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 $h$, 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 $h$. 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, $h$ is single valued, large, and negative, so that neuronal firing rates are suppressed. In region II for a zero or small anesthetic effect, $h$ 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, $h$ 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 transition 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 $h$ 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. [Anesthesiology 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 anaesthetised 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 ~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
distributed current sources and sinks induced by afferent (in-
coming) synaptic activity along these aligned pyramidal den-
drites can be approximated at the cortical surface by a dipole
term. The deviation from rest of the mean excitatory soma
membrane potential \( h_e - h_{e,0} \) 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 ori-
ented 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 ex-
tensive history with significant contributions by Freeman [8],
Wilson and Cowan [9,10], Nunez [7], Robinson and co-
workers [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 ex-
citatory and inhibitory neural populations, and included the
effects of axonal conduction delays.

Liley [17] extended these theories by improving the treat-
ment 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
produce a range of experimental results, they are math-
ematically 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 equa-
tions [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 differen-
tial equations (SDEs), also referred to as Langevin equa-
tions. 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 den-
dritic 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

\[
\frac{d}{dt} \left[ \begin{array}{c} h_e \\ h_i \end{array} \right] = - \left[ \begin{array}{cc} A_e(h_e, h_i) \\ A_i(h_e, h_i) \end{array} \right] + \left[ \begin{array}{c} B_e(\xi_e(t)) \\ B_i(\xi_i(t)) \end{array} \right],
\]

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 \( \delta \)-correlated:

\[
\langle \xi_q(t) \rangle = 0, \quad \langle \xi_q(t) \xi_{q'}(t') \rangle = \delta_{qq'} \delta(t-t').
\]

Setting the noise terms to zero in Eq. (1.1) gives the deter-
ministic equation for \( h_{e,i} \):

\[
\frac{d}{dt} \left[ \begin{array}{c} h_e \\ h_i \end{array} \right] = - \left[ \begin{array}{cc} A_e(h_e, h_i) \\ A_i(h_e, h_i) \end{array} \right].
\]

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: \( h_{e,i} = h_{e,i}^0 + h_{e,i} \).
This decomposition enables us to transform Eq. (1.1) into a
set of linear SDEs,

![Image of a graph showing biphasic effect of propofol anesthetic on 0-5- and 11-15 Hz EEG signals. The x-axis represents time in minutes, and the y-axis represents EEG amplitude in microvolts per second (uV/s). The graph shows two lines: one for 0 to 5 Hz and another for 11 to 15 Hz. The data points for 0 to 5 Hz are relatively low, while those for 11 to 15 Hz are higher, indicating the biphasic effect.](image-url)
where $\bar{h}_{e,i}$ 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 bi-phasic (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

FIG. 2. Schematic representation of the connective topology within a cortical macrocolumn. Only four of the $\approx 100,000$ 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 $e \to e$, $e \to i$ input from distant macrocolumns, and $p_{jk}$ represents input from the subcortex (e.g., thalamus and brainstem). (For clarity, we have omitted $p_{ij}$ and $p_{ij}$ exogenous inputs corresponding to connection types 9 and 10, respectively.)

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

where $\bar{h}_{e,i}$ 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].
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 cm², and thus detects electrical activity averaged across the underlying 5–10 cm² of cerebral cortex. Thus the one-dimensional Laplacian \( \partial^2 \phi / \partial x^2 \) [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):

\[
\begin{align*}
\frac{d}{dt} + \gamma_i & \left( \frac{d}{dt} \right)^2 I_{i,ee} = \left( \left[ \begin{array}{c} I_{i,ee} \end{array} \right] + \left[ \begin{array}{c} I_{i,ii} \end{array} \right] \right) G_{i}, \quad (2.1) \\
\frac{d}{dt} + \gamma_i & \left( \frac{d}{dt} \right)^2 I_{i,ii} = \left( \left[ \begin{array}{c} I_{i,ee} \end{array} \right] + \left[ \begin{array}{c} I_{i,ii} \end{array} \right] \right) G_{i}, \quad (2.2) \\
\frac{d}{dt} + \gamma_i & \left( \frac{d}{dt} \right)^2 \phi_{ee} = \left( \left[ \begin{array}{c} \phi_{ee} \end{array} \right] + \left[ \begin{array}{c} \phi_{ii} \end{array} \right] \right) G_{i}, \quad (2.3) \\
\frac{d}{dt} + \gamma_i & \left( \frac{d}{dt} \right)^2 \phi_{ii} = \left( \left[ \begin{array}{c} \phi_{ee} \end{array} \right] + \left[ \begin{array}{c} \phi_{ii} \end{array} \right] \right) G_{i}, \quad (2.4)
\end{align*}
\]

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 $h_{e,i}^{\text{rest}}$, the second pair describe perturbations to the membrane potential due to synaptic inputs to the neural mass. The $\psi_{jk}$ (where $j,k \in \{e,i\}$) coefficients appearing on the right are normalized weighting functions for these inputs. These coefficients 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

$$
\psi_{ee} = \frac{h_e^{-\text{rev}} - h_e}{|h_e^{-\text{rev}} - h_e^{\text{rest}}|}, \quad \psi_{ei} = \frac{h_i^{-\text{rev}} - h_e}{|h_i^{-\text{rev}} - h_i^{\text{rest}}|},
$$

$$
\psi_{ie} = \frac{h_e^{-\text{rev}} - h_i}{|h_e^{-\text{rev}} - h_i^{\text{rest}}|}, \quad \psi_{ii} = \frac{h_i^{-\text{rev}} - h_i}{|h_i^{-\text{rev}} - h_i^{\text{rest}}|}.
$$

The constant values used for the resting $h_{e,i}^{\text{rest}}$ and reversal $h_{e,i}^{-\text{rev}}$ potentials are listed in Table I. Note that for typical values for $h_e$ and $h_i$, the weights $\psi_{ee}$ and $\psi_{ii}$ for input from excitatory sources are positive, while weights $\psi_{ie}$ and $\psi_{ei}$ 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 $S_i(h_i)$:

$$
S_e(h_e) = [1 + \exp(-g_e(h_e - \theta_e))]^{-1},
$$

$$
S_i(h_i) = [1 + \exp(-g_i(h_i - \theta_i))]^{-1}.
$$

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 constants. 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 ($\phi_e$, $\phi_i$) 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 $\gamma_i$ in Eq. (2.3)]. At concentrations appropriate for surgical anesthesia, the IPSPs are prolonged by a factor of 1.5-4 fold [25–28].

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

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

Here $\lambda$ 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 $\lambda$ corresponds to an increase in anesthetic amount (decrease in $\gamma_i$ rate constant). See Fig. 4. We now describe how the Liley equations are transformed into linearized stochastic differential equations.

\section*{B. Stochastic differential equations (SDEs)}

\subsection*{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) postsynaptic membranes. Curves are normalized to unit height. For application to our model, the heights are scaled by the respective EPSP (excitatory postsynaptic potential) and IPSP (inhibitory postsynaptic potential) amplitudes, $G_{e,i}=0.18$ and 0.37 mV. The rate constants, in $(\text{ms})^{-1}$, for the three curves are $\gamma_e=0.30$, $\gamma_i=0.065$, and $\gamma'_i=0.043$.

We thus adiabatically eliminate these ''fast'' variables by computing observable quantities such as power spectra. Solving them 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., $d/dt\to0$) in which all noise terms have been set to zero. This gives the stationary solution, which we denote by the vector

$$\mathbf{a}^0 = \begin{bmatrix} h_{e,i} & \phi_e & \phi_i \end{bmatrix}^T.$$

Having solved for the equilibrium state $\mathbf{a}^0$, we can linearize Eqs. (2.1), (2.4), (2.7), and (2.8) about $\mathbf{a}^0$, 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 ($I_{e,i}$, $\phi_e$, $\phi_i$) 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):

- relaxation time for $I_{ee}I_{el} = (\gamma_e)^{-1}\approx3.3$ ms,
- relaxation time for $I_{ie}I_{ii} = (\gamma_i)^{-1}\approx15.4$ ms,
- relaxation time for $\phi_e = (\bar{\nu}\lambda_{ee})^{-1}\approx3.6$ ms,
- relaxation time for $\phi_i = (\bar{\nu}\lambda_{ei})^{-1}\approx2.2$ 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 [$I_{ee}I_{ei}I_{el}I_{ii}I_{il}I_{il}$] 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 $d/dt\to0$ in Eqs. (2.2)–(2.4) while retaining the noise terms, allowing us to solve for $I_{ee}I_{ei}I_{el}I_{ii}I_{il}I_{il}$, $\phi_e$, and $\phi_i$ 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 $\mathbf{a}^0$, 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 Gardner [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.)

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 computation 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.,

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

(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

$$\left( \frac{d}{dt} + \gamma_e \right) I_{ee} = \left[ \begin{array}{c} N_{ee} \phi_e \\ N_{ei} \phi_i \end{array} \right] S_e(h_e) + \left[ \begin{array}{c} \langle p_{ee} \rangle \\ \langle p_{ei} \rangle \end{array} \right] G_e \gamma_e e + \left[ \begin{array}{c} \Gamma_1(t) \\ \Gamma_2(t) \end{array} \right],$$

$$\left( \frac{d}{dt} + \gamma_i \right) I_{ii} = \left[ \begin{array}{c} N_{ie} \phi_e \\ N_{ii} \phi_i \end{array} \right] S_i(h_i) + \left[ \begin{array}{c} \langle p_{ie} \rangle \\ \langle p_{ii} \rangle \end{array} \right] G_i \gamma_i e + \left[ \begin{array}{c} \Gamma_3(t) \\ \Gamma_4(t) \end{array} \right],$$

where

$$\Gamma_1(t) = \left[ \begin{array}{c} \langle p_{ee} \rangle \xi_{1}(t) \\ \langle p_{ei} \rangle \xi_{2}(t) \end{array} \right] G_e \gamma_e e,$$

$$\Gamma_2(t) = \left[ \begin{array}{c} \langle p_{ie} \rangle \xi_{2}(t) \\ \langle p_{ii} \rangle \xi_{1}(t) \end{array} \right] G_i \gamma_i e,$$

and the four $\xi_{1,2}(t)$ 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.]
The equations resulting from the adiabatic simplification follow:

\[
\begin{align*}
\left[I_{ie}^0\right] &= \left[\begin{array}{c}
\mathbf{K}^e_{ie} \\
\mathbf{N}_{ie}^e
\end{array}\right] S_e(h_e) + \left[\begin{array}{c}
\mathbf{p}_{ie}^e \\
\mathbf{p}_{ii}
\end{array}\right] G_{ee} \gamma_e \\
+ &\left[\begin{array}{c}
\Gamma_1(t) \\
\Gamma_2(t)
\end{array}\right]/\gamma_e^2, \\
(2.10a)
\end{align*}
\]

\[
\begin{align*}
\left[I_{ie}^0\right] &= \left[\begin{array}{c}
\mathbf{K}^e_{ie} \\
\mathbf{N}_{ie}^e
\end{array}\right] S_e(h_e) + \left[\begin{array}{c}
\mathbf{p}_{ie}^e \\
\mathbf{p}_{ii}
\end{array}\right] G_{ee} \gamma_e + \left[\begin{array}{c}
\Gamma_3(t) \\
\Gamma_4(t)
\end{array}\right]/\gamma_e^2, \\
(2.10b)
\end{align*}
\]

\[
\begin{align*}
\left[I_{ie}^0\right] &= \left[\begin{array}{c}
\mathbf{K}^e_{ie} \\
\mathbf{N}_{ie}^e
\end{array}\right] S_e(h_e) + \left[\begin{array}{c}
\mathbf{p}_{ie}^e \\
\mathbf{p}_{ii}
\end{array}\right] G_{ee} \gamma_e, \\
(2.10c)
\end{align*}
\]

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

\[
\frac{d}{dt} \left[\begin{array}{c}
h_e \\
h_i
\end{array}\right] = \left[\begin{array}{c}
F_1(h_e, h_i) \\
F_2(h_e, h_i)
\end{array}\right] + \left[\begin{array}{c}
\Gamma_1(t) \\
\Gamma_2(t)
\end{array}\right], \\
(2.11a)
\]

where the drift terms are

\[
\begin{align*}
F_1(h_e, h_i) &= \{\mathbf{h}_{ie}^{\text{rest}} - h_e\} + \psi_e \gamma_e \mathbf{N}_{ie}^e S_e(h_e) + \mathbf{p}_{ie}^e G_{ee} \gamma_e \\
+ &\mathbf{p}_{ii} G_{ee} \gamma_e \gamma_i, \\
F_2(h_e, h_i) &= \{\mathbf{h}_{ie}^{\text{rest}} - h_i\} + \psi_i \gamma_i \mathbf{N}_{ei}^e S_e(h_e) + \mathbf{p}_{ei}^e G_{ee} \gamma_e \\
+ &\mathbf{p}_{ii} G_{ee} \gamma_e \gamma_i, \\
(2.11b)
\end{align*}
\]

and the corresponding noise terms are

\[
\begin{align*}
\Gamma_1(t) &= \{\psi_e \mathbf{p}_{ie}^e \xi_1(t) G_{ee} \gamma_e + \psi_i \mathbf{p}_{ii} \xi_3(t) G_{ee} \gamma_e \gamma_i\} / \gamma_e, \\
\Gamma_2(t) &= \{\psi_i \mathbf{p}_{ei}^e \xi_2(t) G_{ee} \gamma_e + \psi_i \mathbf{p}_{ii} \xi_4(t) G_{ee} \gamma_e \gamma_i\} / \gamma_i. \\
(2.11d)
\end{align*}
\]

(11e) (Note that we have replaced \(\gamma_i\) by \(\gamma_i/\lambda\) in the above equations in order to make explicit their dependence on anesthetic "effect" \(\lambda\).)

**C. Fluctuation spectrum: linearized theory**

We linearize SDE’s (2.11) about an equilibrium state \(\mathbf{a}^0\) to obtain the Ito SDE

\[
\frac{d}{dt} \left[\begin{array}{c}
h_e \\
h_i
\end{array}\right] = -\mathbf{A} \left[\begin{array}{c}
h_e \\
h_i
\end{array}\right] + \left[\begin{array}{c}
\Gamma_1(t) \\
\Gamma_2(t)
\end{array}\right] \\
(2.12)
\]

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

\[
\mathbf{A} = -\left[\begin{array}{cc}
\frac{\partial F_1}{\partial h_e} & \frac{\partial F_1}{\partial h_i} \\
\frac{\partial F_2}{\partial h_e} & \frac{\partial F_2}{\partial h_i}
\end{array}\right], \\
(2.13)
\]

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]

\[
\frac{\partial P(\mathbf{h}_e, \mathbf{h}_i)}{\partial t} = \left\{\frac{\partial}{\partial h_e} \left[A_1 \mathbf{h}_e + A_2 \mathbf{h}_i\right] + \frac{\partial}{\partial h_i} \left[A_2 \mathbf{h}_e + A_2 \mathbf{h}_i\right]
+ \frac{1}{2} \left[\frac{\partial^2}{\partial h_e^2} D_{11} + \frac{\partial^2}{\partial h_i^2} D_{22}\right]\right\} P(\mathbf{h}_e, \mathbf{h}_i), \\
(2.14)
\]

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

\[
\langle \Gamma_i(t) \Gamma_j(t') \rangle = D_{ij} \delta(t-t'), \\
(2.15a)
\]

\[
\langle \Gamma_i(t) \Gamma_j(t') \rangle = D_{22} \delta(t-t'), \\
(2.15b)
\]

\[
D_{12} = D_{21} = 0. \\
(2.15c)
\]

(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 \(\mathbf{h}_e\) as

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

then the stationary fluctuation spectrum for \(\mathbf{h}_e\) can be computed from the Fourier transform

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

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

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

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

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

in which \(\mathbf{I}\) is the identity matrix; \text{det} and \text{Tr} are the determinant and trace operators respectively; and where, for example,
The multiple intersections of the isocline curves steady-state values were obtained numerically by locating function of anesthetic "amount".

When we obtain the equilibrium behavior of values for \( \lambda \) will then cause \( h^0_e \) to advance along the \( QC \) subbranch. If instead, the assembly was initially at \( Q_2 \) on the lower branch, then increases in \( \lambda \) 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 (\( h_e = -60 \) 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 \( QC \) 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 \( QC \) 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 \( A_3QC \) 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 necessarily 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 \( I/I_e \), where \( I_i \) (\( I_e \)) 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 \( I/I_e \approx 1 \), i.e., when the inhibitory and excitatory responses were of similar magnitude. However, if the inhibitory response was strongly dominant over excitatory (or vice
versa), they found that the system collapsed to a single steady state.

Relating these findings to our model, their $l_i$ (or $l_e$) ‘‘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 $\gamma_i$) by scaling it with anesthetic factor $\lambda$. Thus, broadly speaking, our $\lambda$ 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 $(h_i^0, h_e^0)$ 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 $\lambda$ 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 $\lambda$ 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 $\lambda$ 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-
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 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 discrepancies 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 $\lambda$ 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 $\lambda_{crit}$ 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):

$$S[h_\epsilon(\omega)] = \frac{1}{2\pi} \frac{D_{11}A_{22}^2 + D_{22}A_{11}^2 + D_{11}\omega^2}{(A_{11}A_{22} - A_{21}A_{12} - \omega^2)^2 + (A_{11} + A_{22})^2 \omega^2}$$

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

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

Similarly, the remaining elements are

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

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

$$\frac{\partial F_1}{\partial \lambda} = \psi_j\pi\gamma_\epsilon S(h_\epsilon) + (p_{ie}) \frac{G_{ie}}{\gamma_\epsilon \tau_\epsilon} = a_1,$$
The diffusion matrix elements $D_{11}$ and $D_{22}$ are nonzero and finite at $\lambda_{\text{crit}}$, as are the values $a_1$ and $a_2$. However, because the $\lambda - h$, and $\lambda - h$ curves of Fig. 5 have turning points at $\lambda_{\text{crit}}$, 

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

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 $1/\omega^2$:

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

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

The peaking of the power spectrum and the divergence of $\Delta h_{\lambda}$ at $\lambda_{\text{crit}}$ 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 $\chi$ 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 ($p_{ee}$, $p_{ei}$, $p_{ie}$, $p_{ii}$) can result in infinite fluctuations in the $h_{\lambda}$ soma potential output: because the $h_{\lambda}$ covariance matrix depends on the stationary-state trajectory which has a turning point at $\lambda_{\text{crit}}$, the variance of $h_{\lambda}$ 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 electrophysiology due to anestheticogenic variations in the inhibitory rate constant $\gamma_i$ 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 $\gamma_i$ 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 $\gamma_i$ 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 ($\theta_{ee}$, $\theta_{ei}$) from other cortical assemblies by assuming that the $\beta$ 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/$h_{\lambda}$ 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 $A_{\lambda}S$ 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 $Q_{\lambda}C$, 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, $Q_{\lambda}C$) 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 (~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 transitioning 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 SA1 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:

\[
H = \bar{h}_c(\text{conc.}) - \bar{h}_c(\text{unconc.})
\]

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

\[
A_{11} = -\frac{\partial F_1}{\partial h_e}
\]

\[
= -\left\{-1 + \left[ \psi^{(1)}_{ee}(N_{ee}^\alpha + N_{ee}^\beta)S_e(h_e) + \langle p_{ee} \rangle + \psi_{ee}(N_{ee}^\alpha + N_{ee}^\beta) + N_{ee}^\alpha S_e^{(1)}(h_e) \right]G_e e / \gamma_e + \lambda \psi^{(1)}_{ei} \right\} \times \left\{N_{ie}^\beta S_i(h_i) + \langle p_{ei} \rangle \right\} G_e e / \gamma_i / \tau_e, \quad (A1)
\]

\[
A_{12} = -\left\{ \lambda \psi_{ie}N_{ie} S_i^{(2)}(h_i) G_e e / \gamma_i \right\} / \tau_e, \quad (A2)
\]

\[
A_{21} = -\left\{ \psi_{ie}(N_{ie}^\alpha + N_{ie}^\beta)S_e^{(1)}(h_e) G_e e / \gamma_e \right\} / \tau_i, \quad (A3)
\]

\[
A_{22} = -\left\{ -1 + \psi^{(2)}_{ei}(N_{ee}^\alpha + N_{ee}^\beta)S_e(h_e) + \langle p_{ei} \rangle \right\} G_e e / \gamma_e \times \left\{ \lambda \psi^{(2)}_{ii}[N_{ii}^\beta S_i(h_i) + \langle p_{ii} \rangle] \right\} \times \left\{ \psi_{ii}N_{ii} S_i^{(2)}(h_i) \right\} G_e e / \gamma_i / \tau_i, \quad (A4)
\]

where

\[
\psi^{(1)}_{ee,ie} = \frac{\partial \psi_{ee,ie}}{\partial h_e} \Bigg|_{\text{eq.}}, \quad \psi^{(2)}_{ii,ii} = \frac{\partial \psi_{ii,ii}}{\partial h_i} \Bigg|_{\text{eq.}},
\]

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

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

\[
\langle \Gamma_c(t) \Gamma_c(t') \rangle = \frac{1}{\tau_e^2} \left\{ \langle \psi_{ee}(p_{ee}) \xi_i(t) \rangle G_e e / \gamma_e \right. \\
+ \lambda \psi_{ie}(p_{ie}) \xi_i(t) G_e e / \gamma_i \right. \\
\times \left\{ \langle \psi_{ee}(p_{ee}) \xi_i(t') \rangle G_e e / \gamma_e \right. \\
+ \lambda \psi_{ie}(p_{ie}) \xi_i(t') G_e e / \gamma_i \right\}.
\]

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

\[
\langle \Gamma_c(t) \Gamma_c(t') \rangle = \frac{1}{\tau_e^2} \left\{ \langle \psi_{ee}(p_{ee}) G_e e / \gamma_e \right\}^2 \\
+ \lambda^2 \langle \psi_{ie}(p_{ie}) G_e e / \gamma_i \rangle^2 \right\} \delta(t-t'),
\]

so that

\[
D_{11} = \frac{1}{\tau_e^2} \left\{ \langle \psi_{ee}(p_{ee}) G_e e / \gamma_e \right\}^2 + \lambda^2 \langle \psi_{ie}(p_{ie}) G_e e / \gamma_i \rangle^2 \right\} \epsilon.
\]

(A5)

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

\[
D_{22} = \frac{1}{\tau_i^2} \left\{ \langle \psi_{ie}(p_{ie}) G_e e / \gamma_e \right\}^2 + \lambda^2 \langle \psi_{ii}(p_{ii}) G_e e / \gamma_i \rangle^2 \right\} \epsilon.
\]

(A6)