linear stochastic theory fails because the eigenmodes for high spatial frequencies $q > q_s$ grow without bound. To identify the new state formed in this regime, we ran simulations of the nonlinear stochastic equations and found that, for particular values of the anesthetic effect λ_{GABA} and wiring ratio R, the cortical rod would spontaneously fall away from the homogeneous state, evolving into a stable, pseudoperiodic pattern of large-amplitude voltage excursions. It is interesting to note that in 1980 Ermentrout and Cowan [26] identified the possibility of stationary periodic spatial patterns in their idealized 1D neural net models. They remarked that strong inhibitory influences are necessary in order to generate stable spatial structures; their finding is consistent with the present work, since, in our system, the $c_3 < 0$ condition arises when the long-range $e \rightarrow i$ inhibitory diffusivity κ_i dominates the $e \rightarrow e$ excitatory diffusivity κ_e.

The biological significance of dissipative structures with respect to anesthesia is unclear. One might be tempted to suggest that the existence of these persistent pseudoperiodic spatial structures shows that, at a given point in time, spatially separated cortical regions can appear to participate in the same neural process with zero time lag. But it seems unlikely that any such ''neural process'' could be associated with normal consciousness since this is a *frozen pattern* of brain activity (a neural ''crystal'') that would severely constrain dynamic communication attempts from nonparticipating neuronal assemblies. Perhaps the structures might ''unfreeze'' into temporal oscillations (hard-mode instabilities) or traveling-wave patterns if we were to relax the spatioadiabatic requirement. This possibility will be investigated in future work.

APPENDIX: SPATIOADIABATIC APPROXIMATION

As discussed in the Introduction, the spatially homogeneous adiabatic approximation adopted in papers [1–3] enables a valid description of the gross EEG changes associated with loss of consciousness. In these earlier papers we assumed that the I_{jk} and ϕ_j inputs equilibrate on time scales much faster than the average soma voltages h_e and h_i, and that the cerebral cortex is spatially uniform (this is the mean-field approximation). These assumptions were effected by setting $\partial/\partial t \rightarrow 0$ (adiabaticity) and $\partial^2/\partial x^2 \rightarrow 0$ (spatial homogeneity) in Eqs. (2.2) and (2.3) and then substituting the resulting steady-state values for the cortical inputs (plus noise) back into the (2.1) equations of motion for the excitatory and inhibitory soma voltages. The result was a pair of coupled stochastic DEs in h_e and h_i alone.

In the present paper, we wish to allow for the development of spatial variability in long-range corticocortical firing rate, with consequent variations in space for the macrocolumn soma voltage. We accomplish this, to first order, by adopting the ''spatioadiabatic'' limit which sets $\partial/\partial t \rightarrow 0$ in Eqs. (2.2) and (2.3), while retaining the $\partial^2/\partial x^2$ terms in Eqs. (2.3). Thus the cortical inputs of Eqs. (2.2) reduce to

$$I_{ee}(h_e) = [N_{ee}^{\beta} S_e(h_e) + \phi_e + \langle p_{ee} \rangle] G_e e/\gamma_e + \Gamma_1(x,t)/\gamma_e^2, \tag{A1a}$$

$$I_{ei}(h_e) = [N_{ei}^{\beta} S_e(h_e) + \phi_i + \langle p_{ei} \rangle] G_e e/\gamma_e + \Gamma_2(x,t)/\gamma_e^2, \tag{A1b}$$

$$I_{ie}(h_i)=\lambda_{\mathrm{GABA}}[N_{ie}^{\beta}S_i(h_i)+\langle p_{ie}\rangle]G_i e/\gamma_i + \lambda_{\mathrm{GABA}}^2\Gamma_3(x,t)/\gamma_i^2, \tag{A1c}$$

$$I_{ii}(h_i)=\lambda_{\mathrm{GABA}}[N_{ii}^{\beta}S_i(h_i)+\langle p_{ii}\rangle]G_i e/\gamma_i + \lambda_{\mathrm{GABA}}^2\Gamma_4(x,t)/\gamma_i^2. \tag{A1d}$$

We set to zero the sigmoid time derivatives $\partial S_e/\partial t$ in Eqs. (2.3) to obtain the spatioadiabatic approximation for the $\phi_{i,e}$ long-range inputs,

$$\phi_e=\frac{1}{\Lambda_{ee}^2}\frac{\partial^2\phi_e}{\partial x^2}+N_{ee}^{\alpha}S_e(h_e), \tag{A2a}$$

$$\phi_i=\frac{1}{\Lambda_{ei}^2}\frac{\partial^2\phi_i}{\partial x^2}+N_{ei}^{\alpha}S_e(h_e). \tag{A2b}$$

This neglect of the sigmoid time derivatives is equivalent to assuming that the impulse response (Green's function) $\Phi_{e,i}$ for the $S_e(h_e)$ source is felt instantaneously throughout the cortical rod: whereas the Green's functions for the original Eqs. (2.3) represent pulses decaying exponentially in space and propagating at speed v,

$$\Phi_e(x,t)=\frac{1}{2}N_{ee}^{\alpha}\Lambda_{ee}e^{-\Lambda_{ee}|x|}\delta(t-|x|/v), \tag{A3a}$$

$$\Phi_i(x,t)=\frac{1}{2}N_{ei}^{\alpha}\Lambda_{ei}e^{-\Lambda_{ei}|x|}\delta(t-|x|/v), \tag{A3b}$$

our adiabatic approximation replaces these impulse functions with their $v\to\infty$ limit,

$$\Phi_e(x,t)=\tfrac{1}{2}N_{ee}^{\alpha}\Lambda_{ee}e^{-\Lambda_{ee}|x|}\delta(t), \tag{A4a}$$

$$\Phi_i(x,t)=\tfrac{1}{2}N_{ei}^{\alpha}\Lambda_{ei}e^{-\Lambda_{ei}|x|}\delta(t), \tag{A4b}$$

with the spatioadiabatic forms of Eqs. (A2) being recovered after convolving, over time and space, the Eq. (A4) adiabatic Green's functions,

$$\phi_{e,i}(x,t)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Phi_{e,i}(x-x',t-t')S_e(x',t')dx'\,dt'. \tag{A5}$$

Substituting Eqs. (A1) and (A2) back into Eqs. (2.1) gives us a pair of coupled spatioadiabatic equations for h_e and h_i. However, Eqs. (A2) still contain terms involving $\phi_{e,i}$ on the right-hand side, so the resulting equations for $h_{e,i}$ will also be dependent on $\phi_{e,i}$.

We now remove this ϕ dependence by making a small-perturbation expansion about the homogeneous steady state. First, we observe that in the homogeneous $(\partial^2/\partial x^2\to 0)$ steady-state limit, Eqs. (A2) predict

$$\phi_e^{\mathrm{eq}}=N_{ee}^{\alpha}S_e(h_e)|_{\mathrm{eq}}, \tag{A6a}$$

$$\phi_i^{\mathrm{eq}}=N_{ei}^{\alpha}S_e(h_e)|_{\mathrm{eq}}. \tag{A6b}$$

This means that in the spatially homogeneous steady-state cortex, $\phi_e(h_e)$ and $\phi_i(h_i)$ are both functions of h_e. Therefore, provided that the perturbations away from the homogeneous steady state always remain small (i.e., h_e is a *weak* function of position x, and never far from equilibrium), ϕ_e and ϕ_i can still be written as functions of h_e:

$$\phi_e\equiv\phi_e[h_e(x)],$$

$$\phi_i\equiv\phi_i[h_e(x)].$$

Close to homogeneous equilibrium we may make a Taylor expansion to nth order,

$$\phi_e[h_e(x)]=\phi_e^{\mathrm{eq}}+[h_e(x)-h_e^{\mathrm{eq}}]\left.\frac{d\phi_e}{dh_e}\right|_{\mathrm{eq}} + \frac{1}{2!}[h_e(x)-h_e^{\mathrm{eq}}]^2\left.\frac{d^2\phi_e}{dh_e^2}\right|_{\mathrm{eq}} + \cdots+\frac{1}{n!}[h_e(x)-h_e^{\mathrm{eq}}]^n\left.\frac{d^n\phi_e}{dh_e^n}\right|_{\mathrm{eq}}. \tag{A7}$$

We obtain an expression for $d\phi_e^n/dh_e^n$ by operating with d/dh_e on Eq. (A2a), giving

$$\frac{d\phi_e}{dh_e}=\frac{1}{\Lambda_{ee}^2}\frac{d}{dh_e}\left[\frac{\partial^2\phi_e}{\partial x^2}\right]+N_{ee}^{\alpha}\frac{d}{dh_e}S_e(h_e), \tag{A8}$$

which, at homogeneous equilibrium, simplifies to

$$\left.\frac{d\phi_e}{dh_e}\right|_{\mathrm{eq}}=N_{ee}^{\alpha}\left.\frac{dS_e}{dh_e}\right|_{\mathrm{eq}} \tag{A9}$$

since $(\partial^2\phi_e/\partial x^2)|_{\mathrm{eq}}=0$. Consequently,

$$\left.\frac{d^n\phi_e}{dh_e^n}\right|_{\mathrm{eq}}=N_{ee}^{\alpha}\left.\frac{d^nS_e}{dh_e^n}\right|_{\mathrm{eq}}. \tag{A10}$$

Substituting Eq. (A10) back into Eq. (A7) gives a near-equilibrium Taylor expansion for ϕ_e in terms of derivatives of S_e, the voltage-to-pulse-rate sigmoidal transfer function,

$$\phi_e[h_e(x)]=\phi_e^{\mathrm{eq}}+[h_e(x)-h_e^{\mathrm{eq}}]N_{ee}^{\alpha}\left.\frac{dS_e}{dh_e}\right|_{\mathrm{eq}} + \frac{1}{2!}[h_e(x)-h_e^{\mathrm{eq}}]^2N_{ee}^{\alpha}\left.\frac{d^2S_e}{dh_e^2}\right|_{\mathrm{eq}}+\cdots + \frac{1}{n!}[h_e(x)-h_e^{\mathrm{eq}}]^nN_{ee}^{\alpha}\left.\frac{d^nS_e}{dh_e^n}\right|_{\mathrm{eq}}. \tag{A11a}$$

Similarly, if ϕ_i can also be assumed to be a function of $h_e(x)$ only, we can write a corresponding Taylor expansion for $\phi_i[h_e(x)]$,

$$\phi_i[h_e(x)] = \phi_i^{\mathrm{eq}} + [h_e(x) - h_e^{\mathrm{eq}}] N_{ei}^{\alpha} \left.\frac{dS_e}{dh_e}\right|_{\mathrm{eq}} + \frac{1}{2!}[h_e(x) - h_e^{\mathrm{eq}}]^2 N_{ei}^{\alpha} \left.\frac{d^2 S_e}{dh_e^2}\right|_{\mathrm{eq}} + \cdots + \frac{1}{n!}[h_e(x) - h_e^{\mathrm{eq}}]^n N_{ei}^{\alpha} \left.\frac{d^n S_e}{dh_e^n}\right|_{\mathrm{eq}}. \tag{A11b}$$

The theoretical calculations for spatial covariance in Sec. IV of this paper utilize a linearized paradigm, so only the first-order terms of Eqs. (A11) need be retained. Therefore, when we operate with d^2/dx^2 on both sides of Eqs. (A11), we retain only the first-order terms, giving

$$\frac{\partial^2 \phi_e}{\partial x^2} \approx N_{ee}^{\alpha} \left.\frac{dS_e}{dh_e}\right|_{\mathrm{eq}} \frac{\partial^2 h_e}{\partial x^2}, \tag{A12a}$$

$$\frac{\partial^2 \phi_i}{\partial x^2} \approx N_{ei}^{\alpha} \left.\frac{dS_e}{dh_e}\right|_{\mathrm{eq}} \frac{\partial^2 h_e}{\partial x^2}. \tag{A12b}$$

Hence, close to homogeneous equilibrium, we can replace the long-range inputs of Eqs. (A2) by their linearized approximations,

$$\phi_e \approx \frac{N_{ee}^{\alpha}}{\Lambda_{ee}^2} \left.\frac{dS_e}{dh_e}\right|_{\mathrm{eq}} \frac{\partial^2 h_e}{\partial x^2} + N_{ee}^{\alpha} S_e(h_e), \tag{A13a}$$

$$\phi_i \approx \frac{N_{ei}^{\alpha}}{\Lambda_{ei}^2} \left.\frac{dS_e}{dh_e}\right|_{\mathrm{eq}} \frac{\partial^2 h_e}{\partial x^2} + N_{ei}^{\alpha} S_e(h_e). \tag{A13b}$$

Equations (A13a) and (A13b) can now be used to approximate ϕ_e and ϕ_i in Eqs. (A1a) and (A1b), which are then substituted into Eqs. (2.1a) and (2.1b) to obtain the final spatioadiabatic reduced equations (2.10a) and (2.10b).

[1] M.L. Steyn-Ross, D.A. Steyn-Ross, J.W. Sleigh, and D.T.J. Liley, Phys. Rev. E **60**, 7299 (1999).

[2] M.L. Steyn-Ross, D.A. Steyn-Ross, J.W. Sleigh, and L.C. Wilcocks, Phys. Rev. E **64**, 011917 (2001).

[3] D.A. Steyn-Ross, M.L. Steyn-Ross, L.C. Wilcocks, and J.W. Sleigh, Phys. Rev. E **64**, 011918 (2001).

[4] N.P. Franks and W.R. Lieb, Nature (London) **367**, 607 (1994).

[5] H.R. Wilson and J.D. Cowan, Biophys. J. **12**, 1 (1972).

[6] D.A. Steyn-Ross, Ph.D. thesis, University of Waikato, Hamilton, New Zealand, 2002.

[7] K. Kuizenga, C.J. Kalkman, and P.J. Hennis, Br. J. Anaesth. **80**, 725 (1998).

[8] H.E. Viertiö-Oja *et al.*, Int. J. Clin. Monit Comput. **16**, 60 (2000).

[9] E.H. Stullken, Jr., J.H. Milde, J.D. Michenfelder, and J.H. Tinker, Anesthesiology **46**, 28 (1977).

[10] E.R. John *et al.*, Conscious Cogn **10**, 165 (2001).

[11] L.E. Reichl, *A Modern Course in Statistical Physics*, 2nd ed. (Wiley, New York, 1998).

[12] P.A. Robinson, C.J. Rennie, and J.J. Wright, Phys. Rev. E **56**, 826 (1997).

[13] P.L. Nunez, Math. Biosci. **21**, 279 (1974).

[14] P. L. Nunez, in *Neocortical Dynamics and Human EEG Rhythms*, edited by P. L. Nunez (Oxford University Press, New York, 1995), Chaps. 9 and 11.

[15] V.K. Jirsa and H. Haken, Phys. Rev. Lett. **77**, 960 (1996).

[16] V.K. Jirsa and H. Haken, Physica D **99**, 503 (1997).

[17] V.K. Jirsa and J.A.S. Kelso, Phys. Rev. E **62**, 8462 (2000).

[18] D.T.J. Liley, P.J. Cadusch, and J.J. Wright, Neurocomputing **26-27**, 795 (1999).

[19] H.C. Tuckwell, *Introduction to Theoretical Neurobiology: Nonlinear and Stochastic Theories* (Cambridge University Press, Cambridge, England, 1988), Vol. 2.

[20] A.L. Hodgkin and A.F. Huxley, J. Physiol. (London) **117**, 500 (1952).

[21] S. Chaturvedi, C.W. Gardiner, I.S. Matheson, and D.F. Walls, J. Stat. Phys. **17**, 469 (1977).

[22] M.R. Spiegel, *Mathematical Handbook of Formulas and Tables*, *Schaum's Outline Series* (McGraw-Hill, New York, 1968).

[23] C.W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, Vol. 13 of *Springer Series in Synergetics* (Springer-Verlag, Berlin, 1985).

[24] K. Kuizenga, J.M.K.H. Wierda, and C.J. Kalkman, Br. J. Anaesth. **86**, 354 (2001).

[25] K. Kuizenga, J.H. Proost, J.M.K.H. Wierda, and C.J. Kalkman, Anesthesiology **95**, 607 (2001).

[26] G.B. Ermentrout and J.D. Cowan, SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **38**, 1 (1980).