



Published in final edited form as:

*Phys Rev Lett.* 2015 February 27; 114(8): 088101.

## Transition to chaos in random networks with cell-type-specific connectivity

Johnatan Aljadeff<sup>1,2,a</sup>, Merav Stern<sup>3,4,a</sup>, and Tatyana Sharpee<sup>1,2,b</sup>

<sup>1</sup>Computational Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, California, USA

<sup>2</sup>Center for Theoretical Biological Physics and Department of Physics, University of California, San Diego, USA

<sup>3</sup>Department of Neuroscience, Columbia University, New York, New York, USA

<sup>4</sup>The Edmond and Lily Safra Center for Brain Sciences, Hebrew University, Jerusalem, Israel

### Abstract

In neural circuits, statistical connectivity rules strongly depend on cell-type identity. We study dynamics of neural networks with cell-type specific connectivity by extending the dynamic mean field method, and find that these networks exhibit a phase transition between silent and chaotic activity. By analyzing the locus of this transition, we derive a new result in random matrix theory: the spectral radius of a random connectivity matrix with block-structured variances. We apply our results to show how a small group of hyper-excitable neurons within the network can significantly increase the network's computational capacity by bringing it into the chaotic regime.

---

The theory of random matrices has diverse applications in nuclear [1] and solid-state [2, 3] physics, number theory and statistics [4] and models of neural networks [5–8]. The increasing use of boolean networks to model gene regulatory networks [9–11] suggests that random matrix theory may advance our understanding of those biological systems as well. Most existing theoretical results pertain to matrices with values drawn from a single distribution, corresponding to randomly connected networks with a single connectivity rule and cell-type. Recent experimental studies describe in increasing detail the heterogeneous structure of biological networks where connection probability depends strongly on cell-type [12–17]. As a step towards bridging this gap between theory and experiment, we extend here mean-field methods used to analyze conventional randomly connected networks to networks with multiple cell-types and allow for cell-type-dependent connectivity rules. We focus here on neural networks.

Randomly connected networks of one cell-type were shown to have two important properties. First, they undergo a phase transition from silent to chaotic activity as the variance of connection strength is increased [7, 8]. Second, such networks reach optimal computational capacity near the critical point [18, 19] in a weakly chaotic regime. We find

---

<sup>a</sup>Equal contribution  
<sup>b</sup>sharpee@salk.edu

both phenomena in networks with multiple cell-types. Importantly, the effective gain of multi-type networks deviates strongly from predictions obtained by averaging across the cell types, and in many cases these networks show greater computational capacity compared to networks with cell-type independent connectivity.

The starting point for our analysis of recurrent activity in neural networks is a firing-rate model where the activation  $x_i(t)$  of the  $i$ th neuron determines its firing-rate  $\phi_i(t)$  through a nonlinear function  $\phi_i(t) = \tanh(x_i)$ . The activation of the  $i$ th neuron depends on the firing-rate of all  $N$  neurons in the network:

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1}^N J_{ij} \phi_j(t), \quad (1)$$

where  $J_{ij}$  describes the connection weight from neuron  $j$  to  $i$ . Previous work [7] considered a recurrent random network where all connections are drawn from the same distribution. There, the matrix elements was drawn from a Gaussian distribution with mean zero and variance  $g^2/N$ , where  $g$  defines the average synaptic gain in the network. According to Girko's circular law, the spectral density of the random matrix  $\mathbf{J}$  in this case is uniform on a disk with radius  $g$  [20, 21]. When the real part of some of the eigenvalues of  $\mathbf{J}$  exceeds 1, the quiescent state  $x_i(t) = 0$  becomes unstable and the network becomes chaotic [7]. Thus, for networks with one cell-type the transition to chaotic dynamics occurs when  $g = 1$ . The chaotic dynamics persist even in the presence of noise, but the critical point  $g_{\text{crit}}$  shifts to values  $> 1$ , with  $g_{\text{crit}} = 1 - \sigma^2 \log \sigma^2$  for small noise intensities  $\sigma^2$  and  $g_{\text{crit}} = \sqrt{\pi/2\sigma}$  for large noise [8].

We now consider networks with  $D$  cell-types, each with a fraction  $\alpha_d$  of neurons in it. The mean connection weight is  $\langle J_{ij} \rangle = 0$ . The variances  $N \langle J_{ij}^2 \rangle = g_{c_i d_j}^2$  depend on the cell-type of the input ( $c$ ) and output ( $d$ ) neurons; where  $c_i$  denotes the group neuron  $i$  belongs to. In what follows, indices  $i, j = 1, \dots, N$  and  $c, d = 1, \dots, D$  correspond to single neurons and neuron groups, respectively. Averages over realizations of  $\mathbf{J}$  are denoted by  $\langle \cdot \rangle$ . It is convenient to represent the connectivity structure using a synaptic gain matrix  $\mathbf{G}$ . Its elements  $G_{ij} = g_{c_i d_j}$  are arranged in  $D^2$  blocks of sizes  $N \alpha_c \times N \alpha_d$  (Fig. 1a–c, top insets). The mean synaptic gain,  $\bar{g}$ , is given by  $N^{-1} (\sum_{i,j=1}^N G_{ij}^2)^{\frac{1}{2}} = (\sum_{c,d=1}^D \alpha_c \alpha_d g_{cd}^2)^{\frac{1}{2}}$ . Defining  $J_{ij}^0 \sim \mathcal{N}(0, N^{-1})$  (but see [22] for discussion of non-Gaussian entries) and  $n_d = N \sum_{c=1}^D \alpha_c$  allows us to rewrite Eq. (1) in a form that emphasizes the separate contributions from each group to a neuron:

$$\dot{x}_i = -x_i + \sum_{d=1}^D g_{c_i d} \sum_{j=n_{d-1}+1}^{n_d} J_{ij}^0 \phi_j(t). \quad (2)$$

We use the dynamic mean field approach [5, 7, 23] to study the network behavior in the  $N \rightarrow \infty$  limit. Averaging Eq. (2) over the ensemble from which  $\mathbf{J}$  is drawn implies that only neurons that belong to the same group are statistically identical. Therefore, to represent the

network behavior it is enough to look at the activities  $\xi_d(t)$  of  $D$  representative neurons and their inputs  $\eta_d(t)$ .

The stochastic mean field variables  $\xi$  and  $\eta$  will approximate the activities and inputs in the full  $N$  dimensional network provided that they satisfy the dynamic equation

$$\dot{\xi}_d(t) = -\xi_d(t) + \eta_d(t), \quad (3)$$

and provided that  $\eta_d(t)$  is drawn from a Gaussian distribution with moments satisfying the following conditions. First, the mean  $\langle \eta_d(t) \rangle = 0$  for all  $d$ . Second, the correlations of  $\eta$  should match the input correlations in the full network, averaged separately over each group.

Using Eq. (3) and the property  $N \langle J_{ij}^0 J_{kl}^0 \rangle = \delta_{ik} \delta_{jl}$  we get the self-consistency conditions:

$$\langle \eta_c(t) \eta_d(t+\tau) \rangle = \sum_{a,b=1}^D \sum_{j=n_{a-1}+1}^{n_a} \sum_{l=n_{b-1}+1}^{n_b} g_{ca} g_{db} \langle J_{ij}^0 J_{kl}^0 \rangle \langle \phi[x_j(t)] \phi[x_l(t+\tau)] \rangle = \delta_{cd} \sum_{b=1}^D \alpha_b g_{cb}^2 C_b(\tau), \quad (4)$$

where  $\langle \cdot \rangle$  denotes averages over  $i = n_{c-1} + 1, \dots, n_c$  and  $k = n_{d-1} + 1, \dots, n_d$  in addition to average over realizations of  $\mathbf{J}$ . The average firing rate correlation vector is denoted by  $\mathbf{C}(\tau)$ . Its components (using the variables of the full network) are

$$C_d(\tau) = \frac{1}{N \alpha_d} \sum_{i=n_{d-1}+1}^{n_d} \langle \phi[x_i(t)] \phi[x_i(t+\tau)] \rangle, \text{ translating to } C_d(\tau) = \langle \phi[\xi_d(t)] \phi[\xi_d(t+\tau)] \rangle$$

using the mean field variables. Importantly, the covariance matrix  $\mathcal{H}(\tau)$  with elements  $\mathcal{H}_{cd}(\tau) = \langle \eta_c(t) \eta_d(t+\tau) \rangle$  is diagonal, justifying the definition of the vector  $\mathbf{H} = \text{diag}(\mathcal{H})$ . With this in hand we rewrite Eq. (4) in matrix form

$$\mathbf{H}(\tau) = \mathbf{M} \mathbf{C}(\tau), \quad (5)$$

where  $\mathbf{M}$  is a constant matrix reflecting the network connectivity structure:  $M_{cd} = \alpha_d g_{cd}^2$

A trivial solution to this equation is  $\mathbf{H}(\tau) = \mathbf{C}(\tau) = 0$  which corresponds to the silent network state:  $x_i(t) = 0$ . Recall that in the network with a single cell-type, the matrix  $\mathbf{M} = g^2$  is a scalar and Eq. (5) reduces to  $H(\tau) = g^2 C(\tau)$ . In this case the silent solution is stable only when  $g < 1$ . For  $g > 1$  the autocorrelations of  $\eta$  are non-zero which leads to chaotic dynamics in the  $N$  dimensional system [7].

In the general case ( $D > 1$ ), Eq. (5) can be projected on the eigenvectors of  $\mathbf{M}$  leading to  $D$  consistency conditions, each equivalent to the single group case. Each projection has an effective scalar given by the eigenvalue in place of  $g^2$  in the  $D = 1$  case. Hence, the trivial solution will be stable if all eigenvalues of  $\mathbf{M}$  have real part  $< 1$ . This is guaranteed if  $\Lambda_1$ , the largest eigenvalue of  $\mathbf{M}$ , is  $< 1$  [24]. If  $\Lambda_1 > 1$  the projection of Eq. (5) on the leading eigenvector of  $\mathbf{M}$  gives a scalar self-consistency equation analogous to the  $D = 1$  case for which the trivial solution is unstable. As we know from the analysis of the single cell-type network, this leads to chaotic dynamics in the full network. Therefore  $\Lambda_1 = 1$  is the critical point of the multiple cell-type network.

Another approach to show explicitly that  $\Lambda_1 = 1$  at the critical point is to consider first order deviations in the network activity from the quiescent state. Here  $\mathbf{C}(\tau) \approx \langle \mathbf{x}(t) \mathbf{x}(t+\tau) \rangle$  where  $\mathbf{x}(t)$  is the autocorrelation vector of the activities with elements  $x_d(t) = \langle \xi_d(t) \xi_d(t+\tau) \rangle$ . By invoking Eq. (3) we have

$$\mathbf{H}(\tau) = \left( 1 - \frac{d^2}{d\tau^2} \right) \mathbf{\Delta}(\tau). \quad (6)$$

Substituting Eq. (6) into Eq. (5) leads to an equation of motion for a particle with coordinates  $\mathbf{x}(\tau)$ :

$$\frac{d^2 \mathbf{\Delta}(\tau)}{d\tau^2} = (\mathbb{I} - \mathbf{M}) \mathbf{\Delta}(\tau). \quad (7)$$

The particle's trajectories depend on the eigenvalues of  $\mathbf{M}$ . The first bifurcation (assuming the elements of  $\mathbf{M}$  are scaled together) occurs when  $\Lambda_1 = 1$ , in the direction parallel to the leading eigenvector. Physical solutions should have  $\|\mathbf{x}(\tau)\| < \infty$  as  $\tau \rightarrow \infty$  because  $\mathbf{x}(\tau)$  is an autocorrelation function. When all eigenvalues of  $\mathbf{M}$  are smaller than 1 the trivial solution  $\mathbf{x}(\tau) = 0$  is the only solution (in the neighborhood of  $x_i(t) = 0$  where our approximation is accurate). At the critical point ( $\Lambda_1 = 1$ ) a non trivial solution appears, and above it finite autocorrelations lead to chaotic dynamics in the full system.

The eigenvalue spectrum of  $\mathbf{J}$  is circularly symmetric in the absence of correlation between matrix entries as is evident from numerical simulations and direct calculations using random matrix theory techniques [25]. To derive the radius  $r$  of the support of its spectral density, one can use the following scaling relationship. If all elements of the matrix  $g_{cd}$  are multiplied by a constant  $\kappa$ , the radius  $r$  will scale linearly with  $\kappa$ . At the same time,  $M_{cd} \propto g_{cd}^2$ , so  $\Lambda_1 \propto \kappa^2$ . Thus,  $r \propto \sqrt{\Lambda_1}$ . The proportionality constant can be determined by noting that for both single and multiple cell-type networks this transition occurs when a finite mass of the spectral density of  $\mathbf{J}$  has real part  $> 1$ , which can also be verified by direct computation of the largest Lyapunov exponent [22]. The transition occurs at  $\Lambda_1 = 1$ , meaning that for  $\Lambda_1 = 1$  the eigenvalues of  $\mathbf{J}$  are bounded in the unit circle  $r = 1$ , so in general:

$$r(\alpha, \mathbf{g}) = \sqrt{\Lambda_1} = \sqrt{\max[\lambda(\mathbf{M})]}. \quad (8)$$

Predictions for the radius according to Eq. (8) matched numerical simulations for a number of different matrix configurations (Fig. 1a,b). Eq. (8) also holds for networks with cell-type independent connectivity, in which case  $\Lambda_1 = g^2$  and  $r = g$ . Importantly,  $r$  differs qualitatively from the mean synaptic gain  $\bar{g}$ . The inequality  $\sqrt{\Lambda_1} \neq \bar{g}$  is a signature of the block structured variances. It is not observed in the case where the variances have columnar structure [26] or when the  $J_{ij}$ 's are randomly permuted.

Next we analyze the network dynamics above the critical point. In the chaotic regime the persistent population-level activity is determined by the structure matrix  $\mathbf{M}$ . Consider the decomposition  $\mathbf{M} = \sum_{c=1}^D \Lambda_c |u_c^R\rangle \langle u_c^L|$  where  $|u_c^R\rangle, \langle u_c^L|$  are the right and left eigenvectors ordered by the real part of their corresponding eigenvalues  $\Re\{\Lambda_c\}$ , satisfying  $\langle u_c^L | u_d^R \rangle = \delta_{cd}$ . We find, with analogy to the analysis of the scalar self consistency equation in [7] that the trivial solution to Eq. (5) is unstable in the subspace  $\mathcal{U}_M = \text{span}\{|u_1^R\rangle, \dots, |u_{D^*}^R\rangle\}$ , where  $D^*$  is the number of eigenvalues of  $\mathbf{M}$  with real part  $> 1$ . In that subspace the solution to Eq. (5) is a combination of  $D^*$  different autocorrelation functions. In the  $D - D^*$  dimensional orthogonal complement subspace  $\mathcal{U}_M^\perp$  the trivial solution is stable. Consequently, the vectors  $\mathbf{H}(\tau), \mathbf{q}(\tau)$  are significant in  $\mathcal{U}_M$  with  $\approx 0$  projection on any vector in  $\mathcal{U}_M^\perp$  (Fig. 2). Note that for asymmetric  $\mathbf{M}$ ,  $|u_c^R\rangle$  are not orthogonal and  $\mathcal{U}_M^\perp$  is spanned by the left rather than the right eigenvectors:  $\mathcal{U}_M^\perp = \text{span}\{\langle u_{D^*+1}^L|, \dots, \langle u_D^L|\}$ .

In the special case  $D^* = 1$  we can write  $\mathbf{u}_1^R q_H(\tau)$  and  $\mathbf{\Delta}(\tau) = \mathbf{u}_1^R q_\Delta(\tau)$  where  $q_H(\tau), q_\Delta(\tau)$  are scalar functions of  $\tau$  determined by the nonlinear self-consistency condition. Therefore, neurons in all groups have the same autocorrelation function with different amplitudes. The ratio of amplitudes is determined by the components  $u_{1c}^R$  of the leading right eigenvector of  $\mathbf{M}$  (see Fig. 2a,b) as  $\Delta_c(\tau)/\Delta_d(\tau) = u_{1c}^R/u_{1d}^R$ . This ratio is independent of  $\tau$  and the firing rate nonlinearity. The latter affects only the overall amount of activity in the network but not the ratio of activity between the subgroups.

We illustrate how these results give insight into a perplexing question in computational neuroscience - how can a small number of neurons have a large effect on the representational capacity of the whole network? In adults, newborn neurons continuously migrate into the existing neural circuit in the hippocampus and olfactory bulb regions [27]. Impaired neurogenesis results in strong deficits in learning and memory. This is surprising since the young neurons, although hyperexcitable, constitute only a very small fraction ( $< 0.1$ ) of the total network. To better understand the role young neurons may play, we analyzed a network with  $D = 2$  groups of neurons: group 1 of young neurons that is significantly smaller than group 2 of mature neurons ( $\alpha_1 \ll \alpha_2$ ). The connectivity within the existing neural circuit is such that by itself that subnetwork would be in the quiescent state:  $g_{22} = 1 - \varepsilon < 1$ . To model the increased excitability of the young neurons all connections of these neurons were set to:  $g_{12} = g_{21} = g_{11} = \gamma > 1 - \varepsilon$ .

We analyzed the network's capacity to reproduce a target output pattern  $f(t)$ . The activity of the neurons serves as a "reservoir" of waveforms from which  $f(t)$  is composed. The learning algorithm in [28] allows us to find the vector  $\mathbf{w}$  such that  $z(t) = \sum_{i=1}^N w_i \phi_i(t) = f(t)$ , where the modified dynamics have  $J_{ij} \rightarrow J_{ij} + u_i w_j$  and  $\mathbf{u}$  is a random vector with  $O(1)$  entries. For simplicity we choose periodic target functions  $f(t) = \sin(\Omega t)$ , and define the learning index as the fraction of power that the output function  $z(t)$  has at the target frequency. The index varies from 0 to 1, and is computed by averaging over 50 cycles.

Performance depends primarily on  $\Lambda_1$  and not on the network structure, peaking for  $\sqrt{\Lambda_1} \approx 1.5$  (Fig. 3). This is directly related to the maximal learning capacity observed at  $g \approx 1.5$  in networks with a single cell-type [28], further supporting the identification of  $\sqrt{\Lambda_1}$  as the effective gain. Importantly, because of the block structured connectivity, the effective gain is larger than the average gain ( $\sqrt{\Lambda_1} > \bar{g}$ ), for all values of  $\gamma$  and  $\alpha_1$  [22]. In other words, for the same average connection strength, networks with block-structured connectivity have a higher effective gain that can place them in a regime with larger learning capacity compared to networks with shuffled connections, demonstrating that a small group of neurons could place the entire network in a state conducive to learning. Moreover, since increases in average connection strength are generally associated with increased metabolic cost, networks with block-structured connectivity can provide a more metabolically efficient way to perform computation compared to statistically homogeneous networks.

Outgoing connections from any given neuron are typically all positive or all negative, obeying Dale's law [29]. Within random networks, this issue was addressed by Rajan and Abbott [26] and Tao [30] who computed the bulk spectrum and the outliers of a model where columns of  $\mathbf{J}$  are separated to two groups, each with its offset and element variance. The dynamics of networks with cell-type-dependent connectivity that is offset to respect Dale's law were addressed in [31] with some limitations, and remain an important problem for future research.

Ultimately, neural network dynamics need to be considered in relation to external inputs. The response properties of networks with  $D = 1$  have been recently worked out [19, 32]. The analogy between the mean field equations suggests that our results can be used to understand the non-autonomous behavior of multiple cell-type networks.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

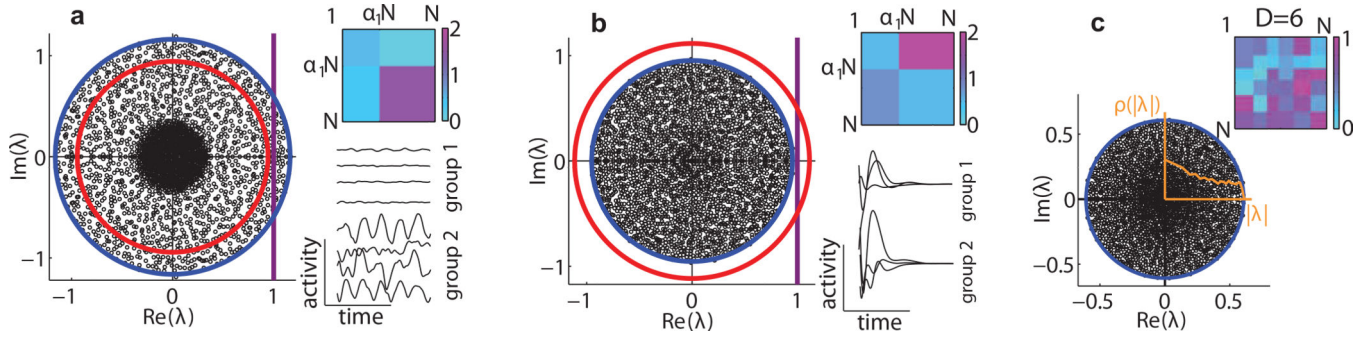
## ACKNOWLEDGMENTS

The authors would like to thank Larry Abbott for his support, including comments on the manuscript, and Ken Miller for many useful discussions. This work was supported by grants R01EY019493 and P30 EY019005, NSF Career award (IIS 1254123). MS was supported by the Gatsby Foundation.

## References

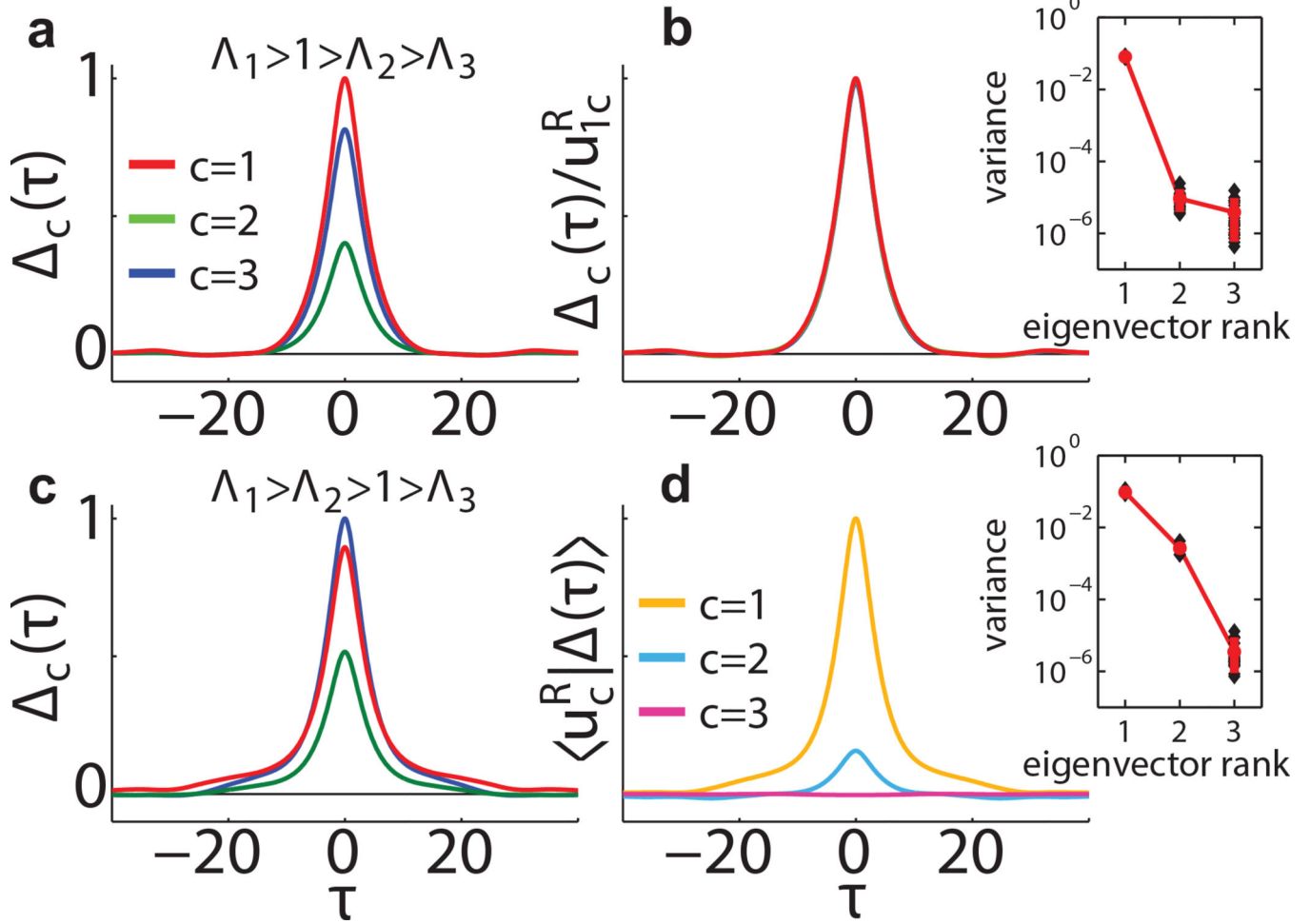
1. Mehta, ML. Random Matrices. Elsevier Academic Press; 2004.
2. Izyumov YA. Adv. Phys. 1965; 14:569.
3. Kosterlitz JM, Thouless DJ, Jones RC. Phys. Rev. Lett. 1976; 36:1217.
4. Wishart J. Biometrika. 1928; 20A:32.
5. Amari S-I. Systems, Man and Cybernetics, IEEE Transactions on. 1972:643.
6. Sommers HJ, Crisanti A, Sompolinsky H, Stein Y. Phys. Rev. Lett. 1988; 60:1895. [PubMed: 10038170]
7. Sompolinsky H, Crisanti A, Sommers HJ. Phys. Rev. Lett. 1988; 61:259. [PubMed: 10039285]
8. Molgedey L, Schuchhardt J, Schuster HG. Phys. Rev. Lett. 1992; 69:3717. [PubMed: 10046895]
9. Kauffman S. Nature. 1969; 224:177. [PubMed: 5343519]

10. Shmulevich I, Kauffman SA, Aldana M. Proceedings of the National Academy of Sciences. 2005; 102:13439.
11. Pomerance A, Ott E, Givan M, Losert W. PNAS. 2009; 106:8309.
12. Schubert D, Kitter R, Zilles K, Luhmann HJ, Staiger JF. J. Neurosci. 2003; 23:2961. [PubMed: 12684483]
13. Yoshimura Y, Dantzker JL, Callaway EM. Nature. 2005; 433:868. [PubMed: 15729343]
14. Yoshimura Y, Callaway EM. Nature Neurosci. 2005; 8:1552. [PubMed: 16222228]
15. Suzuki N, Bekkers JM. J. Neurosci. 2012; 32:919. [PubMed: 22262890]
16. Franks KM, Russo MJ, Sosulski DL, Mulligan AA, Siegelbaum SA, Axel R. Neuron. 2011; 72:49. [PubMed: 21982368]
17. Levy RB, Reyes AD. J. Neurosci. 2012; 32:5609. [PubMed: 22514322]
18. Bertschinger N, Natschläger T. Neural Comp. 2004; 16:1413.
19. Toyozumi T, Abbott LF. Phys. Rev. E. 2011; 84:051908.
20. Girko V. Theory Probab. Appl. 1984; 29:694.
21. Bai Z. The Annals of Probability. 1997; 25:494.
22. see Supplemental Material [url], which includes Refs. [33–34] and (i) calculation of the largest Lyapunov exponent, (ii) extension to matrices with non-Gaussian entries, and (iii) comparison of effective gain of networks with multiple cell-types and that of a network with cell-type independent connectivity..
23. Sompolinsky H, Zippelius A. Phys Rev. B. 1982; 25:6860.
24. Note that  $\mathbf{M}$  has strictly positive elements, so by the Perron-Frobenius theorem its largest eigenvalue (in absolute value) is real and positive and the corresponding eigenvector has strictly positive components.
25. Aljadeff J, Renfrew D, Stern M. arXiv preprint arXiv:1411.2688. 2014
26. Rajan K, Abbott LF. Phys. Rev. Lett. 2006; 97:188104. [PubMed: 17155583]
27. Zhao C, Deng W, Gage FH. Cell. 2008; 132:645. [PubMed: 18295581]
28. Sussillo D, Abbott LF. Neuron. 2009; 63:544. [PubMed: 19709635]
29. Eccles J. Notes and records of the Royal Society of London. 1976; 30:219. [PubMed: 12152632]
30. Tao T. Prob. Theory and Related Fields. 2013; 155:231.
31. Cabana T, Touboul J. J. of Stat. Phys. 2013; 153:211.
32. Rajan K, Abbott LF, Sompolinsky H. Phys. Rev. E. 2010; 82:011903.
33. Tao T, Vu V, Krishnapur M. Ann. Probab. 2010; 38:2023.
34. Wolf A, Swift JB, Swinney HL, Vastano JA. Physica D. 1985; 16:285.

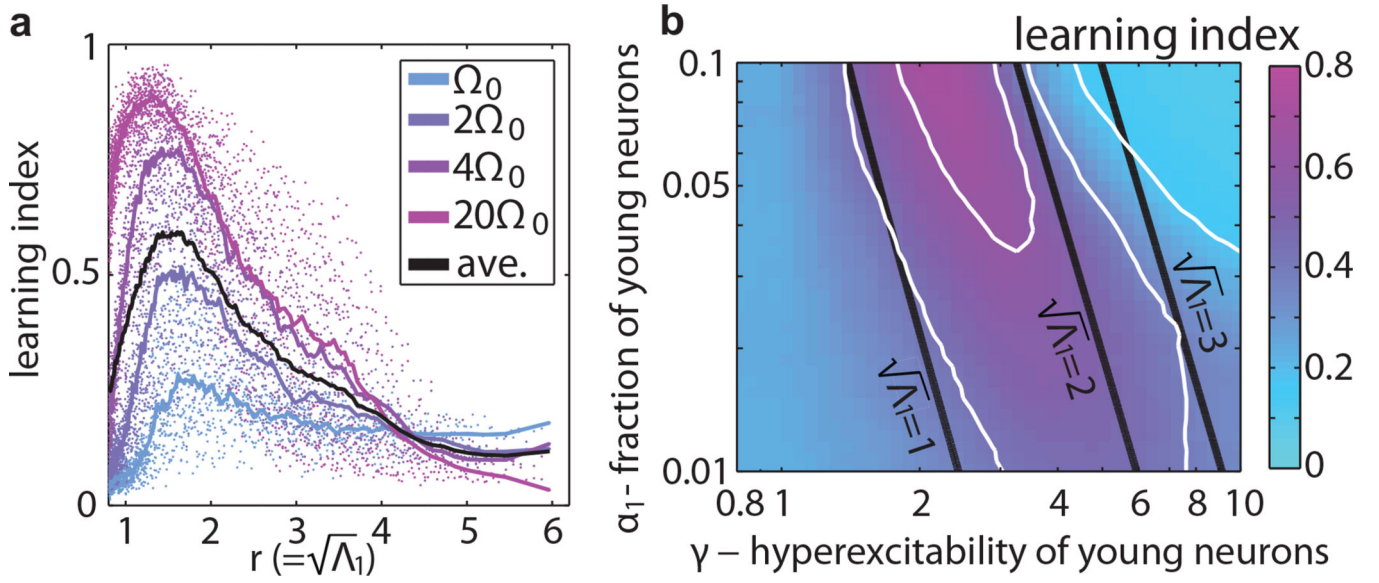
**FIG. 1.**

Spectra and dynamics of networks with cell-type dependent connectivity ( $N = 2500$ ). The support of the spectrum of the connectivity matrix  $\mathbf{J}$  is accurately described by  $\sqrt{\Lambda_1}$  (radius of blue circle) for different networks. Top insets - the synaptic gain matrix  $\mathbf{G}$  summarizes the connectivity structure. Bottom insets - activity of representative neurons from each type. The line  $\Re\{\lambda\} = 1$  (purple) marks the transition from quiescent to chaotic activity. (a) An example chaotic network with two cell-types. The average synaptic gain  $\bar{g}$  (radius of red circle) incorrectly predicts this network to be quiescent. (b) An example silent network. Here  $\bar{g}$  incorrectly predicts this network to be chaotic. (c) An example network with six cell-types. In all examples the radial part of the eigenvalue distribution  $\rho(|\lambda|)$  (orange line) is not uniform [22].



**FIG. 2.**

Autocorrelation modes. Example networks ( $N = 1200$ ) have 3 equally sized groups with  $\alpha, \mathbf{g}$  such that  $\mathbf{M}$  is symmetric. (a) When  $D^* = 1$  autocorrelations maintain a constant ratio independent of  $\tau$ . (b) Rescaling by the components  $u_{1c}^R$  collapses the autocorrelation functions (Here  $\Lambda_1 = 20, \Lambda_2 = 0.2, \Lambda_3 = 0.1$ ). (c) When  $D^* = 2$ , the autocorrelation functions are linear combinations of two autocorrelation “modes” that decay on different timescales. Projections of these functions  $\langle u_c^R | \Delta(\tau) \rangle$  are shown in (d). Only projections on  $|u_1^R\rangle, |u_2^R\rangle$  are significantly different from 0 (Here  $\Lambda_1 = 20, \Lambda_2 = 16, \Lambda_3 = 0.1$ ). Insets show the variance of  $\langle \tau \rangle$  projected on  $|u_c^R\rangle$  averaged over 20 networks in each setting.



**FIG. 3.**

Learning capacity is primarily determined by  $\sqrt{\Lambda_1}$ , the effective gain of the network. (a) The learning index for four pure frequency target functions ( $\Omega_0 = \pi/120$ ) plotted as a function of the radius  $r = \sqrt{\Lambda_1}(\alpha_1, \gamma)$ . The training epoch lasted approximately 100 periods of the target signal. Each point is an average over 25 networks with  $N = 500$ ,  $\varepsilon = 0.2$  and different values of  $\alpha_1$  and  $\gamma$ . The line is a moving average of these points for each frequency. (b) The same data averaged over the target frequencies shown as a function of  $\gamma$  and  $\alpha_1$ . Contour lines of  $l_\Omega$  (white) and of  $\sqrt{\Lambda_1}$  (black) coincide approximately in the region where  $l_\Omega$  peaks.