# Robust entropy requires strong and balanced excitatory and inhibitory synapses

Vidit Agrawal, Andrew B. Cowley, Qusay Alfaori, Daniel B. Larremore, Juan G. Restrepo, and Woodrow L. Shew

Citation: Chaos **28**, 103115 (2018); doi: 10.1063/1.5043429
View online: https://doi.org/10.1063/1.5043429
View Table of Contents: http://aip.scitation.org/toc/cha/28/10
Published by the American Institute of Physics

# Robust entropy requires strong and balanced excitatory and inhibitory synapses

Vidit Agrawal,[1,a)] Andrew B. Cowley,[2,a)] Qusay Alfaori,[1] Daniel B. Larremore,[3,4] Juan G. Restrepo,[2] and Woodrow L. Shew[1,b)]

[1]*Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701, USA*
[2]*Department of Applied Mathematics, University of Colorado, Boulder, Colorado 80309, USA*
[3]*Department of Computer Science, University of Colorado, Boulder, Colorado 80309, USA*
[4]*BioFrontiers Institute, University of Colorado, Boulder, Colorado 80303, USA*

(Received 8 June 2018; accepted 1 October 2018; published online 18 October 2018)

It is widely appreciated that balanced excitation and inhibition are necessary for proper function in neural networks. However, in principle, balance could be achieved by many possible configurations of excitatory and inhibitory synaptic strengths and relative numbers of excitatory and inhibitory neurons. For instance, a given level of excitation could be balanced by either numerous inhibitory neurons with weak synapses or a few inhibitory neurons with strong synapses. Among the continuum of different but balanced configurations, why should any particular configuration be favored? Here, we address this question in the context of the entropy of network dynamics by studying an analytically tractable network of binary neurons. We find that entropy is highest at the boundary between excitation-dominant and inhibition-dominant regimes. Entropy also varies along this boundary with a trade-off between high and robust entropy: weak synapse strengths yield high network entropy which is fragile to parameter variations, while strong synapse strengths yield a lower, but more robust, network entropy. In the case where inhibitory and excitatory synapses are constrained to have similar strength, we find that a small, but non-zero fraction of inhibitory neurons, like that seen in mammalian cortex, results in robust and relatively high entropy. *Published by AIP Publishing.* https://doi.org/10.1063/1.5043429

**In many social and biological networks, there exists a competition between two opposing types of nodes. Excitatory nodes tend to activate other nodes, while inhibitory nodes tend to suppress the activation of other nodes. The collective behavior of such networks is very sensitive to whether excitatory or inhibitory nodes are dominant. One important property of collective network dynamics is entropy, which quantifies fluctuations and diversity of activity states. Here, we examine many different configurations of balanced and imbalanced networks, focusing on neural networks. We find that imbalanced networks have low entropy compared to balanced networks. For balanced networks, there is a trade-off. When weak inhibition balances weak excitation, the dynamics have very high entropy, but this condition is rather fragile to small perturbations of system parameters. In contrast, when strong inhibition balances strong excitation, entropy is lower, but more robust to perturbations.**

## I. INTRODUCTION

The network of neurons in cerebral cortex displays rich and complex dynamics even when not engaged by any particular sensory or motor interaction with the external world.[1,2] From one point of view, such ongoing internal dynamics are thought to mediate memory consolidation and other internal cognitive processes.[3–7] On the other hand, ongoing fluctuations in cortical network dynamics have often been considered a nuisance, imposing noisy fluctuations in neural response to sensory input.[8–10] In both of these contexts, it is important to understand the mechanisms that govern the fluctuations of ongoing cortical network dynamics. Here, we investigate the Shannon entropy of the network spike rate. In the context of internal cognitive processes, high entropy might be beneficial, corresponding to a larger repertoire of internal states to mediate internal information transfer.[11] When considered as noise, high entropy can be a hindrance to effective sensory coding.[8–10] Indeed, in principle, encoding of sensory input would be most reliable if the cortex was totally silent (low entropy) until the stimulus excited it. However, real cortex does not operate this way; it has many jobs to do beyond encoding sensory input and is never silent. Previous studies have shown that ongoing cortical dynamics with high entropy occurs together with high mutual information between stimulus and response,[11,12] suggesting that a large repertoire of ongoing dynamical states may be necessary for a large repertoire of stimulus-evoked states.[3,5]

A crucial factor for determining the entropy of network dynamics in the cortex is the competition between two types of neurons: excitatory (E) and inhibitory (I). This is most apparent in previous experiments that pharmacologically manipulated the E/I balance.[11–14] Enhanced inhibition (GABA agonists) often results in a dynamical regime characterized by low firing rates and weak population-level correlations, while decreased inhibition (GABA antagonists)

[a)] V. Agrawal and A. B. Cowley contributed equally to this work.
[b)] Electronic mail: shew@uark.edu

tends to result in a regime with higher firing rates and strong correlations. Two studies in particular have shown that entropy can be increased by tuning the E/I balance to the tipping point between these two distinct dynamical regimes.[11,12] However, a more systematic understanding of how E/I balance impacts entropy is difficult to obtain experimentally because pharmacological manipulations are rather difficult to precisely control. Moreover, with a few interesting exceptions,[15,16] experiments do not vary the numbers of excitatory or inhibitory neurons. Computational models offer an alternative approach in which the number of excitatory and inhibitory neurons, as well as strength of excitatory and inhibitory synapses, can easily be controlled. Previous computational studies have addressed similar topics but typically have neglected inhibition[12,17] or have not considered the effects of changing the E/I ratio.[18,19] Thus, theoretical and experimental understanding of the relationship between the entropy of ongoing dynamics and the balance of excitation and inhibition—mediated by both relative strengths of excitatory and inhibitory synapses and relative numbers of excitatory and inhibitory cells—remains unresolved.

Here, we attempt to improve the theoretical understanding of entropy of ongoing dynamics by studying a network model of binary neurons in detail. We consider how entropy of the population firing rate depends on the fraction of inhibitory neurons $\alpha$ and the strengths of E and I interactions, $W_E$ and $W_I$, respectively. We find maximal entropy near the tipping point between the low and high firing rate dynamical regimes, as seen in experiments.[12] We also find that, for a given choice of $W_E$ and $W_I$, the tipping point can be achieved by adjusting the value of $\alpha$. This raises the question: among the different possible parameter configurations that place the system at the tipping point, why should one be favored over another? We find that there is a trade-off between high and robust network entropy: networks with weak synapses can achieve a high entropy when excitation and inhibition are balanced, but the entropy degrades significantly upon small deviations from the balanced state. On the other hand, networks with stronger synapses have a lower maximum entropy, but they are more robust to parameter changes. We also find that if E and I synaptic strengths are proportional to each other, as found in many experiments,[20–22] then robust, high entropy requires a small fraction of I neurons ($\alpha$ near 0.1). In mammalian cortex, $\alpha$ has been found to be near 0.2 with remarkable consistency over the lifetime of an organism[23] and over different regions of cortex.[24,25] Our results suggest that mammalian cortex strikes a compromise with intermediate but robust entropy.

In what follows, we introduce and analyze the binary neuron model which both predicts and provides insight into the results of model numerical simulations.

## II. MODELS AND THEORY

### A. Binary neuron model

We explore the effects of excitation and inhibition balance on entropy using a simple, analytically tractable model. The model, studied previously in Ref. 26, consists of a network of $N$ stochastic binary neurons, indexed $i = 1, 2, \ldots, N$. The state of neuron $i$ at time $t$ is denoted by $x_i^t$, which can take the values $x_i^t = 0$ if the neuron is resting and $x_i^t = 1$ if the neuron is spiking. Time is assumed to evolve in discrete steps $t = 0, 1, 2, \ldots$. The evolution of each neuron's state is stochastic and depends on the states of other neurons at the previous time step,

$$x_i^{t+1} = \begin{cases} 1 & \text{with probability } \eta + (1-\eta)\sigma\left(\sum_{j=1}^{N} \epsilon_j w_{ij} x_j^t\right), \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $\epsilon_j = 1$ if neuron $j$ is excitatory and $\epsilon_j = -1$ if neuron $j$ is inhibitory. The strength of the synapse from neuron $j$ to neuron $i$ is $w_{ij} > 0$ and $w_{ij} = 0$ if neuron $j$ does not connect to neuron $i$. The transfer function $\sigma(x) = \min[1, \max(0, x)]$ converts the input to neuron $i$ into a probability. The constant $\eta = 1/(100N)$ represents independent spontaneous activation due to noise or external sources, resulting in one spike per 100 time steps among all neurons, on average. We note that other choices of $\eta$ could cause quantitative changes in our results below, but we expect that our qualitative conclusions are not sensitive to moderate changes in $\eta$. For example, it is well known that noise tends to smooth out the sharpness of phase transitions like the one discussed below.[27]

We consider Erdős-Rényi networks where a directed link is made independently from neuron $j$ to neuron $i$ with probability $k/(N-1)$ for all $i \neq j$. The parameter $k$ is the expected number of outgoing connections from each neuron. To control the relative number of excitatory and inhibitory neurons, we assign each neuron to be inhibitory with probability $\alpha$ and excitatory otherwise. Finally, we assume for simplicity that $w_{ij} = w_E$ for excitatory synapses (i.e., if $\epsilon_j = 1$) and $w_{ij} = w_I$ for inhibitory synapses (i.e., if $\epsilon_j = -1$) and define the effective excitatory weight as $W_E = kw_E$ and the effective inhibitory weight as $W_I = kw_I$. We interpret our model to represent a small patch of cortex, 100 $\mu$m in scale, like a single cortical column. At these scales, it is a reasonable approximation to neglect distance-dependent differences in connectivity for excitatory and inhibitory neurons.[28]

The model is characterized by the parameters $N$, $k$, $W_E$, $W_I$, and $\alpha$. For definiteness, in all simulations, we will consider, unless otherwise indicated, only the parameters $N = 10\,000$ and $k = 100$ and study the population firing dynamics of the model as a function of $(W_E, W_I, \alpha)$. As a measure of collective network dynamics, we study the fraction of spiking neurons, or *network activity*, given by

$$S^t = \frac{1}{N}\sum_{i=1}^{N} x_i^t. \tag{2}$$

In Ref. 26, it was found that the collective dynamics of the network is determined by the largest eigenvalue $\lambda$ of the connection strength matrix $A$ with entries $\{\epsilon_j w_{ij}\}_{i,j=1}^{N}$. Network activity saturates at a high value for $\lambda > 1$ and dies out or reaches a steady low value for $\lambda < 1$. At the tipping point between these two regimes, defined by $\lambda = 1$, excitation and inhibition are balanced such that network activity is characterized by large fluctuations that are effectively ceaseless (their lifetime scales exponentially with $N$).[26] Figure 1(a) shows an example of the time series of network activity for these three

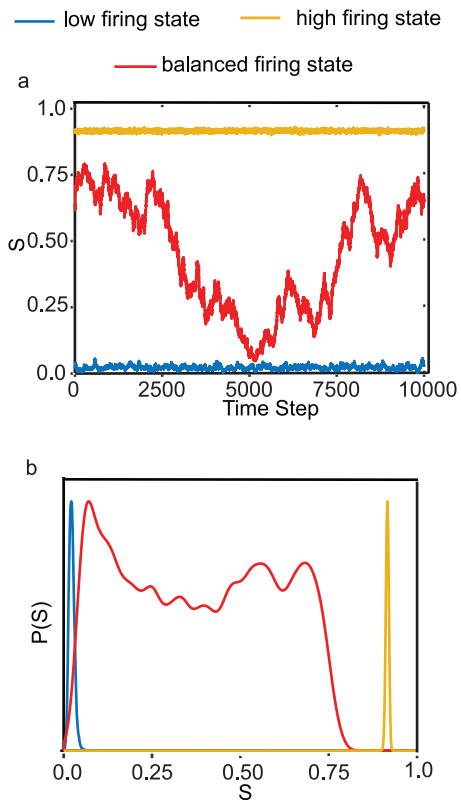


FIG. 1. Network activity and dynamics of binary model. Time series of network activity (a) show diverse fluctuations when excitation and inhibition are balanced ($\lambda = 1$). Similarly, probability distributions (b) of network activity are broadest when $\lambda = 1$. All probability distributions have been normalized by their peak probability to facilitate comparison of their shapes. Dynamical parameters: $\alpha = 0.11$ (Blue), 0.1 (Red), 0.09 (Yellow); $W_E$=$W_I$=1.25.

regimes. For the Erdős-Rényi networks considered here, $\lambda$ can be approximated by the expected row sum of $A$,

$$\lambda \approx kw_E(1-\alpha) - kw_I\alpha = W_E(1-\alpha) - W_I\alpha. \quad (3)$$

With this approximation, then, the parameters that give $\lambda = 1$ form a 2-dimensional surface in the $(W_E, W_I, \alpha)$ parameter space.

### B. Entropy

We consider the Shannon entropy of the time-series of network activity, which quantifies the size of the repertoire of accessible population firing rates. The network activity is discrete (i.e., $0, 1/N, 2/N \ldots, 1$). For a given set of network parameters $(W_E, W_I, \alpha)$, we consider the steady-state probability distribution of network activity $P(S)$ and the associated entropy,

$$H = -\sum_S P(S)\log_2[P(S)], \quad (4)$$

where the sum runs over the allowed values $S = 0, 1/N, 2/N, \ldots, 1$. In practice, we estimate $P(S)$ numerically from a time series of $S^t$ obtained from model simulations [Fig. 1(b)] or from our semi-analytical theory, presented below, that treats the evolution of $S^t$ as a biased random walk.

### C. Simulation-free theory

Here, we present a semi-analytical approach to compute the entropy for a given set of parameter values in the binary model. The main idea of our approach is to treat the evolution of the macroscopic variable $S^t$ as a biased random walk. Although in principle the dynamics of the system depends on the microscopic states $\{x_n\}_{n=1}^N$, for large homogeneous networks, one can describe the evolution of the system in terms of the macroscopic variable $S^t$. To analyze this random variable, one should determine if at any given time it is expected to decrease or increase. This information is encapsulated in the *branching function* introduced in Ref. 26 as the ratio $\Lambda(S) = E[S^{t+1}|S^t = S]/S$, where the expected value is taken over realizations of the stochastic dynamics and microscopic configurations with activity $S$. In our case, the branching function can be approximated by

$$\Lambda(S) = \frac{1}{S}E_P[\sigma(w_E n_E - w_I n_I)], \quad (5)$$

where the random variables $n_E$ and $n_I$ represent the number of active E and I inputs to a single neuron, respectively.[26] Because we consider random networks, $n_E$ and $n_I$ are given by binomial random variables with $N$ trials and probability of success $k(1-\alpha)S/N$ and $k\alpha S/N$, respectively, corresponding to the probability of finding a link $(k/N)$ that is excitatory $(1-\alpha)$ or inhibitory $(\alpha)$ and that is active $(S)$. $E_P[\cdot]$ is an expected value over the random variables $n_E$ and $n_I$. We have neglected the $\eta$ term in Eq. (1), which is only included to prevent activity from dying out. By assuming that the statistics of the macroscopic dynamics depend only on $S$, one can then write a random walk model for $S^t$ as

$$S^{t+1} = S^t\Lambda(S^t) + r(S^t), \quad (6)$$

where $r$ represents statistical noise which, by the definition of $\Lambda$, has mean zero. To obtain a tractable model, we assume that $r(S^t)$ is normally distributed and has variance $V(S^t) = S^t(1-S^t)/N$, as estimated in Ref. 26. This approximation is what one would obtain if each of the $N$ neurons is independently assumed to be active with probability $S$ and inactive with probability $1-S$. In this approximation, the probability that the system makes a transition from a state with activity $S'$ to a state with activity $S$ is given by

$$T(S \mid S') = \frac{1}{\sqrt{2\pi V(S')}}\exp\left(-\frac{[S'\Lambda(S') - S]^2}{2V(S')}\right). \quad (7)$$

The distribution $P^t(S)$ of $S$ at time $t$ evolves following the master equation

$$P^{t+1}(S) = \int_0^1 T(S \mid S')P^t(S')dS', \quad (8)$$

and as $t \to \infty$, it converges to a steady-state, which may be calculated numerically as the Perron-Frobenius eigenvector (with eigenvalue 1) of the linear operator

$$\mathcal{L}\{P\}(S) = \int_0^1 T(S \mid S')P(S')dS'. \quad (9)$$

The eigenvector can be calculated numerically by discretization of the integral in Eq. (9) or as the limit of repeated

iterations of Eq. (8). The entropy is then calculated directly from Eq. (4).

## III. RESULTS

Our primary goal is to determine how the entropy of a network varies with the relative numbers of E and I neurons and the relative strength of E and I synapses. We first describe our results from numerical simulations of the binary model and then describe results from the theory.

First, we show in Fig. 1 that the system network activity visits the widest variety of states when excitation and inhibition are balanced at the tipping point between high and low firing rate regimes. This is visible in time series [Fig. 1(a)] as well as empirical distributions $P(S)$ of network activity (based on $10^4$ time steps of simulation). Correspondingly, entropy $H$ is greatest along the boundary between low and high firing regimes (Fig. 2). In the three-dimensional $(W_E, W_I, \alpha)$ parameter space, this boundary forms a curved surface, which we henceforth refer to as the *maximum entropy surface*.

As discussed in Sec. II A, we expect that the transition from the low to the high firing regimes occurs at the *critical surface* of parameters where $\lambda = 1$. While we find this is usually an excellent approximation to our numerical results, the maximum entropy and critical surfaces differ slightly for high values of $\alpha$, and therefore, we will only use the critical surface as a qualitative guide to the location of the maximum entropy surface.

To numerically identify the maximum entropy surface, for each fixed value of $(W_E, W_I)$, we compute entropy across a wide range of values of $\alpha$, finding the value $\alpha^*$ that maximizes $H(W_E, W_I, \alpha)$. In Fig. 3(a), we show $\alpha^*$ as a function of $W_E$ and $W_I$. As one might expect, higher values of $W_E$ require a larger number of I neurons (higher $\alpha^*$) in order to maintain a balanced network and vice versa. This agrees qualitatively with the estimate using the critical surface, $\alpha^* \approx (W_E - 1)/(W_E + W_I)$ obtained from Eq. (3) with $\lambda = 1$.

Having identified the parameters that characterize the maximum entropy surface, we next ask two questions. First, where on the surface is entropy highest? Second, where on the surface is entropy most robust? We consider the entropy to be robust if it does not drop dramatically when we make a small perturbation in $W_E$, $W_I$, and $\alpha$ away from the peak entropy surface. This approach is similar to other ways to quantify sensitivity to model parameters, such as Fisher information.[29] To quantify how much the entropy decreases if parameters are perturbed away from the maximum entropy surface, we define *fragility* $F(W_E, W_I)$ as follows. For a given pair of $(W_E, W_I)$ values, we first calculate the entropy at the corresponding point on the maximum entropy surface, $H^* = H(W_E, W_I, \alpha^*)$. Then, we calculate the entropy at two points at a small distance $\delta$ above and below the surface, $H_{\text{up}} = H(W_E + \Delta W_E, W_I + \Delta W_I, \alpha + \Delta\alpha)$ and $H_{\text{down}} = H(W_E - \Delta W_E, W_I - \Delta W_I, \alpha - \Delta\alpha)$. The perturbations $\pm(\Delta W_E, \Delta W_I, \Delta\alpha)$ are defined to be normal to the maximum entropy surface, which will give the largest drop in entropy for a given perturbation size. The size of the perturbation was chosen to be small (Euclidean norm $\delta = 0.01$, about 1% variation in parameters) to emphasize that entropy can be quite sensitive to these parameter changes in certain parts of $(W_E, W_I)$ space. Finally, we define *fragility* $F(W_E, W_I)$ as the mean of the entropy difference,

$$F(W_E, W_I) = \frac{(H^* - H_{\text{up}}) + (H^* - H_{\text{down}})}{2}. \tag{10}$$

Our main results are in Figs. 3(b) and 3(c). Figure 3(b) shows the entropy $H^*$ on the maximum entropy surface as a function of the effective E and I weights $W_E$ and $W_I$. Networks with weak effective synapse strengths (low values of $W_E$ and $W_I$) can achieve a higher entropy $H^*$ than networks with strong effective synapse strengths. However, as shown in Fig. 3(c), high entropy comes at the cost of high fragility: networks with weak effective synapse strengths have the highest fragility,

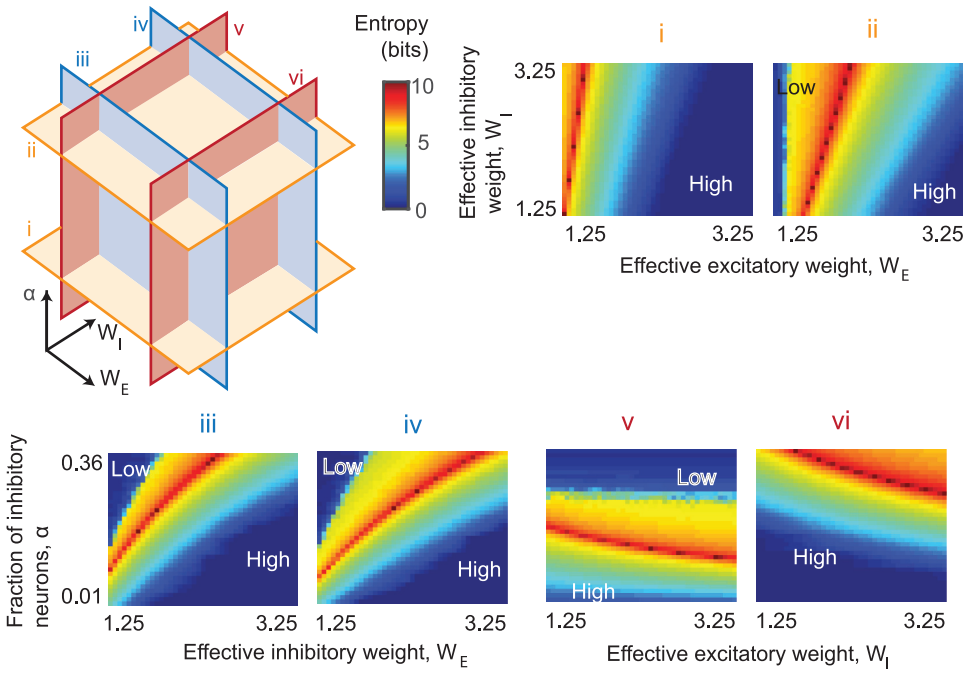


FIG. 2. High entropy at the boundary between high and low firing regimes. Each panel shows how entropy (color) varies across a two-dimensional section of the three-dimensional $W_E$-$W_I$-$\alpha$ parameter space. The relative orientation of the six different sections is illustrated and labeled [(i)–(vi)] in the cartoon (left). For (i) and (ii), $\alpha$ is fixed at 0.1 and 0.2. For (iii) and (iv), $W_I$ is fixed at 1.5 and 2.5. For (v) and (vi), $W_E$ is fixed at 1.5 and 2.5. A curved critical surface in $W_E$-$W_I$-$\alpha$ space separates the high firing regime (H) from a low firing regime (L). Entropy is high along this regime boundary. Note that as I or E synapse strength increases, the width of the peak in entropy also increases, indicating increased robustness (decreased fragility).

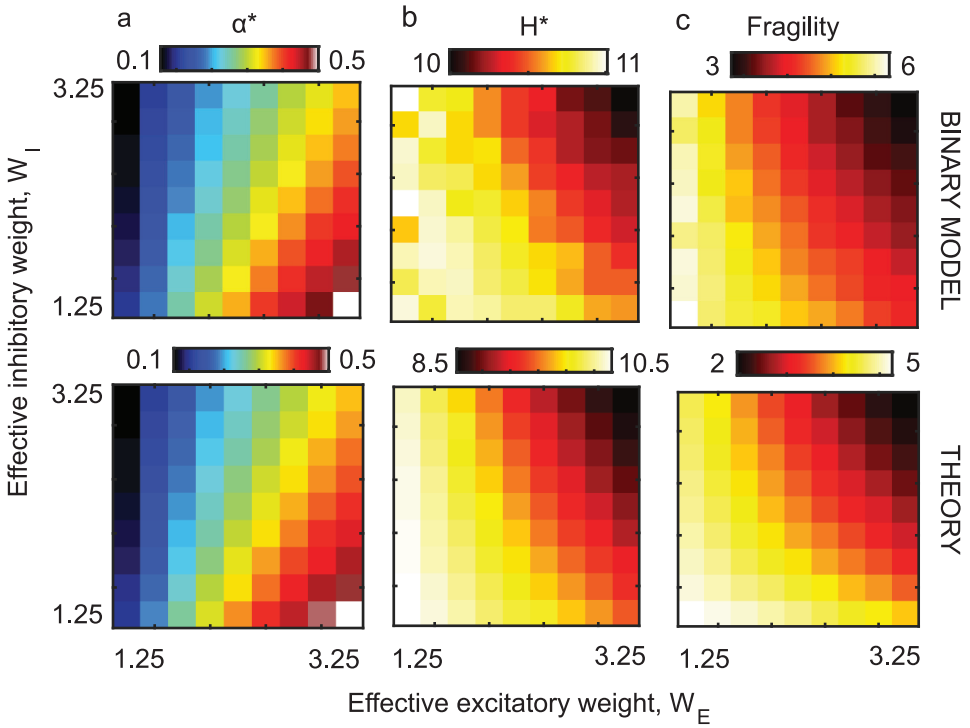


FIG. 3. Trade-off between high entropy and robust entropy. (a) For each combination of $W_E$ and $W_I$ effective synaptic weights, we identify the critical fraction of inhibitory neurons ($\alpha^*$) with the highest entropy. (b) Comparing all critical entropy $H^*$ across the entire critical surface, entropy was highest for low $W_E$ and $W_I$. (c) Highest fragility was also found for low $W_E$ and $W_I$.

while networks with strong effective synapse strengths are the most robust. We note that while the variation in entropy $H^*$ is relatively moderate across the range studied (approximately 10%), the fragility ranges from 3 to 6, indicating that our 1% perturbation of parameters results in a dramatic drop in entropy of approximately $30\% - 60\%$. One could argue that what matters are the final values of entropy after perturbation (i.e., $H_{\rm up}$ and $H_{\rm down}$) rather than how much entropy drops due to perturbation (i.e., $F$). From this perspective, strong synapses are also better; $H_{\rm up}$ and $H_{\rm down}$ are lower for weak synapses than for strong synapses. This can be seen by subtracting Figs. 3(c) from 3(b). We conclude that there is a trade-off between high and robust entropy, with stronger effective synapse strengths promoting lower but more robust entropy, and weaker effective synapse strengths promoting a high but fragile entropy.

Finally, we address the role of the fraction $\alpha$ of I neurons in promoting entropy robustness. We note that if the choices of E and I synapse strengths are constrained to be proportional to each other, as experiments suggest,[20–22] then $W = W_E = bW_I$ and the estimate $\alpha^* \approx (W_E - 1)/(W_E + W_I)$ becomes $\alpha^* = (1 + 1/b)^{-1}(1 - 1/W)$. Thus, $\alpha^*$ is a monotonically increasing function of synapse strength $W$. Therefore, for such constrained networks, entropy and fragility decrease with the fraction of I neurons $\alpha$. Thus, a small non-zero $\alpha$, similar to that found in mammalian cortex, is needed to obtain high and robust entropy.

In the following, we present an interpretation of our results based on the branching function formalism presented above and studied in previous work.[26] If one treats the time series of network activity $S^t$ as a random walk, its bias, or expected velocity, is given by $S^t[\Lambda(S^t) - 1]$. Therefore, when $\Lambda(S^t) > 1$ [$\Lambda(S^t) < 1$], $S^t$ tends to increase (decrease). Since $\Lambda(0) \geq 1$ and $\Lambda(1) \leq 1$ [Ref. 26], the long-time distribution of $S^t$ will be concentrated around the region where $\Lambda(S^t) \approx 1$. The wider this region is, the wider the distribution of $S$ will be and the larger its associated entropy. To understand how the size of this region depends on the weights $W_E$ and $W_I$, we note that, at the tipping point between low and high firing regimes, the branching function deviates from 1 in an interval $[0, S_0)$ on which it is appreciably larger than 1 and in an interval $(S_1, 1]$ on which it is less than 1 [see Fig. 4(a)]. The branching function deviates from 1 in these intervals because the distribution of the random variable $w_E n_E - w_I n_I$ in Eq. (5) extends below 0 or above 1 when $S$ is too close to 0 or 1, respectively. In these cases, the nonlinearity in the transfer function $\sigma$ causes the expected value in Eq. (5) to be different from 1. The larger the values of $w_E$ and $w_I$, the wider the distribution of $w_E n_E - w_I n_I$ and therefore the larger these intervals are. More precisely, we can estimate the scaling of $S_0$ and $S_1$ as follows. The variance of the variable $w_E n_E - w_I n_I$ is $V(S) = w_E^2(1-\alpha)kS[1 - \frac{k}{N}(1-\alpha)S] + w_I^2\alpha kS[1 - \frac{k}{N}\alpha S]$. For sparse networks with $k/N \ll 1$, we have $V(S) = w_E^2(1-\alpha)kS + w_I^2\alpha kS$. Estimating $S_0$ and $S_1$ as the values where $S_0^2 \sim V(S_0)$ and $(1 - S_1)^2 \sim V(S_1)$, we obtain $S_0 \sim w_E^2(1-\alpha)k + w_I^2\alpha k$ and $S_1 \sim 1 - \sqrt{\left(\frac{1}{2}S_0 + 1\right)^2 - 1} + \frac{1}{2}S_0$. Using the approximation that in the balanced state $\alpha = (kw_E - 1)/(kw_E + kw_I)$, this gives closed expressions for the estimates of $S_0$ and $S_1$ as a function of $w_E$ and $w_I$. For low values of $w_E$ and $w_I$, $S_0 \ll 1$ and $1 - S_1 \ll 1$, and therefore, the branching function will be close to 1 over a large region in $[0, 1]$. This is illustrated in the left panel of Fig. 4, which shows the branching function $\Lambda(S)$ and associated probability distribution $P(S)$ for the balanced state (red lines), high-firing (yellow lines), and low-firing (blue lines) cases. While the wide region over which the branching function is approximately one results in a relatively large entropy, a perturbation away from the balanced state displaces the branching function so that it is below or

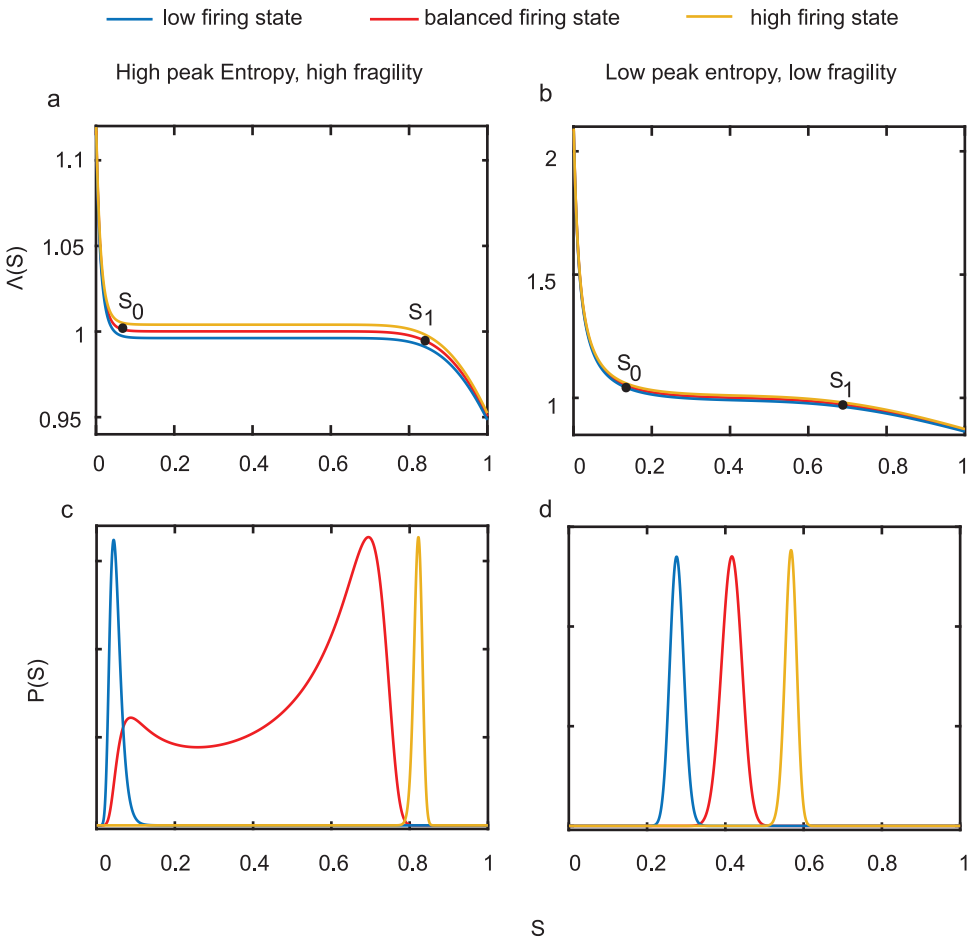


FIG. 4. Interpretation of results based on Branching function formalism. Branching functions $\Lambda(S)$, for (a) low effective excitatory and inhibitory weight ($W_E = W_I = 1.25$) with $S_0 \approx 0.015$ and $S_1 \approx 0.883$ and (b) high effective excitatory and inhibitory weight ($W_E = W_I = 3.25$) with $S_0 \approx 0.105$ and $S_1 \approx 0.724$. The probability distributions (c) low effective weights and (d) high effective weights. All probability distributions have been normalized by their peak probability to facilitate comparison of their shapes.

above 1 over this large region and is close to 1 over a much smaller region. Thus, the entropy decreases substantially. On the other hand, if the weights $w_E$ and $w_I$ are larger, both $S_0$ and $1 - S_1$ will be closer to $1/2$. This is illustrated in the right panel of Fig. 4. While the region over which $\Lambda(S)$ is close to 1 is smaller, resulting in a smaller entropy, it does not change substantially in the low-firing or high-firing cases, resulting in lower fragility.

## IV. DISCUSSION

Here, we have shown that Shannon entropy of neural network dynamics is sensitive to the structure of excitatory and inhibitory interactions. Generally, high entropy is obtained by balancing E and I synaptic efficacy such that the system operates near the tipping point between two phases of network dynamics. Entropy is high all along this boundary, i.e., for a wide range of properly balanced E/I combinations. However, the regions within this boundary with the highest entropy are not robust; small variations in the synaptic strengths $W_E$, $W_I$ and in the fraction of inhibitory neurons $\alpha$ could cause entropy to plummet, drastically reducing the accessible states and disrupting the functioning of the network. We found that entropy is more robust when the effective synaptic strengths are larger. Given that $W_E$, $W_I$, and $\alpha$ are inevitably somewhat variable during development, across brain regions, and across individuals,[23–25] robustness to $W_E$, $W_I$, and $\alpha$ variability may be important. For networks constrained such that $W_E \sim W_I$,[20–22] our findings imply that a small, nonzero fraction $\alpha > 0$ of inhibitory neurons would result in a more robust network entropy. Our results suggest that a population of organisms with reliable and high entropy brains requires that small, nonzero fraction of neurons be inhibitory, which is consistent with what exists in mammalian cortex.[23–25]

Different parts of the space of models we explore here relate to several other models studied previously. The parts of parameter space with relatively weak $W_E$ and $W_I$ and with $\alpha = 0.2$ are similar to models previously studied in the context of "criticality" in the cortex.[26] The parts of parameter space where $W_E$ and $W_I$ are stronger may be related to the widely studied set of models referred to as "chaotic balanced" networks.[30–32] A more detailed comparison of our model dynamics to previous models could bridge the study of the criticality hypothesis with that of chaotic balanced networks.

How might one experimentally test the results of our work? One way would be to measure changes in firing rate fluctuations in response to acute manipulation of excitatory or inhibitory synapses. Such manipulations can be made pharmacologically, for example.[11–14] Our work predicts two testable phenomena. First, if the cortex is on the high entropy surface discussed here, then any manipulation of excitation or inhibition will result in a drop in firing rate fluctuations. Conversely, if either excitatory or inhibitory manipulation results in an increase in firing rate entropy, this would suggest that the cortex is not operating on the high entropy surface.

A second prediction from our work is that size of the drop in entropy due to a manipulation of inhibition or excitation will be correlated with the entropy before the manipulation. This prediction supposes that the cortex is sometimes operating with a weak-synapse E/I balance where entropy is higher and the drop in entropy would be greater and at other times is operating with a strong-synapse E/I balance where entropy is lower and the drop in entropy would be less.

Although high entropy is likely to be beneficial for certain functions of cerebral cortex, other functions might be better served by a low entropy condition. For example, as discussed in the introduction, lower entropy might improve sensory signal processing by increasing the signal-to-noise ratio. In this context, a small shift toward the lower firing side of the phase transition might be beneficial. Such temporary shifts can occur due to neuromodulation; for example, attention is known to shift cortical dynamics toward a regime with smaller collective fluctuations.[33] However, a shift toward the high firing regime or too large a shift toward the extremely inhibition-dominant regime would likely be bad for function. Indeed, extreme deviation from well-balanced excitation and inhibition is implicated in a variety of brain disorders. For instance, when inhibition is sufficiently weak relative to excitation, seizures occur, as in epilepsy.[34] Too much inhibition is associated with Down's syndrome.[35] Autism is also associated with imbalanced excitation and inhibition,[36,37] both in terms of abnormal numbers of inhibitory neurons and strengths of synapses.[38] Our work suggests that the dysfunction associated with these disorders may be, in part, due to abnormal entropy of cortical network dynamics.

If high entropy is a beneficial property for brain circuits, then the robust maximization of entropy could be a phenotypic target of evolution in the nervous system. Our results suggest that hitting this target requires neural circuits that include some inhibitory neurons and operate near the tipping point of a phase transition.

## ACKNOWLEDGMENTS

Numerical simulations were performed on Trestles at the Arkansas High Performance Computing Center, which is funded through multiple National Science Foundation grants and the Arkansas Economic Development Commission.

[1]M. D. Fox and M. E. Raichle, "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging," Nat. Rev. Neurosci. **8**, 700 (2007).
[2]A. Arieli, A. Sterkin, A. Grinvald, and A. Aertsen, "Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses," Science **273**, 1868–1871 (1996).
[3]A. Luczak, P. Barthó, and K. D. Harris, "Spontaneous events outline the realm of possible sensory responses in neocortical populations," Neuron **62**, 413–425 (2009).
[4]F. Han, N. Caporale, and Y. Dan, "Reverberation of recent visual experience in spontaneous cortical waves," Neuron **60**, 321–327 (2008).
[5]P. Berkes, G. Orbán, M. Lengyel, and J. Fiser, "Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment," Science **331**, 83–87 (2011).
[6]S. A. Romano, T. Pietri, V. Pérez-Schuster, A. Jouary, M. Haudrechy, and G. Sumbre, "Spontaneous neuronal network dynamics reveal circuit's functional adaptations for behavior," Neuron **85**, 1070–1085 (2015).
[7]J.-e. K. Miller, I. Ayzenshtat, L. Carrillo-Reid, and R. Yuste, "Visual stimuli recruit intrinsically generated cortical ensembles," Proc. Natl. Acad. Sci. **111**, E4053–E4061 (2014).
[8]A. S. Ecker, P. Berens, R. J. Cotton, M. Subramaniyan, G. H. Denfield, C. R. Cadwell, S. M. Smirnakis, M. Bethge, and A. S. Tolias, "State dependence of noise correlations in macaque primary visual cortex," Neuron **82**, 235–248 (2014).
[9]D. Lee, N. L. Port, W. Kruse, and A. P. Georgopoulos, "Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex," J. Neurosci. **18**, 1161–1170 (1998).
[10]B. B. Averbeck, P. E. Latham, and A. Pouget, "Neural correlations, population coding and computation," Nat. Rev. Neurosci. **7**, 358 (2006).
[11]E. D. Fagerholm, G. Scott, W. L. Shew, C. Song, R. Leech, T. Knöpfel, and D. J. Sharp, "Cortical entropy, mutual information and scale-free dynamics in waking mice," Cereb. Cortex **26**, 3945–3952 (2016).
[12]W. L. Shew, H. Yang, S. Yu, R. Roy, and D. Plenz, "Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches," J. Neurosci. **31**, 55–63 (2011).
[13]B.-Q. Mao, F. Hamzei-Sichani, D. Aronov, R. C. Froemke, and R. Yuste, "Dynamics of spontaneous activity in neocortical slices," Neuron **32**, 883–898 (2001).
[14]S. H. Gautam, T. T. Hoang, K. McClanahan, S. K. Grady, and W. L. Shew, "Maximizing sensory dynamic range by tuning the cortical state to criticality," PLoS Comput. Biol. **11**, e1004576 (2015).
[15]X. Chen and R. Dzakpasu, "Observed network dynamics from altering the balance between excitatory and inhibitory neurons in cultured networks," Phys. Rev. E **82**, 031907 (2010).
[16]R. F. Hunt, K. M. Girskis, J. L. Rubenstein, A. Alvarez-Buylla, and S. C. Baraban, "Gaba progenitors grafted into the adult epileptic brain control seizures and abnormal behavior," Nat. Neurosci. **16**, 692–697 (2013).
[17]M. S. A. Ferraz, H. L. C. Melo-Silva, and A. H. Kihara, "Optimizing information processing in neuronal networks beyond critical states," PloS One **12**, e0184367 (2017).
[18]S. Scarpetta and A. de Candia, "Neural avalanches at the critical point between replay and non-replay of spatiotemporal patterns," PLoS One **8**, e64162 (2013).
[19]D.-P. Yang, H.-J. Zhou, and C. Zhou, "Co-emergence of multi-scale cortical activities of irregular firing, oscillations and avalanches achieves cost-efficient information capacity," PLoS Comput. Biol. **13**, e1005384 (2017).
[20]S. Denève and C. K. Machens, "Efficient codes and balanced networks," Nat. Neurosci. **19**, 375–382 (2016).
[21]M. Wehr and A. M. Zador, "Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex," Nature **426**, 442–446 (2003).
[22]B. Haider, A. Duque, A. R. Hasenstaub, and D. A. McCormick, "Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition," J. Neurosci. **26**, 4535–4545 (2006).
[23]S. Sahara, Y. Yanagawa, D. D. O'Leary, and C. F. Stevens, "The fraction of cortical gabaergic neurons is constant from near the start of cortical neurogenesis to adulthood," J. Neurosci. **32**, 4755–4761 (2012).
[24]S. Hendry, H. Schwark, E. Jones, and J. Yan, "Numbers and proportions of gaba-immunoreactive neurons in different areas of monkey cerebral cortex," J. Neurosci. **7**, 1503–1519 (1987).
[25]D. L. Meinecke and A. Peters, "Gaba immunoreactive neurons in rat visual cortex," J. Comp. Neurol. **261**, 388–404 (1987).
[26]D. B. Larremore, W. L. Shew, E. Ott, F. Sorrentino, and J. G. Restrepo, "Inhibition causes ceaseless dynamics in networks of excitable nodes," Phys. Rev. Lett. **112**, 138103 (2014).
[27]R. V. Williams-García, M. Moore, J. M. Beggs, and G. Ortiz, "Quasi-critical brain dynamics on a non-equilibrium widom line," Phys. Rev. E **90**, 062714 (2014).
[28]S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, "Highly nonrandom features of synaptic connectivity in local cortical circuits," PLoS Biol. **3**, e68 (2005).
[29]E. Lehmann and G. Casella, *Theory of Point Estimation*, Springer Texts in Statistics (Springer-Verlag, New York, 1998).
[30]C. van Vreeswijk and H. Sompolinsky, "Chaotic balanced state in a model of cortical circuits," Neural Comput. **10**, 1321–1371 (1998).
[31]R. Rubin, L. F. Abbott, and H. Sompolinsky, "Balanced excitation and inhibition are required for high-capacity, noise-robust neuronal selectivity," Proc. Natl. Acad. Sci. U.S.A. **114**, E9366–E9375 (2017).
[32]S. Denève and C. K. Machens, "Efficient codes and balanced networks," Nat. Neurosci. **19**, 375–382 (2016).
[33]K. D. Harris and A. Thiele, "Cortical state and attention," Nat. Rev. Neurosci. **12**, 509 (2011).
[34]M. A. Dichter and G. Ayala, "Cellular mechanisms of epilepsy: A status report," Science **237**, 157–164 (1987).

[35]F. Fernandez and C. C. Garner, "Over-inhibition: A model for developmental intellectual disability," Trends Neurosci. **30**, 497–503 (2007).
[36]J. Rubenstein and M. M. Merzenich, "Model of autism: Increased ratio of excitation/inhibition in key neural systems," Genes Brain Behav. **2**, 255–267 (2003).
[37]S. B. Nelson and V. Valakh, "Excitatory/inhibitory balance and circuit homeostasis in autism spectrum disorders," Neuron **87**, 684–698 (2015).
[38]N. Gogolla, J. J. LeBlanc, K. B. Quast, T. C. Südhof, M. Fagiolini, and T. K. Hensch, "Common circuit defect of excitatory-inhibitory balance in mouse models of autism," J. Neurodev. Disord. **1**, 172 (2009).