Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: Evidence for a general anesthetic-induced phase transition

Moira L. Steyn-Ross and D. A. Steyn-Ross
Department of Physics and Electronic Engineering, Private Bag 3105, University of Waikato, Hamilton, New Zealand

J. W. Sleigh
Department of Anaesthetics, Waikato Hospital, Hamilton, New Zealand

D. T. J. Liley
School of Biophysical Sciences and Electrical Engineering, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia
(Received 30 April 1999; revised manuscript received 30 August 1999)

We present a model for the dynamics of a cerebral cortex in which inputs to neuronal assemblies are treated as random Gaussian fluctuations about a mean value. We incorporate the effect of general anesthetic agents on the cortex as a modulation of the inhibitory neurotransmitter rate constant. Stochastic differential equations are derived for the state variable h_e, the average excitatory soma potential, coherent fluctuations of which are believed to be the source of scalp-measured electroencephalogram (EEG) signals. Using this stochastic approach we derive a stationary (long-time limit) fluctuation spectrum for h_e. The model predicts that there will be three distinct stationary (equilibrium) regimes for cortical activity. In region I (''coma''), corresponding to a strong inhibitory anesthetic effect, h_e is single valued, large, and negative, so that neuronal firing rates are suppressed. In region II for a zero or small anesthetic effect, h_e can take on three values, two of which are stable; we label the stable solutions as ''active'' (enhanced firing) and ''quiescent'' (suppressed firing). For region III, corresponding to negative anesthetic (i.e., analeptic) effect, h_e again becomes single valued, but is now small and negative, resulting in strongly elevated firing rates (''seizure''). If we identify region II as associated with the conscious state of the cortex, then the model predicts that there will be a rapid transit between the active-conscious and comatose unconscious states at a critical value of anesthetic concentration, suggesting the existence of phase transitions in the cortex. The low-frequency spectral power in the h_e signal should increase strongly during the initial stage of anesthesia induction, before collapsing to much lower values after the transition into comatose-unconsciousness. These qualitative predictions are consistent with clinical measurements by Bührer *et al.* [Anaesthesiology **77**, 226 (1992)], MacIver *et al.* [*ibid.* **84**, 1411 (1996)], and Kuizenga *et al.* [Br. J. Anaesthesia **80**, 725 (1998)]. This strong increase in EEG spectral power in the vicinity of the critical point is similar to the divergences observed during thermodynamic phase transitions. We show that the divergence in low-frequency power in our model is a natural consequence of the existence of turning points in the trajectory of stationary states for the cortex. [S1063-651X(99)08312-9]

PACS number(s): 87.19.La, 05.10.Gg, 05.70.Fh

I. INTRODUCTION

A standard method for following the anesthetic induction of a patient into unconsciousness is to monitor the electroencephalogram (EEG) signals detected by electrodes attached to the scalp. We aim to develop a theory which models the dominant electrorhythmogenic processes occurring in the cerebral cortex as general anesthetic is administered. Such a theory would be useful not only for quantifying at what point a patient might be considered to be sufficiently anethetised to safely undergo surgery, but also to give better understanding of cortical function and dynamics in general. A reasonable test of the theory would ask the following: Does it predict the kinds of changes in EEG spectral distribution and power which are observed in patients during induction of general anesthesia?

It is well known within the anesthesiology community than many commonly used general-anesthetic agents exhibit what is referred to as a ''biphasic'' or activation-depression response: at low (sedative) anesthetic concentrations there is a significant increase above baseline values in both the total EEG power and in the frequency at which peak power occurs; as concentration is further increased to hypnotic (surgical anesthesia) levels, the total power and median frequency fall away to levels below baseline. This ''biphasic'' response has been observed on human volunteers dosed with thiopental [1] and the widely-used propofol [3]. It has also been measured in rats dosed with thiopental [2,4]. Figure 1 shows a typical activation/depression response from one of the patients in the Kuizenga *et al.* [3] study.

The EEG signal originates from organized assemblies of excitatory and inhibitory neural cells (neurons) acting cooperatively within a small volume of the cortex [5]. Figure 2 is a schematic representation of such an assembly which can be thought of as occupying a cylindrical column of diameter ~ 0.3–1 mm and containing 40 000–100 000 neurons. The excitatory (pyramidal) cells make up $\sim 85\%$ of the total number of neurons [6].

The EEG is generated by the longitudinal flow of current along the apical (superficial layer) dendrites of pyramidal neurons which are aligned with an axial symmetry perpendicular to the cortical surface [7]. The potential due to the

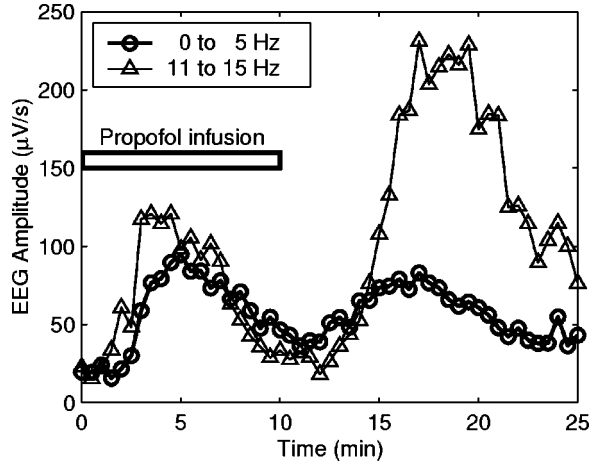

FIG. 1. Biphasic effect of propofol anesthetic on 0–5- and 11–15-Hz EEG signals. During the 10 min of propofol infusion, the anesthetic concentration increases steadily. At low concentrations, the EEG signal shows an initial increase in power (activation). EEG power then falls away (inhibition) as the concentration is further increased and the patient becomes deeply unconscious. A second EEG activation peak is observed as the anesthetic concentration declines and the patient begins to emerge from unconsciousness. (Data supplied courtesy of K. Kuizenga, and reported as ''patient 7'' in Kuizenga *et al.* [3].)

distributed current sources and sinks induced by afferent (incoming) synaptic activity along these aligned pyramidal dendrites can be approximated at the cortical surface by a dipole term. The deviation from rest of the mean excitatory soma membrane potential $h_e - h_e^{\text{rest}}$ has been demonstrated to be proportional to the mirror (i.e., sign-reversed) image of the extracellular local field potential (LFP) [8]. Because the EEG is a spatially smoothed version of the LFP, it is reasonable to assume that it will be proportional to h_e.

In contrast, the inhibitory neurons, comprising 15% of the neural population, are smaller and have their dendrites oriented at random with approximately spherical symmetry, so their equivalent dipole term will be vanishingly small. The resulting synaptic currents induced in the dendrites of the inhibitory cells make negligible contribution to the EEG and ECoG (electrocorticogram) signal.

Because cooperative neural activity is maintained via dense synaptic interconnections, one assumes that cortical parameters can be expressed as values averaged over the assembly. This approach of treating assemblies (also called centers or macrocolumns) of correlated cells is referred to as the mean-field or mass-action formalism, and has an extensive history with significant contributions by Freeman [8], Wilson and Cowan [9,10], Nunez [7], Robinson and coworkers [11,12], Wright and Liley [13], Rotterdam *et al.* [14], Amit [15], and Jirsa and Haken [16]. Robinson and co-workers [11,12] used a mean-field approach when deriving a set of nonlinear equations to describe the generation of electrical waves in the cortex responsible for the EEG signal. Their two-dimensional (2D) continuum model contained excitatory and inhibitory neural populations, and included the effects of axonal conduction delays.

Liley [17] extended these theories by improving the treatment of excitatory and inhibitory neurotransmitter kinetics. He derived a set of integrodifferential equations which give the time variation of the mean excitatory and inhibitory soma potentials of an assembly responding to external inputs and local feedbacks. While these equations have been found to reproduce a range of experimental results, they are mathematically formidable, making it difficult to extract physical insight into the underlying neural processes. To remedy this, Liley simplified the model by reducing its dimensionality and size to represent a 1D neural assembly whose activity can be taken as approximately constant over spatial scales of the order of the intracortical (submillimetric) connectivity. The result was a set of eight coupled partial differential equations [18,19] which give the time development of h_e and h_i for a neural aggregate whose inputs are defined in terms of sigmoidal nonlinear functions. A complete solution of these equations for a specified input yields, as a function of time, the mean soma membrane potential of excitatory neurons, interpreted as the scalp-recordable EEG signal.

In this paper we transform Liley's deterministic partial differential equations (PDEs) into a set of stochastic differential equations (SDEs), also referred to as Langevin equations. This is done by incorporating noise terms, assumed to originate from random fluctuations in the subcortical inputs, into the equations of motion for $h_{e,i}$. This enables us to derive a stationary spectrum for h_e. The Langevin formalism is used in many areas of physics, e.g., in quantum optics [20], to predict emission spectra of atoms interacting with electromagnetic radiation. Jirsa and Haken [21] and Frank *et al.* [22] also use this approach in their modeling of dendritic currents in the cortex.

Expressed in general form, the Langevin equations of state for the excitatory and inhibitory soma potentials $h_{e,i}$ can be written

$$\frac{d}{dt}\begin{bmatrix} h_e \\ h_i \end{bmatrix} = -\begin{bmatrix} A_e(h_e,h_i) \\ A_i(h_e,h_i) \end{bmatrix} + \begin{bmatrix} B_e(\xi_e(t)) \\ B_i(\xi_i(t)) \end{bmatrix}, \tag{1.1}$$

in which $A_{e,i}$ are *drift* terms describing the mean or average behavior of the $h_{e,i}$, and $B_{e,i}$ are the corresponding *diffusion* terms which describe the response of the system to random fluctuations. $\xi(t)$ is a Gaussian white-noise source which has zero mean and is δ-correlated:

$$\langle \xi_\eta(t) \rangle = 0, \quad \langle \xi_\eta(t)\xi_{\eta'}(t') \rangle = \delta_{\eta\eta'}\delta(t-t'). \tag{1.2}$$

Setting the noise terms to zero in Eq. (1.1) gives the deterministic equation for $h_{e,i}$:

$$\frac{d}{dt}\begin{bmatrix} h_e \\ h_i \end{bmatrix} = -\begin{bmatrix} A_e(h_e,h_i) \\ A_i(h_e,h_i) \end{bmatrix}. \tag{1.3}$$

After a sufficiently long time, the system is assumed to settle into its equilibrium state so that the time derivatives on the left of Eq. (1.3) can be set to zero. Thus solving for $A_{e,i} = 0$ gives the equilibrium state values $h_{e,i}^0$ of the cortex.

Having found the stationary state, we linearize the system about this state by writing $h_{e,i}$ as the sum of its ''dc'' (low-frequency or equilibrium) component plus small amplitude ''ac'' fluctuations about this mean value: $h_{e,i} = h_{e,i}^0 + \tilde{h}_{e,i}$. This decomposition enables us to transform Eq. (1.1) into a set of linear SDEs,

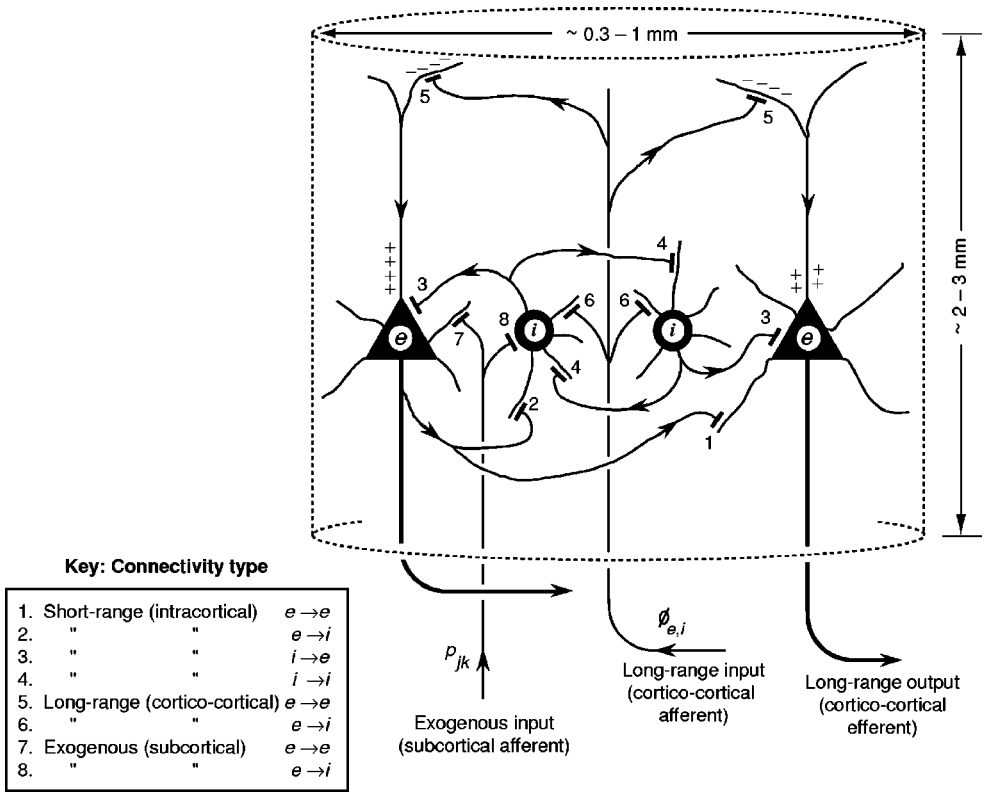

FIG. 2. Schematic representation of the connective topology within a cortical macrocolumn. Only four of the ~100 000 neurons are shown. Triangles are excitatory (pyramidal) cells which receive excitatory input via apical dendrites (e.g., connection type 5) and basal dendrites (1, 7); and inhibitory input directly at the cell body (3). Circles are inhibitory (stellate or basket) cells receiving input from dendritic connections (2, 4, 6) and at the cell body (8). The excitatory output from the macrocolumn is shown bold via trunk lines (axons). The symbol $\phi_{e,i}$ represents $e\rightarrow e$, $e\rightarrow i$ input from distant macrocolumns, and p_{jk} represents input from the subcortex (e.g., thalamus and brainstem). (For clarity, we have omitted p_{ie} and p_{ii} exogenous inputs corresponding to connection types 9 and 10, respectively.)

$$\frac{d}{dt}\begin{bmatrix}\tilde{h}_e\\ \tilde{h}_i\end{bmatrix} = -\mathbf{A}\begin{bmatrix}\tilde{h}_e\\ \tilde{h}_i\end{bmatrix} + \mathbf{B}\begin{bmatrix}\xi_e\\ \xi_i\end{bmatrix}, \tag{1.4}$$

where $\tilde{h}_{e,i}$ are linearized white-noise-driven fluctuations about the stationary solution; **A** and **B** are matrices containing the linearized drift and diffusion coefficients. The power spectrum for fluctuations about the stationary state can then be derived by following standard methods of stochastic calculus [20].

Note that our stochastic approach relies on two fundamental assumptions: (a) that there exists (at least one) well-defined equilibrium state of the cortex; and (b) that an EEG spectrum can be produced by driving this equilibrium state with white noise. We observe that the notion of a stationary state for cortical activity has already been invoked by Robinson and co-workers [11,12], who assumed that such states are meaningful over timescales much longer than dendritic integration times (i.e., $\gg 5-10$ ms). Other workers who have presented white-noise-driven model EEG spectra include Rotterdam *et al.* [14], Nunez [7], Liley [17], Jirsa and Haken [21], and Frank *et al.* [22]. The latter pair of cited references take dendritic current as the state variable, permitting model comparisons with observed magnetoencephalograms (MEGs).

In Sec. II we present the Liley differential equations (DEs) for a cortical assembly, and discuss how anesthetic effect can be modelled in terms of changes to the inhibitory neurotransmitter rate constant. We show how the Liley equations can be transformed into a set of first-order stochastic DEs with the appropriate inclusion of white-noise terms. This allows us to compute both the anesthetic-modulated trajectory of steady states and the corresponding EEG spectrum for small fluctuations about these states.

In Sec. III we give the model predictions and compare these with clinical measurements by other workers. Our model predicts that there will be either one or three stationary (equilibrium) states for h_e as a function of anesthetic amount. For the three-state case, only two are stable; we identify these two states as ''activated'' and ''quiescent.'' The existence of an intermediate, third, state which is unstable to fluctuations allows for transition between the activated and quiescent states at a critical value of anesthetic, suggesting the possibility of an identifiable phase transition in the cortex. The model predictions for anesthetic-modulated changes to EEG spectral power show a clear biphasic (activation followed by inhibition) anesthetic response, with good qualitative agreement with the experimental work of Kuizenger *et al.* [3].

Jirsa and Haken [16] suggested the possibility of a phase transition in the brain after observing MEG patterns of human volunteers taking part in movement coordination experiments. The subjects were required to press a button in response to an acoustical stimulus. When the frequency of stimulus presentation exceeded a critical value, the subjects' movements switched from a deliberated manual action to an

TABLE I. Symbol definitions and given values for model constants.

Symbol	Description	Value	Unit
e, i (as subscripts)	excitatory, inhibitory cell populations		
$h_{e,i}$	population mean soma potential		mV
$\tau_{e,i}$	membrane time constant	40, 40	ms
$h_{e,i}^{\text{rest}}$	cell resting potential	-70, -70	mV
$h_{e,i}^{\text{rev}}$	cell reversal potential (Nernst potential)	45, -90	mV
$I_{ee,ie}$	total $e\to e, i\to e$ ''current'' input to excitatory synapses		mV
$I_{ei,ii}$	total $e\to i, i\to i$ ''current'' 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	1.1, 1.6	$(\text{ms})^{-1}$
$p_{ei,ii}$	exogenous (subcortical) spike input to i population	1.6, 1.1	$(\text{ms})^{-1}$
$\phi_{e,i}$	long-range (cortico-cortical) spike input to e, i populations		$(\text{ms})^{-1}$
$\Lambda_{ee,ei}$	characteristic cortico-cortical inverse-length scale	0.40, 0.65	$(\text{cm})^{-1}$
EPSP, IPSP	excitatory, inhibitory post-synaptic potential		mV
$\gamma_{e,i}$	neurotransmitter rate constant for EPSP, IPSP	0.30, 0.065	$(\text{ms})^{-1}$
$G_{e,i}$	peak amplitude of EPSP, IPSP	0.18, 0.37	mV
e	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	0.7	$\text{cm}\,(\text{ms})^{-1}$
$S_e(h_e), S_i(h_i)$	sigmoid function mapping soma potential to firing rate		$(\text{ms})^{-1}$
$\theta_{e,i}$	inflection-point voltage for sigmoid function	-60, -60	mV
$g_{e,i}$	sigmoid slope at inflection point	0.28, 0.14	$(\text{mV})^{-1}$

involuntary synchronous response. This change in movement response was accompanied by an alteration in the recorded spatiotemporal MEG patterns. To explain these findings, Jirsa and Harken [16] developed a field-theoretical model of the brain, subsequently extended by Frank *et al.* [22] to include white noise. This model predicts that at a critical driving frequency there will be a phase transition in the spatiotemporal distribution of the dendritic currents. We note that their model describes a transition between different *conscious* states of the brain, whereas our present work is concerned with the general-anesthetic-induced phase transition between conscious and unconscious states.

In Sec. IV we discuss the implications of our findings with respect to analogies between classical phase transitions in physics and state changes in the cortex. We define a cortex cooperativity parameter, analogous to the order parameter of a thermodynamic phase transition, and offer some conjectures about how these ideas might relate to ''consciousness.''

II. THEORY

A. Cortical equations

Our starting point is Liley's set of eight coupled PDEs [19,23] in which we have assumed complete spatial homogeneity over the region sampled by the EEG electrode. This is a reasonable approximation, given that a scalp electrode has a contact area of approximately 2 cm^2, and thus detects electrical activity averaged across the underlying 5–10 cm^2 of cerebral cortex. Thus the one-dimensional Laplacian $\partial^2/\partial x^2$ [which would have appeared on the left-hand side of the equation for the long-range potential $\phi(x,t)$; Eq. (2.4) below] is eliminated, and all partial derivatives with time become total derivatives with time. This gives the following set of eight coupled ordinary DEs (the symbols are defined in Table I):

$$\begin{bmatrix}\tau_e & 0\\ 0 & \tau_i\end{bmatrix}\frac{d}{dt}\begin{bmatrix}h_e\\ h_i\end{bmatrix}=\begin{bmatrix}h_e^{\text{rest}}-h_e\\ h_i^{\text{rest}}-h_i\end{bmatrix}+\begin{bmatrix}\psi_{ee}I_{ee}+\psi_{ie}I_{ie}\\ \psi_{ei}I_{ei}+\psi_{ii}I_{ii}\end{bmatrix}, \quad (2.1)$$

$$\left(\frac{d}{dt}+\gamma_e\right)^2\begin{bmatrix}I_{ee}\\ I_{ei}\end{bmatrix}=\left\{\begin{bmatrix}N_{ee}^{\beta}\\ N_{ei}^{\beta}\end{bmatrix}S_e(h_e)+\begin{bmatrix}\phi_e\\ \phi_i\end{bmatrix}+\begin{bmatrix}p_{ee}\\ p_{ei}\end{bmatrix}\right\}G_e\gamma_e e, \quad (2.2)$$

$$\left(\frac{d}{dt}+\gamma_i\right)^2\begin{bmatrix}I_{ie}\\ I_{ii}\end{bmatrix}=\left\{\begin{bmatrix}N_{ie}^{\beta}\\ N_{ii}^{\beta}\end{bmatrix}S_i(h_i)+\begin{bmatrix}p_{ie}\\ p_{ii}\end{bmatrix}\right\}G_i\gamma_i e, \quad (2.3)$$

$$\begin{bmatrix}\left(\frac{d}{dt}+\overline{v}\Lambda_{ee}\right)^2\phi_e\\ \left(\frac{d}{dt}+\overline{v}\Lambda_{ei}\right)^2\phi_i\end{bmatrix}=\overline{v}\begin{bmatrix}\left(\frac{d}{dt}+\overline{v}\Lambda_{ee}\right)\Lambda_{ee}N_{ee}^{\alpha}\\ \left(\frac{d}{dt}+\overline{v}\Lambda_{ei}\right)\Lambda_{ei}N_{ei}^{\alpha}\end{bmatrix}S_e(h_e). \quad (2.4)$$

Equation (2.1) gives the time evolution of h_e and h_i, the excitatory and inhibitory soma potentials averaged over the assembly of cooperating neurons. The neural assembly is assumed to be a single resistance-capacitance (RC) compartment or summing point; in effect, we are defining an average neuron for the mass. The first two terms on the right corre-

spond to an exponential return to a resting potential $h_{e,i}^{\mathrm{rest}}$; the second pair describe perturbations to the membrane potential due to synaptic inputs to the neural mass. The ψ_{jk} (where $j,k \in \{e,i\}$) coefficients appearing on the right are normalized weighting functions for these inputs. These coeffcients represent the facts that excitation and inhibition are mediated by different ionic species and that the corresponding magnitude of the postsynaptic currents will depend on the active state of the neuron [24]; they are defined by

$$\psi_{ee}=\frac{h_e^{\mathrm{rev}}-h_e}{|h_e^{\mathrm{rev}}-h_e^{\mathrm{rest}}|}, \quad \psi_{ie}=\frac{h_i^{\mathrm{rev}}-h_e}{|h_i^{\mathrm{rev}}-h_e^{\mathrm{rest}}|},$$
$$\psi_{ei}=\frac{h_e^{\mathrm{rev}}-h_i}{|h_e^{\mathrm{rev}}-h_i^{\mathrm{rest}}|}, \quad \psi_{ii}=\frac{h_i^{\mathrm{rev}}-h_i}{|h_i^{\mathrm{rev}}-h_i^{\mathrm{rest}}|}. \tag{2.5}$$

The constant values used for the resting $h_{e,i}^{\mathrm{rest}}$ and reversal $h_{e,i}^{\mathrm{rev}}$ potentials are listed in Table I. Note that for typical values for h_e and h_i, the weights ψ_{ee} and ψ_{ei} for input from excitatory sources are positive, while weights ψ_{ie} and ψ_{ii} from inhibitory sources are negative.

The time evolution of the input terms I_{ee}, I_{ie}, I_{ei}, and I_{ii} is governed by Eqs (2.2) and (2.3) which model the variable coupling strength between cells in terms of sigmoid functions $S_e(h_e)$ and $S_i(h_i)$:

$$S_e(h_e)=[1+\exp(-g_e(h_e-\theta_e))]^{-1},$$
$$S_i(h_i)=[1+\exp(-g_i(h_i-\theta_i))]^{-1}. \tag{2.6}$$

These are nonlinear S-shaped transfer functions representing the output pulse rate (in, say, pulses per second) of a homogeneous neural mass in response to a mean field of h_e, h_i. $\theta_{e,i}$ and $g_{e,i}$ are constants: $\theta_{e,i}$ is the soma potential at which the function has maximum gradient, and $g_{e,i}$ determines the ''gain'' at this point of inflection. See Fig. 3, and refer to Table I for values of the contants. For small values of soma potential, the average firing rate is low; as soma potential increases (becomes less negative), firing rate increases rapidly, eventually levelling off at a maximum value of say, 1000 s^{-1}. Thus the strength of the interconnection between neurons is determined by the value of the soma potential at that instant. In addition to sigmoid-modulated spike input from the neural mass, there are exogenous (subcortical) spike input contributions (p_{ee}, p_{ie}, p_{ei}, p_{ii}), plus long-range (cortico-cortical) contributions (ϕ_e, ϕ_i) from distant excitatory assemblies.

It is of interest to note that Eqs. (2.1)–(2.4) have some similarities with those derived by Robinson and co-workers [11,12], and Jirsa and Haken [16]. Robinson and co-workers wrote differential equations for $V_{e,i}$, the neuronal potential at the cell body, in terms of inputs determined by arrival rates of pulses at dendrites, and used a sigmoid function to relate input voltage to neuronal firing rate.

In our present work we wish to modify the Liley equations in order to model the effect of variable anesthetic concentration in the cortex. The primary mechanism of action common to most general anesthetic agents is the prolonging of the duration of the inhibitory postsynaptic potentials (IPSPs) [25] [or, equivalently, the reduction in neurotransmitter rate constant γ_i in Eq. (2.3)]. At concentrations appropriate for surgical anesthesia, the IPSPs are prolonged by a factor of 1.5- to 4-fold [25–28].

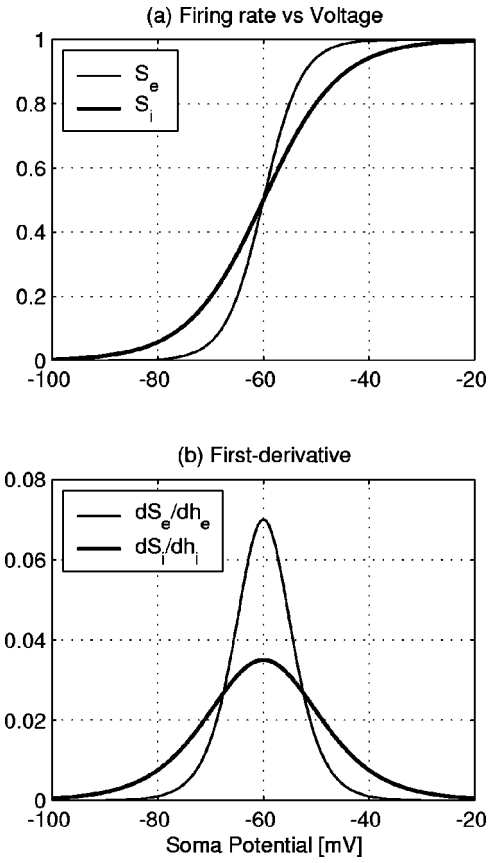

FIG. 3. Sigmoidal functions relating the firing rate to the average soma potential. (a) Sigmoid curves: excitatory sigmoid S_e (light curve); inhibitory sigmoid S_i (bold). (b) First derivative of sigmoid functions: dS_e/dh_e (light curve); dS_i/dh_i (bold). The points of inflection are set at $\theta_{e,i}=-60$ mV; the midpoint ''gains'' are $g_{e,i}=0.28$ and 0.14 $(\mathrm{mV})^{-1}$ [see Eq. (2.6)].

We model this change in inhibitory rate constant by replacing the γ_i appearing on the left-hand side of Eq. (2.3) with $\overline{\gamma}_i$, where

$$\overline{\gamma}_i=\frac{\gamma_i}{\lambda}.$$

Here λ is a multiplicative scaling factor assumed to be proportional to anesthetic concentration, so that $\lambda=1$ corresponds to no anesthetic effect, and an increase in λ corresponds to an increase in anesthetic amount (decrease in γ_i rate constant). See Fig. 4. We now describe how the Liley equations are transformed into linearized stochastic differential equations.

B. Stochastic differential equations (SDEs)

1. System fluctuations

As a first step toward deriving stochastic equations of motion, we need to identify the sources of noise which drive the system. We assume that the noise arises in the subcortical (exogenous) inputs to the assembly, and ignore noise enter-

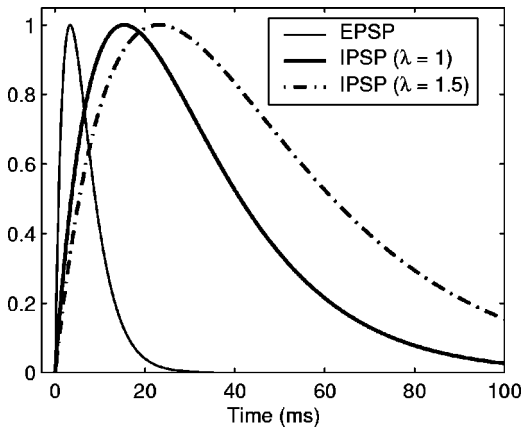

FIG. 4. Impulse response for excitatory (light curve), inhibitory (bold), and anesthetic-modified inhibitory (bold-dashed) post-synaptic membranes. Curves are normalized to unit height. For application to our model, the heights are scaled by the respective EPSP (excitatory post-synaptic potential) and IPSP (inhibitory post-synaptic potential) amplitudes, $G_{e,i}=0.18$ and 0.37 mV. The rate constants, in $(\mathrm{ms})^{-1}$, for the three curves are $\gamma_e=0.30$, $\gamma_i=0.065$, and $\gamma_i'=0.043$.

ing via the long-range (cortico-cortical) connections from distant assemblies. This assumption is modeled by replacing each of the four p_{jk} subcortical sources appearing in Eqs (2.2) and (2.3) by the product of its average value $\langle p_{jk}\rangle$ with a unit-variance white-noise term $(1+\xi_{jk}(t))$, e.g.,

$$p_{ie}\rightarrow\langle p_{ie}\rangle(1+\xi_{ie}(t))$$

(We note that this is akin to the phenomenological inclusion of cortical noise as proposed by Frank *et al.* [22].) Thus Eqs (2.2) and (2.3) are rewritten as

$$\left(\frac{d}{dt}+\gamma_e\right)^2\begin{bmatrix}I_{ee}\\ I_{ei}\end{bmatrix}=\left\{\begin{bmatrix}N_{ee}^{\beta}\\ N_{ei}^{\beta}\end{bmatrix}S_e(h_e)+\begin{bmatrix}\phi_e\\ \phi_i\end{bmatrix}+\begin{bmatrix}\langle p_{ee}\rangle\\ \langle p_{ei}\rangle\end{bmatrix}\right\}G_e\gamma_e e+\begin{bmatrix}\Gamma_1(t)\\ \Gamma_2(t)\end{bmatrix},\tag{2.7}$$

$$\left(\frac{d}{dt}+\gamma_i\right)^2\begin{bmatrix}I_{ie}\\ I_{ii}\end{bmatrix}=\left\{\begin{bmatrix}N_{ie}^{\beta}\\ N_{ii}^{\beta}\end{bmatrix}S_i(h_i)+\begin{bmatrix}\langle p_{ie}\rangle\\ \langle p_{ii}\rangle\end{bmatrix}\right\}G_i\gamma_i e+\begin{bmatrix}\Gamma_3(t)\\ \Gamma_4(t)\end{bmatrix},\tag{2.8}$$

where

$$\begin{bmatrix}\Gamma_1(t)\\ \Gamma_2(t)\end{bmatrix}=\begin{bmatrix}\langle p_{ee}\rangle\xi_1(t)\\ \langle p_{ei}\rangle\xi_2(t)\end{bmatrix}G_e\gamma_e e,$$

$$\begin{bmatrix}\Gamma_3(t)\\ \Gamma_4(t)\end{bmatrix}=\begin{bmatrix}\langle p_{ie}\rangle\xi_3(t)\\ \langle p_{ii}\rangle\xi_4(t)\end{bmatrix}G_i\gamma_i e,\tag{2.9}$$

and the four $\xi_j(t)$ are Gaussian random terms as defined by Eqs (1.2). [We do not include any explicit noise terms in Eqs (2.1) and (2.4), so these remain unaltered.]

2. Adiabatic elimination of fast variables

Our aim is to use Eqs (2.1), (2.4), (2.7), and (2.8) to compute observable quantities such as power spectra. Solving these equations in their full form will require numerical simulation using stochastic-integration techniques, and these can be fraught with stability problems. Instead, for a first approach, we prefer to make some reasonable simplifications which will both permit analytic solution and also give some insights into predicted system behaviors.

The simplification is possible if we perform a linearized analysis which is based on the assumption that an equilibrium state of the cortex exists, and is given by solving Eqs. (2.1), (2.4), (2.7), and (2.8) in the steady-state limit (i.e., $d/dt\rightarrow 0$) in which all noise terms have been set to zero. This gives the stationary solution, which we denote by the vector

$$\boldsymbol{\alpha}^0=[h_e^0 h_i^0 I_{ee}^0 I_{ie}^0 I_{ei}^0 I_{ii}^0 \phi_e^0 \phi_i^0]^{\mathrm{T}}.$$

Having solved for the equilibrium state $\boldsymbol{\alpha}^0$, we can linearize Eqs. (2.1), (2.4), (2.7), and (2.8) about $\boldsymbol{\alpha}^0$, and, by casting them into a set of first-order differential equations, obtain a complete set of stationary statistics such as correlation functions and power spectra [20].

However, the required calculations present a formidable task, since they involve manipulations of several multidimensional matrices. We can reduce the dimensionality of the problem, thereby making it more tractable, by noting that the ''input'' terms (I_{jk}, $\phi_{e,i}$) can vary on time scales that are quite distinct from the time scale of the soma potentials h_e and h_i. This becomes apparent when we compare the various relaxation times (computed from the numerical values listed in Table I):

$$\text{relaxation time for } I_{ee}, I_{ei}=(\gamma_e)^{-1}\approx 3.3 \text{ ms},$$

$$\text{relaxation time for } I_{ie}, I_{ii}=(\gamma_i)^{-1}\approx 15.4 \text{ ms},$$

$$\text{relaxation time for } \phi_e=(\overline{\nu}\Lambda_{ee})^{-1}\approx 3.6 \text{ ms},$$

$$\text{relaxation time for } \phi_i=(\overline{\nu}\Lambda_{ei})^{-1}\approx 2.2 \text{ ms},$$

whereas the $\tau_{e,i}$ time scales for the $h_{e,i}$ soma potentials can be as large as 100 ms [29]. For our present modeling work we set $\tau_e=\tau_i=40$ ms, allowing us to make the working assumption that the six neuronal inputs $[I_{ee}I_{ie}I_{ei}I_{ii}\phi_e\phi_i]$ equilibrate very much faster than the soma potentials $h_{e,i}$ themselves, so that on $h_{e,i}$ equilibration time scales, all time derivatives appearing in Eqs. (2.2)–(2.4) can be set to zero. We thus adiabatically eliminate these ''fast'' variables by setting $d/dt\rightarrow 0$ in Eqs. (2.2)–(2.4) while *retaining* the noise terms, allowing us to solve for I_{ee}, I_{ie}, I_{ei}, I_{ii}, ϕ_e, and ϕ_i as functions of h_e and h_i. The resulting expressions for these six fast variables may then be substituted back into the equations of motion (2.1) for h_e and h_i.

Note that in contrast to the procedure for determining the stationary solutions $\boldsymbol{\alpha}^0$, we do not set the noise terms to zero in the adiabatic elimination, since we wish to allow fluctuations from the fast variables to be incorporated into the $h_{e,i}$ equations. (We note in passing that while Gardiner [20] warns that this method for treating noise is only valid for small fluctuations, it has been used with success by many workers in the field of quantum optics, e.g., by Haken [30,31] in his treatment of the laser, and by Drummond [32] in his work on cooperative fluorescence.)

The equations resulting from the adiabatic simplification follow:

$$\begin{bmatrix} I_{ee}^0 \\ I_{ei}^0 \end{bmatrix} = \left\{ \begin{bmatrix} N_{ee}^\beta \\ N_{ei}^\beta \end{bmatrix} S_e(h_e) + \begin{bmatrix} \phi_e^0 \\ \phi_i^0 \end{bmatrix} + \begin{bmatrix} \langle p_{ee} \rangle \\ \langle p_{ei} \rangle \end{bmatrix} \right\} G_e e/\gamma_e + \begin{bmatrix} \Gamma_1(t) \\ \Gamma_2(t) \end{bmatrix} \Big/ \sqrt{\gamma_e^2}, \tag{2.10a}$$

$$\begin{bmatrix} I_{ie}^0 \\ I_{ii}^0 \end{bmatrix} = \left\{ \begin{bmatrix} N_{ie}^\beta \\ N_{ii}^\beta \end{bmatrix} S_i(h_i) + \begin{bmatrix} \langle p_{ie} \rangle \\ \langle p_{ii} \rangle \end{bmatrix} \right\} G_i e/\overline{\gamma_i} + \begin{bmatrix} \Gamma_3(t) \\ \Gamma_4(t) \end{bmatrix} \Big/ \sqrt{\overline{\gamma_i^2}}, \tag{2.10b}$$

$$\begin{bmatrix} \phi_e^0 \\ \phi_i^0 \end{bmatrix} = \begin{bmatrix} N_{ee}^\alpha \\ N_{ei}^\alpha \end{bmatrix} S_e(h_e). \tag{2.10c}$$

Substituting Eqs (2.10) back into Eqs (2.1), we obtain the stochastic equations of motion for the soma potentials in the adiabatic limit:

$$\frac{d}{dt} \begin{bmatrix} h_e \\ h_i \end{bmatrix} = \begin{bmatrix} F_1(h_e, h_i) \\ F_2(h_e, h_i) \end{bmatrix} + \begin{bmatrix} \Gamma_e(t) \\ \Gamma_i(t) \end{bmatrix}, \tag{2.11a}$$

where the drift terms are

$$F_1(h_e, h_i) = \{(h_e^{\text{rest}} - h_e) + \psi_{ee}[(N_{ee}^\alpha + N_{ee}^\beta) S_e(h_e) + \langle p_{ee} \rangle] G_e e/\gamma_e + \lambda \psi_{ie}[N_{ie}^\beta S_i(h_i) + \langle p_{ie} \rangle] G_i e/\gamma_i\}/\tau_e, \tag{2.11b}$$

$$F_2(h_e, h_i) = \{(h_i^{\text{rest}} - h_i) + \psi_{ei}[(N_{ei}^\alpha + N_{ei}^\beta) S_e(h_e) + \langle p_{ei} \rangle] G_e e/\gamma_e + \lambda \psi_{ii}[N_{ii}^\beta S_i(h_i) + \langle p_{ii} \rangle] G_i e/\gamma_i\}/\tau_i, \tag{2.11c}$$

and the corresponding noise terms are

$$\Gamma_e(t) = \{\psi_{ee} \langle p_{ee} \rangle \xi_1(t) G_e e/\gamma_e + \lambda \psi_{ie} \langle p_{ie} \rangle \xi_3(t) G_i e/\gamma_i\}/\tau_e, \tag{2.11d}$$

$$\Gamma_i(t) = \{\psi_{ei} \langle p_{ei} \rangle \xi_2(t) G_e e/\gamma_e + \lambda \psi_{ii} \langle p_{ii} \rangle \xi_4(t) G_i e/\gamma_i\}/\tau_i. \tag{2.11e}$$

(Note that we have replaced $\overline{\gamma_i}$ by γ_i/λ in the above equations in order to make explicit their dependence on anesthetic ''effect'' λ.)

C. Fluctuation spectrum: linearized theory

We linearize SDE's (2.11) about an equilibrium state $\boldsymbol{\alpha}^0$ to obtain the Ito SDE

$$\frac{d}{dt} \begin{bmatrix} \tilde{h}_e \\ \tilde{h}_i \end{bmatrix} = -\mathbf{A} \begin{bmatrix} \tilde{h}_e \\ \tilde{h}_i \end{bmatrix} + \begin{bmatrix} \Gamma_e(t) \\ \Gamma_i(t) \end{bmatrix} \tag{2.12}$$

where $\tilde{h}_{e,i}$ represent small deviations of the $h_{e,i}$ from the equilibrium state. The drift matrix $\mathbf{A}$ is given by

$$\mathbf{A} = -\begin{bmatrix} \frac{\partial F_1}{\partial h_e} & \frac{\partial F_1}{\partial h_i} \\ \frac{\partial F_2}{\partial h_e} & \frac{\partial F_2}{\partial h_i} \end{bmatrix}_{\text{eq.}}, \tag{2.13}$$

where the eq. subscript means ''evaluate at the equilibrium point.'' Since the SDE is now in Ito form, we may define an equivalent Fokker–Planck equation [20]

$$\frac{\partial P(\tilde{h}_e, \tilde{h}_i)}{\partial t} = \left\{ \frac{\partial}{\partial \tilde{h}_e}[A_{11}\tilde{h}_e + A_{12}\tilde{h}_i] + \frac{\partial}{\partial \tilde{h}_i}[A_{21}\tilde{h}_e + A_{22}\tilde{h}_i] + \frac{1}{2}\left[\frac{\partial^2}{\partial \tilde{h}_e^2} D_{11} + \frac{\partial^2}{\partial \tilde{h}_i^2} D_{22} \right] \right\} P(\tilde{h}_e, \tilde{h}_i), \tag{2.14}$$

where P is the probability distribution function for the $\tilde{h}_{e,i}$. The D_{jj} are the elements of the diffusion (noise) matrix defined via

$$\langle \Gamma_e(t) \Gamma_e(t') \rangle = D_{11} \delta(t - t'), \tag{2.15a}$$

$$\langle \Gamma_i(t) \Gamma_i(t') \rangle = D_{22} \delta(t - t'), \tag{2.15b}$$

$$D_{12} = D_{21} = 0. \tag{2.15c}$$

(The full form of the drift and diffusion matrices is given in the Appendix.)

Equation (2.14) describes a multivariate Ornstein-Uhlenbeck process, the stationary statistics of which have been extensively studied [20]. In particular, if we define the time autocorrelation for $\tilde{h}_e$ as

$$\mathcal{G}(t') = \lim_{T \to \infty} \frac{1}{T} \int_0^T \tilde{h}_e(t) \tilde{h}_e(t + t') dt; \tag{2.16}$$

then the stationary fluctuation spectrum for $\tilde{h}_e$ can be computed from the Fourier transform

$$\mathcal{S}[\tilde{h}_e(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t'} \mathcal{G}(t') dt'. \tag{2.17}$$

Using standard Ornstein–Uhlenbeck analysis [20], we can derive the spectrum $\mathcal{S}[\tilde{h}_e(\omega)]$ and the covariance matrix $\boldsymbol{\sigma}$ in terms of the drift and diffusion matrices $\mathbf{A}$ and $\mathbf{D}$:

$$\mathcal{S}[\tilde{h}_e(\omega)] = \frac{1}{2\pi} (\mathbf{A} + i\omega \mathbf{I})^{-1} \, \mathbf{D} \, (\mathbf{A}^T - i\omega \mathbf{I})^{-1}, \tag{2.18}$$

where the superscript T signifies a matrix transpose. The stationary covariance matrix is

$$\boldsymbol{\sigma} = \begin{bmatrix} \langle \tilde{h}_e, \tilde{h}_e \rangle & \langle \tilde{h}_e, \tilde{h}_i \rangle \\ \langle \tilde{h}_i, \tilde{h}_e \rangle & \langle \tilde{h}_i, \tilde{h}_i \rangle \end{bmatrix} = \frac{\det(\mathbf{A}) \mathbf{D} + [\mathbf{A} - \text{Tr}(\mathbf{A}) \mathbf{I}] \, \mathbf{D} \, [\mathbf{A} - \text{Tr}(\mathbf{A}) \mathbf{I}]^T}{2 \, \text{Tr}(\mathbf{A}) \det(\mathbf{A})}, \tag{2.19}$$

in which $\mathbf{I}$ is the identity matrix; det and Tr are the determinant and trace operators respectively; and where, for example,

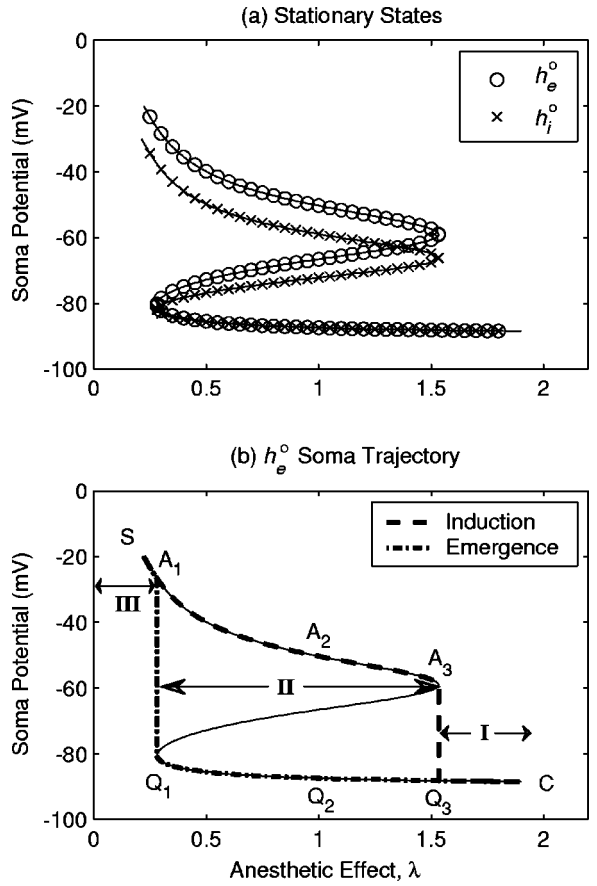

FIG. 5. (a) Model predictions for the stationary states for h_e (circles) and h_i (crosses) as a function of scale factor λ (anesthetic amount). Ordinate axes carry units of mV. (b) In region II bounded by $A_1A_3Q_3Q_1$, for a given value of λ, there are three possible values for h_e, but only two of these are stable: points lying on the upper (''active'': A_1A_3) branch, and points on the lower (''quiescent'': Q_1Q_3) branch. For $\lambda \gtrsim 1.53$ (region I), h_e becomes single valued and neural firing is strongly suppressed (''coma'': Q_3C); for $\lambda \lesssim 0.3$ (region III), h_e is again single valued but now neural firing is maximized (''seizure'': SA_1).

$$\langle \tilde{h}_e, \tilde{h}_e \rangle = \mathrm{var}(\tilde{h}_e) = \langle (\tilde{h}_e - \langle \tilde{h}_e \rangle)^2 \rangle = \langle \tilde{h}_e^2 \rangle - \langle \tilde{h}_e \rangle^2. \tag{2.20}$$

Fluctuations of h_e about its stationary state are thus given by

$$\Delta h_e = \sqrt{\mathrm{var}(\tilde{h}_e)}. \tag{2.21}$$

III. RESULTS

A. Stationary solutions

By solving Eqs. (2.1)–(2.4) in the steady-state, zero-noise limit, we obtain the equilibrium behavior of h_e^0 and h_i^0 as a function of anesthetic ''amount'' λ. See Fig. 5. These steady-state values were obtained numerically by locating the multiple intersections of the isocline curves $dh_e/dt = 0$ and $dh_i/dt = 0$. (See Wilson and Cowan [9], Fig. 4, for an illustration.) Values for our system constants follow closely those of Liley [19], and are listed in Table I. Recall that in our model, the inhibitory neurotransmitter rate constant is assumed to scale inversely with λ, $\overline{\gamma_i} = \gamma_i/\lambda$, thus $\lambda = 1$ corresponds to no anesthetic effect. Figure 5(b) shows that for this value of λ, there are three distinct values for h_e^0: A_2 on the SA_3 upper branch, Q_2 on the Q_1C lower branch, and an intermediate value (unlabeled) on the A_3Q_1 middle branch. From stability analysis, we have established that only the upper and lower branches contain stable equilibrium points; all points along the intermediate branch are unstable with respect to small perturbations.

Suppose the neural assembly is initially at location A_2. As anesthetic effect λ increases, h_e^0 will slide to the right down the upper branch to A_3, whereupon a sudden jump to Q_3 on the lower branch must occur, since the middle branch is unstable and therefore disallowed. Further increases in λ will then cause h_e^0 to advance along the Q_3C subbranch. If instead, the assembly was initially at Q_2 on the lower branch, then increases in λ would lead to smoothly decreasing (more negative) values for soma potential, with no jump discontinuity.

The points on the SA_3 upper branch correspond to very strong neural firing, since along this branch the soma potential exceeds the sigmoidal inflection-point voltage ($\theta_e = -60$ mV; see Fig. 3); thus we refer to the upper-branch states as being ''active.'' Maximum activity will occur at S (upper-left corner) when soma potential is least negative; we refer to the SA_1 subbranch (region III) as ''seizure .'' The Q_1C lower-branch states have large negative soma potentials, and therefore suppressed firing rates, so we label this quiet branch ''quiescent.'' Maximum suppression occurs at C (lower-right corner), so the Q_3C subbranch (region I) is labeled ''coma.''

If the cortex is pictured as a superposition of neural assemblies, some active and some quiescent, then even if only a small proportion are in the activated state, we might expect an anesthetic-driven downwards transition across the A_3Q_3 gap to produce a measurable change in the EEG signal if the active assemblies are acting synchronously. We make some theoretical predictions about the nature of these spectral changes in the following subsection.

The existence of multiple stationary states in the cortex was first suggested by Wilson and Cowan [9]. In their abstract model of populations of inhibitory and excitatory neurons containing sigmoid nonlinearities, they demonstrated that for sigmoid functions with n inflection points, there could be up to $2n+3$ stationary, but not neccessarily concurrent, states. Recently, Robinson *et al.* [12] investigated the nature of the steady-state solutions for a similar mathematical model of the cortex, but, after an extensive parameter space search, rather than five equilibrium states, they found a maximum of either three steady states or a single steady state; and that for the three-state case, only two were stable. This finding is in complete accord with our results reported here.

Examining the results of Robinson *et al.* [12] in more detail, they classified their solutions in terms of a ratio l_i/l_e, where l_i (l_e) is the net response at the cell body per unit concentration of inhibitory (excitatory) neurotransmitter at the synapses. They found that the three-state case occurred when $l_i/l_e \approx 1$, i.e., when the inhibitory and excitatory responses were of similar magnitude. However, if the inhibitory response was strongly dominant over excitatory (or vice

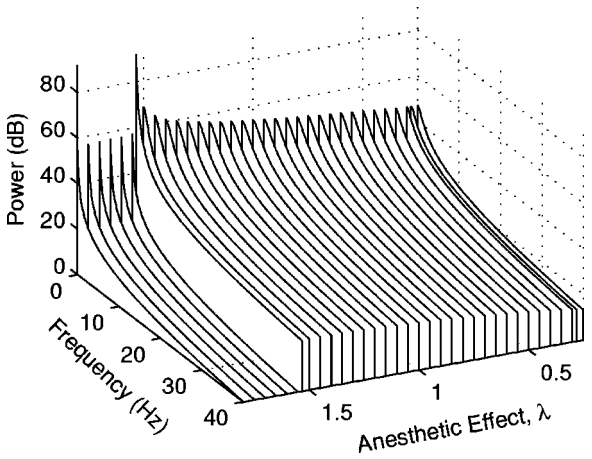

FIG. 6. 3D plot showing variation of spectral power along the anesthesia-induction trajectory $A_1A_3Q_3C$ of Fig. 5. As λ increases from 0.3 to 1.53, low-frequency spectral power increases to reach a strong maximum at the critical point marked A_3 on the upper branch of Fig. 5. A further increase in λ produces an abrupt drop in power as the soma potential transits the $A_3 \rightarrow Q_3$ jump to reach the lower branch. [The power scale is dB relative to the $1.28 \times 10^{-8}(\mathrm{mV})^2/\mathrm{Hz}$ coma minimum at 40 Hz.]

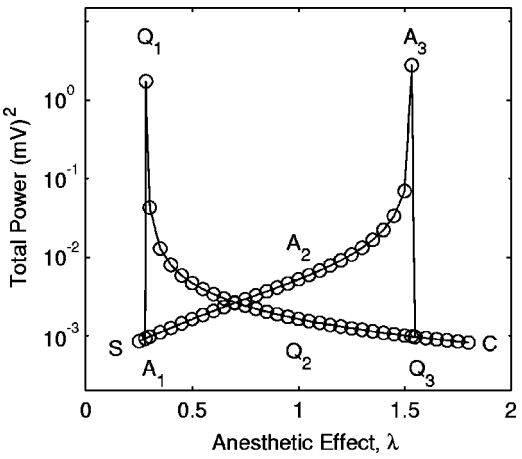

FIG. 7. Total power as a function of anesthetic effect. Power values were obtained by integrating the induction (Fig. 6) and emergence (not shown) spectral power curves over the frequency range 0–40 Hz. The labels correspond to those used in Fig. 5.

versa), they found that the system collapsed to a single steady state.

Relating these findings to our model, their l_i (l_e) ''net response'' concept would seem to correspond to our IPSP amplitude G_i (EPSP amplitude G_e). In our case, we maintained these amplitudes constant (see Fig. 4), and instead increased the inhibitory effectiveness by prolonging the inhibitory neurotransmitter time constant (by reducing its inverse, the rate constant γ_i) by scaling it with anesthetic factor, λ. Thus, broadly speaking, our λ maps to the l_i/l_e ratio of Robinson *et al.* [12] since $\lambda \gg 1$ corresponds to strong inhibition (leading to ''coma''), while at the opposite extreme, $\lambda \ll 1$ corresponds to excessive excitation (leading to ''seizure'').

B. Spectrum for fluctuations about the steady state

For each of the $(h_e^0,\ h_i^0)$ equilibrium states marked (as circles and crosses) on the upper and lower branches of Fig. 5(a), we solved Eq. (2.18) for the fluctuation spectrum over the frequency range 0–40 Hz. Figure 6 shows the predicted variation in spectral power for a macrocolumn whose inhibitory neurotransmitter time constant is multiplied by a λ factor which increases steadily from 0.3 to 1.8. This corresponds to induction of anesthesia via the trajectory $A_1A_3Q_3C$ from the upper (''active'') to the lower (''quiescent'') branch of Fig. 5(b).

Each spectral curve is peaked at zero frequency, with power diminishing smoothly with frequency. There is no suggestion of any cortical resonances (such as the 8–13 Hz alpha rhythm) in these curves; this lack of higher-frequency structure is not unexpected given the approximations we have made (linearization about equilibrium, adiabatic elimination of fast variables).

The interesting feature is the very strong increase in low-frequency power as the turning point at $\lambda \approx 1.53$ [A_3 in Fig. 5(b)] is approached. When λ is increased beyond this critical value, the macrocolumn suddenly collapses to its quiescent state with much reduced spectral power.

In Fig. 7 we show the total power (area under each of the spectral power curves) as a function of anesthetic effect λ for both the induction trajectory ($A_1A_3Q_3C$) and the emergence-from-anesthesia trajectory ($Q_3Q_1A_1$). The two cusps correspond to the two turning points (A_3 and Q_1) in the stationary states trajectory of Fig. 5(b). Figure 8 shows the corresponding steady-state noise amplitude as a function of anesthetic effect. The shapes of Figs 7 and 8 are rather similar because the zero-frequency peak dominates all of the spectral power curves.

How well do these theoretical curves match up with clinical measurements? Kuizenga *et al.* [3] performed a clinical study which examined the ''biphasic'' relationship between the concentration of a general anesthetic agent (propofol) in arterial blood and EEG effects during the transition from the awake state to hypnosis and during subsequent emergence. The subjects were ten healthy male patients who were sched-

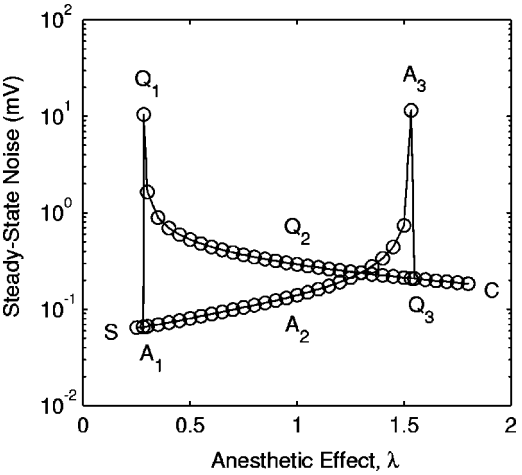

FIG. 8. Steady-state noise amplitude as a function of anesthetic effect. The ordinate is the spectral amplitude at zero frequency, and can be interpreted as a vertical ''error bar'' to be applied to the h_e^0 steady-state trajectory of Fig. 5. Maximum noise occurs at the A_3 and Q_1 critical points.

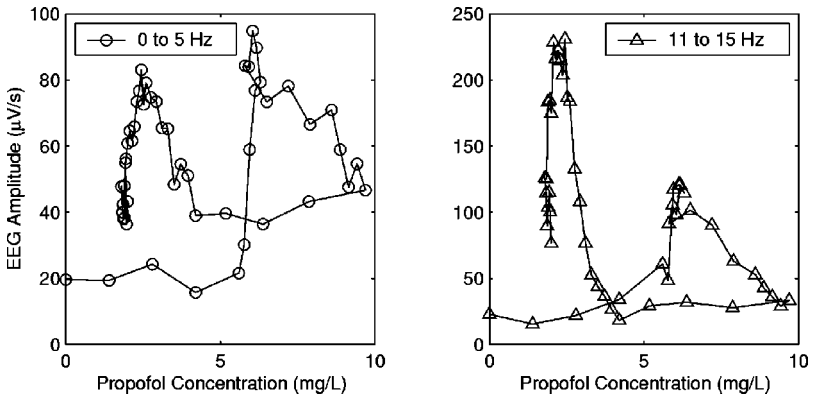

FIG. 9. EEG amplitude data from Fig. 1 plotted as a function of measured propofol blood concentration. (a) 0–5 Hz. (b) 11–15 Hz. Each trajectory commences at the lower-left corner at zero concentration. For the 0–5-Hz band, the activation peak is stronger during the induction phase (right-hand peak); for the 11–15-Hz band, the activation peak is considerably stronger for the emergence phase (left-hand peak). (Data supplied courtesy of K. Kuizenga, and reported as ''patient 7'' in Kuizenga *et al.* [3].)

uled for lower-limb surgery. A scalp electrode pair was placed at the mastoid (bone behind the ear) and the forehead to monitor the differential EEG signal developed across the hemisphere. Each patient received a 10-min infusion of propofol. The EEG was recorded continuously from 5 min before the start of propofol infusion until the patient regained consciousness (approximately 15 min after conclusion of infusion), and thereafter intermittently for 5-min periods, coinciding with blood sampling, until 190 min after start of infusion. Blood samples were drawn from a femoral artery at 2-min intervals during the first 22 min, then at more widely spaced intervals thereafter.

The EEG signal was processed, over 15-s epochs, into of six frequency bands (0–5, 6–10, 11–15, 16–20, 21–25, and 26–30 Hz) using ''aperiodic analysis.'' This technique measures the vertical distance between consecutive peaks and valleys in the voltage trace and computes an effective instantaneous frequency from (half the reciprocal of) the time interval for the peak-to-trough excursion. These voltage excursions are then accumulated, unsigned, into one of the six frequency bins to give a total voltage deviation in each frequency band for the 15-s epoch. Dividing each band total by 15 s then gives a measure of the average amplitude ''slew rate,'' in μV/s, which Kuizenga *et al.* referred to as the ''EEG amplitude.''

Figure 1 shows the time course of EEG activity for the 0–5- and 11–15-Hz bands for patient 7 of the Kuizenga *et al.* study, and Fig. 9 shows the same information, but now plotted as a function of propofol concentration at the femoral artery. Both bands show a pair of pronounced activation peaks: the first peak occurs during the induction phase as the patient becomes unconscious; the second peak occurs some time later as the patient emerges from unconsciousness. For the 0–5-Hz band, the induction peak is stronger, while for the 11–15-Hz band the emergence peak is strongly dominant.

The detection of two activation peaks, one during induction of anesthesia and the second during emergence from anesthesia, provides encouraging qualitative agreement between the clinical results with the steady-state model predictions of Figs 7 and 8. There are two important and not unexpected quantitative differences between theory and experiment, however, which should not go unremarked. First, the model predicts a dynamic range of about 10^4:1 in total power (Fig. 7) and 200:1 in dc amplitude (Fig. 8) while the experiment yields dynamic ranges which are much smaller: 5:1 for the 0–5-Hz band, and 10:1 for the 11–15-Hz band. There is also a scale difference with model voltages being stated in mV, while EEG measurements are in μV. These apparent discrepencies arise because the model is predicting the soma potentials for a single coordinated macrocolumn in the cortex, while the EEG measurement is recording the complex of signals from thousands of macrocolumns in the vicinity of the scalp electrodes, attenuated and filtered by the intervening skull and skin. The fact that the activation peaks can be detected at all suggests that a fraction of the macrocolumns must be acting coherently in the vicinity of the critical point.

The second point of difference concerns the interpretation of anesthetic ''effect.'' In comparing our model with the results of Kuizenga *et al.*, we have implicitly assumed that our λ factor (degree of prolongation of the inhibitory time constant) corresponds to propofol concentration measured in the femoral artery. Strictly speaking, what is needed is the propofol concentration at the cortex, but obtaining this information is a complicated exercise in pharmacokinetics modeling which requires several additional assumptions about multiple-compartment time constants.

C. Power divergence at transition

The theoretical origin of the peaking of the power spectrum at the transition points $\lambda_{\rm crit}$ corresponding to A_3 on the upper branch of Fig. 5, and Q_1 on the lower branch can be seen by examining the terms making up Eq. (2.18):

$$\mathcal{S}[h_e(\omega)]=\frac{1}{2\pi}\frac{D_{11}A_{22}^2+D_{22}A_{12}^2+D_{11}\omega^2}{(A_{11}A_{22}-A_{21}A_{12}-\omega^2)^2+(A_{11}+A_{22})^2\omega^2}. \tag{3.1}$$

From Eq. (2.13), the matrix element A_{11} can be written

$$A_{11}=-\frac{\partial F_1}{\partial h_e}=-\frac{\partial F_1}{\partial\lambda}\frac{\partial\lambda}{\partial h_e}$$

Similarly, the remaining elements are

$$A_{12}=-\frac{\partial F_1}{\partial\lambda}\frac{\partial\lambda}{\partial h_i},\quad A_{21}=-\frac{\partial F_2}{\partial\lambda}\frac{\partial\lambda}{\partial h_e},\quad A_{22}=-\frac{\partial F_2}{\partial\lambda}\frac{\partial\lambda}{\partial h_i}.$$

From Eqs. (2.11b) and (2.11c), we have

$$\frac{\partial F_1}{\partial\lambda}=\psi_{ie}[N_{ie}^{\beta}S_i(h_i)+\langle p_{ie}\rangle]\frac{G_i e}{\gamma_i\tau_e}\equiv a_1,$$

$$\frac{\partial F_2}{\partial \lambda} = \psi_{ii}[N_{ii}^{\beta} S_i(h_i) + \langle p_{ii} \rangle] \frac{G_i e}{\gamma_i \tau_i} \equiv a_2 .$$

The diffusion matrix elements D_{11} and D_{22} are nonzero and finite at λ_{crit}, as are the values a_1 and a_2. However, because the $\lambda - h_e$ and $\lambda - h_i$ curves of Fig. 5 have turning points at λ_{crit},

$$\lim_{\lambda \to \lambda_{\text{crit}}} \frac{\partial \lambda}{\partial h_e} = \lim_{\lambda \to \lambda_{\text{crit}}} \frac{\partial \lambda}{\partial h_i} = 0,$$

all four elements of matrix **A** will be zero, and thus at a critical point Eq. (3.1) predicts that the spectral power will scale as $1/\omega^2$:

$$\lim_{\lambda \to \lambda_{\text{crit}}} \mathcal{S}[h_e(\omega)] = \frac{D_{11}}{\omega^2}.$$

That is, for nonzero D_{11} at a critical point, $\mathcal{S}[h_e(\omega)]$ diverges at low frequencies. Further, examination of Eqs (2.19) and (2.21) shows that Δh_e, the fluctuations in h_e, will become infinite as $\lambda \to \lambda_{\text{crit}}$.

The peaking of the power spectrum and the divergence of Δh_e at λ_{crit} is similar to the singular behavior observed in thermodynamic phase transitions. For example, at the ferromagnetic critical temperature, both heat capacity C_V and magnetic susceptibility χ diverge. The traditional scaling-hypothesis model for critical phenomena asserts that these singularities arise from large-scale correlated fluctuations of magnetic spin alignment which occur at the critical point.

For the case of our 1D cortex, because we have a microscopic model for the interactions within a cortical macrocolumn, we can see how the presence of finite-amplitude white noise in the input terms (p_{ee}, p_{ie}, p_{ei}, p_{ii}) can result in infinite fluctuations in the h_e soma potential output: because the h_e covariance matrix depends on the stationary-state trajectory which has a turning point at λ_{crit}, the variance of h_e tends to infinity as $\lambda \to \lambda_{\text{crit}}$. Essentially, the presence of the turning point provides the required divergent ''gain'' as the anesthetic effect approaches its critical value.

There is an apparent paradox here. How is it that a linearized, first-order, equilibrium theory is able to reproduce the highly nonlinear, nonequilibrium fluctuations and divergences associated with a phase transition? The key would seem to be the inclusion in the model inputs of white noise. These small random fluctuations move the system just far enough away from equilibrium to allow sampling and capture of the essential characteristics of the nonequilibrium behavior: divergent low-frequency power and infinite fluctuations at the critical point.

IV. DISCUSSION

The significant result obtained in this study is that power spectral variations in a linearized stochastic model of cortical electrorhythmogenesis due to anesthetogenic variations in the inhibitory rate constant γ_i show qualitative agreement with clinical observations [33]: there is a sharp increase in low-frequency power in the vicinity of the critical points. As a consequence of the adiabatic elimination of the ''fast'' variables (Sec. II B 2) in Eqs (2.10), we can see that γ_i and G_i (the IPSP ''amplitude'') have reciprocal effects on the stationary points and on the corresponding spectral densities [see Eqs (A1)–(A4)]. Thus reductions in γ_i are equivalent to increases in G_i and vice versa. This means that we can interpret the effect of anesthetic agents on EEG as arising either from augmentation of the IPSP amplitude, or from increases in the time course (reduction in the rate constant) associated with IPSP kinetics.

Each of the exogenous spike-rate inputs (p_{ee}, p_{ei}, p_{ie}, p_{ii}) into the neuronal assembly is assumed to take the form of a white-noise fluctuation about an equilibrium mean. These inputs originate from neural action potentials generated within such subcortical structures as the thalamus and the reticular nuclei of the brainstem. We ignore the long-range contributions (ϕ_{ee}, ϕ_{ei}) from other cortical assemblies by assuming the ϕ terms are spatially homogeneous and constant in time. Because of the simplifications inherent in the adiabatic elimination and subsequent linearization, our theory does not demonstrate cortical resonances such as the 8–13 Hz alpha rhythm. Nevertheless, we believe the low–zero-frequency predictions of our model give some insight into the underlying cortical ''gain'' manifest in the EEG signal.

The neurons within an assembly are coupled via a sigmoid nonlinearity which defines the firing rate as a function of the soma potential (the spike-rate/h_e sigmoid curve; Fig. 3). Below a threshold value the firing rate is low (weak coupling between cells), whereas above threshold many neurons are firing (strong coupling). As a result, the stationary solution of our model predicts two distinct, stable-equilibrium states for the soma potential as shown in Fig. 5. The upper branch A_3S corresponds to the top plateau of the spike-rate sigmoid; we describe this as the ''active'' state of the cortex arising from strong intracortical connectivity and a relatively high (near zero) mean soma potential. Conversely, the lower branch Q_1C, corresponding to low spike rate, is the ''quiescent'' state brought about by weak intracortical connectivity and a lower (more negative) mean soma potential.

If the inhibitory post-synaptic potential decay time is prolonged (thereby moving the cortex into region I, Q_3C) either by application of drugs, or as a result of disease processes, there is a marked decrease in spike rate. This has been observed when neural preparations are exposed to therapeutic concentrations of general anesthetic agents [34], and when patients are in a state of coma. Although the degree of hyperpolarization induced by general anesthetics is minimal ($\sim$4 mV) [35], the spike-rate reduction is dramatic.

If we can assume that the essential requirement of normal cortical function (and presumably of conscious awareness) is the ability of the cortex to make and unmake strong but transient connections between assemblies, then in region II of our steady-state model we can picture individual assemblies transiting momentarily from the quiescent to the active branch. A collection of such active assemblies, firing collectively, should produce a coherent effect, much like the light field generated by a laser. However, unlike the laser analogy, neuronal assemblies do not remain in a state of high excitation for extended periods. Indeed, prolonged high excitation is a feature of the convulsive state, and may be induced by analeptic drugs such as bicuculine which *shortens* the inhibi-

tory post-synaptic decay time-constant [36]. The convulsive state corresponds to our region III [subbranch SA_1 in Fig. 5(b)].

The strong divergence in low-frequency power as the cortex changes state is similar to the divergences observed in thermodynamic phase transitions. In the thermodynamic case, phase changes can be described within the Ising framework which introduces the concept of an order parameter to distinguish between ordered and disordered states. For example, in the ferromagnetic phase transition, the order parameter is the net magnetization which is zero above a critical temperature, and nonzero below this temperature. We postulate that for the cortex, instead of an order parameter, we can define a *cooperativity* parameter H as the whole-cortex mean soma potential relative to its value in the unconscious state:

$$H=\overline{h}_e(\text{consc.})-\overline{h}_e(\text{unconsc.})$$

This parameter will have a large net value in the conscious state, and will be zero in the unconscious state. The phase transition is effected by varying the anesthetic amount. Thus the anesthetic provides the randomizing agent which breaks the connections between coherent subpopulations, transforming the cortex from a strongly-connected, cooperative conscious system to an unconscious system characterized by weak connectivity and negligible cooperativity. Currently, the phenomenologically-derived bispectral index (based on the computation of a limited bispectrum) is the most sophisticated measure used in the clinical practice of anesthesia to determine loss of consciousness and thus depth of anesthesia. However, the use of the theoretically derived parameter H may offer a more rational basis for the assessment of the depth of anesthesia and thus may have considerable clinical utility.

The actual neural mechanisms and dynamic routes by which the cortex may switch between quiescent and active states are not known, and are the subject of ongoing investigation. We speculate that noise-induced transitions may be important for maintaining conscious awareness [37].

We conclude that although the EEG is the spatially and temporally filtered summation of multiple and complex neuronal processes, the fact that our model correctly predicts a strong increase in low-frequency power at the critical points of induction and emergence suggests that the model design and assumptions provide a useful advance toward understanding cortical function.

ACKNOWLEDGMENTS

Figures 1 and 9 were generated from clinical data [3] provided courtesy of Dr Karel Kuizenga (Department of Anesthesiology, Academisch Ziekenhuis Groningen, The Netherlands). We thank Dr Kuizenga for his prompt and courteous response to our request for access to his measurements.

APPENDIX: DRIFT AND DIFFUSION MATRICES

The four elements of the drift matrix **A** are obtained by substituting Eqs. (2.11b) and (2.11c) into Eq. (2.13) and calculating the soma potential partial derivatives $\partial/\partial h_e$ and $\partial/\partial h_i$. The results are

$$\begin{aligned}A_{11}&=-\frac{\partial F_1}{\partial h_e}\\&=-\{-1+[\psi_{ee}^{(1)}((N_{ee}^{\alpha}+N_{ee}^{\beta})S_e(h_e)+\langle p_{ee}\rangle)+\psi_{ee}(N_{ee}^{\alpha}\\&\quad+N_{ee}^{\beta})S_e^{(1)}(h_e)]G_e e/\gamma_e+\lambda\psi_{ie}^{(1)}\\&\quad\times[N_{ie}^{\beta}S_i(h_i)+\langle p_{ie}\rangle]G_i e/\gamma_i\}/\tau_e,\end{aligned}\tag{A1}$$

$$A_{12}=-\{\lambda\psi_{ie}N_{ie}^{\beta}S_i^{(2)}(h_i)G_i e/\gamma_i\}/\tau_e,\tag{A2}$$

$$A_{21}=-\{\psi_{ei}(N_{ei}^{\alpha}+N_{ei}^{\beta})S_e^{(1)}(h_e)G_e e/\gamma_e\}/\tau_i,\tag{A3}$$

$$\begin{aligned}A_{22}&=-\{-1+\psi_{ei}^{(2)}[(N_{ei}^{\alpha}+N_{ei}^{\beta})S_e(h_e)+\langle p_{ei}\rangle]G_e e/\gamma_e\\&\quad+\lambda(\psi_{ii}^{(2)}[N_{ii}^{\beta}S_i(h_i)+\langle p_{ii}\rangle]\\&\quad+\psi_{ii}N_{ii}^{\beta}S_i^{(2)}(h_i))G_e e/\gamma_i\}/\tau_i,\end{aligned}\tag{A4}$$

where

$$\psi_{ee,ie}^{(1)}=\left.\frac{\partial\psi_{ee,ie}}{\partial h_e}\right|_{\text{eq.}},\quad \psi_{ei,ii}^{(2)}=\left.\frac{\partial\psi_{ei,ii}}{\partial h_i}\right|_{\text{eq.}},$$

$$S_e^{(1)}(h_e)=\frac{\partial S_e(h_e)}{\partial h_e},\quad S_i^{(2)}(h_i)=\frac{\partial S_i(h_i)}{\partial h_i}.$$

To compute the diffusion matrix **D**, we first note that the two off-diagonal elements are zero [see Eq. (2.15c)]. To calculate D_{11}, we substitute the $\Gamma_e(t)$ noise term from Eq. (2.11d) into Eq. (2.15a), leading to

$$\begin{aligned}\langle\Gamma_e(t)\Gamma_e(t')\rangle&=\frac{1}{\tau_e^2}\langle[\psi_{ee}\langle p_{ee}\rangle\xi_1(t)G_e e/\gamma_e\\&\quad+\lambda\psi_{ie}\langle p_{ie}\rangle\xi_3(t)G_i e/\gamma_i]\\&\quad\times[\psi_{ee}\langle p_{ee}\rangle\xi_1(t')G_e e/\gamma_e\\&\quad+\lambda\psi_{ie}\langle p_{ie}\rangle\xi_3(t')G_i e/\gamma_i]\rangle.\end{aligned}$$

Recalling that the $\xi(t)$ terms represent δ-correlated Gaussian white noise with zero mean [see Eq. (1.2)], the $\langle\Gamma_e(t)\Gamma_e(t')\rangle$ autocorrelation simplifies to

$$\begin{aligned}\langle\Gamma_e(t)\Gamma_e(t')\rangle&=\frac{1}{\tau_e^2}\{(\psi_{ee}\langle p_{ee}\rangle G_e e/\gamma_e)^2\\&\quad+\lambda^2(\psi_{ie}\langle p_{ie}\rangle G_i e/\gamma_i)^2\}\delta(t-t')\\&=D_{11}\delta(t-t'),\end{aligned}$$

so that

$$D_{11}=\frac{1}{\tau_e^2}\{(\psi_{ee}\langle p_{ee}\rangle G_e e/\gamma_e)^2+\lambda^2(\psi_{ie}\langle p_{ie}\rangle G_i e/\gamma_i)^2\}_{\text{eq.}}.\tag{A5}$$

Similarly, solving Eqs. (2.11e) and (2.15b) for D_{22} yields

$$D_{22}=\frac{1}{\tau_i^2}\{(\psi_{ei}\langle p_{ei}\rangle G_e e/\gamma_e)^2+\lambda^2(\psi_{ii}\langle p_{ii}\rangle G_i e/\gamma_i)^2\}_{\text{eq.}}.\tag{A6}$$

[1] M. Bührer, P. O. Maitre, O. R. Hung, W. F. Ebling, S. L. Shafer, and D. R. Stanski, Anesthesiology **77**, 226 (1992).
[2] M. B. MacIver, J. W. Mandema, D. R. Stanski, and B. H. Bland, Anesthesiology **84**, 1411 (1996).
[3] K. Kuizenga, C. J. Kalkman, and P. J. Hennis, Bri. J. Anaesthesia **80**, 725 (1998).
[4] D. P. Archer and S. H. Roth, Bri. J. Anaesthesia **79**, 744 (1997).
[5] R. Elul, Int. Rev. Neurobiol. **15**, 227 (1972).
[6] V. Braitenberg and A. Schüz, *Anatomy of a Cortex: Statistics and Geometry* (Springer-Verlag, Berlin, 1991).
[7] P. L. Nunez, *Electric Fields of the Brain: The Neurophysics of EEG* (Oxford University Press, New York, 1981).
[8] W. J. Freeman, *Mass Action in the Nervous System* (Academic Press, New York, 1975).
[9] H. R. Wilson and J. D. Cowan, Biophys. J. **12**, 1 (1972).
[10] H. R. Wilson and J. D. Cowan, Kybernetik **13**, 55 (1973).
[11] P. A. Robinson, C. J. Rennie, and J. J. Wright, Phys. Rev. E **56**, 826 (1997).
[12] P. A. Robinson, C. J. Rennie, J. J. Wright, and P. D. Bourke, Phys. Rev. E **58**, 3557 (1998).
[13] J. J. Wright and D. T. J. Liley, Behav. Brain Sci. **19**, 285 (1996).
[14] A. Rotterdam, F. H. Lopez da Silva, J. van der Ende, M. A. Viergever, and A. J. Hermans, Bull. Math. Biol. **44**, 283 (1982).
[15] D. J. Amit, *Modelling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, Cambridge, 1990).
[16] V. K. Jirsa and H. Haken, Phys. Rev. Lett. **77**, 960 (1996).
[17] D. T. J. Liley, in *Spatiotemporal Models in Biological Artificial Systems*, edited by F. L. Silva (IOS Press, Amsterdam, 1997), pp. 89–96.
[18] D. T. J. Liley, P. J. Cadusch, and J. J. Wright, in *A Continuum Theory of Electro-cortical Activity* (California Institute of Technology, Santa Barabara, CA, 1998).
[19] D. T. J. Liley, P. J. Cadusch, and J. J. Wright, Neurocomputing **26-27**, 795 (1999).
[20] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, Springer Series in Synergetics Vol. 13 (Springer-Verlag, Berlin, 1985).
[21] V. K. Jirsa and H. Haken, Physica D **99**, 503 (1997).
[22] T. D. Frank, A. Daffertshofer, P. J. Beck, and H. Haken, Physica D **127**, 233 (1999).
[23] D. T. J. Liley and P. J. Cadusch, Bull. Math. Biol. (to be published).
[24] H. C. Tuckwell, *Introduction to Theoretical Neurobiology* (Cambridge University Press, Cambridge, 1988), Vol. 1.
[25] N. P. Franks and W. R. Leib, Nature (London) **367**, 607 (1994).
[26] P. W. Gage and B. Robertson, Br. J. Pharmacol. **85**, 645 (1985).
[27] B. Antkowiak and H. Hentschke, Neurosci. Lett. **231**, 87 (1997).
[28] M. V. Jones and N. L. Harrison, J. Neurophysiol. **70**, 1339 (1993).
[29] C. Koch, M. Rapp, and I. Segev, Cerebral Cortex **6**, 93 (1996).
[30] H. Haken, *Handbuch der Physik* (Springer-Verlag, Berlin, 1970), Vol. XXV/2c.
[31] H. Haken, *Synergetics: An Introduction: Nonequilibrium Phase Transitions and Self-organization in Physics* (Springer-Verlag, Berlin, 1978).
[32] P. D. Drummond, Ph.D. thesis, University of Waikato, Hamilton, New Zealand, 1979.
[33] J. J. Stockard and R. G. Bickford, in *A Basis and Practise of Neuroanaesthesia II*, edited by E. Gordan (Excerpta Medica, Amsterdam, 1975), pp. 3–46.
[34] B. Antkowiak and C. Helfrich-Forster, Anesthesiology **88**, 1592 (1998).
[35] J. Berg-Johnsen and I. A. Langmoen, Acta Physiol. Scand. **130**, 679 (1987).
[36] R. F. Squires, E. Saederup, J. N. Crawley, P. Skolnick, and S. M. Paul, Life Sci. **35**, 1439 (1984).
[37] D. Forster, Science **273**, 1812 (1996).