Simple spontaneously active Hebbian learning model: Homeostasis of activity and connectivity, and consequences for learning and epileptogenesis

David Hsu,[1,*] Aonan Tang,[2] Murielle Hsu,[1] and John M. Beggs[2]

[1]*Department of Neurology, University of Wisconsin, Madison, Wisconsin 53792, USA*

[2]*Department of Physics, Indiana University, Bloomington, Indiana 47405, USA*

A spontaneously active neural system that is capable of continual learning should also be capable of homeostasis of both firing rate and connectivity. Experimental evidence suggests that both types of homeostasis exist, and that connectivity is maintained at a state that is optimal for information transmission and storage. This state is referred to as the critical state. We present a simple stochastic computational Hebbian learning model that incorporates both firing rate and critical homeostasis, and we explore its stability and connectivity properties. We also examine the behavior of our model with a simulated seizure and with simulated acute deafferentation. We argue that a neural system that is more highly connected than the critical state (i.e., one that is "supercritical") is epileptogenic. Based on our simulations, we predict that the postseizural and postdeafferentation states should be supercritical and epileptogenic. Furthermore, interventions that boost spontaneous activity should be protective against epileptogenesis.

PACS number(s): 87.18.Sn, 87.19.La, 89.75.Fb

INTRODUCTION

The destabilizing effects of associative (Hebbian) learning on neural systems have long been recognized: favorable synapses become stronger, unfavorable synapses become weaker, and there is a strong drive either towards ever escalating, runaway activity or else global network silence [1–7]. Activity-dependent regulation of synaptic strengths and intrinsic membrane currents have been proposed as mechanisms of maintaining homeostasis of neuronal activity [8]. Because of these homeostatic mechanisms, neuronal circuits can behave similarly even when the cellular level constituents of neuronal behavior are different [9,10]. A particularly striking example of a synaptic mechanism of homeostasis is multiplicative postsynaptic up regulation of alpha-amino-3–hydroxy-5–methyl-4–isoxazolepropionic acid (AMPA) receptors as a compensatory mechanism for synaptic blockade [11,12].

At present, studies of neural systems homeostasis primarily focus on activity or firing rate homeostasis, whereby the firing rate of a system is constrained to approach a target level of activity [4,5,13,14]. However, although firing rate homeostasis can maintain stability of activity, it does not guarantee a neural system that is functionally useful. For instance, a trivial solution for firing rate homeostasis is to cut all neuronal connections, and then to adjust the spontaneous (or intrinsic) firing rate for each neuron until the target firing rate is reached. Such a neural system is incapable of associative learning, as every neuron is isolated from every other neuron. Conversely, a connectivity pattern such that activation of one neuron reliably causes activation of all other neurons is equally useless. Neurons that are capable of learning must maintain an intermediate level of connectivity, even in

the face of synaptic changes brought about by Hebbian learning and firing rate homeostasis. Therefore, we expect at least one other homeostatic principle, one for homeostasis of connectivity. Such a mechanism would prevent the neural system from entering a state that is either underconnected or overconnected.

Local field potential measurements have demonstrated evidence for homeostasis of connectivity (reviewed in Refs. [15,16]). In acute cortical slices and in cortical slices cultured on 60 channel microelectrode arrays, activity consists of periods of quiescence broken by bursts of activity of any number of electrodes, which occur in clusters (or "avalanches") of all possible sizes [17–19]. A branching ratio σ can be defined as the number of electrodes that are excited after any other single electrode is excited, averaged over time and over all electrodes. This ratio fluctuates about unity for hours at a time [17,18], with root mean square deviations of about $5–25\%$ [70]. The condition $\sigma=1$ represents the critical point. The return to criticality after fluctuations away from criticality represents critical homeostasis.

The critical point so defined has a number of interesting properties. Chief among them are that critical branching optimizes information throughput [17] and maximizes information storage capacity [18,19]. A related property is that the probability distribution function of avalanche sizes $G_A(n)$ obeys a power law with $G_A(n) \sim n^{-1.5}$ [17]. The power law behavior suggests that the neural system has long range correlations and is able to access the entire repertoire of possible activation states. Such systems are also referred to as being scale-free [20–23]. That the critical point is optimal for information throughput and storage capacity, and that critical systems have scale-free access to the entire range of possible activation patterns, are highly desirable properties for learning and information processing systems. Thus critical homeostasis represents a homeostasis of connectivity optimized for learning and information processing.

That a branching ratio of $\sigma=1$ is optimal for information storage and processing is intuitively plausible. If $\sigma>1$, then

*Corresponding author; present address: Department of Neurology, University of Wisconsin Hospital & Clinics, 600 Highland Ave., Madison, WI 53792 USA; hsu@neurology.wisc.edu

one might expect loss of information during transmission due to its being "whited out" by accelerating network activity. If $\sigma < 1$, then information is lost due to its being damped out. A branching ratio of $\sigma = 1$ is most likely to preserve signal transmission in a faithful way. Similarly, for information storage, one might think of information as being stored in the patterns of network activation. If $\sigma < 1$, then these patterns are limited to small cluster sizes, while if $\sigma > 1$, then only large clusters can be activated. At $\sigma = 1$, however, it is possible to activate clusters of all sizes, from the smallest sizes up to clusters the size of the entire network. It is at $\sigma = 1$ that one has the largest repertory of patterns of activation, and thus the largest information storage capacity.

The biomolecular mechanism for critical homeostasis is not known. One might conjecture that such a mechanism depends on the concentration of intracellular calcium, on trophic factors such as brain derived neurotrophic factor, on backpropagation of action potentials into the proximal dendritic tree, or on competition for cellular resources [2,5,14]. Interneuron interactions are another candidate. In any case, *we postulate* that such a mechanism exists, particularly for brain regions most involved in learning and high order information processing. Real neural systems might conceivably fluctuate about criticality, or might fluctuate about a point not quite at criticality. It may also turn out that components of a neural system may be tuned to a state far from criticality. Nonetheless, at a large enough length scale, we hypothesize that any neural system that is capable of continual learning must be capable of some degree of critical homeostasis.

Here we investigate the properties of a simple model that incorporates both firing rate homeostasis and critical homeostasis. Our purpose is not to propose a model that is quantitatively correct in every detail, but to investigate first whether it is possible to construct such a model, and second whether there are algorithmic consequences of imposing both firing rate and critical homeostasis for this particular model. In the Discussion section, we compare our model with other recent homeostatic models.

METHODS

The model consists of a set of nodes labeled by $i = 1$ to $N = 64$, each of which has an activation level $A(i,t)$. The activation level $A(i,t)$ gives the probability that node i "fires" at some point in the time interval $(t - \delta t, t]$, where δt is the time step. Each node represents the local field average over some number of neurons near one microelectrode. Each firing event represents a population spike consisting of the near-simultaneous action potential discharge of a subpopulation of nearby neurons. In practice, the firing events are taken to be local field potential spikes that are more than three standard deviations above the background amplitude. We speak of nodal firing events to distinguish communal action potential discharges from, for instance, lower amplitude postsynaptic potentials. For the system of Haldeman and Beggs [19], with microelectrode diameter of 0.02 mm and interelectrode distance of 0.2 mm, the number of neurons contributing to a nodal firing event may be as large as 1000.

The nodal local field potential dynamics that we model is not simply related to the underlying neuronal dynamics.

Neuronal dynamics consists of the action of ions, molecules, macromolecular and cellular-level structures, while nodal dynamics consists of the interaction of many principal neurons and local interneurons in one field, with principal neurons and interneurons in another distant field. In general, the time scale of larger scale dynamics is slower than that of smaller scale systems, because of the possibility of the emergence of collective modes involving the mass action of many neurons acting in phase with one another [24]. A similar phenomenon is well known in solid state physics, where the high frequencies of molecular vibrations and translational motion are transformed, when these molecules are condensed into a solid, into low frequency as well as high frequency bands. The low frequency bands correspond to collective vibrations involving many molecules moving in phase relative to each other, while the high frequency bands correspond to intrinsic intramolecular motions. Thus the possibility that nodal time scales may be different from neuronal time scales should not surprise us.

The theoretical bridging of microscopic with macroscopic behavior is a monumental task, beyond the scope of this paper [25–28]. Nonetheless, we expect nodal dynamics to reflect neuronal dynamics in a qualitative way. Thus if neuronal dynamics exhibits firing rate homeostasis, then nodal firing rates must also exhibit firing rate homeostasis. Similarly, Hebbian learning at the neuronal level must also be reflected in correlations in nodal dynamics. We will also assume that every node is equivalent. By equivalent, we mean that every node has the same steady state firing rate and the same steady state branching ratio. Heterogeneous systems may be generalized from what follows.

The brain is immensely complex and structured, and this structure certainly contributes to its powerful information processing capabilities. Understanding exactly how the detailed circuitry of the brain accomplishes its tasks is an important goal and will require more electrophysiological experiments at the synaptic and circuit levels, as well as detailed, biophysically realistic compartmental models. Clearly, this is an ongoing task. In the mean time, we believe that a statistical approach to optimality can yield some insights as to how the brain operates. The constraints revealed by our analysis are only meant to apply to the neocortex at some coarse-grained level, and may or may not be applicable at the microscopic circuit level. By focusing our attention on macroscopic behavior, we do not imply that microscopic and mesoscopic level complexities are irrelevant. The simplifications that we make help us to identify certain properties that we believe may be general to spontaneously active neural systems. Some generalizations of our model are possible and desirable, depending on the questions one wishes to answer. We discuss this issue further in the Discussion section. Examples of other collective or "lumped" neural system models exist [29–34].

To continue, let $F(i;t) = 1$ mean that node i fired at some point in the time interval $(t - \delta t, t]$, and $F(i;t) = 0$ mean that it was quiescent. Let τ_0 be the target average time interval between firings, so that $1/\tau_0$ is the target average firing rate. The value of $F(i;t)$ can only be zero or one, never anything in-between. However, if $\langle \cdots \rangle$ represents a time average over a time period that is much longer than τ_0, then one expects

that $\langle F(i;t)\rangle$ becomes approximately equal to $\langle A(i;t)\rangle$, which itself approaches $\delta t/\tau_0$.

Let $P(i,j;t)$ be the conditional probability that firing at node j in the time interval $(t-\delta t,t]$ causes firing at node i in the next time interval $(t,t+\delta t]$, with $P(i,i;t)=0$. We refer to the $P(i,j;t)$'s as the stimulated firing probabilities, or equivalently as the connection strengths. Let $S(i;t)$ be the probability of spontaneous firing at node i in the time interval $(t,t+\delta t]$. This probability does not depend on whether any other node has recently fired. The total probability that node i will fire in the time interval $(t,t+\delta t]$ is then given by one minus the probability that it will not fire, i.e.,

$$A(i;t+\delta t) = 1 - [1-S(i;t)]\prod_{j=1}^{N}[1-P(i,j;t)F(j;t)]. \quad (1)$$

Here Π_j signifies a product over all j from 1 to the total number of nodes, N. In deriving this equation, we have assumed that the connection strengths $P(i,j;t)$ are mutually independent, so that the probability of multiple mutually independent processes occurring at the same time is given by the product of the individual probabilities. Multinodal events are not ignored but are approximated by the product of one- and two-nodal probabilities. More sophisticated models would have to take into account simultaneous multinodal interdependence, including terms of the form $P(i,j,k;t)$ for simultaneous interactions between three nodes, for instance. We leave these higher order correction terms for future studies. They may prove important but are much more complex to analyze.

For convenience, define a relative firing rate $f(i;t) = \langle F(i;t)\rangle_a \tau_0/\delta t$, where $\langle\cdots\rangle_a$ represents a time average over the prior time interval τ_a with $\tau_a \gg \delta t$. We take $\tau_a = \tau_0$ for simplicity. Thus firing rate homeostasis requires $f(i;t)$ to hover about unity, and the fluctuation $\Delta f(i;t)=f(i;t)-1$ to fluctuate about zero. Physiologically, τ_a is the time scale on which each node, through regulatory feedback mechanisms, perceives what its time-averaged firing rate is, relative to its target firing rate. This time scale has to be on the time scale of τ_0 or longer, because a given feedback time cannot yield an average firing interval longer than itself. We take $\tau_a = \tau_0$ primarily for demonstration purposes, but also note that this time scale is the shortest time scale at which a node can compare actual to target firing rates. This choice thus optimizes nodal responsiveness to fluctuations away from target nodal behavior. In what follows we found it most challenging to achieve stable solutions with smaller values of τ_a; firing rate and critical homeostasis are easier to achieve for larger values of τ_a, presumably because fluctuations about the steady state are smoothed out.

Nodal connectivity can be defined in terms of an input ratio $\eta(i;t)$ as follows:

$$\eta(i;t) = \sum_{j=1}^{N} P(i,j;t). \quad (2)$$

Here Σ_j signifies a sum over all j from 1 to the total number of nodes, N. Analogously, the branching ratio for node i can be defined as

$$\sigma(i;t) = \sum_{j=1}^{N} P(j,i;t). \quad (3)$$

The input ratio is a measure of excitatory input into each node. The branching ratio is a measure of excitatory output from each node. The input ratio averaged over all nodes is equal to the corresponding average of the branching ratio. We define critical homeostasis in terms of homeostasis of the input ratio rather than of the branching ratio, but either method will work. Our choice is influenced by the fact that the input ratio is a postsynaptic attribute, in the same spirit as the postsynaptic scaling mechanism found by Turrigiano and co-workers [12,35]. Let fluctuations of the input ratio about the target input ratio be $\Delta\eta(i;t)=\eta(i;t)-1$. Let the physical distance between nodes i and j be $D(i,j)$.

We next need equations governing the time evolution of $S(i;t)$ and $P(i,j;t)$. We wish to scale $S(i;t)$ and $P(i,j;t)$ down or up depending on whether each node is firing too frequently or too infrequently, and whether connectivity to that node is too high or too low. Our model for dynamical homeostasis then consists of the equations

$$\frac{\partial}{\partial t}S(i;t) = -[k_{11}\Delta f(i;t) + k_{12}\Delta\eta(i;t)]S(i;t), \quad (4)$$

$$\frac{\partial}{\partial t}P(i,j;t) = -[k_{21}\Delta f(i;t) + k_{22}\Delta\eta(i;t) + k_D D(i,j)]P(i,j;t). \quad (5)$$

The rate constants k_{11}, k_{12}, k_{21}, k_{22} set the time scales for scaling of $S(i;t)$ and $P(i,j;t)$ in response to fluctuations of the relative firing rate and input ratio. For convenience we can define a matrix **K**, where the elements of the matrix are the rate constants k_{ij} above. The rate constant k_D controls a distance-dependent cost factor. Larger values of k_D increase the cost of maintaining a connection between two nodes over a period of time, which is greater for nodes that are far apart.

We do not claim that Eqs. (4) and (5) represent the only possible algorithm for regulating the spontaneous and stimulated firing probabilities, and we do not expect these equations to hold for extreme conditions, such as when all connections are severed. We adopt these equations only because they are particularly simple and provide a convenient starting point.

Hebbian learning can be incorporated by increasing $P(i,j;t)$ by a factor $1+C_H$ if firing of node j in the time interval $(t-\delta t,t]$ is followed by firing of node i in the next time interval $(t,t+\delta t]$. We refer to C_H as the Hebbian learning factor. This type of Hebbian learning models long-term potentiation (LTP). Long-term depression (LTD) can be modeled by reducing $P(i,j;t)$ by a factor of $1-C_H$ if firing of node j in the time interval $(t-\delta t,t]$ is not followed by firing of node i in the succeeding time step. The following equations then hold:

$$\text{LTP:} \quad P(i,j;t+\delta t) = P(i,j;t)[1 + C_H F(i;t)F(j;t-\delta t)], \quad (6)$$

TABLE I. Numerical tests of stability. List of all choices k_{ij} such that the relative firing rates and input ratios converged to unity. We allowed each k_{ij} and the Hebbian learning factor C_H to take on the values of 0, 2×10^{-5}, and 0.01 in all possible permutations. A time step of $\delta t = 4$ ms was taken, for a total of 50 million time steps for each set of parameters. Other details are as described in the Methods section. The states that converged are denoted with a "y". All data refer to LTP results except the column marked C_H^*, which denotes identical results for both LTP, LTD, and STDP.

k_{11}	k_{12}	k_{21}	k_{22}	$C_H=0$	$C_H=2\times10^{-5}$	$C_H^*=0.01$
2×10^{-5}	0	0	0.01	y	y	y
2×10^{-5}	0	2×10^{-5}	2×10^{-5}	y	y	
2×10^{-5}	0	2×10^{-5}	0.01	y	y	y
2×10^{-5}	0	0.01	0.01	y	y	y
2×10^{-5}	2×10^{-5}	0	0.01	y	y	y
2×10^{-5}	2×10^{-5}	2×10^{-5}	0.01	y	y	y
2×10^{-5}	0.01	0	0.01	y	y	

LTD:
$$P(i,j;t+\delta t) = P(i,j;t)\{1 - C_H[1 - F(i;t)]$$
$$\times F(j;t-\delta t)\}. \qquad (7)$$

Spike-timing-dependent plasticity (STDP) can be modeled by combining the criteria for both LTP and LTD.

Equations (4) and (5) can be easily integrated numerically. Because $S(i;t)$ and $P(i,j;t)$ represent probabilities, their values are reset to one if their calculated values ever exceed one. Every node is initially allowed to be connected to every other. The total number of nodes is taken to be $N=64$, with 8 rows of 8 nodes in a square lattice with lattice constant a (that is, a is the distance from one node to its nearest neighbor). The time step is taken to be $\delta t = 4$ ms. A refractory period of 20 ms is imposed after every firing at each node. The target firing interval is $\tau_0 = 6.25$ s. These parameters are chosen to be in agreement with experiment [17,18].

RESULTS

Let us first study the stability properties of the dynamical homeostatic equations, Eqs. (4) and (5). We have performed extensive computer simulations, allowing the rate constants k_{ij} and the Hebbian learning factor C_H to take on all possible permutations of the values of 0, 2×10^{-5} and 0.01, using LTP, LTD, and STDP versions of Hebbian learning. There are $3^5 = 243$ possible permutations for each type of Hebbian learning. If convergence is defined numerically to be such that $\langle f(i;t)\rangle$ and $\langle \eta(i;t)\rangle$ both approach unity to within 5% within a total simulation time of 50 million time steps, then the only states that converge are those listed in Table I. By visual inspection, it is generally clear which simulations will converge by 3–5 million time steps, but we continue every simulation for at least 50 million time steps and occasionally up to 300 million time steps. Note that most choices for k_{ij} and C_H resulted in unstable systems. To summarize Table I, it appears that convergence requires the following necessary conditions to be met: (1) the spontaneous firing probability must be greater than zero, (2) scaling of the spontaneous firing probability must be dominated by firing rate homeostasis (the k_{11} term), (3) scaling of the connection strengths must be dominated by critical homeostasis (the k_{22} term), (4)

"off-diagonal" terms are permissible as long as $k_{11}k_{22} > k_{12}k_{21}$, (5) the time scale for critical homeostasis is at least as fast as Hebbian learning ($k_{22} \gtrsim C_H$), (6) critical homeostasis primarily affects scaling of connection strengths, not firing rate ($k_{22} > k_{12}$), and (7) critical homeostasis occurs on a faster time scale than firing rate homeostasis ($k_{22} \gg k_{11}$ converges fastest but sometimes $k_{22} = k_{11}$ may also converge).

We were not able to derive these stability and convergence criteria analytically for the general case. However, if we assume that simultaneous multinodal activations are rare (on average there are 1562 time steps between each firing event for each node), that this solution does not depend on initial conditions for $F(i;t)$ or $A(i;t)$ (so that the steady state is stable to perturbations), that every node has the same steady-state values for firing rate and input ratio (the system is homogeneous), that fluctuations in the zero-time time correlation between the connection strength and activity are small (i.e., it takes finite time for connection strengths to be adjusted up or down), and that $k_D = 0$ (no distance-dependent connectivity cost factor), then we show in the Appendix that convergence to steady state requires that $S(i;t) > 0$ and $\det(\mathbf{K}) > 0$, where $\det(\mathbf{K}) = k_{11}k_{22} - k_{12}k_{21}$. These two requirements are equivalent to the stability criteria (1)–(4) above. Furthermore, the only stable steady state under these conditions is that for which firing rate and critical homeostasis are achieved, i.e., where $\langle f(i;\infty)\rangle = \langle \eta(i;\infty)\rangle = 1$.

The assumptions listed above are all reasonable for our system except for the assumption that $k_D = 0$. Realistic biological systems should have a nonzero distance-dependent connectivity cost factor. However, we will show later in this section that simulation results depend only weakly on this cost factor as long as it is smaller than some particular value. We take up this issue again later in this section; for now, we ignore the term in k_D.

It is instructive to look at sample states that did not converge, and to compare them to states that do converge. In Fig. 1, we show that firing rate homeostasis by itself cannot guarantee critical homeostasis (k_{11} and $k_{21} > 0$, $k_{12}=k_{22}=0$), since $\det(\mathbf{K}) = 0$ in this case. This result is not unexpected, since critical homeostasis represents a constraint distinct from firing rate homeostasis, and there is no reason to expect
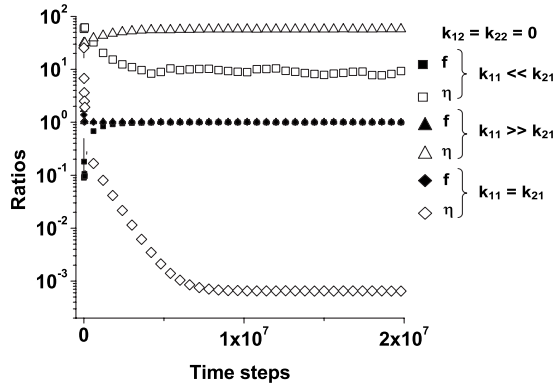
FIG. 1. Relative firing rate f and input ratio η as a function of time, semilogarithmic scale. Each time step is 4 ms. A total of 50 million time steps are taken for each simulation but only the first 20 million time steps are shown. These curves show that firing rate homeostasis by itself (k_{11} and $k_{21} > 0$, $k_{12} = k_{22} = 0$) cannot guarantee critical homeostasis. With $k_{11} = 2 \times 10^{-5}$ and $k_{21} = 0.01$, the average relative firing rate converges to $f = 1$ but the average input ratio converges to $\eta = 9.8$. With $k_{11} = 0.01$ and $k_{21} = 2 \times 10^{-5}$, f converges to one but η converges to 57.2. With $k_{11} = 0.01$ and $k_{21} = 0.01$, f converges to one but η converges to 6.5×10^{-4}. All values are averaged over the prior one hour of data and over all $N = 64$ nodes. The Hebbian learning factor in these calculations is $C_H = 0.01$.

that firing rate homeostasis alone should guarantee critical homeostasis.

In Fig. 2, we show that states with k_{12} and $k_{21} > 0$, and $k_{11} = k_{22} = 0$ are also unstable, because $\det(\mathbf{K}) < 0$ in this case. If the initial $S(i;t)$'s are too low, these states tend to fall into global silence with zero firing rates and high input ratios $\eta(i) \approx N$. If the initial $S(i;t)$'s are too high, these states tend to approach tonic hyperactivity with low input ratios, $\eta(i) \ll 1$. When $\det(\mathbf{K}) < 0$, there appears to be no mechanism for the $S(i;t)$'s to find an optimal value such that firing rates and input ratios are stable about their target values.

In Fig. 3, we show that when $\det(\mathbf{K}) > 0$, stable states are found that maintain relative firing rates and input ratios that fluctuate about unity. Interestingly, not every state with $\det(\mathbf{K}) > 0$ converges to the correct target values. These states appear to depend on initial conditions. It appears from our exploration of numerical examples that additional criteria for convergence to target values are that $k_{22} \geq C_H$, that $k_{22} > k_{12}$, and that $k_{22} \geq k_{11}$ (see Table I). Although we were unable to prove this analytically, it may be that these additional criteria remove the dependence on initial conditions.

The numerical convergence condition $k_{22} \geq C_H$ states that scaling of the connection strengths must not be much slower than the time scale for Hebbian learning, because otherwise Hebbian learning would dominate critical homeostasis and make critical homeostasis impossible. If the condition $k_{22} \geq C_H$ were not satisfied, then repeated Hebbian learning would eventually destabilize the network, causing either network overconnectivity or underconnectivity. Conversely, if $k_{22} \gg C_H$, then the neural network would have stable connectivity but learning would be very slow. Optimal learning with stable connectivity is achieved when $k_{22} \approx C_H$. A caveat, however, is that it is acceptable to allow C_H to be slightly
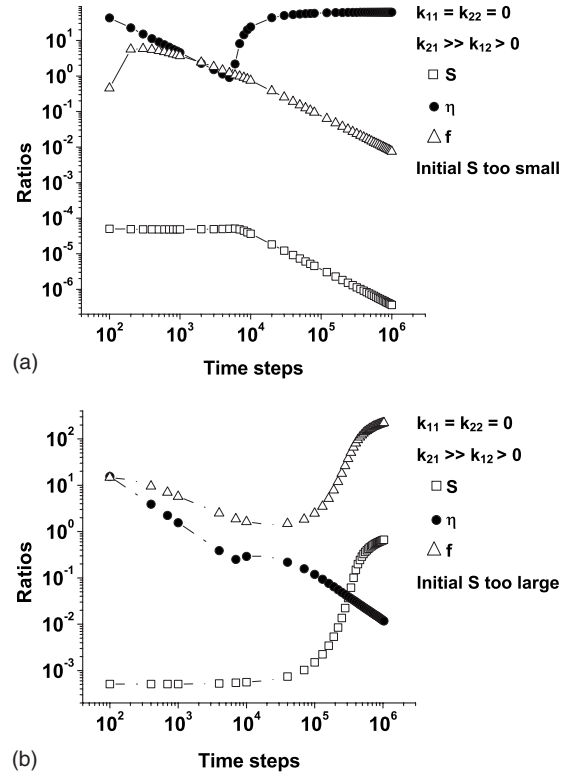

(a)

(b)

FIG. 2. Relative firing rate f and input ratio η as a function of time, log-log scale. Each time step is 4 ms, $C_H = 0.01$. These two examples show that firing rate homeostasis scaling of connection strengths ($k_{21} = 0.01$, $k_{22} = 0$) and critical homeostasis scaling of spontaneous firing probabilities ($k_{12} = 2 \times 10^{-5}$, $k_{11} = 0$) result in an unstable system. In example (a), the initial $S(i;0)$'s are too small to maintain firing rates and input ratios at unity. When the relative firing rates eventually drop below unity, the connection strengths $P(i,j;t)$ are scaled up and the input ratios rise. As the input ratios rise, the spontaneous firing probabilities are scaled down further, which causes the relative firing rates to drop even more, perpetuating a cycle such that the system eventually becomes silent ($f \approx 0$) despite maximal input ratios [$\eta(i) \approx N$, where N is the total number of nodes]. Varying the relative magnitudes of k_{12} and k_{21} does not improve stability. In example (b), the initial $S(i;0)$'s are too large, and the opposite situation arises, where a cycle is entered such that the spontaneous firing probabilities rise as the input ratios drop. This system then enters tonic hyperactivity with minimal connectivity.

larger than k_{22} for a certain period of time (i.e., while the animal is awake during which speed of learning is advantageous) but then turning C_H down or even off for another period of time (i.e., when the animal sleeps), to allow critical homeostasis to catch up to learning (compare Refs. [36,37]).

Why k_{22} must be greater than k_{11} for convergence to occur is not obvious to us. However, an evolutionarily favorable consequence of this criterion is that the dynamics of changes in connectivity is then unconstrained by firing rate homeostasis. That is, if firing rate homeostasis enters on a longer time scale than critical homeostasis, and if the time scale for critical homeostasis and Hebbian learning are similar, then firing rate homeostasis is less likely to interfere with the speed at which an animal can learn.
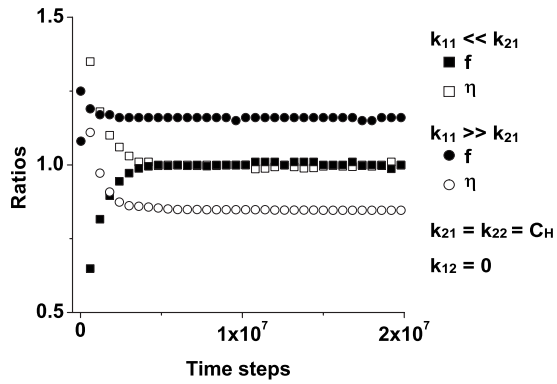
FIG. 3. Relative firing rate f and input ratio η as a function of time, linear scale. Time step is 4 ms. These examples show that the convergence criterion $\det(\mathbf{K}) > 0$ does not guarantee convergence to the target asymptotic firing rate and input ratio. With $k_{11} = 2 \times 10^{-5}$, $k_{12} = 0$, and $k_{21} = k_{22} = C_H = 0.01$, both f and η converge to unity. However, with $k_{11} = 0.01$, $k_{12} = 0$, and $k_{21} = k_{22} = C_H = 2 \times 10^{-5}$, f converges to 1.16 while η converges to 0.84, even through $\det(\mathbf{K})$ is identical in the two cases. All values are averaged over the prior one hour and over all $N = 64$ nodes.

The stability of our model to fluctuations in the firing rate and branching ratio seems to depend strongly on the behavior of the spontaneous firing probability $S(i;t)$, even though steady-state values of $S(i;t)$ are very small. In our examples, the average $S(i;t)$ is on the order of $1 \times 10^{-5} - 8 \times 10^{-5}$. With a time step of $\delta t = 4$ ms and target firing interval of $\tau_0 = 6.25$ s, the proportion of total firing due to spontaneous firing then comes out to $2 - 12\%$. That is, even though most of the activity of our system is due to connectivity-related activity, nonetheless nonzero spontaneous activity is necessary for system stability.

We do not have experimental values for $\mathbf{K}$, C_H, or k_D. These parameters will depend on the size of the electrodes and the number of neurons each electrode overlies. As discussed earlier, these parameters do not have a simple relationship to analogous neuronal properties. For instance, the nodal spontaneous firing probability will depend on both the local neuronal spontaneous firing probabilities and also on the strengths of local neuronal connectivities, while nodal connection strengths will depend on longer-distance connections between neurons underlying different microelectrodes. Further, the time scales of neuronal dynamics will not translate directly to nodal time scales. Nonetheless, we can make a few statements about their relative magnitudes. First, the rate constant k_{22} should be on the same order of magnitude as the Hebbian learning factor C_H, as discussed above, in order that neither dominates the other. Second, all of these time scales must be longer than the simulation time step; if this were not true, then we would need to choose a smaller simulation time step. In effect, this means all the k_{ij}'s must be much smaller than one, in units of $1/\delta t$. Third, we expect the rate constant k_{11} to be much slower than k_{22}, so that firing rate homeostasis infringes as little as possible on the ability of the neural system to react quickly to the environment. We found above that such a choice also appears to be necessary for the firing rate and input ratio to converge to the correct values.

To demonstrate the effect of varying k_D, let $k_{11} = 2 \times 10^{-5}$, $k_{12} = 0$, and $k_{21} = k_{22} = C_H = 0.01$. We integrated Eqs. (4) and (5) using k_D equal to 0, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, and 10^{-2}, with k_D in units of $1/(a\delta t)$, where a is the lattice constant, $\delta t = 4$ ms is the time step, and a total of 50 million time steps are taken per simulation. For the representative case of $k_D = 10^{-5}$, the average input ratio comes out to $\eta = 0.99$ with standard deviation 0.22, and the relative firing rate is 1.01 with standard deviation 0.26.

When k_D is between 10^{-5} and 10^{-3}, the distribution of avalanche sizes $G_A(n)$ as a function of avalanche size n [Fig. 4(a)] is suggestive of a power law with an exponent close to -1.5. At smaller values of k_D, the distribution $G_A(n)$ is still suggestive of a power law but there is an upturn at large avalanche sizes with greater than expected frequency of these larger avalanche sizes. Values of k_D larger than 10^{-3} destroy the power law dependence, replacing it by a decreasing exponential with a cutoff.

In Fig. 4(b), we show the relative strength of connectivity as a function of distance. As k_D increases, the connectivity strength drops off very quickly as a function of distance. By $k_D = 10^{-3}$, next-nearest neighbor connectivities are 5 orders of magnitude weaker than nearest-neighbor interactions, and yet the corresponding plot in Fig. 4(a) shows that avalanche sizes can still span almost the entire network of 64 nodes. However, by $k_D = 10^{-2}$, there are no connections beyond nearest neighbors, and in this case, Fig. 4(a) shows that avalanche sizes are sharply curtailed, to a maximal size of less than 10 nodes. These results suggest that being able to activate large avalanche sizes requires at least some connections beyond nearest neighbors.

The rate constant k_D is related to a length scale $L = 1/(k_D\tau_0)$. This length scale is that at which the increased cost of maintaining a long-distance connection becomes apparent on the time scale of the target firing rate. When k_D is between 10^{-5} and 10^{-3}, L is between 64 and 0.64, respectively, in units of the lattice constant a. This range corresponds to the length scales of our 8×8 system. Thus the results of Figs. 4(a) and 4(b) and suggest that as long as $L \geq a$, connectivities extend beyond nearest neighbors and activation patterns can spread to cover the entire array.

Figure 4(c) shows the averaged values of the input ratio, relative firing rate, and spontaneous firing probability as a function of k_D. As long as $k_D < 10^{-2}$, it is possible to maintain the average input ratio very near unity. That is, critical homeostasis is not very sensitive to the distance-dependent cost factor, until the cost of maintaining connectivity is so high that there are no connections at all beyond nearest neighbors. At larger values of k_D, it is no longer possible to maintain critical homeostasis, although firing rate homeostasis is still maintained.

The weak dependence of critical homeostasis on k_D for $k_D < 10^{-2}$ allowed us to ignore k_D in the stability and convergence analysis performed in the Appendix.

Figure 4(d) shows the distribution of connection strengths as a function of connection strength P. As long as $k_D < 10^{-2}$, the general shape of this distribution is of a tall peak near $P = 0$ with a long, flat tail or a smooth hump extending out to the maximal strength of $P = 1$. For the representative plot of $k_D = 10^{-5}$, about 97% of connection strengths are less
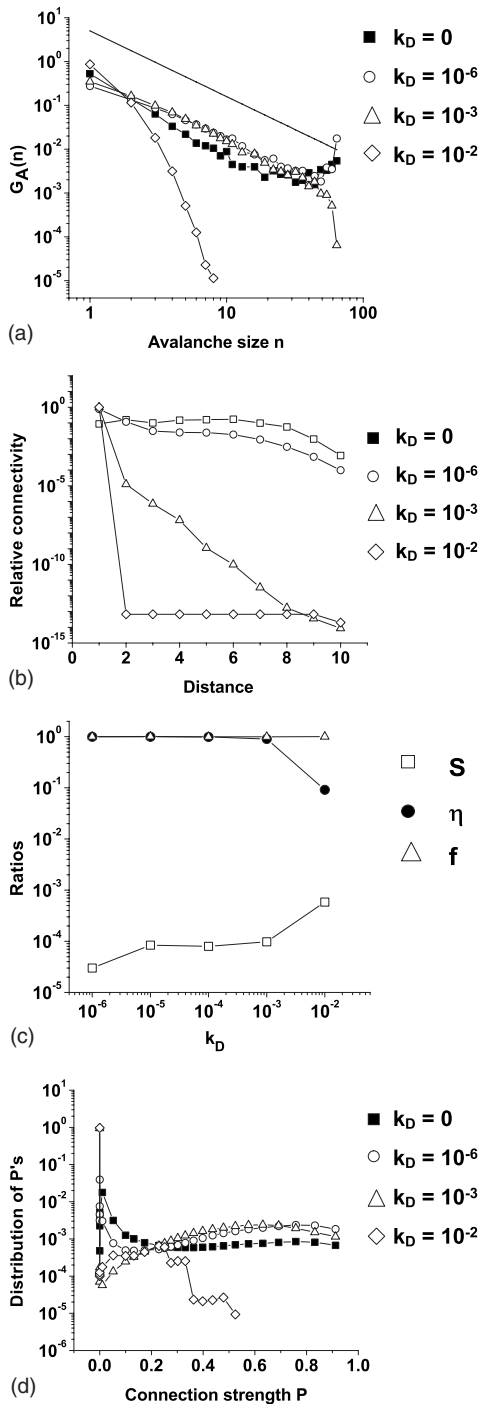
FIG. 4. Dependence on k_D. Example with $k_{11}=2\times10^{-5}$, $k_{12}=0$, $k_{21}=k_{22}=C_H=0.01$, time step of 4 ms and total of 50 million time steps. (a) Log-log plot of the distribution of avalanche sizes $G_A(n)$ vs avalanche size n. The solid line is a power law with an exponent of -1.5 shown for reference. (b) Semilogarithmic plot of relative connectivity vs distance, calculated by adding all $P(i,j;t)$'s at each distance and normalizing area under the curve to unity. Distances are in units of the lattice constant a, rounded to the nearest integral value. (c) Log-log plot of averaged spontaneous firing probability, input ratio, and relative firing rate as a function of k_D. (d) Semilogarithmic plot of distribution of connection strengths $P(i,j;t)$ vs connection strength, with total area under the curve normalized to unity.

than $P=0.001$. Thus the connectivity pattern, for $k_D<10^{-2}$, is generally sparse.

Connectivity that is sparse and consists of mostly nearest-neighbor connections is less costly for biological neural systems to maintain. When such networks are nonetheless capable of activating large-scale clusters spanning almost the entire network, one is able to reduce metabolic cost (by retaining fewer long-distance connections) without giving up much information processing power. This favorable situation is present for our examples with $k_D<10^{-2}$, and is thus relatively insensitive to k_D, as long as the length scale requirement $L\geq a$ is satisfied. Such a situation is suggestive of so-called small world connectivity [20–23]. Small world connectivity has been shown to have precisely this property of being able to maintain global connectivity even with mostly local connections. Furthermore, the ability to maintain global connectivity is relatively insensitive to the percentage of long-distance connections, until this percentage is very nearly zero. This relative insensitivity to the percentage of long-distance connections is also demonstrated by our model qualitatively [Figs. 4(a) and 4(b)].

Finally, in addition to the specific permutations in the values of the k_{ij}'s above, we also randomly sampled other values of k_{ij}. We did not find any exceptions to the stability and convergence criteria discussed above. What appeared most important are not the absolute values of the k_{ij}'s, but rather their relative values. The absolute values control the absolute time scales of the dynamics, but the relative values are key for stability and convergence.

DISCUSSION

We have explored the requirements of a simple biologically plausible neural system model that incorporates (1) spontaneous nodal firing activity, (2) stimulated nodal firing activity, (3) firing rate homeostasis, (4) critical homeostasis, (5) Hebbian learning, and (6) a distance-dependent connectivity cost function. Our goals were to investigate first whether it is possible to construct such a model, and second whether there are algorithmic consequences of imposing both firing rate and critical homeostasis for this particular model. Our primary findings are that we were indeed able to construct such a model, and we found that not every possible set of parameters for our model is stable.

In particular, within the limits of our simple model, we found from extensive computer simulations that stability and convergence of the system to target firing rates and connectivity require that (1) the spontaneous firing probability must be greater than zero, (2) scaling of the spontaneous firing probability must be dominated by firing rate homeostasis (the k_{11} term), (3) scaling of the connection strengths must be dominated by critical homeostasis (the k_{22} term), (4) "off-diagonal" terms are permissible as long as $k_{11}k_{22}>k_{12}k_{21}$, (5) the time scale for critical homeostasis is at least as fast as Hebbian learning ($k_{22}\geq C_H$), (6) critical homeostasis primarily affects scaling of connection strengths, not firing rate ($k_{22}>k_{12}$), and (7) critical homeostasis occurs on a faster time scale than firing rate homeostasis ($k_{22}\geq k_{11}$). The first four stability conditions can also be derived analytically, un-

der certain circumstances discussed as above and in the Appendix.

A secondary finding of our simulations is that by making only the simple, biologically plausible assumptions above, the resulting model neural system is able to activate cluster or avalanche sizes ranging from the smallest to the largest possible sizes (the system is scale-free), and this property is maintained even when most connection strengths are very weak and most are very short distance (i.e., when $k_D > 0$). This finding suggests that neural systems that satisfy our basic assumptions have available to them near-maximal computational power at low metabolic cost, over a fairly wide range of parameters.

We do not know if our conclusions are generalizable to other specific neural system models, nor whether real biological systems have the same stability and convergence requirements. Nonetheless, our results suggest that real biological systems may have similar constraints, which must be satisfied in order for these systems to function optimally. If such constraints exist, then they will be important to discover and characterize, because real life failure to satisfy these constraints will represent disease states. For instance, neural tissue that is persistently underconnected will have difficulty activating large-scale circuits, and may then result in learning disorders and mental retardation. Understanding the nature of these constraints may also help us to manipulate them in times of physiological stress, to protect patients after a brain insult.

As an example, we simulated a seizure on our system by forcing 12 out of 64 nodes to fire at rates far above the target rates for a period of time (Fig. 5). From simulations, we found that, during a seizure, hyperactive firing at one set of nodes drives the spontaneous firing probability of all system nodes to lower values [Fig. 5(a)]. The connectivities are also driven to lower values. At the end of the seizure, both the spontaneous firing probabilities and the connectivities begin to recover, but the connectivities recover faster because $k_{22} > k_{11}$. If it happens that $k_{22} \gg k_{11}$, then the connectivity will actually overshoot for a period of time, and the postseizural state will be supercritical even though activity is depressed [Figs. 5(b) and 5(c)]. In our example, the postseizural supercritical state lasts on the order of 50 million time steps (about 55 h). We conjecture that this postseizural supercritical state plays a role in epileptogenesis by increasing the chance of creating a large-scale hypersynchronous pattern of activation. Such a pattern, if repeated often enough, may be "burned into memory" and thus may be reactivated at some unpredictable time in the future. The possibility of reactivating a spatially extended hypersynchronous state is a prerequisite of epilepsy, and is a separate condition distinct from the more traditional concept of increased neuronal hyperexcitability. If the postseizural state is indeed supercritical and epileptogenic, then we suggest that postseizural interventions that boost spontaneous firing, for example, by high frequency subthreshold electrical stimulation, may relieve the drive towards supercriticality and may be protective against epileptogenesis. Spontaneous activity and connectivity-dependent activity have opposite effects on epileptogenesis, and thus it is important to distinguish between these two types of activity.
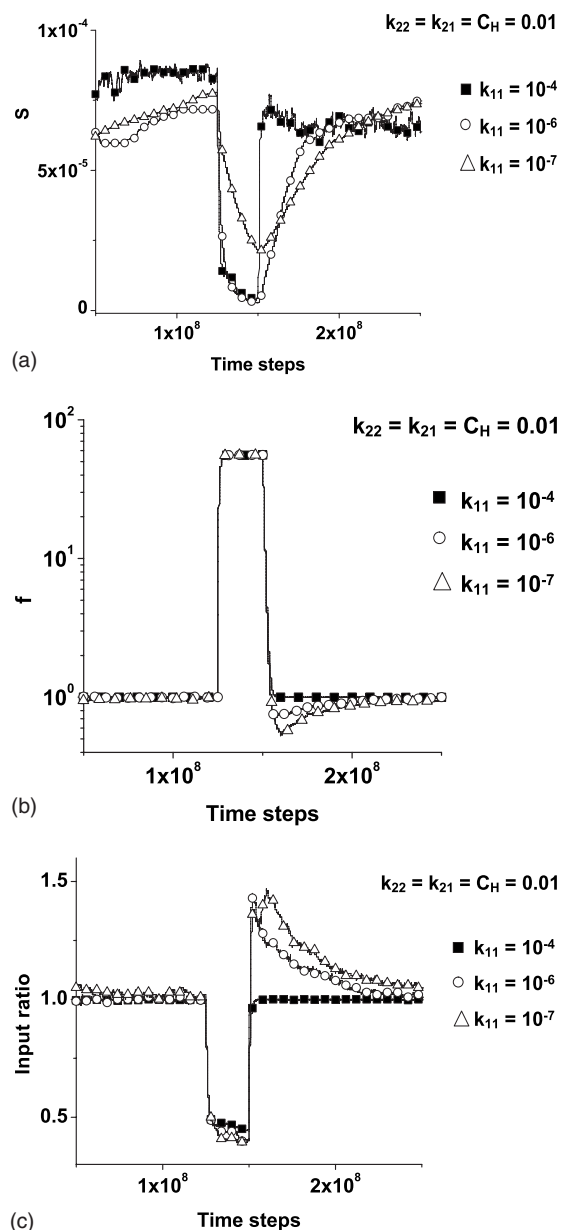

FIG. 5. The spontaneous firing probability S, relative firing rate f, and input ratio η as a function of time. A simulated seizure occurs between time steps 1.25×10^8 and 1.5×10^8 during which the activation levels $A(i;t)$ of 12 nodes were set to 0.8. In all examples, $k_{12} = 0$, $k_{21} = k_{22} = C_H = 0.01$, and the time step is 4 ms. (a) The spontaneous firing probability averaged over all nodes S declines with seizure onset and slowly recovers after the seizure ends. The rate of recovery is slower for smaller values of k_{11}. (b) The relative firing rate f is elevated during the seizure and becomes depressed for a time after a seizure ends, more prominently so for smaller values of k_{11}. Note the semilogarithmic scale. (c) The average input ratio η is depressed during a seizure and becomes elevated (i.e., the system is supercritical) for a time after a seizure. For $k_{11} = 10^{-6}$ and $k_{11} = 10^{-7}$, the input ratio remains supercritical for about 50 million time steps (55 h).

Another example of a disease state is when an area of cortex is suddenly cut off from its neighbors, or "deafferentated" [38]. Deafferentation is a model for traumatic brain
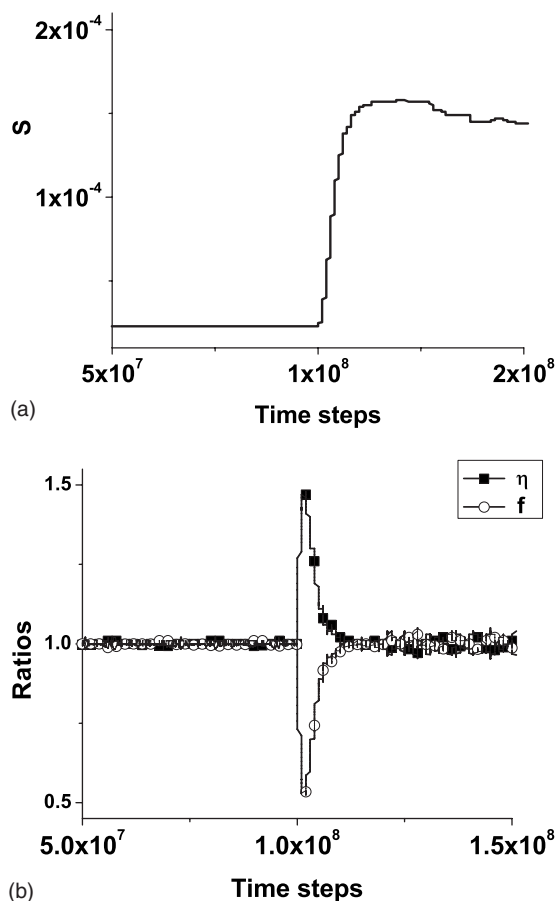
FIG. 6. The spontaneous firing probability S, relative firing rate f, and input ratio η as a function of time. Simulated acute deafferentation occurs at time step 1×10^8, at which time a system of 100 nodes is suddenly reduced to 10 nodes. Parameters are $k_{11}=10^{-6}$, $k_{12}=0$, $k_{21}=k_{22}=C_H=0.01$, and the time step is 4 ms. (a) The spontaneous firing probability S as a function of time rises to a new steady state after acute deafferentation. (b) The relative firing rate f declines with acute deafferentation and does not recover to steady-state baseline until S has reached its new baseline. The input ratio η compensates for the depressed relative firing rate by rising to supercritical levels and remaining there until S and f have both reached their new steady-state levels. The supercritical state in this example lasts for 8 million time steps (9 h).

injury. Immediately after deafferentation, neuronal activity drops precipitously. Epileptiform discharges then later appear. Why does this happen? Assuming that the target firing interval τ_0 is the same in intact brain as in deafferentated brain, our model predicts that the baseline spontaneous firing probability needed to maintain the target firing rate should be smaller for intact brain than for deafferentated brain [Fig. 6(a)]. The reason is that spontaneous activity can propagate to other nodes, and hence, if there are more nodes, lower spontaneous activity is needed to maintain the same level of activity. With acute deafferentation the firing rate therefore immediately drops [Fig. 6(b)]. In response to the drop in firing rate, connectivities increase to supercritical levels and remain supercritical until the spontaneous firing probability reaches its new steady-state value [Fig. 6(b)]. In our example, the postdeafferentation supercritical state lasts for

9 h. Again, we hypothesize that this period of supercriticality is one during which epileptogenesis may occur, and that interventions that either boost spontaneous firing or that inhibit connectivity-dependent firing will decrease the tendency towards supercriticality, and thus be protective against epileptogenesis. In addition, if for some reason there is a ceiling on maximal spontaneous firing probabilities, such that the deafferentated nodes are blocked from reaching these critical spontaneous firing rates, then it is possible to produce an indefinitely persistent supercritical state.

In experimental models of neocortical post-traumatic brain injury, one sees immediate focal slowing of electrical activity with suppression of fast oscillations. After a period of weeks, one then sees increased neuronal hyperexcitability, decreased efficacy of inhibitory systems, and pathological recurrent axonal sprouting of layer V principal neurons with increased functional connectivity [39–44]. Neuronal hyperexcitability can be the result of either increased spontaneous activity or increased connectivity. To distinguish between these two etiologies would require network-level multielectrode correlation analysis, which to our knowledge has not been done. Axonal sprouting may be interpreted as an increased drive towards greater connectivity, in agreement with our prediction, but if the length scale at which axonal sprouting occurs is much smaller than the size of a node, then axonal sprouting may result in activity that is interpreted as increased spontaneous nodal activity. It would be interesting in such systems to distinguish spontaneous from connectivity-related activity, as we predict that there should be a period of time during post-traumatic epileptogenesis where one finds increased connectivity-related activity in the presence of depressed spontaneous activity. Also of interest is the finding that there appears to be a "critical period" in post-traumatic epilepsy, where suppression of neuronal activity with focal application of tetrodotoxin, if applied for three days starting no later than 3 days after injury, results in suppression of epileptogenesis [43,44]. We hypothesize that such suppression allows the drive towards increased spontaneous firing enough time to rise to a new steady state, while preventing connectivity-dependent firing from reaching such levels as to burn into memory spatially extended hypersynchronous patterns of activation. It would be interesting to monitor electrical activity during the time period of tetrodotoxin application, to see if there is indeed suppression of connectivity-dependent activity.

Latham *et al.* [45] have also studied the role of spontaneous activity in neural networks. This model is somewhat different from ours, involving a heterogeneous population of cells, which includes a fraction of endogeneously active cells. They find that increasing the fraction of endogenously active cells produces firing patterns that are more regular and of lower frequencies. Systems with too few endogeneously active cells are either silent or fire in bursts of high frequency. Our system is homogeneous and describes not individual cells but groups of cells, each group represented by a node. In our model, when the spontaneous firing probabilities $S(i;t)$ are too low for a prolonged period, for whatever reason, the firing rates also become too low. In partial compensation, the connection strengths $P(i,j;t)$ then rise to supercritical levels. What tends to happen in this situation is

that the baseline firing rate is still too low, but when there is spontaneous activity, that activity quickly spreads to many nodes, in a burst, because the connectivity is supercritical. During the period of the activity burst, the firing frequency would appear very high. If the $S(i;t)$'s are allowed to recover (that is, as they slowly increase), then the $P(i,j;t)$'s are scaled back down. When the connectivity returns to critical or subcritical levels, then spontaneous firing of a node is less likely to spread quickly to a large number of other nodes. The firing pattern then becomes more regular. If we take our spontaneous firing probability $S(i;t)$ to be analogous to the fraction of endogenously active cells of Latham *et al.* then our conclusions are concordant with theirs.

Regarding spontaneous firing, we also found that spontaneous activity accounts for only 2–12 % of total activity, and yet stability about target firing rate and connectivity requires that spontaneous activity must be greater than zero. Perhaps this requirement for stability is the reason why the brains of higher animals (i.e., those capable of learning and adapting to the environment) are never entirely quiet, even at rest and even in sleep.

Two of us have previously attempted to construct a model exhibiting both firing rate homeostasis and critical connectivity, but assuming only a mechanism for firing rate homeostasis [46]. This earlier model is the same as the model presented here but with no homeostatic mechanism for connectivity (i.e., k_{21} and k_{22} were both equal to zero). In addition, the earlier model was not capable of Hebbian learning ($C_H=0$). It was found that as long as the rate constant for scaling of the spontaneous firing probability is smaller than that for the connection strength ($k_{11}<k_{21}$), then critical connectivity was approximately maintained, with best results for $k_{11}/k_{21}=0.5$. However, when Hebbian learning was turned back on ($C_H>0$), no stable state that converged on target firing rates and critical connectivity could be found despite extensive exploration of parameter space [71]. It was the failure to find such a state, when Hebbian learning was turned on, that led us to consider whether homeostasis of connectivity may be a homeostatic principle in its own right, distinct from that of firing rate homeostasis.

Abbott and Rohrkemper [47] have proposed an elegant "growth" model that essentially allowed connectivity to grow or shrink so as to achieve firing rate homeostasis. Power-law behaviors were then seen suggestive of critical connectivity. That is, critical behavior was seen in this model assuming only firing rate homeostasis, without having to invoke critical homeostasis explicitly. As in our own earlier effort, however, this model does not incorporate Hebbian learning. It would be interesting to know whether this growth model remains capable of critical behavior after Hebbian learning is turned on.

Sullivan and de Sa [48] proposed a simple model capable of both Hebbian learning and firing rate homeostasis. However, they did not investigate whether their model exhibits critical connectivity, and their model is not meant to be spontaneously active. The model of Yeung *et al.* [49] incorporating a calcium-dependent plasticity is appealing for its underlying realism, but it similarly does not investigate issues of critical connectivity. Others have investigated mechanisms of

firing rate homeostasis, usually in the form of constraints [2,3]. A modification of spike-timing-dependent plasticity learning rules can also achieve activity homeostasis [50]. However, this model does not explicitly incorporate critical homeostasis, and it is not clear that critical homeostasis would arise as a natural consequence of the dynamics. There are also models which are capable of self-organized criticality based purely on local learning rules [51]. Such models are attractive in that they do not need to suppose nonlocal learning rules. In contrast, we assume that each node must sense the total input to that node, which is a nodal-wide property and not a strictly local (i.e., synaptic) property. However, the experimental finding of multiplicative synaptic scaling [12] and of active and passive backpropagation of action potentials into the dendritic tree [52–54] show that nonlocal neuron-wide mechanisms do exist, and suggest that nodal-wide mechanisms may also exist. We leave it to future experiments to determine whether critical homeostasis really involves only purely synaptic-level learning rules.

The precise magnitudes of the rate constants k_{ij} are not known, as experiments have not been framed in terms of determining these rate constants. Of particular interest are not simply the absolute magnitudes of these rate constants, but their relative magnitudes, because stability and convergence depends on the relative magnitudes and not so much on the absolute magnitudes. We suggest that it may be interesting to design experiments to look at these time scales, to see if these time scales are constrained in the ways we have described above, and also to see if a simple scaling algorithm is found as in Eqs. (4) and (5).

In terms of orders of magnitudes of time scales, the synaptic scaling mechanism of Wierenga *et al.* [12] is a slow process, with a time scale of hours to days. Since critical homeostasis must occur on the same time scale as Hebbian learning, in order that neither critical homeostasis nor Hebbian learning dominates the other, we must look elsewhere for a fast biomolecular mechanism for critical homeostasis. In addition to being fast, such a mechanism must also be nonlocal, i.e., not restricted to the level of individual synapses, because the input and branching ratios are nonlocal properties requiring simultaneous knowledge of total input and output weights across an entire node. Here we mention three examples of candidate fast, nonlocal mechanisms: (a) When homosynaptic LTP (or LTD) is induced in the intercalated neurons of the amygdala, compensatory heterosynaptic depression (or facilitation) is observed such that the total synaptic weight of a given neuron remains constant [55]. The counterbalancing heterosynaptic response is suggestive of the scaling mechanism of Eqs. (4) and (5). This mechanism depends on the release of intracellular calcium stores and has a time scale of minutes. Similar heterosynaptic LTD has been reviewed by Artola and Singer [56,57]. (b) The phenomenon of backpropagation of action potentials into the dendritic tree, as seen in certain neocortical and hippocampal neurons [52–54], allows widely distributed numbers of synapses to receive nearly simultaneous information about neuronal output. This information is conjectured to play a role in LTP, LTD, and STDP, but might conceivably also be used for critical homeostasis. (c) A third mechanism may involve the interaction of principal output neurons with local interneu-

rons. It may be that a certain subset of local interneurons can sense both the total input into and total output out of a local community of output neurons. This information might then be used to modulate either the input or branching ratio of that group of output neurons. In support of this possibility, blocking interneurons with bicuculline can produce a dramatic increase in the branching ratio within minutes [70].

Which of these three mechanisms, if any, are related to critical homeostasis remains to be seen. Nonetheless, that these mechanisms are all fast and nonlocal suggests that an appropriate mechanism for critical homeostasis is at least biologically plausible. We reiterate that there are theoretical reasons why critical homeostasis should exist in systems that maintain capacity for continual learning, because neural systems that operate far from criticality are inefficient and unreliable. Thus, we hope to encourage future experiments to determine which neural systems are capable of critical homeostasis, how tightly critical homeostasis is maintained, and what mechanisms underlie it.

Desirable improvements in our model include generalization to include higher order multinodal interactions [e.g., terms of the form $P(i,j,k;t)$ as discussed in the Methods section], the inclusion of non-Markovian memory effects in the pairwise connection strengths, incorporating heterogeneous neural subsystems, and including the effect of inhibitory interactions. The inclusion of multinodal effects would be important when simultaneous multinodal discharges are frequent, as in a seizure, or in the case of cells that preferentially activate after correlated input from two different cell populations, for example, the cerebellar Purkinje cells [58–60] and possibly the granule cells of the hippocampal dentate gyrus [61].

By Markovian connectivity, we mean that $P(i,j;t)$ is defined as the conditional probability that firing of node j at time t causes firing of node i in the next time instant, but there is no further effect at time steps beyond the immediately adjacent time step. Non-Markovian connectivity would allow a node to remember from which other nodes it received input during time steps beyond the immediately preceding time step. Such memory effects in the connectivity would allow the formation of more complex spatiotemporal patterns of activity.

To include multinodal interactions and non-Markovian connectivity, we experimented with the following generalizations:

$$A(i;t+\delta t) = 1 - [1 - S(i;t)]\prod_{j=1}^{N}[1 - Q(i,j;t)], \qquad (8)$$

$$Q(i,j;t) = G(i,j;t)\left[1 + \lambda\sum_{k\neq j}^{N} G(i,k;t)\right], \qquad (9)$$

$$G(i,j;t) = \sum_{t'=0}^{\infty} G_0(t')P(i,j;t-t')F(j;t-t'). \qquad (10)$$

Here $G(i,j;t)$ represents the conditional probability that node i fires at time t due to activity from node j at any prior
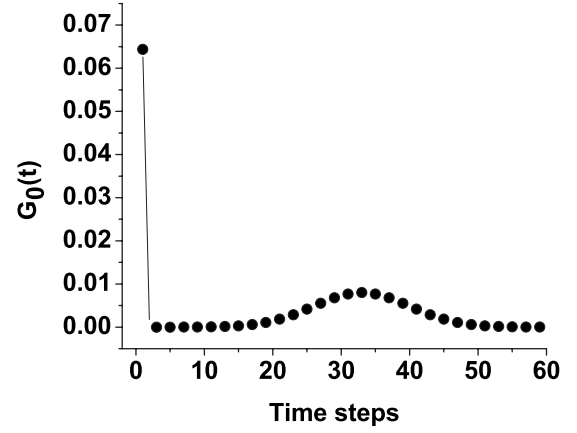

FIG. 7. Non-Markovian weighting function $G_0(t)$ as a function of time. In this example, $G_0(t)$ consists of a step function at short times plus a Gaussian centered at longer times. The areas under the step function and under the Gaussian are approximately equal.

time, summed over all prior times t' with a non-Markovian weighting factor $G_0(t')$. We choose $G_0(t')$ to have the form of a delta function at short times plus a Gaussian centered at longer times (see Fig. 7). The delta function at short times represents the Markovian component of the connectivity, while the Gaussian at longer times represents non-Markovian connectivity. The parameter $\lambda \geq 0$ controls the strength of multinodal interactions. The function $Q(i,j;t)$ is taken to have a maximal value of one, so that it is reset to one if the right-hand side of Eq. (9) ever exceeds one. The input and branching ratios are then given by

$$\eta(i;t) = \left[\sum_{j=1}^{N} P(i,j;t)\left(1 + \lambda\sum_{k\neq j}^{N} G(i,k,t)\right)\right]\left[\sum_{t'=0}^{\infty} G_0(t')\right],$$
$$(11)$$

$$\sigma(i;t) = \left[\sum_{j=1}^{N} P(j,i;t)\left(1 + \lambda\sum_{k\neq j}^{N} G(k,i,t)\right)\right]\left[\sum_{t'=0}^{\infty} G_0(t')\right].$$
$$(12)$$

We have not explored Eqs. (4)–(12) as systematically as Eqs. (1)–(7), but with $\lambda = 0.5$ and $\lambda = 1$, and $\tau_0 = 6.25$ s and $\tau_0 = 120$ ms, we find through computer simulations that the same stability criteria as described above appear to be preserved.

The significance of Markovian vs non-Markovian connectivity is that Markovian systems are not as good for the formation of sequential memory, wherein specific temporal sequences of spatial patterns are to be learned. To see this, consider what happens if exactly the same stimulus is fed into a neural system with Markovian connectivity, at intervals longer than the nodal refractory period. In this case, every stimulus appears exactly the same to the neural system, and the neural system has no way to know if it has just seen the stimulus for the first time, the second time, or the nth time. There is no sense of "history" in a system with Markovian connectivity, no possibility of "historical con-

text." Although such a system is capable of long-term memory, it will respond to the same stimulus in the same way, every time that stimulus is presented, regardless of what may have happened in the meantime. In contrast, a system with non-Markovian connectivity will see each succeeding stimulus a little differently, as modified by recent history. To consider a visuospatial example, an animal that has Markovian connectivity may be capable of taking in a visual scene and remembering it, as of a photographic still. An animal that has non-Markovian connectivity, on the other hand, may be capable of stringing multiple photographic stills together to create a movie. Similarly, an animal with Markovian connectivity may be trained to recognize the letters of the alphabet, but being able to read and understand a novel requires non-Markovian connectivity. We predict that animals that are capable of historical context, and that can understand a movie or a novel, have non-Markovian connectivity.

An additional feature of non-Markovian connectivity is that one can imagine taking a given memorized temporal sequence of spatial activation patterns, and adding a tail such that the temporal sequence returns to the beginning, thus creating a recurrent spatiotemporal loop. If this loop is activated and the brain cannot break out of it for an abnormally long period of time, then such a loop may represent a seizure. Thus, epileptogenesis may involve not just the creation of a hyperexcitable state (the traditional conception) and the existence of supercritical connectivity (as described earlier in this section), but it may also require non-Markovian recurrent loops. If any of these three prerequisites for spontaneous seizures can be disrupted, then epileptogenesis might be prevented. It is interesting in this regard that animals such as reptiles and insects are much less likely to develop epilepsy, in their natural states, than presumably more intelligent animals such as humans and dogs. Non-Markovian connectivity, if it exists, is both a gift and a curse.

The brains of even lower animals are quite complex systems consisting of many subsystems, each possibly with its own target firing rate, Hebbian learning rate, and homeostatic scaling constants. Our model is easily generalized to allow the simulation of such heterogeneous systems simply by allowing these parameters to vary among the different subpopulations.

We have ignored inhibitory interactions in our model in the interest of simplicity. In the neocortex, on the order of 70% of all neurons are excitatory pyramidal cells [62]. Furthermore, most inhibitory cells are interneurons and have shorter ranged projections than the excitatory cells. Thus, one may regard the inhibitory interneurons in the neocortex as being mostly involved in modifying local circuits. These functions are likely to be highly complex [63–69], including the modulation of local excitability (the most obvious function), the timing and synchronization of firing of excitatory cells, activity-dependent disinhibition, and possibly even critical homeostasis as discussed above. We do not know the upper and lower ranges of the length scales at which critical homeostasis is active, and thus it is not clear whether neglecting inhibitory interactions in our model is a serious omission. Determining these length scales is crucial. If inhibitory interactions are indeed active at the nodal level of description, then the next question would be whether inhibitory interactions obey an as yet unidentified, new constraint, distinct from the homeostasis of activity and connectivity considered here.

APPENDIX

We were not able to find a general closed form expression for a steady-state solution of Eqs. (4) and (5). However, if we make certain assumptions, then a steady-state solution can be found. The assumptions are suggestive, but we do not know if they are necessary conditions for a steady-state solution to exist.

The assumptions we make are (a) that simultaneous multinodal activations are rare, (b) that these solutions are independent of the initial condition, (c) that the steady-state average firing rate and input ratio are the same for every node, and (d) that fluctuations in the zero-time time-correlation between the connection strength and activity are small. Assumption (a) is satisfied when the system is not persistently supercritical and when $\delta t \ll \tau_0$, where δt is the time step and τ_0 is the target time interval between firings. Assumption (b) states that the steady-state solutions are stable to perturbations. Assumption (c) states that the system is homogeneous. Assumption (d) states that it takes finite time for connection strengths to be adjusted up or down. We will ignore the distance-dependent cost factor k_D. This factor is discussed separately in the text. To begin, the dynamical homeostatic equations can be written as follows:

$$\frac{\partial}{\partial t}S(i;t) = -[k_{11}\Delta f(i;t) + k_{12}\Delta\eta(i;t)]S(i;t), \quad \text{(A1)}$$

$$\frac{\partial}{\partial t}\eta(i;t) = -[k_{21}\Delta f(i;t) + k_{22}\Delta\eta(i;t)]\eta(i;t), \quad \text{(A2)}$$

where $\Delta f(i;t) = f(i;t) - 1$ and $\Delta\eta(i;t) = \eta(i;t) - 1$. The latter equation is obtained by summing both sides of Eq. (5) over all j. For later convenience, let us define a matrix **K** as follows:

$$\mathbf{K} = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}. \quad \text{(A3)}$$

Next, define the steady-state values of $\langle f(i;\infty)\rangle = f_0$, $\langle \eta(i;\infty)\rangle = \eta_0$, and $\langle S(i;\infty)\rangle = S_0$, where $\langle \cdots \rangle$ denotes a time average over a time much longer than τ_0. These steady-state values must be the same for each node because the system is homogeneous. For later convenience, define $\delta f(i;t) = f(i;t) - f_0$, $\delta\eta(i;t) = \eta(i;t) - \eta_0$, and $\delta S(i;t) = S(i;t) - S_0$. Recall that

$$A(i;t+\delta t) = 1 - [1 - S(i;t)]\prod_{j=1}^{N}[1 - P(i,j;t)F(j;t)].$$

$$(A4)$$

When simultaneous multinodal activations are rare, as we have assumed in assumption (a), then $A(i;t+\delta t)$ can be expanded to lowest orders in $S(i;t)$ and $F(j;t)$ as follows:

$$A(i;t+\delta t) \approx S(i;t) + \sum_{j=1}^{N} P(i,j;t)F(j;t). \qquad (A5)$$

Note that Eq. (A5) represents a recursion relationship, which generates a trajectory for the vector $\mathbf{A}(t)$. The firing function $F(i;t)$ is essentially a digitized representation of $A(i;t)$. The parameters that control the trajectory for $\mathbf{A}(t)$ are the vector $\mathbf{S}(t)$, the matrix $\mathbf{P}(t)$, and the initial values for $\mathbf{F}(i;0)$. Since we are only interested in stable steady-state solutions that do not depend on initial conditions, we will regard the trajectory of $\mathbf{A}(t)$ as being dependent only on $\mathbf{S}(t)$ and $\mathbf{P}(t)$. The variation of $A(i,t+\delta t)$, as $S(i;t)$ and $P(i,j;t)$ are varied, is then given by

$$\delta A(i;t+\delta t) \approx \delta S(i;t) + \sum_{j=1}^{N} \delta P(i,j;t)F(j;t). \qquad (A6)$$

Note that $\langle F(i;t)\rangle = \langle A(i;t)\rangle$, and that $f(i;t) = \langle F(i;t)\rangle_a \tau_0/\delta t$, where $\langle \cdots \rangle_a$ denotes a time average over a time equal to τ_0. Applying assumption (d) above, we have

$$\langle \delta P(i,j;t)F(j;t)\rangle \approx \langle \delta P(i,j;t)\rangle\langle F(j;t)\rangle. \qquad (A7)$$

Combining Eqs. (A6) and (A7) we then have

$$\langle \delta f(i;t)\rangle \approx (\tau_0/\delta t)\langle \delta S(i;t)\rangle + f_0\langle \delta\eta(i;t)\rangle. \qquad (A8)$$

Taking the time derivative of both sides in Eq. (A8), combining that equation with Eqs. (A1) and (A2), and expanding

to lowest order in $\langle \delta S(i;t)\rangle/S_0$ and $\langle \delta\eta(i;t)\rangle/\eta_0$, we have

$$\frac{\partial}{\partial t}\langle \delta f(i;t)\rangle = B_{11}(1-f_0) + B_{12}(1-\eta_0) - B_{11}\langle \delta f(i;t)\rangle$$

$$- B_{12}\langle \delta\eta(i;t)\rangle, \qquad (A9)$$

$$\frac{\partial}{\partial t}\langle \delta\eta(i;t)\rangle = B_{21}(1-f_0) + B_{22}(1-\eta_0) - B_{21}\langle \delta f(i;t)\rangle$$

$$- B_{22}\langle \delta\eta(i;t)\rangle, \qquad (A10)$$

where the matrix $\mathbf{B}$ is defined as

$$B_{11} = k_{11}S_0(\tau_0/\delta t) + k_{21}f_0\eta_0,$$

$$B_{12} = k_{12}S_0(\tau_0/\delta t) + k_{22}f_0\eta_0,$$

$$B_{21} = k_{21}\eta_0,$$

$$B_{22} = k_{22}\eta_0. \qquad (A11)$$

A little manipulation then shows that

$$\begin{bmatrix} \langle \delta f(i;t)\rangle \\ \langle \delta\eta(i;t)\rangle \end{bmatrix} = \begin{bmatrix} 1-f_0 \\ 1-\eta_0 \end{bmatrix} + \exp(-\mathbf{B}t)\begin{bmatrix} \delta f(i;0) + f_0 - 1 \\ \delta\eta(i;0) + \eta_0 - 1 \end{bmatrix}.$$

$$(A12)$$

The steady state corresponds to $\langle \delta f(i;t)\rangle \approx 0$ and $\langle \delta\eta(i;t)\rangle \approx 0$. Equation (A12) thus shows that a stable steady state exists only if both eigenvalues of $\mathbf{B}$ are positive and if $f_0 = \eta_0 = 1$. It is a simple matter to show that the eigenvalues of $\mathbf{B}$ are $\lambda = b_0 \pm \sqrt{b_0^2 - d}$, where $b_0 = (B_{11}+B_{22})/2$ and $d = \det(\mathbf{B}) = S_0(\tau_0/\delta t)\eta_0 \det(\mathbf{K})$. Therefore, under assumptions (a)–(d) above, the existence of steady-state solutions requires $f_0 = \eta_0 = 1$, $S_0 > 0$, and $k_{11}k_{22} > k_{12}k_{21}$.

[1] E. L. Bienenstock, L. N. Cooper, and P. W. Munro, J. Neurosci. **2**, 32 (1982).
[2] K. D. Miller, Neuron **17**, 371 (1996).
[3] K. D. Miller and D. J. C. MacKay, Neural Comput. **6**, 100 (1994).
[4] G. G. Turrigiano and S. B. Nelson, Curr. Opin. Neurobiol. **10**, 358 (2000).
[5] G. G. Turrigiano and S. B. Nelson, Nat. Rev. Neurosci. **5**, 97 (2004).
[6] C. von der Malsburg, Kybernetik **14**, 85 (1973).
[7] L. F. Abbott and S. B. Nelson, Nat. Neurosci. **3**(Suppl.), 1178 (2000).
[8] G. LeMasson, E. Marder, and L. F. Abbott, Science **259**, 1915 (1993).
[9] D. J. Schulz, J. M. Goaillard, and E. Marder, Nat. Neurosci. **9**, 356 (2006).
[10] A. A. Prinz, D. Bucher, and E. Marder, Nat. Neurosci. **7**, 1345 (2004).

[11] G. G. Turrigiano *et al.*, Nature (London) **391**, 892 (1998).
[12] C. J. Wierenga, K. Ibata, and G. G. Turrigiano, J. Neurosci. **25**, 2895 (2005).
[13] A. Destexhe and E. Marder, Nature (London) **431**, 789 (2004).
[14] E. Marder and A. A. Prinz, BioEssays **24**, 1145 (2002).
[15] D. Plenz and T. C. Thiagarajan, Trends Neurosci. **30**, 101 (2007).
[16] J. M. Beggs, Philos. Trans. R. Soc. London, Ser. A (to be published).
[17] J. M. Beggs and D. Plenz, J. Neurosci. **23**, 11167 (2003).
[18] J. M. Beggs and D. Plenz, J. Neurosci. **24**, 5216 (2004).
[19] C. Haldeman and J. M. Beggs, Phys. Rev. Lett. **94**, 058101 (2005).
[20] R. Albert and A. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[21] V. M. Eguiluz *et al.*, Phys. Rev. Lett. **94** 018102 (2005).
[22] S. H. Strogatz, Nature (London) **410**, 268 (2001).
[23] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[24] W. J. Freeman, *Mass Action in the Nervous system* (Academic Press, New York, 1975).

[25] L. Ingber, Phys. Rev. A **44**, 4017 (1991).

[26] E. De Schutter *et al.*, Trends Neurosci. **28**, 562 (2005).

[27] W. J. Freeman, *Neurodynamics: An exploration in Mesoscopic Brain Dynamics* (Springer Verlag, London, 2000).

[28] P. D. Roberts, Phys. Rev. E **69**, 031910 (2004).

[29] F. H. Lopes da Silva, *et al.*, Kybernetik **15**, 27 (1974).

[30] P. Suffczynski, S. Kalitzin, and F. H. Lopes Da Silva, Neuroscience **126**, 467 (2004).

[31] P. Suffczynski *et al.*, Int. J. Psychophysiol **43**, 25 (2001).

[32] P. Suffczynski *et al.*, J. Clin. Neurophysiol. **22**, 288 (2005).

[33] P. Suffczynski *et al.*, IEEE Trans. Biomed. Eng. **53**, 524 (2006).

[34] S. Kalitzin, B. W. van Dijk, and H. Spekreijse, Biol. Cybern. **83**, 139 (2000).

[35] G. G. Turrigiano and S. B. Nelson, Neuron **21**, 933 (1998).

[36] G. Tononi and C. Cirelli, Sleep Med. Rev. **10**, 49 (2006).

[37] B. A. Pearlmutter, C. J. Houghton, and S. A. Tilbury, http://www.maths.tcd.ie/report_series/tcdmath/tcdm0603.pdf

[38] A. R. Houweling *et al.*, Cereb. Cortex **15**, 834 (2005).

[39] K. M. Jacobs and D. A. Prince, J. Neurophysiol. **93**, 687 (2005).

[40] H. Li and D. A. Prince, J. Neurophysiol. **88**, 2 (2002).

[41] K. M. Jacobs *et al.*, Epilepsia **41**(Suppl. 6), S153 (2000).

[42] P. Salin *et al.*, J. Neurosci. **15**, 8234 (1995).

[43] K. D. Graber and D. A. Prince, Ann. Neurol. **55**, 860 (2004).

[44] K. D. Graber and D. A. Prince, Ann. Neurol. **46**, 234 (1999).

[45] P. E. Latham *et al.*, J. Neurophysiol. **83**, 808 (2000).

[46] D. Hsu and J. M. Beggs, Neurocomputing **69**, 1134 (2006).

[47] L. F. Abbott and R. Rohrkemper, Prog. Brain Res. (to be published).

[48] T. J. Sullivan and V. R. de Sa, Neural Networks **19**, 734 (2006).

[49] L. C. Yeung *et al.*, Proc. Natl. Acad. Sci. U.S.A. **101**, 14943 (2004).

[50] M. C. van Rossum, G. Q. Bi, and G. G. Turrigiano, J. Neurosci. **20**, 8812 (2000).

[51] S. Bornholdt and T. Rohl, Phys. Rev. E **67**, 066118 (2003).

[52] D. Johnston *et al.*, Philos. Trans. R. Soc. London, Ser. B **358**, 667 (2003).

[53] G. Stuart *et al.*, Trends Neurosci. **20**, 125 (1997).

[54] J. Waters, A. Schaefer, and B. Sakmann, Prog. Biophys. Mol. Biol. **87**, 145 (2005).

[55] S. Royer and D. Pare, Nature (London) **422**, 518 (2003).

[56] A. Artola, S. Brocher, and W. Singer, Nature (London) **347**, 69 (1990).

[57] A. Artola and W. Singer, Trends Neurosci. **16**, 480 (1993).

[58] M. Ito, Prog. Neurobiol. **78**, 272 (2006).

[59] M. Ito, M. Sakurai, and P. Tongroach, J. Physiol. (London) **324**, 113 (1982).

[60] R. Llinas, E. J. Lang, and J. P. Welsh, Learn. Memory **3**, 445 (1997).

[61] D. Hsu, Prog. Brain. Res **163**, 601 (2007).

[62] R. Douglas and D. Martin, *The Synaptic Organization of the Brain*, edited by G. M. Shepherd, (Oxford University Press, New York, 1998), pp. 459–509.

[63] T. F. Freund and G. Buzsaki, Hippocampus **6**, 347 (1996).

[64] S. M. Thompson and B. H. Gahwiler, J. Neurophysiol. **61**, 512 (1989).

[65] S. M. Thompson and B. H. Gahwiler, J. Neurophysiol. **61**, 501 (1989).

[66] S. M. Thompson and B. H. Gahwiler, J. Neurophysiol. **61**, 524 (1989).

[67] D. D. Mott *et al.*, J. Neurophysiol. **69**, 674 (1993).

[68] R. A. Deisz and D. A. Prince, J. Physiol. (London) **412**, 513 (1989).

[69] M. Farrant and K. Kaila, Prog. Brain Res. **160**, 59 (2007).

[70] J. M. Beggs (unpublished).

[71] D. Hsu and J. M. Beggs (unpublished).