Edge of Chaos and Avalanches in Neural Networks with Heavy-Tailed Synaptic Weight Distribution

Łukasz Kuśmierz,[1,*] Shun Ogawa,[1] and Taro Toyoizumi[1,2,†]
[1]*Laboratory for Neural Computation and Adaptation, RIKEN Center for Brain Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
[2]*Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan*

(Received 12 November 2019; revised 3 March 2020; accepted 26 May 2020; published 7 July 2020)

We propose an analytically tractable neural connectivity model with power-law distributed synaptic strengths. When threshold neurons with biologically plausible number of incoming connections are considered, our model features a continuous transition to chaos and can reproduce biologically relevant low activity levels and scale-free avalanches, i.e., bursts of activity with power-law distributions of sizes and lifetimes. In contrast, the Gaussian counterpart exhibits a discontinuous transition to chaos and thus cannot be poised near the edge of chaos. We validate our predictions in simulations of networks of binary as well as leaky integrate-and-fire neurons. Our results suggest that heavy-tailed synaptic distribution may form a weakly informative sparse-connectivity prior that can be useful in biological and artificial adaptive systems.

DOI: 10.1103/PhysRevLett.125.028101

Scale-free neuronal avalanches, commonly associated with criticality, have been observed in cortical networks in various settings, including cultured and acute slices from rat somatosensory cortex [1], eye-attached *ex vivo* preparation of turtle visual cortex [2], visual cortex in anesthetized rats [3], primary visual cortex in anesthetized monkeys [3], and premotor, motor, and somatosensory cortex in awake monkeys [4]. Criticality implies the existence of a continuous transition between two distinct collective phases. In the context of neuronal avalanches, most commonly studied transitions are between quiescent and active states [1,5,6] or synchronuous and asynchronuous states [3,7]. In addition to providing a plausible generating mechanism for the neuronal avalanches, the existence of a continuous transition would have important functional implications, as it has been shown that computation is most efficient around a critical point [6,8,9], often associated with the edge of chaos [10–14]. However, the relation between neuronal avalanches, criticality, and the edge of chaos is not fully understood [9,15,16].

Different scenarios of the transition to chaos in randomly connected neural networks were extensively studied over the last 30 years [14,17–27]. According to the prevailing assumption rooted in the central limit theorem, the total synaptic input current of each neuron can be modeled as a Gaussian random variable (*Gaussian assumption*). Here we argue that the Gaussian assumption cannot account for some of the experimentally observed features of neuronal circuits.

In particular, the continuous nature of the phase transition observed in the conventional models is sensitive to theoretical assumptions that are not biologically grounded. Most works that study transition to chaos employ rate models with continuous nonthresholded activation functions, often of a sigmoidal shape [14,17,18,21,23,24,27]. Sometimes thresholds are introduced, but often the analysis is restricted to the suprathreshold regime [25,28]. But most neurons in the brain spike only when driven by strong enough excitatory synaptic input above a threshold [29–31]. Thus, we model a self-sustained (autonomous) activity in a network of individually subthreshold neurons. Other models exhibiting continuous transition to chaos [11] or neuronal avalanches [6,32] rely on extremely sparse ($\sim$10 connections per neuron) networks. However, many neurons in the vertebrate brain receive a large number of inputs from other neurons ($\sim 10^4$) [33]. We observed that the transition to chaos becomes discontinuous when densely connected subthreshold units are used in tandem with the Gaussian assumption (Fig. 3) [20,34]. This discontinuous character of the transition makes it hard for the network to robustly exhibit the edge of chaos, low activity levels, or avalanches. Although a discontinuous transition can be smoothed by noise, leading again to critical behavior away from the edge of chaos if the noise level is appropriate [35], such noise-induced criticality requires extra fine-tuning. We explore instead the possibility that an autonomous network exhibits critical behavior at the edge of chaos.

To fix this issue, we draw on the experimental works reporting heavy-tailed distributions of synaptic weights in various areas of the brain [36–41]. Multiple theoretical mechanisms have been suggested to realize such distributions, e.g., modified spike-timing-dependent plasticity (STDP) rule [42] or STDP combined with homeostatic plasticity [43]. Notably, recent studies have suggested that

 © 2020 American Physical Society

experimentally observed activity-independent intrinsic spine dynamics can straightforwardly explain the heavy-tailed distributions of synaptic weights [44–48].

Although extensively studied, the computational role of synaptic heavy tails is still not fully understood. A log-normal distribution is often assumed and the results are obtained by means of computer simulations [40,49–52]. Unfortunately, the log-normal distribution is not a stable distribution [53]. Consequently, the corresponding distribution of the membrane potential depends in a nontrivial way on the details of the connectivity, including number of incoming connections. Moreover, if the number of incoming connections is scaled linearly with the number of neurons, the Gaussian assumption is recovered in the thermodynamic limit. This hinders theoretical approaches to study the effects of heavy tails in these models. Therefore, a simple model that robustly predicts the effects of synaptic heavy tails is needed. We fill this gap by assuming random, power-law distributed synaptic weights.

Our aim is to inspect how the distribution of synaptic efficacies, modulated by the activation function, affects the transition to chaos and the associated avalanches. To this end, in our calculations we focus on the network effects and hence simplify the dynamics of individual neurons by considering the following discrete-time network dynamics

$$x_i(t+1) = \sum_{j=1}^{N} J_{ij}\phi(x_j(t)), \tag{1}$$

where $\phi(x)$ is the activation function, assumed to be identical across the network, and $\boldsymbol{J}$ is the connectivity matrix. The network is fully connected and the synaptic weights are independently drawn from the common Cauchy distribution (Fig. 1)

$$\rho(J_{ij}) = \frac{1}{\pi}\frac{g/N}{(g/N)^2 + J_{ij}^2}, \tag{2}$$

with the characteristic function

$$\Phi_J(k) = e^{-\gamma|k|}, \tag{3}$$

where $\gamma = g/N$ defines the width of the distribution. We refer to the model prescribed by (1) and (2) as the Cauchy network.

Due to the generalized central limit theorem [53], in the thermodynamic limit of $N \to \infty$ results obtained for the Cauchy model are applicable to networks with connections drawn independently from any symmetric distribution with $1/x^2$ tails that are scaled with the number of neurons as $1/N$. In contrast, in the more commonly used Gaussian networks, the synaptic weights are independently drawn from the normal distribution $J_{ij} \sim \mathcal{N}(0, g^2/N)$. In the thermodynamic limit this corresponds to connectivity matrices with entries independently drawn from any distribution with zero mean and a finite variance, as long as the weights are scaled as $1/\sqrt{N}$.

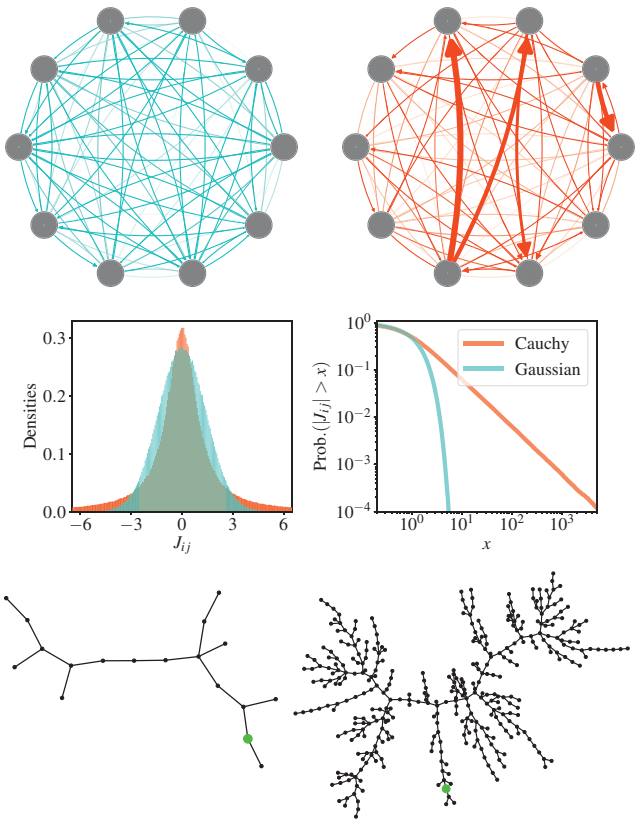

FIG. 1. (Top) Visualizations of neural networks with Gaussian (cyan) and Cauchy (orange) distribution of weights. Thickness and color saturation of edges correspond to the (nonlinearly transformed) strengths of the connections. (Middle) Probability density functions (left) and cumulative distribution functions (right) of Cauchy and Gaussian random variables. The Cauchy distribution features much thicker tails than the Gaussian distribution. (Bottom) Sample realizations of the Poisson critical branching process with the duration $T = 11$, size $S = 18$ (left), and $T = 35$, $S = 367$ (right). The initial seeds are marked with the green color. In this Letter, we show that activity of a fully connected Cauchy (but not Gaussian) network around the critical point can be mapped to the critical branching process.

The natural order parameter in the system at hand is the mean network activity, defined as

$$m_t = \frac{1}{N}\sum_{i=1}^{N} |\phi(x_i(t))|. \tag{4}$$

The state at time $t+1$ depends on $\boldsymbol{J}$ and $\boldsymbol{x}(t)$. We fix the activity vector at time t and treat $\boldsymbol{x}(t+1)$ as a function of $\boldsymbol{J}$, which allows us to characterize the distribution of $x_i(t+1)$ using $\Phi_{x_i(t+1)}(k)$ as

$$\begin{aligned}\langle e^{ikx_i(t+1)}\rangle_{\boldsymbol{J}} &= \exp\left(-g|k|N^{-1}\sum_{j=1}^{N}|\phi(x_j(t))|\right) \\ &= \exp(-gm_t|k|).\end{aligned} \tag{5}$$

The activity of a neuron at time $t+1$, as a function of synaptic weights, is a Cauchy random variable whose

width depends on the activity at time t only through its mean value.

To proceed we assume self-averaging, i.e., that the mean activity is the same for each realization of the network. Since in our model synaptic weights are statistically the same for all neurons, in the limit of $N \to \infty$ the mean activity can alternatively be expressed as

$$m_t = \langle |\phi(x_i(t))| \rangle_{\boldsymbol{J}} \quad (\forall\ i). \tag{6}$$

We use the result of (5), i.e., that $x_i(t)$ averaged over $\boldsymbol{J}$ is a Cauchy variable with $\gamma = g m_t$, together with (6) and arrive at the evolution of the mean activity in a simple integral form

$$m_{t+1} = \int_{-\infty}^{\infty} Dz |\phi(g m_t z)|, \tag{7}$$

where $Dz = \pi^{-1} dz/(1+z^2)$ denotes that the integral is calculated with respect to the standard Cauchy measure. The steady-state mean activity can be obtained from (7) in a self-consistent manner.

We are now in the position to analyze the dependence of the dynamics of the Cauchy network on the activation function. For $\phi(x) = x$, the integral on the right-hand side of (7) diverges, suggesting that the network is unstable. Indeed, it is easy to understand why this is the case. For linear networks the dynamics is fully determined by the eigenvalues of the connectivity matrix $\boldsymbol{J}$. It is known that, in contrast to random matrices with Gaussian entries, a Cauchy random matrix features an unbounded support of the eigenvalues density, even in the limit of $N \to \infty$ [54–56]. Thus, we can conclude that, regardless of the value of the g, the dynamics of a Cauchy neural network is in this case divergent. For the same reason, any $\phi(x)$ that is linear around $x \approx 0$ and grows sufficiently slow for large x leads to a self-sustained, active dynamics for any g [57]. However, in the biologically relevant regime neurons exhibit saturation and thresholding at, respectively, large and low values of total synaptic input. The corresponding Cauchy network generically exhibits two phases: quiescent and active, and an associated transition between them [57]. In general the nature of the active phase will depend on the details of the activation function.

To further simplify the calculations and to simultaneously model edge of chaos and avalanches, in the following we focus our attention on the binary activation function $\phi(x) = \Theta(x - \theta)$, where Θ denotes the Heaviside function and θ denotes the threshold. In this case the mean-field equation (7) simplifies to

$$m_{t+1} = \frac{1}{\pi} \arctan\left(m_t g/\theta\right). \tag{8}$$

The stability of the trivial fixed point can be checked by expanding the rhs of (8) around $m_t = 0$: $m_{t+1} = (g/\pi\theta) m_t + O(m_t^3)$. The fixed point at $m_t = 0$, corresponding to the quiescent phase, is unstable for $g > \pi\theta$. Since $\arctan(x)/\pi$ is saturating and concave for all $x > 0$, another stable fixed point m^* close to 0 appears, through the supercritical pitchfork bifurcation, exactly when the trivial fixed point loses its stability [$m^* \approx \sqrt{3}(g/\theta)^{-3/2}\sqrt{(g/\theta) - \pi}$ near the transition point]. Due to the quenched, asymmetric disorder of the connectivity matrix we can expect this fixed point to represent a chaotic attractor of the network [61], with a large sensitivity to small perturbations. Our computer simulations confirm this prediction [57].

The transition from the quiescent to the chaotic phase can be understood from the underlying structure of connections. Due to the power-law connectivity density, we can expect that only a small fraction of the connections contribute to the activity profile of the network. Indeed, as we show in the following, the transition to chaos is driven by the percolation transition of *autocrat* connections for which $J_{ij} > \theta$; i.e., an active presynaptic neuron will activate the postsynaptic neuron in the absence of other inputs. Around the critical point the mean activity of the network is infinitesimal and thus the higher order interaction events (e.g., two neurons activating another neuron) are negligible. In other words, to a good approximation, a neuron can only be activated by another single neuron through an autocrat connection, independently from other neurons. This suggests that the transition to chaos in the neural network model is related to the critical branching processes [64] (Fig. 1).

In the Cauchy case the probability that a given connection is an autocrat reads

$$\mathrm{Prob}(J_{ij} > \theta) = \frac{1}{\pi} \arctan\left(\frac{g}{N\theta}\right). \tag{9}$$

For a given neuron, the number of outgoing (or incoming) autocrat connections is a binomial random variable with N trials and the probability of success given by (9). In the limit of $N \to \infty$ it converges to the Poisson random variable with intensity

$$\lambda = \lim_{N\to\infty} \frac{N}{\pi} \arctan\left(\frac{g}{N\theta}\right) = \frac{g}{\theta\pi}. \tag{10}$$

Now, let the initial state of the network be such that only a single neuron (seed) is active. The number of active neurons (descendants) in the next step is given by the Poisson distribution and the mean number of active neurons is given by (10). The theory of branching processes predicts that the population will eventually die out almost surely for $\lambda \leq 1$ and has a finite survival probability for $\lambda > 1$. At $\lambda = 1$ the process is critical and features scale-free avalanches. The critical point predicted by the branching process formulation of the network dynamics, $g^* = \pi\theta$, is the same as the mean-field critical point predicted by (8).

The mapping to the branching process explains many features of the Cauchy neural network around the critical

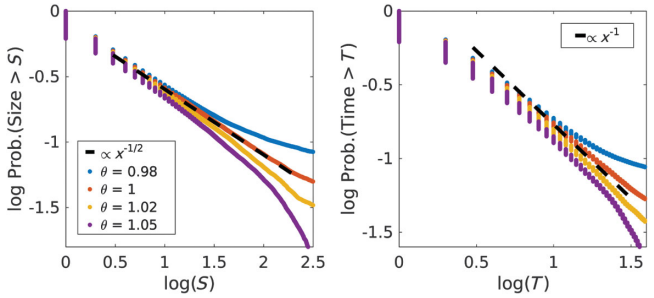

FIG. 2. Avalanche size and lifetime distributions in the networks of binary units. As expected from our theoretical predictions, at the transition point these distributions are described by power laws, and the critical exponents match those of the critical branching process.

point. Below the critical point the steady state is quiescent, and a bit-flip perturbation corresponds to a single neuron (seed) being activated. The local expansion rate of such perturbation is given by λ. Above the critical point ($g > \pi\theta$) each bit flip contributes in the same manner as a single seed and, additionally, interacts with other active neurons to activate and deactivate other descendants. Thus, in the vicinity of the transition point λ gives a lower bound on the local expansion rate of a perturbation in the steady state, and for $\lambda > 1$ the network is expected to be chaotic in the thermodynamic limit [57]. Moreover, the transition to chaos belongs to the mean-field directed percolation universality class [65–67]. The propagation of the corresponding avalanches is characterized by [57] power-law distributed sizes S: $\mathrm{Prob}(S > s) \sim s^{-1/2}$, and power-law distributed lifetimes T: $\mathrm{Prob}(T > t) \sim t^{-1}$. These theoretical predictions were corroborated by our computer simulations of the Cauchy network, as shown in Fig. 2.

For a comparison, we have also studied Gaussian networks of threshold units with a fixed number of connections per neuron K [57,68]. While extremely sparsely connected Gaussian networks ($K \lessapprox 12$) behave qualitatively similar to the Cauchy network, the transition to chaos becomes discontinuous in the biologically relevant regime of $K \gtrapprox 13$. With a biologically realistic K and finite N, the network activity jumps between two metastable states near the transition point, and cannot be robustly posed at the edge of chaos. The discontinuous property is due to the emergence of a metastable active state by the saddle-node bifurcation as g/θ increases.

Importantly, although our theoretical predictions were derived assuming simplistic threshold neural units, they translate directly to networks of more biologically plausible leaky integrate-and-fire (LIF) neurons. The difference of continuous and discontinuous transition is confirmed by the presence or absence of a hysteresis loop in more realistic networks of LIF neurons (Fig. 3). Hence, unlike the Gaussian networks with realistic K, Cauchy networks demonstrate critical phenomena and can reproduce experimentally observed scale-free avalanches at the critical point. Moreover, a large Cauchy network can exhibit arbitrarily low, self-sustained activity levels. In contrast, the lowest possible activity level that can be achieved by the Gaussian network with realistic K is about 11% in the binary case and 40 Hz in the LIF case (Fig. 3).

For clarity we chose to limit our presentation to the Cauchy distribution of J_{ij}, but our results naturally extend to other power-law distributions. Indeed, let the synaptic efficacy density asymptotically behave like a power law $\rho(J_{ij}) \sim C\alpha g^{\alpha} N^{-1} |J_{ij}|^{-1-\alpha}$ [69]. We then have $\mathrm{Prob}(J_{ij} > \theta) = CN^{-1}(g/\theta)^{\alpha}$, which holds for large enough N. The branching parameter is calculated as in (10) and reads $\lambda = C(g/\theta)^{\alpha}$. A continuous transition takes place at $\lambda = 1$ and its features

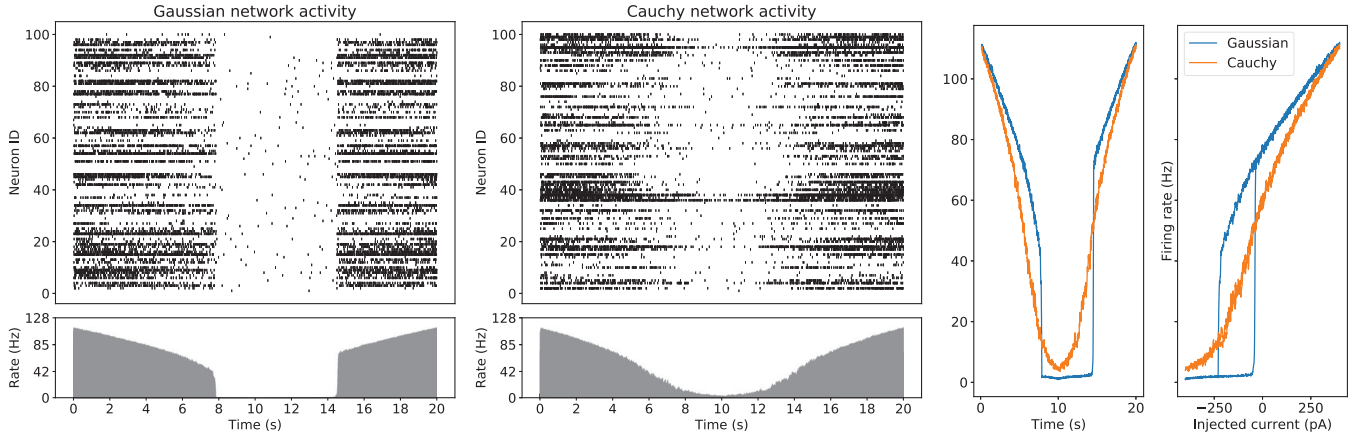

FIG. 3. Continuous vs discontinuous transition in networks of leaky integrate-and-fire neurons. A slowly changing current was injected and an average firing rate of the network was recorded as a function of time and the injected current amplitude. As predicted by our theory, a network with Gaussian weights exhibits a discontinuous transition between active and inactive states, which generates a characteristic hysteresis loop. In contrast, the Cauchy network exhibits a continuous transition and thus shows no signs of the hysteresis loop.

are, as before, described by the directed percolation universality class.

The connectivity in the current model is unstructured. It would be interesting to combine power-law synaptic weight distributions with structured networks, for example, with hierarchical modules [70] or oscillations [3,7,71].

Incidentally, the same Cauchy distribution of membrane potential was found in the quadratic integrate and fire neuron model [72] due to single-neuron dynamics. Nontrivial effects may arise from a combination of single-neuron and network-driven heavy-tail statistics.

Power-law distributions of synaptic weights feature many very weak synapses, that do not directly contribute to the computation. Even though this may seem wasteful, we think that such architectures are not only biologically plausible [73] but may be beneficial. One possibility is that even weak connections can activate a neuron once contextual input from another part of the brain increases the baseline membrane potential close to its spiking threshold. Such contextual input can also raise the spiking probability of nearby neurons so that synchronous activation of weak connections is more likely. Weak synapses have also been reported to play a role in unsupervised features extraction [74]. In the context of reservoir computing [75,76] high computational capabilities were achieved using nonbiologically sparse connectivity [77]. Our model provides a more biologically plausible solution that the connectivity can be anatomically dense but effectively sparse due to heavy-tailed synaptic weight distribution. More generally, the optimal degree of sparsity depends on the role of a given brain structure and the type of the employed plasticity [78]. Power-law distributed synaptic weights may in this context provide a weakly informative [79,80] sparse connectivity prior, with weak and effectively silent synapses providing a pool of potential connections that can be recruited when and if needed, as observed in the brain during development [81,82].

Our results demonstrate that the shape of the synaptic weight distribution can dramatically affect dynamics of neural networks. A biological distribution of synaptic weights can give distinct predictions from the frequently assumed Gaussian distribution. The proposed mathematical framework with power-law synaptic weights can easily be adapted to other scenarios in future studies.

We thank Francesco Fumarola and anonymous referees for invaluable discussions and comments on the manuscript. Supported by RIKEN Center for Brain Science, Brain/MINDS from AMED under Grant No. JP20dm020700, and JSPS KAKENHI Grant No. JP18H05432.

*nalewkoz@gmail.com
†taro.toyoizumi@riken.jp

[1] J. M. Beggs and D. Plenz, J. Neurosci. **23**, 11167 (2003).
[2] W. L. Shew, W. P. Clawson, J. Pobst, Y. Karimipanah, N. C. Wright, and R. Wessel, Nat. Phys. **11**, 659 (2015).
[3] A. J. Fontenele, N. A. P. de Vasconcelos, T. Feliciano, L. A. A. Aguiar, C. Soares-Cunha, B. Coimbra, L. Dalla Porta, S. Ribeiro, A. J. Rodrigues, N. Sousa *et al.*, Phys. Rev. Lett. **122**, 208101 (2019).
[4] T. Petermann, T. C. Thiagarajan, M. A. Lebedev, M. A. Nicolelis, D. R. Chialvo, and D. Plenz, Proc. Natl. Acad. Sci. U.S.A. **106**, 15921 (2009).
[5] C. Haldeman and J. M. Beggs, Phys. Rev. Lett. **94**, 058101 (2005).
[6] O. Kinouchi and M. Copelli, Nat. Phys. **2**, 348 (2006).
[7] S. di Santo, P. Villegas, R. Burioni, and M. A. Muñoz, Proc. Natl. Acad. Sci. U.S.A. **115**, E1356 (2018).
[8] R. Stoop and F. Gomez, Phys. Rev. Lett. **117**, 038102 (2016).
[9] M. A. Munoz, Rev. Mod. Phys. **90**, 031001 (2018).
[10] C. G. Langton, Physica (Amsterdam) **42D**, 12 (1990).
[11] N. Bertschinger and T. Natschläger, Neural Comput. **16**, 1413 (2004).
[12] R. Legenstein and W. Maass, in *New Directions in Statistical Signal Processing: From Systems to Brain*, edited by S. Haykin, J. C. Principe, T. Sejnowski, and J. McWhirter (MIT Press, Cambridge, 2007), pp. 127–154.
[13] T. Toyoizumi and L. F. Abbott, Phys. Rev. E **84**, 051908 (2011).
[14] J. Schuecker, S. Goedeke, and M. Helias, Phys. Rev. X **8**, 041029 (2018).
[15] M. Benayoun, J. D. Cowan, W. van Drongelen, and E. Wallace, PLoS Comput. Biol. **6**, e1000846 (2010).
[16] J. M. Beggs and N. Timme, Front. Physiol. **3**, 163 (2012).
[17] H. Sompolinsky, A. Crisanti, and H.-J. Sommers, Phys. Rev. Lett. **61**, 259 (1988).
[18] L. Molgedey, J. Schuchhardt, and H. G. Schuster, Phys. Rev. Lett. **69**, 3717 (1992).
[19] D. J. Amit and N. Brunel, Cereb. Cortex **7**, 237 (1997).
[20] N. Brunel, J. Comput. Neurosci. **8**, 183 (2000).
[21] K. Rajan, L. F. Abbott, and H. Sompolinsky, Phys. Rev. E **82**, 011903 (2010).
[22] M. Massar and S. Massar, Phys. Rev. E **87**, 042809 (2013).
[23] M. Stern, H. Sompolinsky, and L. F. Abbott, Phys. Rev. E **90**, 062710 (2014).
[24] J. Aljadeff, M. Stern, and T. Sharpee, Phys. Rev. Lett. **114**, 088101 (2015).
[25] J. Kadmon and H. Sompolinsky, Phys. Rev. X **5**, 041030 (2015).
[26] G. Wainrib and M. N. Galtier, Neural Netw. **76**, 39 (2016).
[27] A. Crisanti and H. Sompolinsky, Phys. Rev. E **98**, 062120 (2018).
[28] O. Harish and D. Hansel, PLoS Comput. Biol. **11**, e1004266 (2015).
[29] P. Dayan and L. F. Abbott, *Theoretical Neuroscience* (The MIT Press, Cambridge, 2001).
[30] M. Häusser and P. Monsivais, Neuron **40**, 449 (2003).
[31] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski, *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition* (Cambridge University Press, Cambridge, England, 2014).
[32] D. Millman, S. Mihalas, A. Kirkwood, and E. Niebur, Nat. Phys. **6**, 801 (2010).
[33] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (The MIT Press, 2005).

[34] J. H. Marshel, Y. S. Kim, T. A. Machado, S. Quirin, B. Benson, J. Kadmon, C. Raja, A. Chibukhchyan, C. Ramakrishnan, M. Inoue *et al.*, Science **365**, eaaw5202 (2019).

[35] S. Scarpetta, I. Apicella, L. Minati, and A. de Candia, Phys. Rev. E **97**, 062305 (2018).

[36] R. Sayer, M. Friedlander, and S. Redman, J. Neurosci. **10**, 826 (1990).

[37] D. Feldmeyer, V. Egger, J. Lübke, and B. Sakmann, J. Physiol. **521**, 169 (1999).

[38] S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, PLoS Biol. **3**, e68 (2005).

[39] S. Lefort, C. Tomm, J.-C. F. Sarria, and C. C. Petersen, Neuron **61**, 301 (2009).

[40] Y. Ikegaya, T. Sasaki, D. Ishikawa, N. Honma, K. Tao, N. Takahashi, G. Minamisawa, S. Ujita, and N. Matsuki, Cereb. Cortex **23**, 293 (2013).

[41] Y. Loewenstein, A. Kuras, and S. Rumpel, J. Neurosci. **31**, 9481 (2011).

[42] M. Gilson and T. Fukai, PLoS One **6**, e25339 (2011).

[43] P. Zheng, C. Dimitrakakis, and J. Triesch, PLoS Comput. Biol. **9**, e1002848 (2013).

[44] N. Yasumatsu, M. Matsuzaki, T. Miyazaki, J. Noguchi, and H. Kasai, J. Neurosci. **28**, 13592 (2008).

[45] A. Nagaoka, H. Takehara, A. Hayashi-Takagi, J. Noguchi, K. Ishii, F. Shirai, S. Yagishita, T. Akagi, T. Ichiki, and H. Kasai, Sci. Rep. **6**, 26651 (2016).

[46] J. Humble, K. Hiratsuka, H. Kasai, and T. Toyoizumi, Front. Comput. Neurosci. **13**, 38 (2019).

[47] K. Ishii, A. Nagaoka, Y. Kishida, H. Okazaki, S. Yagishita, H. Ucar, N. Takahashi, N. Saito, and H. Kasai, eNeuro **5**, 0282 (2018).

[48] H. Okazaki, A. Hayashi-Takagi, A. Nagaoka, M. Negishi, H. Ucar, S. Yagishita, K. Ishii, T. Toyoizumi, K. Fox, and H. Kasai, Neurosci. Lett. **671**, 99 (2018).

[49] J.-n. Teramae, Y. Tsubo, and T. Fukai, Sci. Rep. **2**, 485 (2012).

[50] G. Buzsáki and K. Mizuseki, Nat. Rev. Neurosci. **15**, 264 (2014).

[51] J.-n. Teramae and T. Fukai, Proc. IEEE **102**, 500 (2014).

[52] Y. Omura, M. M. Carvalho, K. Inokuchi, and T. Fukai, J. Neurosci. **35**, 14585 (2015).

[53] W. Feller, *An Introduction to Probability and Its Applications* (Wiley, New York, 1971), Vol. II.

[54] P. Cizeau and J.-P. Bouchaud, Phys. Rev. E **50**, 1810 (1994).

[55] Z. Burda, R. A. Janik, J. Jurkiewicz, M. A. Nowak, G. Papp, and I. Zahed, Phys. Rev. E **65**, 021106 (2002).

[56] E. Gudowska-Nowak, M. A. Nowak, D. R. Chialvo, J. K. Ochab, and W. Tarnowski, Neural Comput. **32**, 395 (2020).

[57] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.125.028101 for details, which includes Refs. [58–60].

[58] E. Goles, Discrete Appl. Math. **13**, 97 (1986).

[59] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, SIAM Rev. **59**, 65 (2017).

[60] C. Linssen *et al.*, Nest 2.16.0, 2018, https://doi.org/10.5281/zenodo.1400175.

[61] In the binary case the system has a finite number of states for any finite N and the attractor has to be periodic. Thus, formally, irregular aperiodic behavior can only be observed in the limit of $N \to \infty$. Chaoslike signatures can nonetheless be observed for finite N; e.g., the typical lengths of transients and cycles rapidly change around the predicted transition point, and grow exponentially with N in the "chaotic" phase [62,63].

[62] C. van Vreeswijk and H. Sompolinsky, Neural Comput. **10**, 1321 (1998).

[63] B. Luque and R. V. Solé, Physica (Amsterdam) **284A**, 33 (2000).

[64] T. E. Harris, *The Theory of Branching Processes* (Springer, Berlin, 1963).

[65] P. Alstrøm, Phys. Rev. A **38**, 4905 (1988).

[66] M. A. Munoz, R. Dickman, A. Vespignani, and S. Zapperi, Phys. Rev. E **59**, 6175 (1999).

[67] G. Ódor, Rev. Mod. Phys. **76**, 663 (2004).

[68] B. Derrida, E. Gardner, and A. Zippelius, Europhys. Lett. **4**, 167 (1987).

[69] Note that we scale J_{ij} with the number of neurons as $N^{-1/\alpha}$, which assures the existence of a nontrivial limit $N \to \infty$. For $\alpha > 2$ another choice is to scale the synaptic strengths as $N^{-1/2}$, which in the limit of $N \to \infty$ corresponds to the Gaussian network.

[70] E. J. Friedman and A. S. Landsberg, Chaos **23**, 013135 (2013).

[71] S.-S. Poil, R. Hardstone, H. D. Mansvelder, and K. Linkenkaer-Hansen, J. Neurosci. **32**, 9817 (2012).

[72] E. Montbrió, D. Pazó, and A. Roxin, Phys. Rev. X **5**, 021028 (2015).

[73] L. Cossell, M. F. Iacaruso, D. R. Muir, R. Houlton, E. N. Sader, H. Ko, S. B. Hofer, and T. D. Mrsic-Flogel, Nature (London) **518**, 399 (2015).

[74] H. Huang, J. Phys. A **51**, 08LT01 (2018).

[75] H. Jaeger, German National Research Center for Information Technology, Bonn, Germany, GMD Technical Report No. 148, p. 13 2001.

[76] W. Maass, T. Natschläger, and H. Markram, Neural Comput. **14**, 2531 (2002).

[77] L. Büsing, B. Schrauwen, and R. Legenstein, Neural Comput. **22**, 1272 (2010).

[78] A. Litwin-Kumar, K. D. Harris, R. Axel, H. Sompolinsky, and L. Abbott, Neuron **93**, 1153 (2017).

[79] A. Gelman, A. Jakulin, M. G. Pittau, Y.-S. Su, Ann. Appl. Stat. **2**, 1360 (2008).

[80] S. Van Dongen, J. Theor. Biol. **242**, 90 (2006).

[81] D. Liao, N. A. Hessler, and R. Malinow, Nature (London) **375**, 400 (1995).

[82] G. A. Kerchner and R. A. Nicoll, Nat. Rev. Neurosci. **9**, 813 (2008).