The spectrum of covariance matrices of randomly connected recurrent neuronal networks

Yu Hu[1] and Haim Sompolinsky[2,3]

[1]Department of Mathematics and Division of Life Science, The Hong Kong University of Science and Technology, Hong Kong SAR, China
[2]Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel
[3]Center for Brain Science, Harvard University, Cambridge, Massachusetts, United States of America

Abstract

A key question in theoretical neuroscience is the relation between connectivity structure and collective dynamics of a network of neurons. Here we study the connectivity-dynamics relation as reflected in the distribution of eigenvalues of the covariance matrix of the dynamic fluctuations of the neuronal activities, which is closely related to the network's Principal Component Analysis (PCA) and the associated effective dimensionality. We consider the spontaneous fluctuations around a steady state in randomly connected recurrent network of spiking neurons. We derive an exact analytical expression for the covariance eigenvalue distribution in the large network limit. We show analytically that the distribution has a finitely supported smooth bulk spectrum, and exhibits an approximate power law tail for coupling matrices near the critical edge. Effects of adding connectivity motifs and extensions to EI networks are also discussed. Our results suggest that the covariance spectrum is a robust feature of population dynamics in recurrent neural circuits and provide a theoretical predictions for this spectrum in simple connectivity models that can be compared with experimental data.

1 Introduction

Collective dynamics in networked systems are of great interest with numerous applications in many fields including neuroscience, spin glasses, and ecological networks. Many studies have focused on how certain statistics of dynamics depend on the connectivity structure of the network, such as the population average [22] and variance [25, 10] of pairwise correlations. Although powerful and directly comparable with experimental data, these are *local* features of dynamics meaning that it can be estimated just from the local measurements of the activity of involved neurons. In contrast, certain *global* features of dynamics rely on simultaneously recorded activity of a population of neurons and cannot be inferred from just local descriptions of dynamics.

One important example of such global aspects of population dynamics is the eigenvalues of the covariance matrix which are complicated nonlinear functions of all matrix elements. These eigenvalues arise naturally when performing the widely used principle component analysis (PCA) of population activity, where they correspond to the amount of variance contained in each activity PC. Another example that received substantial recent interest [28, 37, 41, 35] is the dimensionality which can be defined based on the moments of the covariance eigenvalues. Many recent experimental studies have observed a low dimensional dynamics of neurons in the

brain [36, 13], and theoretical investigations have illustrated the importance of having a low or high dimensionality on brain function and computation [12], such as encoding information [3, 28].

As the experimental techniques of making joint observations of the activity large population of neurons in networks become increasingly available, it is crucial to study how the connectivity structure of the network affects these joint aspects of population dynamics to take full advantage of such data.

In this work, we study the eigenvalue distribution (i.e. spectrum) of the covariance matrix of spontaneous activity in a large recurrent network of spiking neurons with random connectivity. We study several basic and widely used models of random connectivity, including independent and identical Gaussian distributed connectivity [39] (Section 2.1), networks with connectivity motifs [40, 31, 47, 22] (Section 2.3), and Excitation-Inhibition (EI) network (Section 2.4). Random connectivity has been a fundamental model in theoretical studies of neuronal network dynamics[28, 39, 24]. They can be motivated as minimal models to capture the highly complex, disordered connections observed in many neuronal circuits such as in cortex. Some aspects of these covariance spectra might be distinct from those under ordered, deterministic connectivity (Section 3.1).

The dynamics we consider here is simple where the co-fluctuations between neurons are described by linearizing around the steady state [27, 44]. Despite the simple dynamics and minimal connectivity model, we found the resulting spectrum has a continuous bulk of non-trivial shape and shows interesting features such as a power law long tail of large eigenvalues (Section 2.2), and strong effect due to the non-normality of connectivity matrix (Section 2.3.3). These covariance spectra highlight non-trivial population-level structures of neuronal co-fluctuations shaped by recurrent interactions that was previously unexplored.

The covariance eigenvalue distribution we focused on conveys rich and robust information of the joint population dynamics. In particular, it allows a derivation of closed form expressions for the effective dimension (previously known for the simple random iid Gaussian connectivity [10]) Furthermore, we show that the continuous bulk spectrum has the advantage over low order statistics such as the dimension measure in that it is robust to outlier eigenvalues (Sections 2.3 and 2.4 and Fig. 7).

Our analytically derived eigenvalue distributions can be readily compared to real data of population activity of recurrent neural circuits, or to simulations of more sophisticated computational models. We provide ready-to-use code to facilitates such application (Section 4.8).

1.1 Neuronal networks with random recurrent connectivity

We consider a recurrent network of generalized integrate-and-fire neurons [44] and under the common dynamic regime and assumptions used in the literature [44, 27]: we assume the network is at a steady state of asynchronous firing, that is, the neurons fires irregularly without coherent timings across the population. Under the asynchronous firing regime, linear response theory can be used to approximate closely the fluctuation of neuron spiking [27, 44].

$$\Delta y_i(t) \approx \Delta y_i^0(t) + A_i(t) * \left(\sum_{j=1}^{N} W_{ij} F_{ij}(t) * \Delta y_j(t) \right), \quad i = 1, 2, \ldots, N. \tag{1}$$

Here $a(t) * b(t) = (a * b)(t) = \int_0^t a(s)b(t-s)ds$ denotes convolution. $y_i(t) = \sum_k \delta(t - t_{i,k})$ and $\Delta y_i(t) = y_i(t) - r_i$ is the mean subtracted spike train. $y_i^0(t)$ is the "unperturbed" spike train in absence of recurrent connections W_{ij} (as a matrix of connection weights $W = \{W_{ij}\}$). $A_i(t)$ is the linear response kernel for neuron i typically having a rise-and-decay shape. $F_{ij}(t)$ is the temporal kernel of the synapse. For simplicity, we assume the neurons and synapses have identical dynamics, that is, $A_i(t) = A(t)$, $F_{ij}(t) = F(t)$ and $y_i^0(t)$ are independent and identically distributed for all i. Further details of the model can be found in [22] and do not affect our results in this work as explained below.

We will focus on the structure of long time scale co-fluctuations in the network, which are described by the *long time window* covariance $C_{ij} = \lim_{\Delta T\to\infty} \frac{1}{\Delta T} C_{ij,\Delta T}$, where $C_{ij,\Delta T}$ is the covariance of the spike counts summed over a window of length ΔT: $C_{ij,\Delta T} = \langle \Delta s_i(t) \Delta s_j(t)\rangle$, $\Delta s_i(t) = \int_t^{t+\Delta T} \Delta y_i(t')dt'$. For spiking neurons, $C_{ij,T}$ typically settles to its limiting value when $\Delta T > 50$ms [5]. As shown in [44], the covariance matrix $C = (C_{ij})$ can be derived from Eq. (1) using Fourier transform and has a particularly simple form,

$$C = \sigma^2 (I - aW)^{-1} (I - aW)^{-T} = \sigma^2 (I - J)^{-1} (I - J)^{-T}. \tag{2}$$

Here I is the identity matrix, and A^{-1}, A^T are matrix inverse and transpose ($A^{-T} = (A^{-1})^T$). The scalar $a = \left(\int_0^\infty A(t)dt\right)\left(\int_0^\infty F(t)dt\right)$ summarizes the cellular and synaptic dynamics. We define $J = aW$ for notation simplicity and refer to J as connectivity matrix. $\sigma^2 = \lim_{\Delta T\to\infty} \frac{1}{\Delta T}\langle(\Delta s_i^0(t))^2\rangle$, $\Delta s_i^0(t) = \int_t^{t+\Delta T} \Delta y_i^0(t')dt'$, is the long time window variance of the spike trains of an isolated neuron (assumed same for all neurons). The covariance matrix C can also be easily calculated from simultaneously recorded neuron activity data (Method). We consider generalizations beyond the long time window covariance in Sec. 2.5.

Our analysis and results starts from the covariance-connectivity relation Eq. (2), which also appears in other scenarios including alternative spiking neuron networks (Hawkes model [19, 32]), fixed point activity averaged over whitened inputs, OU process [14, 23], and structural equation modeling in statistics [1] (see details in Section 4.2).

For many biological neural networks, such as cortical local circuits, the recurrent connectivity is complex and disordered. Random connectivity is a widely used minimal model to gain theoretical insights on the dynamics of neuronal networks [39, 45]. We first and mainly consider a random connectivity where $J_{ij} \sim \mathcal{N}(0, g^2/N)$ are drawn as independent and identically distributed (iid) Gaussian variables with zero mean and variance g^2/N (referred subsequently as the *iid Gaussian connectivity*). We will later show how to generalize to other types of random connectivity, including Erdős-Rényi random connectivity, networks with excitation and inhibition , obeying Dale's law (Section 2.4), and with connectivity motifs (Section 2.3). The theory we derived assumes the network is large and is exact as $N \to \infty$, and we verify their applicability to finite size networks numerically.

1.2 PCA and dimensionality

Principle component analysis (PCA) is a widely used analysis of population dynamics, where the activity is decomposed along orthogonal patterns or modes. The modes can be calculated as the eigenvectors of the covariance matrix C (Eq. (33)) and the associated eigenvalues λ_i are positive and shows the amount of activity or variance distributed along the modes. In this work, we focus on the distribution of these covariance eigenvalues, described by the (empirical) probability density function (pdf) $p_C(x)$ which is defined as $\int_a^b p_C(x)(x)dx = \frac{1}{N}\#\{\lambda_i \in (a, b]\}$. We also refer $p_C(x)$ as the spectrum (do not confuse with the frequency spectrum in Fourier transform). We will derive the limit of $p_C(x)$ as $N \to \infty$ and study how it depends on parameters of the connectivity model such as connection variance g.

The shape of $p_C(x)$ can provide important theoretical insights on interpreting PCA. For example, it can be used to separate outlying eigenvalues corresponding to low dimensional externally driven signals from small eigenvalues corresponding to fluctuations amplified by recurrent connectivity interactions [15] (Fig. 7). $p_C(x)$ is also closely related to the (participation ratio) *dimension* [28, 34]:

$$D = \frac{(\sum_{i=1}^N \lambda_i)^2}{\sum_{i=1}^N \lambda_i^2}. \tag{3}$$

To see D is a dimensionality measure, consider when there are d non-zero λ_i with same value and the rest eigenvalues are zero. Importantly, D can be calculated from the first two moments $p_C(x)$. We will also derive formulas of D in the random connectivity models.

2 Results

2.1 Continuous bulk spectrum with finite support

For networks with iid Gaussian connectivity (Section 1.1), there is one parameter g describing the overall connection strength. For stability of the fixed point and the validity of the linear response theory around it, g is required to be less than 1 [39]. The parameter σ in Eq. (2) just scales all λ_i and thus is hereafter set to 1 for simplicity. The main result of this paper is the derivation of the probability density function of the covariance eigenvalues in the large N limit. In (Supplementary Materials) we show that the eigenvalue pdf is

$$p_C(x) = \frac{3^{\frac{1}{6}}}{2\pi g^2 x^2}\left[\sum_{\xi=1,-1}\xi\left((1+\frac{g^2}{2})x - \frac{1}{9} + \xi\sqrt{\frac{(1-g^2)^3 x(x_+ - x)(x - x_-)}{3}}\right)^{\frac{1}{3}}\right], x_- \le x \le x_+. \tag{4}$$

where

$$x_\pm = \frac{2 + 5g^2 - \frac{g^4}{4} \pm \frac{1}{4}g(8+g^2)^{\frac{3}{2}}}{2(1-g^2)^3}, \tag{5}$$

and $p_C(x) = 0$ for $x > x_+$ and $x < x_-$. The distribution has a smooth, unimodal shape and is skewed towards the left (Fig. 1C). Near both support edges, the density scales as $|x - x_\pm|^{\frac{1}{2}}$ (Supplementary Materials). The derivation of the formula is based on techniques similar to those used previously in deriving the eigenvalue distribution of J [4, 38] However, we emphasize that our result on the covariance spectrum does not have a simple relation to the well-known spectrum of the connectivity J because J is a non-normal matrix (i.e. $J^T J \neq J J^T$). This point is further elaborated in Section 2.3.3. Although the theory is derived in the large N limit, for networks of sizes of several hundreds, our results still match well with numerical simulations (Fig. 1AB). Interestingly, the derived pdf is different from the well known Marchenko–Pastur law ([29], Eq. (28) in Methods) for the pdf of the covariance matrix of sampled iid Gaussian noise, highlighting the crucial role of the nontrivial correlations generated by the recurrent dynamics.

2.2 Long tail of large eigenvalues near the critical coupling

As g approaches the critical value of 1, the upper limit of the support x_+ diverges as $(1-g^2)^{-3}$ (Section 4.3 in Methods). This means there is an activity mode with diverging variance and is consistent with the stability requirement of $g < 1$. Note that lower edge x_- is always bounded away from 0, and has a limit of $\frac{4}{27}$ as $g \to 1$. Analyzing the shape of $p_C(x)$ for large x in the critical regime $g \to 1$ indicates a long tail of large eigenvalues, with a power law shape (Fig. 2AB)

$$p_C(x) \approx \frac{\sqrt{3}}{2\pi} x^{-\frac{5}{3}}. \tag{6}$$

Moreover, the power law approximation does not require g be very close to 1 and is valid as long as $x_- \ll x \ll x_+$ is far away from the support edges (Fig. 2C, see details in Methods and Supplementary Materials).

Because the probability density is small at the power law long tail, large eigenvalues can appear to be sparsely located (Fig. 1A) and potentially mistaken for statistical outliers. This underscores the importance of knowing the exact distribution and support edges for interpreting PCA results of population activity, which we revisit later at (Fig. 7).

Note that such a long tail is a distinct feature of correlations arising from the recurrent network dynamics. For example, for the Marchenko–Pastur law that is often used for modeling empirical covariance spectra , the upper edge of the tail relative to the mean is bounded by 4 (Methods), whereas for $p_C(x)$ (Eq. (4)) the edge relative to the mean can be arbitrarily large as $O((1-g^2)^{-2})$ (see below for the evaluation of the mean). This highlights the difference between covariance generated by finite samples of noise and correlations generated by the recurrent dynamics.

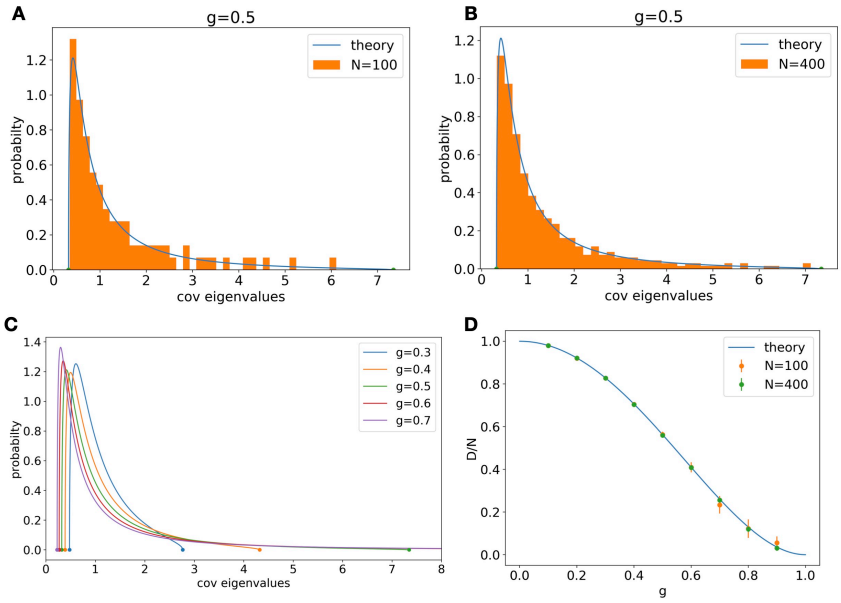

Figure 1: **Covariance spectrum under random Gaussian connectivity. A**. Compare theory (Eq. (4)) with finite size simulation using Eq. (2) at $N = 100$, $g = 0.5$. Histogram shows eigenvalues from one network realization. **B**. Same as A at $N = 400$. **C**. Covariance eigenvalue distribution at various value of g. As g increases the distribution develops a long tail of large eigenvalues. **D**. Dimension (normalized by network size) vs g. The dots and error bars are mean and sd over repeated trials from finite size simulations (Eq. (2) and use Eq. (3)). Note some error bars are smaller than the dots

The long tail of the eigenvalue distribution is also reflected by a low dimension (Eq. (3)). We show that the mean and second moment of the eigenvalues λ for large N are

$$\mathbf{E}\lambda = (1-g^2)^{-1}, \quad \mathbf{E}\lambda^2 = (1-g^2)^{-4}. \tag{7}$$

From Eq. (7), the dimension is

$$D = N(1-g^2)^2. \tag{8}$$

In particular, the *relative dimension* with respect to the network size $\hat{D} = D/N$ vanishes as g approaches 1 (Fig. 1D). In comparison, $\hat{D}$ for the Marchenko–Pastur law (Eq. (28)) is at least $\frac{1}{2}$.

While these low order moments can be derived from previous methods (see e.g., [10] and Supplementary Materials), our method allows for the first time the derivation of the full pdf, from which higher order moments can also be derived, for example,

$$\mathbf{E}\lambda^3 = (1-g^2)^{-7}(1+2g^2), \quad \mathbf{E}\lambda^4 = (1-g^2)^{-10}(1+g^2)(1+5g^2). \tag{9}$$

2.3 Impact of connectivity motifs

We next consider generalizations of random connectivity beyond the iid Gaussian model (Section 1.1). One important feature of biological neural networks is having motif structures [40, 31], which means the frequencies of certain subgraphs or motifs are significantly different from that in an edge-shuffled network (i.e. an iid random graph with matching connection probability). We can introduce *second order motifs* (Fig. 3A), i.e. those composed of two edges, to the Gaussian random connectivity using an additive model (Eq. (29) in Methods, [21]) which introduce

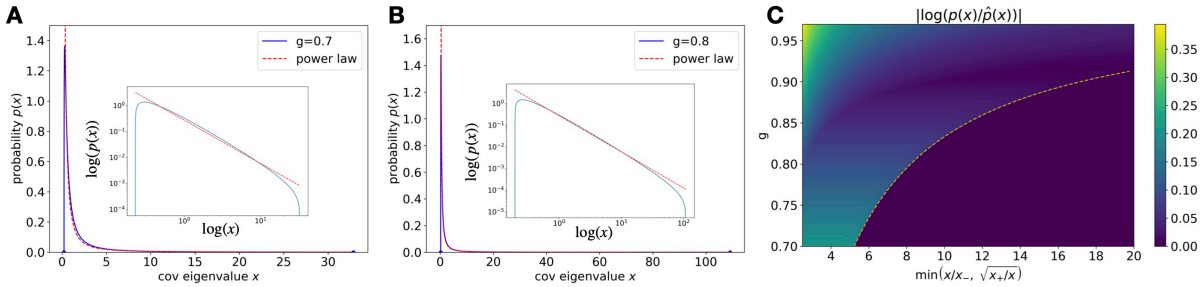

Figure 2: **Approximate power law tail. A**. The exact pdf (solid line) of covariance spectrum compared with the power law approximation (dashed line, Eq. (6)) at $g = 0.7$. Inset shows in the log-log scale. bf B. Same as A for $g = 0.8$. The approximation improves as g approaches the critical value 1. **C**. The log error between the exact pdf and approximation $|\log(p(x)) - \log(\hat{p(x)})|$ as a function of g and "distance" from the support edges. We quantify this "distance" as the minimum ratio of x/x_- and $\sqrt{x_+/x}$ (see Supplementary Materialsfor explanation of the square root). The plot shows the log error is small when this ratio is large which means x being far away from the edges. The dashed line shows the attainable region of the ratio which increases with g.

the following correlations between the edge variables

$$\hat{\kappa}_{div} = \rho(J_{ik}, J_{jk}), \;\; \hat{\kappa}_{con} = \rho(J_{ki}, J_{kj}), \;\; \hat{\kappa}_{ch} = \rho(J_{ik}, J_{kj}), \;\; \hat{\kappa}_{re} = \rho(J_{ij}, J_{ji}), \;\; 1 \leq i \neq j \neq k \leq N. \tag{10}$$

Here $\hat{\kappa}_*$ measures the strength of diverging, converging, chain, and reciprocal motifs respectively (see Methods). The previous iid model of J (no motifs) corresponds to all $\hat{\kappa}_*$ being 0. Under the new Gaussian random connectivity (Eq. (29) in Methods), the diverging, converging and chain motifs are generated by adding random correlated low rank components to a full rank matrix $\tilde{J}$, where $\tilde{J}$ is a Gaussian matrix with only reciprocal correlations. We have shown that a low rank perturbation to $\tilde{J}$ will not affect the continuous bulk spectrum of the covariance in the large network limit, but introduce at most a finite number of outliers (Supplementary Materials, Fig. 3BC). Therefore, for the continuous bulk covariance spectrum, we can focus on the study the case of a Gaussian coupling matrix with only reciprocal motifs. Note that the motif measures defined in Eq. (10) are correlations that are normalized by $\mathrm{var}(J_{ij})$. When focusing on the $\tilde{J}$ component, we need to account for the change from $\mathrm{var}(J_{ij})$ to $\mathrm{var}(\tilde{J}_{ij})$ by using

$$\hat{\kappa}_{re}(\tilde{J}) = \frac{\hat{\kappa}_{re}(J) - 2\hat{\kappa}_{ch}(J)}{1 - \hat{\kappa}_{div}(J) - \hat{\kappa}_{con}(J)}. \tag{11}$$

So all the other types of second order motifs do not directly affect the bulk spectrum, but channel their impact via the effective $\hat{\kappa}_{re}$. We will from here on focus on the reciprocal only case to study the shape of the bulk spectrum and write $\kappa = \hat{\kappa}_{re}$ for notation simplicity.

2.3.1 Symmetric and anti-symmetric random networks

First, we look at two special, extreme cases for the reciprocal motifs (as a correlation $-1 \leq \kappa \leq 1$): J being symmetric ($J_{ij} = J_{ji}$, $\kappa = 1$) or anti-symmetric matrix ($J_{ij} = -J_{ji}$, $\kappa = -1$). These cases are much simpler, because J is a *normal matrix* so $p_C(x)$ can be derived from the well known eigenvalue distribution of J ([38]). For symmetric random connectivity,

$$p_{C,g,\kappa=1}(x) = \frac{\sqrt{(4g^2-1)x - 1 + 2\sqrt{x}}}{4\pi g^2 x^2}, \;\; x \in (x_-, x_+), \; x_\pm = (1 \mp 2g)^{-2}. \tag{12}$$

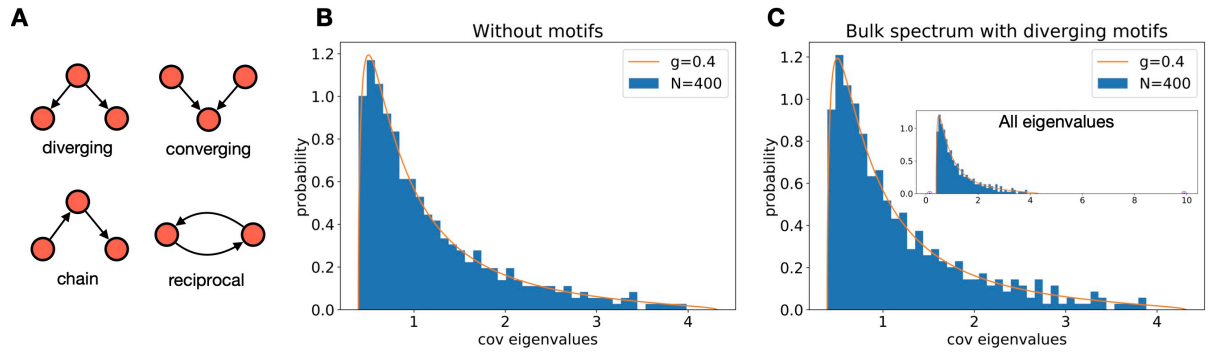

Figure 3: **Connectivity motifs and invariance of the bulk spectrum**. **A** Diagrams of all second order motifs (composed of two edges). **B** The histogram of covariance eigenvalues of a network with iid Gaussian connectivity J ($N = 400$, $g = 0.4$). **C** Now add diverging motifs to (the same) J as in Eq. (29) with b_j such that $\hat{\kappa}_{div} = 0.25$. The bulk histogram of eigenvalues has little change and remains well described by the theory Eq. (4) except for two outlying eigenvalues (inset, highlighted by semicircles). As shown in Supplementary Materials converging and chain motifs similarly do not *directly* affect (see Eq. (11)) the bulk spectrum in large networks.

Here stability requires that $g < \frac{1}{2}$. For anti-symmetric random connectivity,

$$p_{C,g,\kappa=-1}(x) = \frac{\sqrt{(4g^2+1)x-1}}{2\pi g^2 x^2 \sqrt{1-x}}, \quad x \in ((1+4g^2)^{-1}, 1). \tag{13}$$

Here the network is stable for all g. The derivations are given in the Supplementary Materials.

From the above equations, we see that $p_C(x)$ of the symmetric random network has a power law tail analogous to Eq. (6) as $g \to 1/2$ for large x, but with a different coefficient,

$$p_{C,g,\kappa=1}(x) \approx \frac{\sqrt{2}}{\pi} x^{-\frac{7}{4}}. \tag{14}$$

The $p_C(x)$ of the anti-symmetric random network does not have a long tail as the upper limit of the support is always 1.

2.3.2 General reciprocal motifs

Now, for the Gaussian random connectivity with reciprocal motifs $\kappa = \rho(J_{ij}, J_{ji})$, $-1 < \kappa < 1$, we derived an equation for $p_{C,g,\kappa}(x)$ in the large N limit based on results in [38] similarly as the iid case (Supplementary Materials). Although a closed form expression is possible using the root formula for quartic equations, it seems quite involved and is thus not pursued but left implicit to the code that we use to numerically calculated $p_{C,g,\kappa}(x)$. For a fixed g, as κ decreases, the distribution narrows in from both sides (Fig. 4B). Intuitively, these effects are mainly due to the change of stability constrained critical $g_c = (1+\kappa)^{-1}$ with κ (based on the spectrum of J [38]). So the same g becomes *relatively* weaker comparing to g_c as *kappa* decreases. This motivates us to compare the distributions $p_{C,g,\kappa}(x)$ with the same *relative coupling strength* $g_r = g/g_c = g(1+\kappa)$, which is also the maximum real part of J's eigenvalues [38]. As shown in Fig. 4C, the effects of κ are now smaller comparing to Fig. 4B since the main effect on relative coupling is controlled, and the changes to the shape is the opposite: the distribution narrows as κ increases. As strong connections, we also observe a highly skewed long tail of large eigenvalues. Interestingly, we show that (Supplementary Materials) for $-1 < \kappa_{re} < 1$ reciprocal motifs, the critical covariance spectrum has an approximate power law tail with the same exponent as the

iid random case (Eq. (6))

$$p_C(x) \approx \frac{\sqrt{3}}{2\pi}(1-\kappa)^{\frac{1}{3}}(1+\kappa)x^{-\frac{5}{3}}, \quad \text{as } x \to \infty,\ g \to g_c = (1+\kappa)^{-1} \tag{15}$$

These shape changes of $p_{C,g,\kappa}(x)$ with reciprocal motifs are also reflected by the dimension measure, for which we derived a closed form expression (Supplementary Materials)

$$D = N\frac{\mu_1(2g^2\mu_1+1) - 2\theta\mu_1(\theta\mu_1+1)}{(\theta\mu_1+1)^2(g^2\mu_1+1)}, \quad \mu_1 = \frac{2\theta - 1 + \sqrt{1+4(g^2-\theta)}}{2(g^2-\theta^2)}, \ \theta = g^2(1+\kappa). \tag{16}$$

Here μ_1 is the mean of the distribution. Comparing with Eq. (8), this shows the non-trivial dependence of dimension on the reciprocal motif strength κ. Consistent with the shape changes, the dimension increases with reciprocal motifs (Fig. 4D).

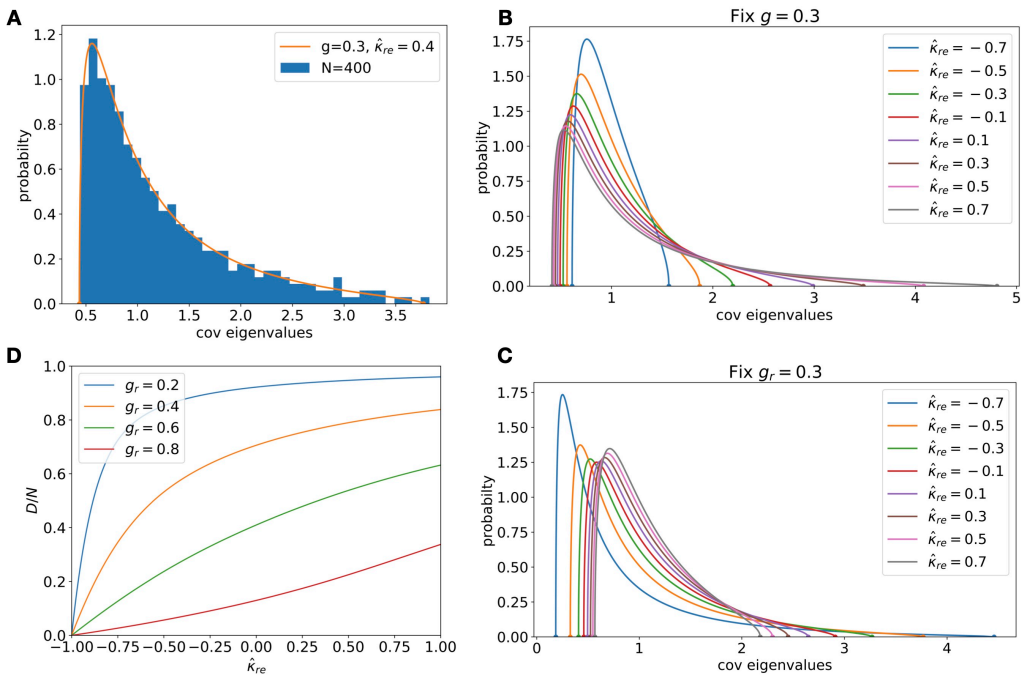

Figure 4: **Impact of reciprocal motifs**. **A**. Compare theoretical covariance spectrum for random connectivity with reciprocal motifs and simulation ($g = 0.4$, $\hat{\kappa}_{re} = 0.4$, $N = 400$). **B**. The spectra at various $\hat{\kappa}_{re}$ while fixing $g = 0.3$. **C** Same as B but fixing relative $g_r = g/g_c$ to control the main effect (see text). The changes in shape are now smaller and the support narrows with increasing $\hat{\kappa}_{re}$. **C**. The impact of reciprocal motifs on dimension for various g_r (Eq. (16)). For small g_r, the dimension increases sharply with $\hat{\kappa}_{re}$.

2.3.3 Strong non-normal effects at $\kappa_{re} < 0$

Here we revisit the implications of the fact that the random J is a non-normal matrix. In particular, we consider the shape of $p_{C,g,\kappa}(x)$ in networks with negative reciprocal motifs, $\kappa < 0$. One may naively expect that as κ decreases towards -1, the shape of $p_{C,g,\kappa}(x)$ will become similar to that of the anti-symmetric network $p_{C,g,\kappa=-1}(x)$ (Eq. (13)). Note that $p_{C,g,\kappa=-1}(x)$ has a diverging density at the upper edge of 1 (Supplementary Materials), and can also be bimodal for g in certain range (i.e. has another peak in addition to the divergence at 1, Fig. 5A).

One may thus expect $p_{C,g,\kappa}(x)$ inherit some of these features and is bimodal when κ is close to -1.

However, we find that for all $\kappa > -1$ and $g < g_c$, there is no divergence in $p_{C,g,\kappa}(x)$, and the pdf is unimodal for all but a minuscule combinations of κ and g where κ is very close to -1 (Fig. 5B-E, the main figure in panel E is a highly zoomed-in view in order to show the bimodal region which is hard to see within the whole attainable parameter region (the inset of panel E)).

This intriguing observation is due the non-normal connectivity J. To show this, we compare $p_{C,g,\kappa}(x)$ to the covariance eigenvalue distribution $p^n_{C,g,\kappa}(x)$ from a network with a *normal* connectivity matrix J^n that has a matching eigenvalue distribution to J. Similarly we consider how $p^n_{C,g,\kappa}(x)$ changes as $\kappa < 0$ approaches -1. Note that at $\kappa = -1$, J itself is normal and $p_{C,g,\kappa=-1}(x) = p^n_{C,g,\kappa=-1}(x)$. Interestingly, the normal matching $p^n_{C,g,\kappa}(x)$ shows similar features as the $\kappa = -1$ (Fig. 5F): the diverging density of $p_{C,g,\kappa=-1}(x)$ at the upper edge corresponds to a non-differentiable peak and the distribution is bimodal, both phenomena occurs in a large region of (g, κ) (Fig. 10 in Supplementary Materials). The absence of such features in the actual $p_C(x)$ thus highlights the strong effect on the covariance eigenvalue distribution due to connectivity being non-normal.

2.4 Erdos-Renyi connectivity and E-I networks

The Gaussian random connectivity we have considered so far has a non-zero connection weight for all pairs of neurons with probability 1. A more biologically plausible alternative is a network with random sparse connections (Erdős-Rényi (ER) random connectivity) where each pair of neurons is independently chosen to be connected with a weight $w_0 > 0$ at probability $0 < p < 1$ (and unconnected with probability $1 - p$). Here we compare the dynamics of these two types of networks (dense and sparse) following the framework of [24] but in terms of the covariance spectrum.

For the all excitatory ER network, we found that the covariance matrix spectrum from the iid Gaussian connectivity (Eq. (4)) also describes the bulk spectrum of an ER network once substituting $g = aw_0\sqrt{Np(1-p)}$ to match the variance of connections $\mathrm{var}(J_{ij})$ (similar to a replacement in [24]). The result can be understood by viewing J as $(J - w_0 p\mathbf{1}) + w_0 p\mathbf{1}$, where $\mathbf{1}$ is a matrix of all 1 entries. The first matrix has *zero mean* iid entries. In light of universality results in random matrix theory [43], we hypothesize it to have the same spectrum as the matching iid Gaussian matrix as $N \to \infty$ and the expected degree $pN \to \infty$, which is verified numerically (Fig. 6A). The second part of J is a rank-1 matrix. As mentioned in the case with motifs, adding a low rank term to the connectivity does not affect the continuous bulk spectrum (Supplementary Materials), therefore J has a same bulk covariance spectrum described by the iid Gaussian theory (see Fig. 11 in Supplementary Materials). However, for all excitatory ER network the low rank term does affect stability [24] and creates a large outlying eigenvalue (Fig. 11A in Supplementary Materials). This means either w_0 or p has to be small to ensure stability. Such a limitation can be overcome by adding inhibition [24] in subsequently discussed networks.

Adding inhibition Having inhibitory connections (negative connection weights) is an important feature in neural circuits. In the iid Gaussian connectivity, this is implicitly modeled as the entries having zero mean on average showing the cancellation of inhibition and excitation (positive connection weights) [33, 24]. In the ER network, we can add a *global inhibition*, where the summed output of all neurons is fed back to each neuron with a strength of $-w_I < 0$. This is equivalent to have a connectivity $J - w_I\mathbf{1}$. Since this is again a rank-1 perturbation, this ER network with inhibition has the same bulk spectrum described by Eq. (4). Moreover, the added inhibition cancels the overall excitation and allows for larger w_0 and effective g while keeping the network stable [24]. In particular, for $w_I = w_0 p$ this is the same zero mean matrix consider above (Fig. 6A). Similarly, stability can be attained at stronger connections in other connectivity models with inhibition (according to the general conditions of [24] simplified at linear dynamics) such as (i) all-inhibitory ER network ($w_0 < 0$, Supplementary Materials) and

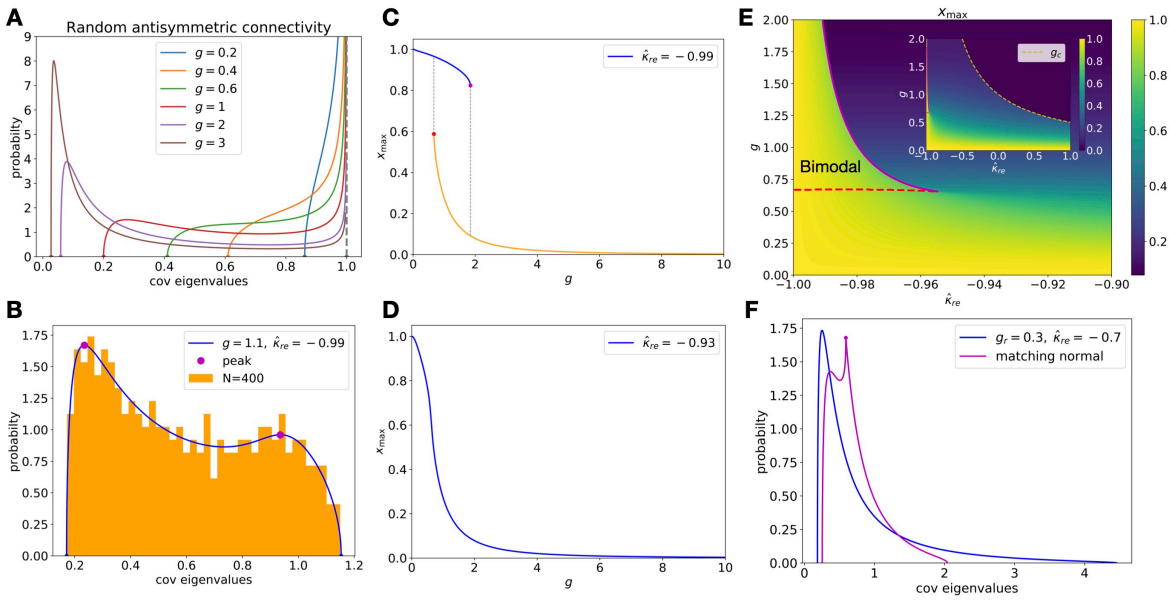

Figure 5: **Strong non-normal effect at $\hat{\kappa}_{re} < 0$.** **A**. The covariance spectrum for anti-symmetric random connectivity ($\hat{\kappa}_{re} = -1$, Eq. (13)). The density diverges as the right edge $x = 1$ (dashed line). For larger g, the distribution becomes bimodal (counting the divergence as a peak). **B**. An example bimodal distribution at $\hat{\kappa}_{re} > -1$ (-0.99) compared with $N = 400$ simulation. **C**. Location of the peak(s) in the covariance spectrum as a function of g at $\hat{\kappa}_{re} = -0.99$. As g increases, a left peak emerges at $g = \frac{2}{3}$, and the right peak disappears after $g \approx 1.85$. For a range of g between the bifurcations, the distribution is bimodal. **D**. Same as C but for $\hat{\kappa}_{re} = -0.93$. There is no bifurcations of peaks and the distribution is always unimodal. **E**. The location of the peak (right most if multiple) as a function of g and $\hat{\kappa}_{re}$. The red dashed line and magenta solid lines label the location of the peak bifurcations as in C (same colored dots) and the triangular region enclosed is when the covariance spectrum is bimodal. Note the $\hat{\kappa}_{re}$-axis is highly enlarged and the bimodal region is minuscule within the whole parameter space only for $\hat{\kappa}_{re} \lesssim -0.95$ (inset, the yellow dashed line shows the stability bound g_c). **F**. A typical unimodal spectrum (blue) for $\hat{\kappa}_{re} < 0$ compared with the covariance spectrum predicted from the spectrum of connectivity J assuming it being normal (magenta). The normal prediction is very different and bears similarity to the anti-symmetric case (see A) being bimodal with a non-differential point (dot).

(ii) *mixed E-I network* where the entries are independent and identically distributed as

$$J_{ij} = \begin{cases} w_e, & \text{with prob. } p_e \\ -w_i, & \text{with prob. } p_i \\ 0, & \text{with prob. } 1 - p_e - p_i \end{cases} \tag{17}$$

where $w_e, w_i > 0$ and satisfies the balance $w_e p_e - w_i p_i = 0$ to ensure zero-mean. By similar hypothesis of universality describe above, the covariance spectrum of J is again described by Eq. (4) with an effective g matching $\mathrm{var}(J_{ij})$, which agrees well with the numerics (Fig. 6B).

Dale's law E-I network Biological neuronal networks largely obey the so called Dale's law, where each neuron can send out only excitatory connections, or only inhibitory connections depending the cell type. This requires a two population network with E and I neurons, where columns of J is constrained in sign. Here we consider a simple case [33, 24]: $N/2$ E neurons sending connections each with strength $w_e > 0$ to other neurons at probability p, and $N/2$ I

neurons sending $-w_e < 0$ connections with probability p. The overall strength is balanced in this network and $var(J_{ij})$ is the same for all entries. Again, we found that the bulk spectrum is well described by the iid Gaussian theory (Fig. 6C) with an effective $g = we\sqrt{Np(1-p)}$ [24]. This can be generalized to unequal fractions of E and I neurons as long as the variance of connections is kept equal (Fig. 6D). For general EI networks with unconstrained connection weights and probability for E, I neurons, the covariance spectrum differs from the iid Gaussian case, analogously as their connectivity spectra differ [33], and its characterization is left for future work.

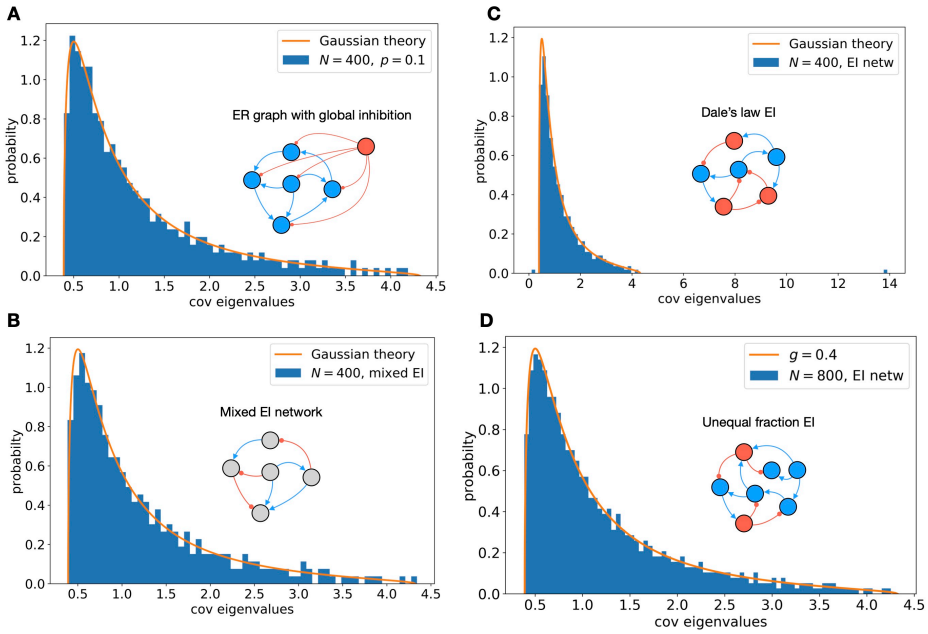

Figure 6: **Erdős-Rényi and EI networks**. **A**. Simulation based on Eq. (2) of an ER network with global inhibition to balance the mean connection to 0 (see text) compared with the Gaussian random connectivity theory (Eq. (4)). $N = 400$, connection probability $p = 1$, the effective $g = 0.4$ (see text). **B**. Similar as A for an Mixed EI network with randomly assigned E or I edges (Eq. (17)). $p_e = 0.025$, $p_i = 0.075$, $N = 400$, the effective $g = 0.4$ (see text). **C**. An EI network obeying the Dale's law with equal number of E and I neurons and balanced connection strength (see text). $N = 400$, connection probability $p = 0.1$, effective $g = 0.4$. Except for two outliers, the covariance eigenvalues are well described by a Gaussian random connectivity theory. **D**. The bulk spectrum (excluding two outlying eigenvalues similar as in C) for an Dale's law EI network with 70% E and 30% I neurons ($N = 800$). The connection probabilities and weights are chosen to ensure the variance of E and I connection strengths are equal: $p_e = 0.028$, $p_i = 0.15$, $w_e = 0.0864$, $w_i = 0.0396$. The effective $g = 0.4$.

2.5 Frequency dependent covariance

We have focused on the long time window covariance matrix. This would be especially suitable to neural activity recordings that have limited temporal resolution such as calcium imaging [46]. Temporal structures of correlation beyond the slow time scale can be described by the frequency covariance matrix (or coherence matrix)

$$C_{ij}(\omega) = \lim_{T\to\infty} \langle \Delta y_i(\omega) \Delta y_j^*(\omega) \rangle, \tag{18}$$

where $\Delta y_i(\omega) = \frac{1}{\sqrt{T}} \int_0^T \Delta e^{-\mathrm{i}\omega} \Delta y_i(t) dt$ is the Fourier transform of the spike train and z^* is the complex conjugate. Importantly, the long time window covariance corresponds to $C(\omega = 0)$ (Wiener-Khinchin theorem, Supplementary Materials). Non-zero frequency covariance matrix can be calculated from neural activity data either using the formula above or by Fourier transform of the time-lagged cross-correlation functions. Importantly, there is a simple C-J relation analogous to Eq. (2) for $C(\omega)$ (Supplementary Materials).

It turns out for our main model of iid Gaussian random connectivity, the spectrum of $C(\omega)$ is simply obtained replace g with an effective

$$g(\omega) = A(\omega)F(\omega)\sqrt{N\mathrm{var}(W_{ij})} \tag{19}$$

which takes into account frequency dependency of neuronal ($A(\omega)$, Section 1.1) and synaptic ($F(\omega)$) dynamics. Note, however, such a replacement by the effective g may not apply to other cases, for example networks with non-zero reciprocal motifs and there can be qualitative changes with frequency (Supplementary Materials).

2.6 Fitting the theoretical spectrum to data

Our theory on the bulk covariance spectrum can be fitted to neural activity whenever the covariance eigenvalues can be calculated. In many settings, the value of the unperturbed neuronal variance σ^2 is not known. But this can be easily addressed by scaling both the observed eigenvalues and theory to have mean equal to 1. After fitting the connectivity parameter g for the normalized eigenvalues, σ^2 can then be easily estimated using the original mean. For our theoretical spectra, the mean μ of covariance eigenvalues is available in closed form (Eqs. (7) and (16)), and the scaled pdf is easily found as $p_R(x) = \mu p_C(\mu x)$.

Furthermore, the recorded neural activity is sometimes normalized for each neuron (e.g. by converting to z-scores), in which case the eigenvalues of a *correlation matrix* instead of a covariance matrix needs to be analyzed. Interestingly, we found that for our Gaussian random connectivity models, the correlation eigenvalue distribution in the large N limit is the same as the rescaled $p_R(x)$ above. This is because the diagonal entries of C becomes uniform (thus converge to μ) for large N (Supplementary Materials).

The best fitting theoretical spectrum can be found by minimizing the L^2 or L^∞ error between the empirical and theoretical cumulative distributions (motivated by methods in hypothesis testing, see Methods) with respect to connectivity parameters such as g. We note that the availability of closed form or analytic solutions of the theoretical distributions makes this optimization highly efficient.

We conclude by demonstrating fit to data with a simple example while the full application to neural data will be explored elsewhere. Consider a covariance matrix that is a sum of two components: one from the iid Gaussian random recurrent network (Eq. (2)) and another low rank part $\sigma_1^2 u_1 u_1^T + \sigma_2^2 u_2 u_2^T$ representing contributions from e.g., external inputs from two other brain areas, each with magnitude σ_i^2 and projection weights to neurons in the recurrent network described by column vector u_i. Here u_1 and u_2 are chosen as random vectors with unit 2-norm and σ_1 and σ_2 are set to be sufficiently large such that the low rank part corresponds two outlying eigenvalues [7]. We then fit the theoretical spectrum (Eq. (4)) to *all* eigenvalues of the total covariance matrix. The inferred g is highly accurate despite the presence of (unknown to the algorithm) outliers, and can in fact often correctly separate them out (Fig. 7).

3 Discussion

In this work, we studied the eigenvalue distribution (i.e. spectrum) of the covariance matrix of a randomly connected neuronal network during spontaneous fluctuations around a steady state. We derived the distribution in the large network limit analytically in terms of the connectivity parameters such as coupling strength and motif statistics. Our results also include closed form

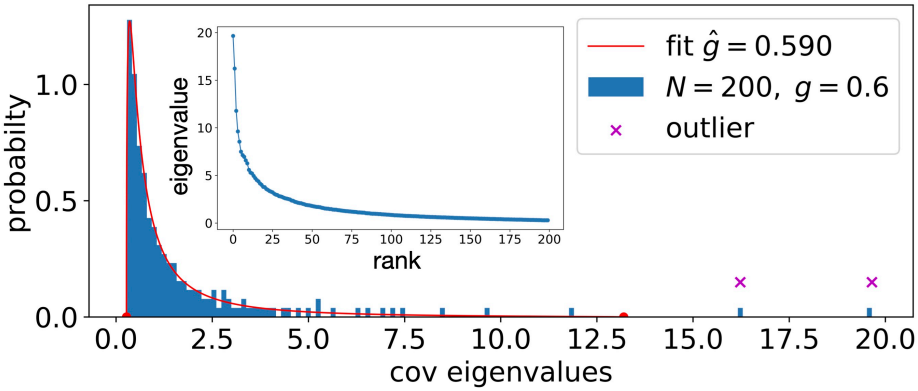

Figure 7: **Separating outlying eigenvalues by fitting to the theoretical spectrum**. A rank 2 component is added to a covariance matrix of a network with random connectivity (Eq. (2), $N = 200$, $g = 0.6$, $\sigma_1^2 = 17$, $\sigma_2^2 = 15$, $\sigma^2 = 1$, see text). We then use *all* eigenvalues to find the best fit to theory Eq. (4) (see text). The eigenvalues outside the theoretical support are then considered outliers (magenta cross). We then re-fit using the non-outlier eigenvalues and the process is repeated a few times till convergence. The fitting successfully identifies the two outlying eigenvalues in 86% of 100 trials and accurately infers g with a root mean square error of 0.01.

expressions for the dimension measure generalizing known results [10]. These results are to the best of our knowledge, novel. The new spectrum exhibits characteristic long tails of large eigenvalues at strong connections that are approximated by a power law. Knowing the exact shape and support of the bulk eigenvalue distribution can facilitate separating out outlying eigenvalues corresponding to low dimensional structure (coming from other un-modeled effect such as external input).

The shape of the bulk spectrum reflects structured amplification of the neuronal noise by the random recurrent interactions and is robust to low rank components in the connectivity or from external inputs (Supplementary Materials). The interpretation bulk spectrum corresponds to smaller eigenvalues in the PCA analysis of neural activity data, their meaning and relation to circuit connectivity, Unlike the large eigenvalues [30], the interpretation of the bulk spectrum of PCA of neural activity data has received little attention. A notable exception is a recent work [8] which studied the power law of covariance spectrum of data near criticality based on renormalization group analysis. Our theory thus provides an important benchmark to compare with experimental data and advocates the bulk covariance spectrum as a powerful joint description of collective dynamics in neuronal networks.

One limitation of The main limitation of the work is the assumed dynamic regime where fluctuations of the neuronal activity are described by linear response theory [27, 44] around a fixed point. Extensions and comparisons to highly nonlinear activity such as chaotic dynamics [39] is left for future research. Future work could also consider more general network architectures such as multiple populations of EI networks [24] and incorporating distant dependent connectivity patterns based on known cortical microcircuit architectures [6, 9, 26].

3.1 Ordered vs. disorder connectivity

We have studied the covariance spectrum under random connectivity which are used to model for complex recurrent networks. Here we ask whether certain features of these spectra are distinct results of the connectivity being random. To address this question we briefly discuss here the covariance spectra from a few widely known examples of ordered connectivity for comparison.

The simplest nonrandom connectivity model is the uniform all-to-all connected network where $J_{ij} = g/N$ for all i, j. The covariance matrix has one eigenvalue being $1/(1-g^2)$ and $N-1$ eigenvalues being 1. Thus, for large networks, the spectrum consists a delta function at 1 ('bulk') and a single outlier. A similar situation arises when we consider a ring network [6], with long-range connections, where J_{ij} follows a smooth shape depending on the distance of the pair of neurons on the ring (Fig. 8, see Methods). In the large network limit, the covariance spectrum has delta distribution at 1 with a number of discretely located larger eigenvalues.

We next consider the ring network with short-range connections, in particular, nearest-neighbor connections. The covariance spectrum now has a continuous bulk on an interval, but the density diverges at both edges as $(\Delta x)^{-\frac{1}{2}}$ (Fig. 8B).

To seek further if we can find an ordered connectivity with a qualitatively similar covariance bulk spectrum as the random connectivity, we consider the multi-dimensional generalizations of the nearest-neighbor ring network, i.e., a torus (Methods). As the dimension of the torus increases, the smoothness within and at the edges of the support increases, and the covariance spectrum becomes qualitatively similar to the random case [16] (Fig. 8F). This indicates that the overall smooth density and long tail shape of the covariance spectrum is a shared property in highly connected networks with high rank connectivity matrices, including random networks and high dimensional short-range spatially invariant networks.

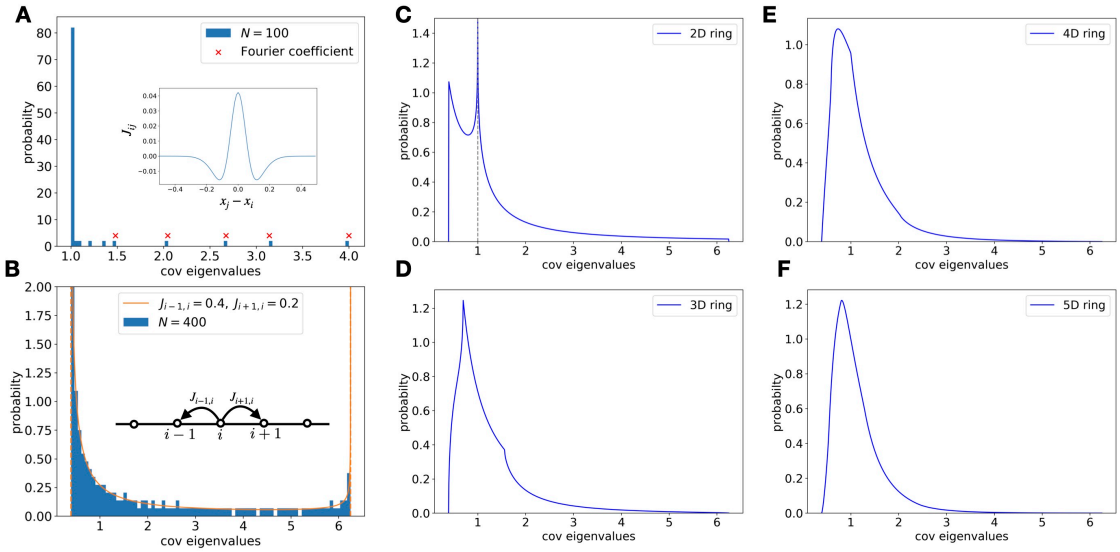

Figure 8: Covariance spectra under some deterministic connectivity models. **A**. The covariance eigenvalues of a ring network with a long-range connection profile (inset, $N = 100$). Most eigenvalues are close to 1 and the rest eigenvalues appear at the locations predicted by top Fourier coefficients of the connection profile (cross, see Methods). These eigenvalues will thus remain at these discrete locations even for large N. **B**. A ring network with nearest-neighbor connections: $J_{i-1} = 0.4$, $J_{i+1,i} = 0.2$. The solid line are theoretical spectrum in large N limit which has two diverging singularities at both support endpoints. The effect of such singularities is also evident in the finite size network at $N = 400$. **C-F**. Higher dimensional generalizations of the nearest-neighbor ring network ($ad = 0.6$, see Methods). As the dimension increases, the singularity in the pdf becomes milder and less evident, and the overall shape becomes qualitatively similar to the random connectivity case (Fig. 1).

4 Methods

Notation	Description
N	number of neurons
C	covariance matrix Eq. (2)
$p_C(x)$	pdf of eigenvalues of C Section 1.2
$x_\pm$	support edges of $p_C(x)$
W	matrix of connection weights
J	connectivity matrix $J = aW$ Section 1.1
σ^2	unperturbed neuronal variance Section 1.1
g	variance of connections $\mathrm{var}(J_{ij})$
g_c	maximum g constrained by stability
g_r	g/g_c
$\hat{\kappa}_{re}$	normalized reciprocal motif cumulant, κ, Eq. (10)
μ	mean of eigenvalues $\frac{1}{N}\sum_{i=1}^{N}\lambda_i$
D	dimension Eq. (3)

4.1 Models of random connectivity

Here is a summary of results on various random connectivity models.

- iid Gaussian random connectivity $J_{ij} \sim \mathcal{N}(0, g^2/N)$: closed form pdf and endpoints Eq. (4), including frequency dependent covariance spectrum Section 2.5, and power law tail approximation Eq. (6).

- Gaussian random connectivity with second order motifs Eq. (10): analytic solution and endpoints (quartic root, Supplementary Materials) and power law tail approximation (Eq. (15)). For special case of symmetric and ant-symmetric connectivity, closed form pdf Eqs. (12) and (13), including frequency dependent covariance spectrum (Supplementary Materials).

- Erdős-Rényi and certain EI network Section 2.4: same bulk spectrum as the iid Gaussian case.

For all cases, the mean μ and the dimension D are derived in closed form (Eqs. (7), (8) and (16).

For simplicity, we do not require J_{ii} to be zero (i.e. no self-coupling), but allow it, for example in iid Gaussian model, to be distributed in the same way as other entries J_{ij}. In large network limit, since individual connections are weak (e.g. $O(1/\sqrt{N})$), allowing this self-coupling or setting $J_{ii} = 0$ does not affect the covariance spectrum (Supplementary Materials).

4.2 Applications to alternative neuronal models and signal covariance

Although the relation between C and J (Eq. (2)) is derived based on the linear response theory of integrate-and-fire neurons, it also arises in other networked systems.

Linearly interacting Poisson neurons This is also called a multivariate Hawkes model [19]. This is a simple model for spiking neuron networks, but is versatile to capture for example the temporal spiking correlations seen in other more sophisticated nonlinear spiking neuron

networks [32, 18]. A time dependent Poisson firing rate is calculated as a filtered input spike trains $s_j(t)$ (sum of delta functions), and spikes are then drawn as a Poisson process given $y_i(t)$,

$$y_i(t) = y_0 + \int_0^\infty A(t-\tau)\left(\sum_j W_{ij} s_j(t-\tau)\right) d\tau. \tag{20}$$

Here we consider a homogeneous network where the baseline firing rate y_0 and response filter $A(t)$ is the same for all neurons.

The exact long time window spike count covariance matrix of this network can be shown to be [19]

$$C = (I - aW)^{-1} C^0 (I - aW)^{-T}, \quad C^0 = \mathrm{diag}\{Y_1, Y_2, , \ldots, Y_N\}, \quad Y = (I - aW)^{-1} Y^0, \tag{21}$$

which is valid if the time varying $y_i(t)$ do not often becomes negative (for example when any negative connections W_{ij} are small compare to y_0). Here $a = \int_0^\infty A(t)dt$, Y^0 and Y are vectors of baseline and perturbed (with recurrent connections) firing rates of the neurons respectively. If we assume that the effective connection strength aW is weak so that we can approximate Y with Y^0, then (21) becomes

$$C = y_0 (I - aW)^{-1} (I - aW)^{-T},$$

the same as Eq. (2) (note that for Poisson process $\mathrm{var}(\int_t^{t+\Delta t} s_i(u)du) = \int_t^{t+\Delta t} y_i(u)du$).

Another condition that ensures a uniform Y and does not restrict weak connection is a *fine-scale balance* condition of W sometimes assumed in EI networks [**?**],

$$\sum_{j=1}^N J_{i_1 j} = \sum_{j=1}^N J_{i_2 j}. \tag{22}$$

This is not unreasonable to assume, for example, considering the homeostatic mechanisms of neurons.

Ornstein-Uhlenbeck process The multivariate Ornstein-Uhlenbeck process is a classic minimal model for stochastic dynamical systems [14],

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1}^N J_{ij} x_j(t) + \xi_i(t), \quad i = 1, \ldots, N. \tag{23}$$

Here $\xi_i(t)$ is iid Gaussian white noise with variance σ^2. It can be shown that the long time window (or zero-frequency, see Sec.2.5) covariance of the system is given by Eq. (2).

Fixed point over whitened input The covariance we considered so far describes the structure of fluctuations of spontaneous dynamics without or under fixed external input, often referred as *noise covariance* [2]. We can also consider a *signal covariance* in a network of firing rate neurons

$$\tau \frac{dr_i}{dt} = -r_i + f\left(\sum_{j=1}^N J_{ij} r_j + u_i\right), \quad i = 1, 2, \ldots, N. \tag{24}$$

Here u_i is the external input to neuron i. Assume the network settles to a steady state where all neurons have firing rate $r_i > 0$, and approximate the nonlinearity as $f(x) \approx x$, then the fixed point firing rates are

$$\vec{r} = (I - J)^{-1} \vec{u}. \tag{25}$$

Now consider the network activity across an ensemble of input patterns, which has a whitened statistics [11],

$$\mathrm{var}(u_i) = \sigma^2, \quad \mathrm{cov}(u_i, u_j) = 0. \tag{26}$$

It is easy to see that the covariance of firing rates $\vec{r}$ is given by $\sigma^2(I-J)^{-1}(I-J)^{-T}$ that is the same as Eq.(2).

We note that Eq. (25) or equivalently $\vec{r} = J\vec{r} + \vec{u}$ appears in broader contexts beyond neuroscience and is studied in the field of linear structural equation modeling (SEM) [1].

4.3 Power law approximation of the eigenvalue distribution

First note the limits of the support edges. As $g \to 1^-$, $(1-g^2)^3 x_+ \to \frac{27}{4}$. For the lower edge, $x_- \to \frac{27}{4}$ can be found by Taylor expansion in $(1-g^2)$ or note that $(1-g^2)^3 x_+ x_- = 1$. Consider a x that is far away from the support edges as $g \to 1$, given the above, this means,

$$x \to \infty, \quad x(1-g^2)^3 \to 0. \tag{27}$$

Note that since $x_+/x_- \sim (1-g^2)^{-3}$, there is plenty range of x to satisfies the above for strong connections when g is close to 1. Under these limits, Eq. (4) greatly simplifies as various terms vanish leading to (Supplementary Materials)

$$\lim_{g\to 1^-,\, x_- \ll x \ll x_+} p_C(x) / \left(\frac{\sqrt{3}}{2\pi} x^{-\frac{5}{3}} \right) \to 1.$$

This explains the validity of power law approximation away from support edges. If we are only interested in the leading order power law tail in the critical distribution (i.e. $g \to 1^-$ and *then* $x \to \infty$), there is a simpler alternative derivation that we also use for other connectivity models (see Supplementary Materials).

4.4 Comparison with the Marchenko–Pastur distribution

The Marchenko–Pasturdistribution is widely used for modeling covariance eigenvalues arising from noise [29, 15]. It is a family with one shape parameter α. We focus on the case when the covariance is positive definite which restrict $0 < \alpha < 1$ (otherwise there is a delta distribution at 0) and the pdf is

$$p_{MP}(x) = \frac{\sqrt{(\alpha_+ - x)(x - \alpha_-)}}{2\pi\alpha x}, \quad \alpha_\pm = (1 \pm \sqrt{\alpha})^2, \tag{28}$$

The first two moments are 1 and $1+\alpha$, from which we know the dimension is $1/(1+\alpha)$ has a lower limit 1/2. The upper edge α_+ is bounded by 4.

4.5 Gaussian random connectivity with second order motifs

Following [21], we can construct a connectivity matrix with jointly Gaussian distributed entries as

$$J_{ij} = a_i + b_j + \tilde{J}_{ij}, \;\; 1 \le i, j \le N, \tag{29}$$

where a_i, b_j, $\tilde{J}_{ij}$ are zero mean Gaussian variables that have the same variance across i, j

$$\mathrm{var}(a_i) = a^2, \;\; \mathrm{var}(b_j) = b^2, \;\; \mathrm{var}(\tilde{J}_{ij}) = c^2, \;\; 1 \le i, j \le N, \tag{30}$$

and are independent except for

$$\mathrm{cov}(a_i, b_i) = k, \;\; \mathrm{cov}(\tilde{J}_{ij}, \tilde{J}_{ji}) = r, \;\; i \ne j. \tag{31}$$

Such a structure and correlation introduces the second order motifs to J,

$$\begin{aligned} &\kappa_{div} = \mathrm{cov}(J_{ik}, J_{jk}) = b^2, \;\; \kappa_{con} = \mathrm{cov}(J_{ki}, J_{kj}) = a^2, \;\; \kappa_{ch} = \mathrm{cov}(J_{ik}, J_{kj}) = k, \\ &\kappa_{re} = \mathrm{cov}(J_{ij}, J_{ji}) = 2k + r, \quad i \ne j \ne k \end{aligned} \tag{32}$$

The relation between these covariances and the motifs can be seen in network where J_{ij} take binary values of w_0 or 0 [47, 22]. In large networks $\frac{N^3}{w_0^2}\mathbf{E}J_{ik}J_{jk}$ is the number of diverging motifs. Note that $\mathbf{E}J_{ij}/w_0$ is the connection probability and $\frac{N^3}{w_0^2}\mathbf{E}J_{ik}\mathbf{E}J_{jk} = Np^2$ is expected number of diverging motifs in a matching Erdős-Rényi (ER) random graph. Therefore κ_{div} (called motif cumulant in [23]) is proportional to the deviations the diverging motif frequency from the (iid) ER graph. κ_* can be normalized to correlation $\hat{\kappa}_*$ as in Eq. (10) by dividing $\mathrm{var}(J_{ij}) = a^2 + b^2 + c^2$. So we can achieve various second order motif cumulants by adjusting $\{a^2, b^2, c^2, k, r\}$.

From Eq. (29) it is clear that J can be viewed as $\tilde{J}$ perturbed by a rank-2 matrix $ae^T + eb^T$, $a = (a_i)$, $b = (b_i)$, $e = (1, 1, \ldots, 1)^T$. According to Lemma 7.1 in Supplementary Materials, the continuous spectrum of the covariance based on J is the same as that for $\tilde{J}$ for large N. Note, however, the normalized value of $\hat{\kappa}_{re}(X) = \frac{\kappa_{re}(X)}{\mathrm{var}(X_{ij})}$ is different from J and $\hat{J}$ and needs to be mapped using Eq. (11).

4.6 Deterministic connectivity

4.6.1 Ring network with short and long range connections

In a ring network, neurons are equally spaced on a circle (can be physical or functional space) and neuron i is associated with a location $x_i = i/N$, $i = 0, \ldots, N-1$. The connection between two neurons j and i only depends on the location difference $x_i - x_j$ thus is translation invariant.

For long-range connections, the connectivity has a shape determined by a fixed smooth periodic function $f(x)$ on $[0, 1)$,

$$J_{ij} = \frac{1}{N} f(x_i - x_j) = \frac{1}{N} f\left(\frac{i-j}{N}\right).$$

In the large network limit, the eigenvalues of the covariance has an approximate delta distribution at 1 except for a finite number of discretely located larger eigenvalues (Fig. 8A). A precise statement of this result is described in Supplementary Materials. The outlying eigenvalues corresponds to the top Fourier coefficients of $f(x)$.

For the nearest-neighbor (NN) connectivity, only $J_{i-1,i}$ and $J_{i+1,i}$ are non-zero and remain fixed as $N \to \infty$.

4.6.2 Multi-dimensional ring network

For d-dimensional ring, the neurons are equally spaced on a d-dimensional lattice

$$x_{\vec{i}} = (i_1/N, i_2/N, \ldots, i_d/N),$$

which is periodic along each coordinate. We focus on the NN connectivity where each neuron is connected to $2d$ neighboring neurons with strength $J^k_{i_k-1,i_k}$ and $J^k_{i_k+1,i_k}$ along direction k. We show that the probability density function at both support edges scales as $(\Delta x)^{\frac{d}{2}-1}$ (for comparison, the random network edges scale as $(\Delta x)^{\frac{1}{2}}$). This means for dimension $d \geq 2$, there is no singularity at the support edges (Fig. 8).

To characterize the shape of the covariance spectrum (Fig. 8C-F), we further simplify by setting $J^k_{i_k-1,i_k} = J^k_{i_k+1,i_k} = a$ (see also Supplementary Materialsfor motivations based on 1D ring) and analytically derived $p_C(x)$ (Supplementary Materials). For small dimensions $d \leq 3$, there are distinct "inflection points" within the support. As d increases, this non-smooth feature becomes less evident and becomes hard to tell in empirical eigenvalue distributions from a finite size network (not shown).

4.7 Fitting the theoretical spectrum to data

For neural activity data, C can be calculated from large number of time samples of binned spike count $s_i(t)$ (assuming bin size is ΔT large),

$$C_{ij} = \frac{1}{\Delta T}\frac{1}{M-1}\sum_{t=1}^{M-1}(s_i(t)-\bar{s}_i)(s_j(t)-\bar{s}_j), \quad \bar{s}_i = \frac{1}{M}\sum_{t=1}^{M} s_i(t). \tag{33}$$

For calcium imaging data, the fluorescence is approximately integrating the spikes over the indicator time constant. So we can still apply Eq. (33) by plugging in the fluorescence signal in place of $s_i(t)$ to calculate the covariance C (omit the constant factor ΔT which does not affect fitting to theory, Section 2.6).

We fit the theoretical spectrum to empirical eigenvalues by finding the connectivity parameter g that minimize the error between the *cumulative distribution functions* (cdf) $F(x) = \int_{-\infty}^{x} p(x)dx$. This avoids issues such as binning when estimating the probability density function from empirical eigenvalues. We numerically integrate theoretical pdf (Eq. (4)) to get its pdf. As seen below, the theoretical cdf only needs to be calculated at the empirical eigenvalues.

The cdf error can be measured under the L^2 norm using the Cramer-von Mises statistic

$$D^2_{CvM} = \int (F(x) - F_n(x))^2 dF_n(x) = \frac{1}{12n^2} + \frac{1}{n}\sum_{i=1}^{n}\left(F(x) - \frac{2i-1}{2n}\right)^2$$

Here n is the number of samples. Alternatively, the error can also be measure under L^∞ norm based on the Kolmogorov-Smirnov statistic

$$D_{KS} = \sup_x |F_n(x) - F(x)|.$$

where x_i are samples. Our code implements both measures.

4.8 Code

Codes for simulations, fitting, and making all figures are available at
`https://github.com/huyu00/netw_cov_spectrum` .

Acknowledgements

This work has been partially supported by the Swartz Program in Theoretical Neuroscience at Harvard, the NIH grant from the NINDS (1U19NS104653) and the Gatsby Charitable Foundation. YH acknowledges support from HKUST. The authors also thank Eric Shea-Brown and Stefano Recanatesi for helpful discussions.

References

[1] Carlos Améndola, Philipp Dettling, Mathias Drton, Federica Onori, and Jun Wu. Structure Learning for Cyclic Linear Causal Models. *arXiv*, pages 1–19, 2020.

[2] Bruno B Averbeck, Peter E Latham, and Alexandre Pouget. Neural correlations, population coding and computation. *Nature reviews. Neuroscience*, 7(5):358–366, 2006.

[3] Baktash Babadi and Haim Sompolinsky. Sparseness and Expansion in Sensory Representations. *Neuron*, 83(5):1–14, 2014.

[4] Z. D. Bai. Circular law. *Annals of Probability*, 25(1):494–529, 1997.

[5] W Bair, E Zohary, and W T Newsome. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 21(5):1676–1697, 2001.

[6] R. Ben-Yishai, R. Lev Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92(9):3844–3848, 1995.

[7] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494–521, 2011.

[8] Serena Bradde and William Bialek. PCA Meets RG. *Journal of Statistical Physics*, 167(3-4):462–475, 2017.

[9] Yoram Burak and Ila R. Fiete. Accurate path integration in continuous attractor network models of grid cells. *PLoS Computational Biology*, 5(2), 2009.

[10] David Dahmen, Sonja Grün, Markus Diesmann, and Moritz Helias. Second type of criticality in the brain uncovers rich multiple-neuron dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 116(26):13051–13060, 2019.

[11] David J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4(12):2379, 1987.

[12] Stefano Fusi, Earl K. Miller, and Mattia Rigotti. Why neurons mix: High dimensionality for higher cognition. *Current Opinion in Neurobiology*, 37:66–74, 2016.

[13] Juan A. Gallego, Matthew G. Perich, Stephanie N. Naufel, Christian Ethier, Sara A. Solla, and Lee E. Miller. Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nature Communications*, 9(1):1–13, 2018.

[14] CW Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Springer-Verlag, Berlin, 2009.

[15] Matan Gavish and David L. Donoho. The Optimal Hard Threshold for Singular Values is 4/sqrt(3). *IEEE Transactions on Information Theory*, 60(8):5040–5053, 2014.

[16] Joshua a Goldberg, Uri Rokni, and Haim Sompolinsky. Patterns of Ongoing Activity and the Functional Architecture of the Primary Visual Cortex. *Neuron*, 42:489–500, 2004.

[17] Friedrich Götze and Alexander Tikhomirov. The circular law for random matrices. *Annals of Probability*, 38(4):1444–1491, 2010.

[18] Dmytro Grytskyy, Tom Tetzlaff, Markus Diesmann, and Moritz Helias. A unified view on weakly correlated recurrent networks. *Frontiers in Computational Neuroscience*, 7(October):1–19, 2013.

[19] Alan G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[20] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[21] Yu Hu, Steven L Brunton, Nicholas Cain, Stefan Mihalas, J Nathan Kutz, and Eric Shea-brown. Feedback through graph motifs relates structure and function in complex networks. *Physical Review E*, 062312:1–25, 2018.

[22] Yu Hu, James Trousdale, Krešimir Josić, and Eric Shea-Brown. Motif statistics and spike correlations in neuronal networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(03):P03012, 2013.

[23] Yu Hu, James Trousdale, Krešimir Josić, and Eric Shea-Brown. Local paths to global coherence: Cutting networks down to size. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 89(3):1–16, 2014.

[24] Jonathan Kadmon and Haim Sompolinsky. Transition to chaos in random neuronal networks. *Physical Review X*, 5(4):1–28, 2015.

[25] Scott Kirkpatrick and David Shcrrington. Infinite-ranged models of sinn-glasses. *Physical Review B*, 17(11), 1978.

[26] Robert B. Levy and Alex D. Reyes. Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. *Journal of Neuroscience*, 32(16):5609–5619, 2012.

[27] Benjamin Lindner, Brent Doiron, and André Longtin. Theory of oscillatory firing induced by spatially correlated noise and delayed inhibitory feedback. *Physical Review E*, 72(6):061919, 2005.

[28] Ashok Litwin-Kumar, Kameron Decker Harris, Richard Axel, Haim Sompolinsky, and L F Abbott. Optimal Degrees of Synaptic Connectivity. *Neuron*, 93(5):1153–1164.e7, mar 2017.

[29] V. A. Marchenko and L. A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457, 1967.

[30] Michael Okun, Nicholas a. Steinmetz, Lee Cossell, M. Florencia Iacaruso, Ho Ko, Péter Barthó, Tirin Moore, Sonja B. Hofer, Thomas D. Mrsic-Flogel, Matteo Carandini, and Kenneth D. Harris. Diverse coupling of neurons to populations in sensory cortex. *Nature*, 521(7553):511–515, 2015.

[31] R. Perin, T.K. Berger, and H. Markram. A synaptic organizing principle for cortical neuronal groups. *P Natl Acad Sci USA*, 108(13):5419–5424, 2011.

[32] Volker Pernice, Benjamin Staude, Stefano Cardanobile, and Stefan Rotter. How Structure Determines Correlations in Neuronal Networks. *PLoS Computational Biology*, 7(5):e1002059, 2011.

[33] K Rajan and L. F. Abbott. Eigenvalue spectra of random matrices for neural networks. *Phys Rev Lett*, 97:188104, 2006.

[34] Kanaka Rajan, Larry F. Abbott, and Haim Sompolinsky. *Interactions between Intrinsic and Stimulus-Evoked Activity in Recurrent Neural Networks*, chapter The Dynamic Brain: An Exploration of Neuronal Variability and Its Functional Significance, pages 65–82. Oxford University Press, 2011.

[35] Stefano Recanatesi, Gabriel Koch Ocker, Michael A. Buice, and Eric Shea-Brown. Dimensionality in recurrent spiking networks: Global trends in activity and local origins in connectivity. *PLoS Computational Biology*, 15(7):1–29, 2019.

[36] Patrick T. Sadtler, Kristin M. Quick, Matthew D. Golub, Steven M. Chase, Stephen I. Ryu, Elizabeth C. Tyler-Kabara, Byron M. Yu, and Aaron P. Batista. Neural constraints on learning. *Nature*, 512(7515):423–426, 2014.

[37] João D. Semedo, Amin Zandvakili, Christian K. Machens, Byron M. Yu, and Adam Kohn. Cortical Areas Interact through a Communication Subspace. *Neuron*, 102(1):249–259.e4, 2019.

[38] H. J. Sommers, a. Crisanti, H. Sompolinsky, and Y. Stein. Spectrum of large random asymmetric matrices. *Physical Review Letters*, 60(19):1895–1898, 1988.

[39] Haim Sompolinsky, a. Crisanti, and H. J. Sommers. Chaos in random networks. *Phys. Rev. Lett.*, 61(3):259–262, 1988.

[40] 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(3):e68, 2005.

[41] Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Matteo Carandini, and Kenneth D Harris. High-dimensional geometry of population responses in visual cortex. *Nature*, 2019.

[42] Terence Tao. Outliers in the spectrum of iid matrices with bounded rank perturbations. *Probability Theory and Related Fields*, 155(1-2):231–263, 2013.

[43] Terence Tao and Van Vu. Random matrices: Universality of local eigenvalue statistics. *Acta Math.*, 206(1):127–204, 2011.

[44] James Trousdale, Yu Hu, Eric Shea-Brown, and Krešimir Josić. Impact of network structure and cellular response on spike time correlations. *PLoS Comput Biol*, 8(3):e1002408, 2012.

[45] C van Vreeswijk and H Sompolinsky. Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity. *Science*, 274(5293):1724 LP – 1726, dec 1996.

[46] Nikita Vladimirov, Yu Mu, Takashi Kawashima, Davis V Bennett, Chao-Tsung Yang, Loren L Looger, Philipp J Keller, Jeremy Freeman, and Misha B Ahrens. Light-sheet functional imaging in fictively behaving zebrafish. *Nature Methods*, 11(9):883–884, 2014.

[47] Liqiong Zhao, Bryce Beverlin, Theoden Netoff, and Duane Q Nykamp. Synchronization from second order network connectivity statistics. *Front Comput Neurosci*, 5:1–16, Jan 2011.