RESEARCH ARTICLE

The spectrum of covariance matrices of randomly connected recurrent neuronal networks with linear dynamics

Yu Hu[1]*, **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

* mahy@ust.hk (YH); haim@fiz.huji.ac.il (HS)

Abstract

A key question in theoretical neuroscience is the relation between the connectivity structure and the 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 dynamics' Principal Component Analysis (PCA) and the associated effective dimensionality. We consider the spontaneous fluctuations around a steady state in a randomly connected recurrent network of stochastic neurons. An exact analytical expression for the covariance eigenvalue distribution in the large-network limit can be obtained using results from random matrices. The distribution has a finitely supported smooth bulk spectrum and exhibits an approximate power-law tail for coupling matrices near the critical edge. We generalize the results to include second-order connectivity motifs and discuss extensions to excitatory-inhibitory networks. The theoretical results are compared with those from finite-size networks and the effects of temporal and spatial sampling are studied. Preliminary application to whole-brain imaging data is presented. Using simple connectivity models, our work provides theoretical predictions for the covariance spectrum, a fundamental property of recurrent neuronal dynamics, that can be compared with experimental data.

Author summary

Here we study the distribution of eigenvalues, or spectrum, of the neuron-to-neuron covariance matrix in recurrently connected neuronal networks. The covariance spectrum is an important global feature of neuron population dynamics that requires simultaneous recordings of neurons. The spectrum is essential to the widely used Principal Component Analysis (PCA) and generalizes the dimensionality measure of population dynamics. We use a simple model to emulate the complex connections between neurons, where all pairs of neurons interact linearly at a strength specified randomly and independently. We derive a closed-form expression of the covariance spectrum, revealing an interesting long

tail of large eigenvalues following a power law as the connection strength increases. To incorporate connectivity features important to biological neural circuits, we generalize the result to networks with an additional low-rank connectivity component that could come from learning and networks consisting of sparsely connected excitatory and inhibitory neurons. To facilitate comparing the theoretical results to experimental data, we derive the precise modifications needed to account for the effect of limited time samples and having unobserved neurons. Preliminary applications to large-scale calcium imaging data suggest our model can well capture the high dimensional population activity of neurons.

1 Introduction

Collective dynamics in networked systems are of great interest, with numerous applications in many fields, including neuroscience, spin glasses, social and ecological networks [1]. Many studies of neuronal networks have focused on how certain statistics of dynamics depend on the network's connectivity structure [2–4], including the population average [5] and variance [6] of pairwise correlations. These dynamic features can be estimated experimentally by repeatedly recording a small number of neurons at each time. In this sense, they may be regarded as *local* or *marginal* features of dynamics. On the other hand, certain *global* or *joint* dynamic features may only be possible or efficient to estimate by recording a population of neurons *simultaneously*. An important example is the eigenvalues of the covariance matrix, which are complicated nonlinear functions of all the matrix elements. These eigenvalues arise naturally when performing the widely used Principal Component Analysis (PCA) of population activity, where they correspond to the amount of variance contained in each principal component of the activity. Although one can in principle fill out the covariance matrix through repeated pairwise recordings, the matrix is much more efficiently calculated from simultaneously recorded data. Furthermore, a sample covariance matrix calculated from simultaneously recorded data also requires particular methods to address the effect due to the finite recording length (Section 3.7) which would be different from repeated recordings. Another example of global dynamic features that has recently received substantial interest [7–11] is the effective dimensionality of neural population activity. When describing the data distribution in terms of linear subspaces, the dimensionality 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 [12, 13], and theoretical investigations have illustrated the importance of having a low dimensionality for brain function and computation [14], such as when representing stimuli [15] and generating motor outputs [13].

As the experimental techniques of measuring the activity of large population of neurons in biological networks become increasingly available, new opportunities arise for studying how the network's connectivity structure affects these global aspects of population dynamics. In this work, we study the eigenvalue distribution (i.e., spectrum) of the covariance matrix of spontaneous activity in a large recurrent network of stochastic neurons with random connectivity. We focus on several basic and widely used models of random connectivity, including independent and identical Gaussian distributed connectivity [2] (Section 3.1), networks with second-order connectivity motifs [5, 16–20] (Section 3.3), and random Excitation-Inhibition (EI) networks (Section 3.5). Random connectivity has been a fundamental model in theoretical studies of neuronal network dynamics [2, 8, 21]. It can be motivated as a minimal model to capture the complex, disordered connections observed in many neuronal circuits, such as in

the cortex. Some aspects of these covariance spectra might be distinct from those under ordered, deterministic connectivity (Section 4.2).

The dynamics considered here is simple where the activity fluctuations around the steady-state are described by a linear response [22, 23], which experimentally is related to spontaneous or persistent neural activity in absence of structured spatial-temporal stimuli. Despite the simple dynamics and minimal connectivity model, we find the resulting spectrum has a continuous bulk of nontrivial shape exhibiting interesting features such as a power-law long tail of large eigenvalues (Section 3.2), and strong effects due to the non-normality of the connectivity matrix (Section E.2 in S1 Text). These covariance spectra highlight interesting population-level structures of neuronal variability shaped by recurrent interactions that were previously unexplored.

Using the theory of the covariance spectrum, we derive closed-form expressions for the effective dimensionality (previously known for the simple random i.i.d. Gaussian connectivity [6]) We show that the continuous bulk spectrum has the advantage over low-order statistics such as the dimensionality thanks to its robustness to low rank perturbations (Section 3.3 to 3.5 and 3.8). Our analytically derived eigenvalue distributions can be readily compared to real activity data of recurrent neural circuits or simulations of more sophisticated computational models. We provide ready-to-use code to facilitate such applications (see **Data Availability Statement**). An example of such an application for a whole-brain calcium imaging data is presented in Section 3.8.

2 Model

2.1 Neuronal networks with random recurrent connectivity

We consider a recurrent network of linear rate neurons driven by noise

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

Here $x_i(t)$ is the firing rate of neuron i. J_{ij} describes the recurrent interaction from neuron j to i. τ is a time constant describing how quickly the firing rates changes in response to inputs. The network is driven by independent Gaussian white noise $\xi_i(t)$ with variance σ^2, that is, the expectation $\langle \xi_i(t) \xi_j(t + \tau) \rangle = \sigma^2 \delta_{ij} \delta(\tau)$.

We focus on the structure of long time scale covariation in the network, which are described by the *long time window* covariance $C_{ij} = \lim_{\Delta T \to \infty} \frac{1}{\Delta T} C_{ij,\Delta T}$. $C_{ij,\Delta T}$ is the covariance of the summed activity over a window of ΔT: $C_{ij,\Delta T} = \langle \Delta s_i(t) \Delta s_j(t) \rangle$, $\Delta s_i(t) = \int_t^{t+\Delta T} \Delta x_i(t') dt'$. For biophysical neurons, $C_{ij,\Delta T}$ typically settles to its limiting value when $\Delta T > 50$ms [24]. It can be shown [25] that the long time window covariance C (also the zero-frequency covariance, see Section 3.6) is

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

Here I is the identity matrix, and A^{-1}, A^T are the matrix inverse and transpose $(A^{-T} = (A^{-1})^T)$. For simplicity we will set $\sigma^2 = 1$ unless stated otherwise. The covariance matrix C can also be estimated from experimental data consisting of simultaneously recorded neurons (Methods). We consider generalizations beyond the long time window covariance in Section 3.6.

Our analysis and results start from the covariance-connectivity relation Eq (2), which also describes, or closely approximates, the network dynamics in other models (Section 5.2) including networks of integrate-and-fire or inhomogeneous Poisson neurons [23, 26–28], fixed

point activity averaged over whitened inputs, and structural equation modeling in statistics [29]. These models provide additional motivations for the covariance (Eq (2)), which may allow interpreting our results in experiments where the neural activity is driven by stimuli [10].

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 [2, 4]. We first consider a random connectivity where

$$J_{ij} \sim \mathcal{N}(0, g^2/N) \tag{3}$$

are drawn as independent and identically distributed (i.i.d.) Gaussian variables with zero mean and variance g^2/N (referred subsequently as the *i.i.d. Gaussian connectivity*). The covariance spectrum follows directly from results in random matrices [30, 31]. We then show how to generalize to other types of random connectivity, including: those with connectivity motifs (Section 3.3), Erdős-Rényi random connectivity, networks with excitation and inhibition (Section 3.5). 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. A list of variable notations is given in Table 1 for ease of reference.

2.2 Covariance eigenvalues and dimensionality

Principal Component Analysis (PCA) is a widely used analysis of population dynamics, where the activity is decomposed along orthogonal patterns or Principal Components (PCs). The PCs are the eigenvectors of the covariance matrix C (Eq (37)), and the associated eigenvalues λ_i are nonnegative and show 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 through the equality $\int_a^b p_C(x)(x)dx = \frac{1}{N}\#\{\lambda_i \in (a, b]\}$ for all a, b. We also refer to $p_C(x)$ as the spectrum (which should not be confused with the frequency spectrum obtained via Fourier transform). We will

Table 1. List of notations.

Notation	Description
N	number of neurons
C	covariance matrix, Eq (2)
$p_C(x)$	pdf of eigenvalues of C, Section 2.2
$x_\pm$	support edges of $p_C(x)$
J	matrix of connection weights Section 2.1
σ^2	variance of white noise input, Section 2.1
g	connection strength var$(J_{ij}) = g^2/N$, (3)
g_c	maximum g constrained by stability
g_r	g/g_c
κ	$\rho(J_{ij}, J_{ji})$ reciprocal motif cumulant, Eq (13)
μ	mean of eigenvalues $\frac{1}{N}\sum_{i=1}^N \lambda_i$
D	dimension, Eq (4)
M	number of time samples, Eq (37)
α	N/M, Section 3.7
N_s	number of sampled neurons, Section 3.7
f	N_s/N, Section 3.7

derive the limit of $p_C(x)$ as $N \to \infty$ and study how it depends on the connectivity parameters such as $g = N\mathrm{var}(J_{ij})$.

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 [32] (Section 3.8). the spectrum is also closely related to the effective *dimension* of the population activity. In many cases, the linear span of the activity fluctuations is full rank, N. Nevertheless, most of the variability is embedded in a much lower dimensional subspace. A useful measure of the effective dimension, known as the participation ratio [8, 33] is given by

$$D \equiv \frac{\left(\sum_{i=1}^{N} \lambda_i\right)^2}{\sum_{i=1}^{N} \lambda_i^2}. \tag{4}$$

which can be calculated from the first two moments of $p_C(x)$. We will also derive explicit expressions for D in random connectivity models.

3 Results

3.1 Continuous bulk spectrum with finite support

For networks with i.i.d. Gaussian connectivity (Section 2.1), 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 [2]. The parameter σ in Eq (2) just scales all λ_i and thus is hereafter set to 1 for simplicity. Our main theoretical result is the following expression for the probability density function (pdf) of the covariance eigenvalues in the large N limit (Section A in S1 Text),

$$p_C(x) = \frac{3^{\frac{1}{6}}}{2\pi g^2 x^2} \left[\sum_{\xi=1,-1} \xi \left(\left(1 + \frac{g^2}{2}\right)x - \frac{1}{9} + \xi \sqrt{\frac{(1-g^2)^3 x(x_+ - x)(x - x_-)}{3}} \right)^{\frac{1}{3}} \right], \tag{5}$$

$$x_- \leq x \leq x_+.$$

where

$$x_{\pm} = \frac{2 + 5g^2 - \frac{g^4}{4} \pm \frac{1}{4}g(8+g^2)^{\frac{3}{2}}}{2(1-g^2)^3}, \tag{6}$$

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 $|x - x_{\pm}|^{\frac{1}{2}}$ (Section A.1 in S1 Text).

The above result for the distribution $p_C(x)$ follows from the derivation of the circular law distribution of the eigenvalues of the random matrix J [30, 31, 34, 35]. However, to the best of our knowledge, this is the first exposition of the explicit expression for the spectrum of C, (Eq (5), which is essential for fitting to empirical data Section 3.8) and for the study of network dynamics. We emphasize that $p_C(x)$ does not have a *simple* relation to the spectrum of J because J is a non-normal matrix (i.e., $J^T J \neq JJ^T$). This point is further elaborated in Section 3.3.2. Although the above result is derived in the large N limit, it matches accurately the spectrum of C in networks of sizes of several hundred, as demonstrated in our numerical results, Fig 1A and 1B (see also Fig A in S1 Text for additional realizations and trial-averages). In PCA
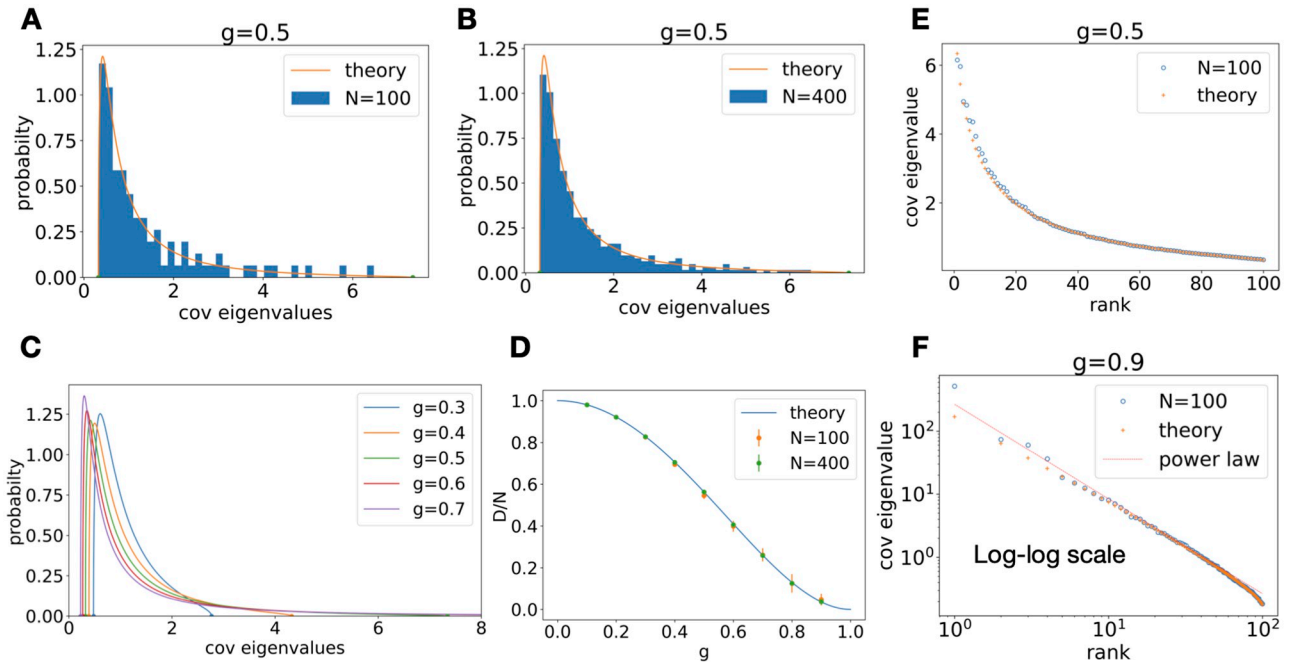
Fig 1. Covariance spectrum under random Gaussian connectivity. A. Compare theory (Eq (5)) with finite-size network covariance using Eq (2) at $N = 100$, $g = 0.5$. The histogram of eigenvalues is a single realization of the random connectivity. **B**. Same as A. at $N = 400$. **C**. Covariance eigenvalue 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 networks (Eq (2) and use Eq (4)). Note some error bars are smaller than the dots **E**. Covariance eigenvalues vs. their rank (in descending order). The circles are covariance eigenvalues from a single realization of the random connectivity with $N = 100$ (Eq (2)). The crosses are predictions based on the theoretical pdf (Eq (5)). **F**. Same as E. but for $g = 0.9$ and on the log-log scale. The red dashed line is the power law with exponent $-3/2$ derived from Eq (5), see Section 3.2.

and other analyses, the covariance eigenvalues are plotted in descending order vs. their rank [10, 36]. We can use the theoretical pdf Eq (5) to predict this *rank plot* by numerically solving the inverse cumulative distribution function (cdf), i.e., quantile function, at probability $\frac{N-\frac{1}{2}}{N}, \frac{N-1-\frac{1}{2}}{N}, \ldots, \frac{\frac{1}{2}}{N}$. The closed form pdf (Eq (5)) allows for using the highly efficient Newton's method to compute the quantiles. Fig 1E and 1F show a good agreement between the theory Eq (5) and a single realization of a $N = 100$ random network.

3.2 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 5.3 in Methods). This corresponds to an activity PC with diverging variance and is consistent with the stability requirement of $g < 1$. Note that the lower edge x_- is always bounded away from 0 and has a limit of $\frac{4}{27}$ as $g \rightarrow 1$. Analyzing the shape of $p_C(x)$ for large x in the critical regime $g \rightarrow 1$ yields a long tail of large eigenvalues, following a power law (Fig 2A and 2B, Methods)

$$p_C(x) \approx \frac{\sqrt{3}}{2\pi} x^{-\frac{5}{3}}. \tag{7}$$

As shown in Fig 2A, the power-law shape of $p_C(x)$ is apparent for $g = 0.7$ and does not require a g very close to 1.
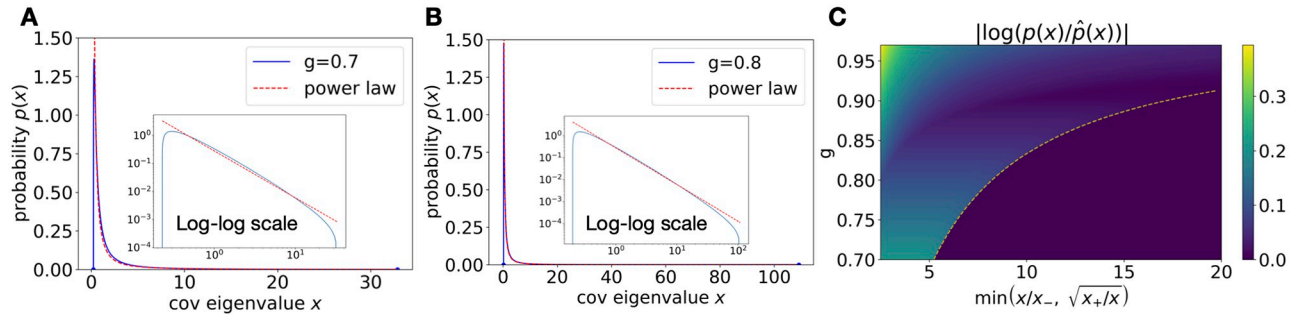
Fig 2. Approximate power-law tail. A. The exact pdf (solid line) of the covariance spectrum compared with the power-law approximation (dashed line, Eq (7)) at $g = 0.7$. Inset shows the log-log scale. **B**. Same as A. for $g = 0.8$. The approximation improves as g approaches the critical value 1. **C**. 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 $\sqrt{x_+/x}$ (more details and motivations in Section A.2 in S1 Text. 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.

To better elucidate the range of validity of the above power law, we consider the regime where $1 - g^2 \ll 1$ and $x \gg 1$. Define $z = \sqrt{1 - x/x_+}$ where $x_+ \propto (1 - g^2)^{-3}$ is the upper edge of the support of $p_C(x)$. Then,

$$p_C(x) \approx \frac{\sqrt{3}}{2\pi} x^{-\frac{5}{3}} 2^{-\frac{1}{3}} F(z) \tag{8}$$

where $F(z) = (1 + z)^{1/3} - (1 - z)^{1/3}$. Thus, far from the spectrum upper edge, $z \to 1$ and we obtain Eq (7), whereas near the upper edge $z \to 0$ and $p_C(x) \approx \frac{3^{-\frac{1}{2}}2^{-\frac{1}{3}}}{\pi} x_+^{-\frac{5}{3}} \sqrt{1 - x/x_+}$, which is the expected square-root singularity near the edge.

The power-law approximation of the probability density function Eq (5) translates to an approximation for the cumulative distribution function $F_C(x) \approx 1 - \frac{3\sqrt{3}}{4\pi} x^{-\frac{2}{3}}$ This also means a power law in the rank plot has an exponent of $-3/2$ when connection strength g is close to the critical value (Fig 1F), providing an alternative mechanism based on recurrent circuits for the experimental observations of power-law distributed covariance eigenvalues in the literature [10, 36] (see also the model Eq 31 and discussions in Methods).

Because the probability density is small in the power-law 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, topics which we revisit later (Section 3.8). Note that a long tail in the spectrum is a distinct feature of correlations arising from the recurrent network dynamics (see also a heuristic explanation in Section E.3 in S1 Text). For example, for the Marchenko–Pastur law that is often used for modeling empirical covariance spectra, the upper edge of its support relative to the mean is bounded by 4 (Methods). In contrast, the same ratio for $p_C(x)$ (Eq (5)) can be arbitrarily large as $O((1 - g^2)^{-2})$ (see below for calculating the mean). This highlights the difference between covariance generated by finite samples of noise and correlations generated by the recurrent dynamics.

The long tail of the eigenvalue distribution is also reflected by a low effective dimension (Eq 4). In Section B in S1 Text we show that the mean and second moment of the eigenvalue distribution above, are given by

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

which yields for the dimension

$$D = N(1 - g^2)^2. \tag{10}$$

In particular, the *relative dimension* with respect to the network size D/N vanishes as g approaches 1 (Fig 1D). In comparison, D/N for the Marchenko–Pastur law (Eq 35) is at least $\frac{1}{2}$.

While these low-order moments can be derived from previous methods (see [6] and Section B in S1 Text, and also a parallel work [37]), our method allows for the derivation of higher-order moments, such as,

$$\mathbf{E}(\lambda^3) = (1 - g^2)^{-7}(1 + 2g^2), \quad \mathbf{E}(\lambda^4) = (1 - g^2)^{-10}(1 + g^2)(1 + 5g^2). \tag{11}$$

and in general,

$$\mathbf{E}(\lambda^n) \propto (1 - g)^{-3(n-1)-1}, \quad \text{as } g \to 1^- \tag{12}$$

3.3 Impact of the asymmetry in connectivity

We next consider generalizations of random connectivity beyond the i.i.d. Gaussian model (Sction 2.1). An important feature of biological neural networks is the presence of motif structures at various scales [16–18, 38], which correspond to an overabundance of certain subgraphs, relative to their frequency in an edge-shuffled network (i.e., an i.i.d. random graph with matching connection probability). Using a random connectivity model with Gaussian distributed entries [38] Section C.1 in S1 Text, we can study the effects of the four types of second order motifs (i.e., consisting of two edges). Among them, the diverging, converging, and chain motifs correspond to a low-rank component L of the connectivity $J = L + \tilde{J}$, where the remaining part $\tilde{J}$ contains only reciprocal motifs. Based on the results in Section 3.4, we can focus below on studying the covariance spectrum under the simpler connectivity $\tilde{J}$ with only reciprocal motifs, because the bulk spectrum of $J = L + \tilde{J}$ is remains the same when adding the low-rank L (see examples in Section 3.4 and Section C in S1 Text). Note that the magnitude of diverging, converging, and chain motifs in the original connectivity still indirectly affect the bulk covariance spectrum by affecting the parameters g and κ (see below) of $\tilde{J}$ (see Section C.1 in S1 Text for the details of this relation).

For ease of notation, we still use J to represent a random connectivity matrix with potentially reciprocal motifs ($\tilde{J}$ described above). This is equivalent to varying the degree of asymmetry of J [31]. In this case, each component of J is $J_{ij} \sim \mathcal{N}(0, g^2/N)$ but J_{ij} and J_{ji} are correlated,

$$\kappa = \rho(J_{ij}, J_{ji}), \tag{13}$$

with $-1 \leq \kappa \leq 1$. All other correlations are zero.

3.3.1 Symmetric and anti-symmetric random networks. First, we consider two extreme cases for the reciprocal motifs: $\kappa = 1$ corresponding to $J_{ij} = J_{ji}$, and $\kappa = -1$ corresponding to anti-symmetric matrix (or skew-symmetric $J_{ij} = -J_{ji}$). These cases are much simpler to analyze, because J is a *normal matrix* so $p_C(x)$ can be derived from the well known eigenvalue distribution of J [31]. For symmetric random connectivity,

$$p_{C,g,\kappa=1}(x) = \frac{\sqrt{(4g^2 - 1)x - 1 + 2\sqrt{x}}}{4\pi g^2 x^2}, \quad x \in (x_-, x_+), \ x_{\pm} = (1 \mp 2g)^{-2}. \tag{14}$$
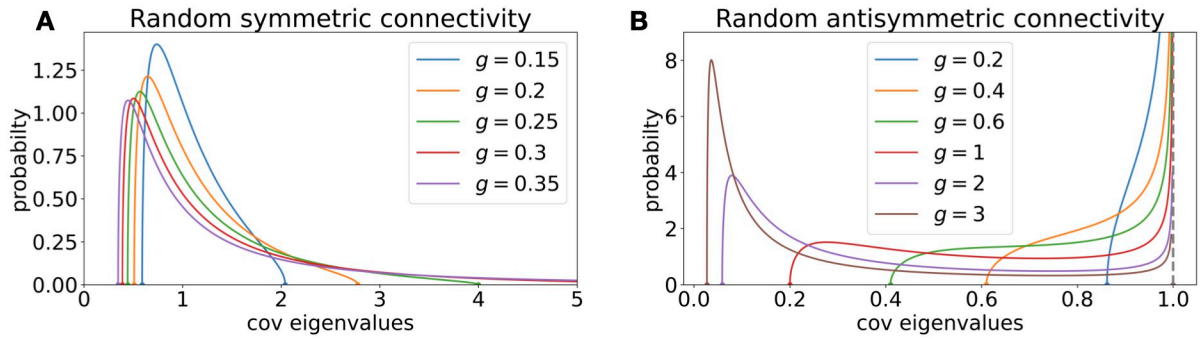
Fig 3. Covariance spectrum for the symmetric and anti-symmetric random connectivity. A. The pdf of a covariance spectrum with random symmetric J with different g (note $g < \frac{1}{2}$ for stability). **B.** Same as A., but for random anti-symmetric $J_{ij} = -J_{ji}$. The pdf diverges at $x = 1$ as $|1 - x|^{-\frac{1}{2}}$.

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

$$p_{C,g,\kappa=-1}(x) = \frac{\sqrt{(4g^2 + 1)x - 1}}{2\pi g^2 x^2 \sqrt{1 - x}}, \quad x \in ((1 + 4g^2)^{-1}, 1). \tag{15}$$

Here the network is stable for all g. The derivations are given in the Section D in S1 Text.

Since the general shape and trend depending on g of the spectrum in the simpler symmetric case (Fig 3A) is qualitatively similar to the i.i.d. case (Fig 1C), one can use it to gain intuition, for example, of the thin tail of large eigenvalues as g approaches its critical value.

Quantitatively using the equations above, we see that $p_C(x)$ of the symmetric random network (Fig 3A) has a power-law tail analogous to Eq (7) as $g \to 1/2$ (i.e., large x) but with a different exponent from the i.i.d. case (Eq 7),

$$p_{C,g,\kappa=1}(x) \approx \frac{\sqrt{2}}{\pi} x^{-\frac{7}{4}}. \tag{16}$$

The $p_C(x)$ of the anti-symmetric random network (Fig 3B) does not have a long tail as the upper limit of the support is always 1. Since J here is a normal matrix, this qualitative difference from the i.i.d. random connectivity can be understood considering the eigenvalues of J [31]. The eigenvalues of J all lie on the imaginary axis and therefore never approach 1. Thus, the eigenvalues of the covariance matrix do not develop a long tail when increasing g.

3.3.2 Connectivity with general asymmetry. For the Gaussian random connectivity with $\kappa = \rho(J_{ij}, J_{ji})$, $-1 < \kappa < 1$, we have derived an implicit equation for $p_{C,g,\kappa}(x)$ in the large N limit based on the results in [31] (Section E in S1 Text). Although a closed-form expression can be derived using the root formula for quartic equations, it seems quite cumbersome, hence we show here the numerical solutions of this equation. For a fixed g, as κ increases, the distribution broadens on both sides (Fig 4C). Intuitively, these effects (also Fig 4B) can be understood due to the change of the critical g for stability, which is now given by $g_c = (1 + \kappa)^{-1}$ (based on the spectrum of J [31]). As κ increases, the *relative coupling strength* $g_r = g/g_c = g(1 + \kappa)$, which is also the maximum real part of J's eigenvalues [31] increases, and the shape of the spectrum changes similar to increasing g in the case of i.i.d. connectivity (Fig 1C). This motivates us to further compare the distributions $p_{C,g,\kappa}(x)$ with the same g_r to study effects due to κ beyond changing g_r. As shown in Fig 4D, when fixing the relative strength g_r the distribution narrows as κ increases.
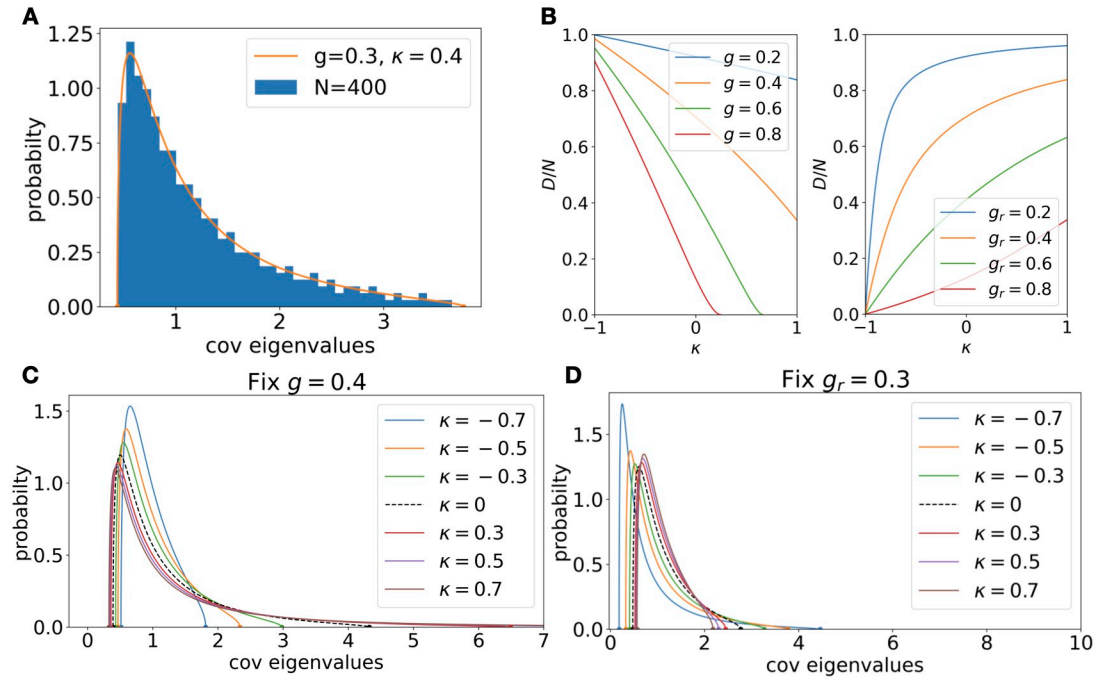
Fig 4. Impact of reciprocal motifs. A. Compare theoretical covariance spectrum for random connectivity with reciprocal motifs and a finite-size network covariance using Eq (2)($g = 0.4$, $\kappa = 0.4$, $N = 400$). **B**. The impact of reciprocal motifs on dimension for various $g_r = g/g_c$ (Eq (18)). For small g_r, the dimension increases sharply with κ. **C**. The spectra at various κ while fixing $g = 0.4$. The black dashed line is the i.i.d. random connectivity ($\kappa = 0$). **D**. Same as C. but fixing relative $g_r = 0.4$ to control the main effect (see text). The changes in shape are now smaller and the support narrows with increasing κ.

Consistent with the above results, we have shown (Section E.1 in S1 Text) that for all intermediate values of $-1 < \kappa < 1$, the critical covariance spectrum has an asymptotic power-law tail with the same exponent as the i.i.d. random case (Eq (7))

$$p_c(x) \approx \frac{\sqrt{3}}{2\pi}(1-\kappa)^{\frac{1}{3}}(1+\kappa)x^{-\frac{5}{3}}, \quad \text{as } x \to \infty, \; g \to g_c = (1+\kappa)^{-1} \qquad (17)$$

In comparison, the $\kappa = \pm 1$ are singular cases in Eq 17 and have a different limiting power-law behavior (an exponent of $-\frac{4}{7}$ and no long tail). This is intuitively consistent with the spectrum of J which becomes confined on a line for $\kappa = \pm 1$ rather than in an ellipses for $-1 < \kappa < 1$ [31] (see Section E.1 in S1 Text for more discussion).

The 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 (Section E in S1 Text)

$$D = N\frac{\mu_1\sqrt{1+4(g^2-\theta)}}{(\theta\mu_1+1)^2(g^2\mu_1+1)}, \quad \mu_1 = \frac{2\theta-1+\sqrt{1+4(g^2-\theta)}}{2(g^2-\theta^2)}, \; \theta = g^2(1+\kappa). \qquad (18)$$

Here μ_1 is the mean of the distribution. Comparing with Eq 10, this shows the nontrivial dependence of dimension on the reciprocal motif strength κ. As $g \to g_c = (1+\kappa)^{-1}$, $g^2 - \theta^2 = g^2(1-g_r^2) \to 0$. The numerator of μ_1 is at least $2\theta - 1 \geq 2/(1+\kappa) - 1 > 0$. Therefore μ_1 diverges as $O((1-g_r^2)^{-1})$. Since $1+4(g^2-\theta) \geq 1+4(g^2-g) \geq (1-2/(1+\kappa))^2 > 0$, we have D/N vanishes as $O(\mu_1^{-2}) = O((1-g_r^2)^2)$. The above limits under the critical g are similar to the

$\kappa = 0$ case, Eq (10). Consistent with the shape changes, the dimension decreases with reciprocal motifs when fixing g and increases when fixing g_r (Fig 4B).

The general asymmetric random connectivity also provides an example of the strong effect of J being a non-normal matrix on the covariance spectrum. By continuity, one may 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 $p_{C,g,\kappa = -1}(x)$, which is bimodal for sufficiently large g (i.e., has another peak in addition to the divergence at 1, Fig 3B). Indeed, assuming a normal J predicts a covariance spectrum that is bimodal with a non-smooth peak in a large region of $-1 < \kappa < 0$ and g (Fig H in S1 Text). Intriguingly, the actual spectrum $p_{C,g,\kappa}(x)$ is unimodal for all but a minuscule region of (κ, g) where $\kappa < -0.95$, indicating a suppression of a bimodal spectrum under negative κ due to the non-normal J (Section E.2 in S1 Text).

3.4 Adding low rank connectivity structure

An important property of the spectrum of C is the robustness of its bulk component to the addition of low rank structured connectivity. Many connectivity structures that are important to the dynamics and function of a recurrent neuronal network can be described by a full rank random component plus a low rank component [39, 40]. For example, such components may arise from Hebbian learning [41] and by training neural networks by gradient decent [42]. A simple case is where we add a rank k structured matrix that is deterministic or independent to the random component [43, 44]. As shown in Section C in S1 Text, in large networks, the bulk covariance spectrum remains unchanged, but the low rank component may give rise to at most $2k$ outlying eigenvalues. This is illustrated by the example of rank-1 perturbation to J with i.i.d. Gaussian entries in Fig 5C and 5D, where the expected location of the outliers in the covariance spectrum can be predicted analytically (Fig 5E and 5F and Section C.3 in S1 Text). This is in contrast to the spectrum of J, where the same perturbations can lead to an unbounded number of randomly located eigenvalues [43, 45] (Fig 5A and 5B). In sum, the bulk spectrum of covariance is robust against low rank perturbations to the connectivity. Note, however, the relevance of the bulk spectrum for the network dynamics depends on the location of outliers. Outliers to the right of the bulk spectrum may indicate potential instability of the dynamics even for $g < 1$, as discussed in the example below.

3.5 Sparse excitatory-inhibitory networks

The Gaussian random connectivity has a non-zero connection weight for all pairs of neurons with probability 1, where many biological networks are sparsely connected. In addition, each neuron has both excitatory (positive) and inhibitory (negative) weights, in contrast to many neuronal networks that obey Dale's Law, namely all neurons are either excitatory (with all outgoing weights positive) or inhibitory (with negative outgoing weights). We consider here a simple model of E-I network, consisting of $N/2$ excitatory and $N/2$ inhibitory neurons. The probability of each connection J_{ij} to be nonzero, which may depend on the types of neurons i and j, is $K_{\alpha\beta}/N$, $\alpha, \beta = E, I$. Thus, the mean number of inputs to a neuron of type α from a population of type β is $K_{\alpha\beta}/2$. All excitatory non-zero connections are of strength $w_0/\sqrt{K_{\alpha\beta}}$ and the inhibitory connections are $-w_0/\sqrt{K_{\alpha\beta}}$. We assume that $K_{\alpha\beta} = k_{\alpha\beta}K$ where $k_{\alpha\beta} = O(1)$ and $K \ll N$.

To map this architecture on to the one studied above, we adopt the framework of [21] and consider the equivalent Gaussian connectivity with matching variance for each J_{ij}. Importantly, the choice of connection probabilities and weights ensures that var(J_{ij}) is w_0^2/N (to the leading order for $N \gg 1$) regardless of the cell type of neuron i, j. This allows us to define the effective
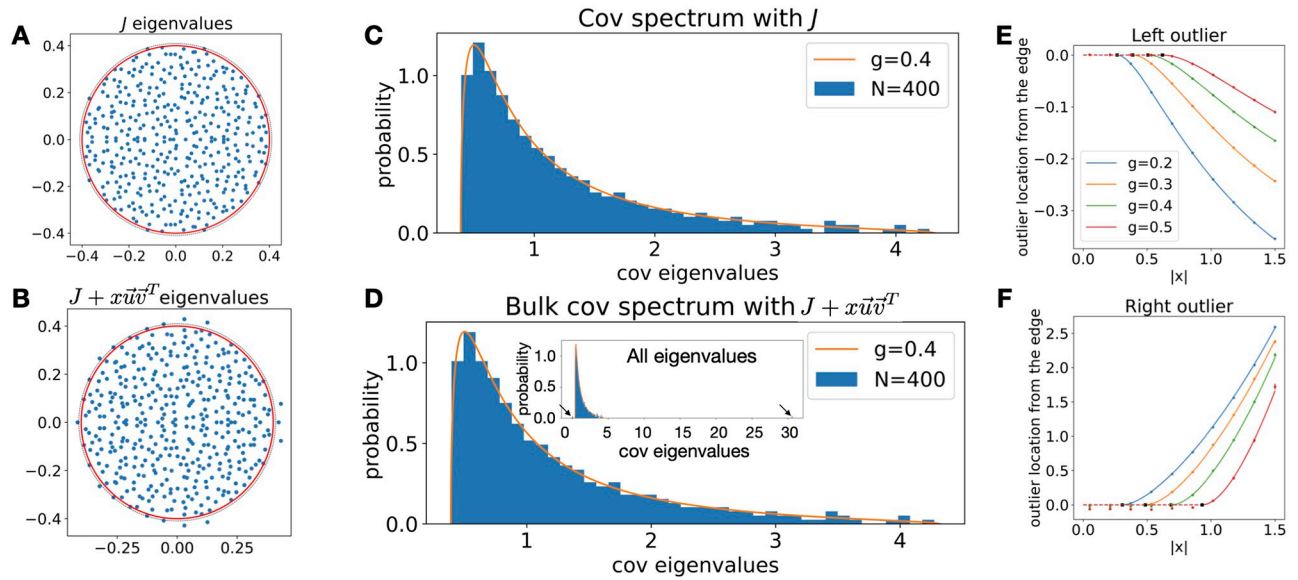
Fig 5. Robustness of the covariance spectrum to low rank perturbations of the connectivity. A. Eigenvalues of a Gaussian random connectivity J (Eq (3)), $g = 0.4$, $N = 400$. As $N \to \infty$, the limiting distribution of eigenvalues is uniform in the circle with radius g ([34] red solid line). The black dashed line is the 0.995 quantile of the eigenvalue radius calculated from 1000 realizations. **B**. Same as A. but for the rank-1 perturbed $J + x\vec{u}\vec{v}^T$. $u = (1, 1, \ldots, 1)^T / \sqrt{N}$, $v_i \sim$ i.i.d. $\mathcal{N}(0, 1/N)$ and $x = 4.03$. This example also corresponds to adding diverging motifs (Section 3.3 and Section C.1 in S1 Text). **C**. The histogram of covariance eigenvalues (Eq (2)) under the J in A. **D**. The bulk histogram of eigenvalues with $J + x\vec{u}\vec{v}^T$ has little change and remains well described by the Gaussian connectivity theory (red line, Eq (5)). Besides the bulk, there are two outlier eigenvalues to the left and right (inset, arrows) **E,F**, Analytical predictions (solid and dashed lines) of the outlier locations given g and $|x|$ when u, v are (asymptotically) orthogonal unit vectors that are independent of J (see other cases in Section C.3 in S1 Text). The y-axis is the outlier location subtracting the corresponding edge x_+, Eq (6), so it is zero for small $|x|$ before the outlier emerges (dashed line). The dots are the mean of the smallest (for the left outlier) or largest (right outlier) eigenvalues averaged across 100 realizations of the random J, $N = 4000$. The errorbars are the standard error of the mean (SEM, many are smaller than the dots).

synaptic gain as $g^2 = w_0^2$ for all neurons. The mean of the connections between a presynaptic neuron of type β and postsynaptic α is $\mathbf{E}(J_{ij}) = \sqrt{K_{\alpha\beta}} w_\beta / N$ where $w_E = w_0$ and $w_I = -w_0$. Thus, we can write J_{ij} as a zero-mean i.i.d. Gaussian matrix with uniform variance w_0^2 / N and a rank-2 matrix of the means. As stated in Section 3.4, in such a case the bulk spectrum of the neurons' covariance matrix is the same as Eq (5). In addition there are at most 4 outlier eigenvalues. For $K \gg 1$, from the analysis of [21], the stability of the recurrent dynamics of a linear network with the above connectivity amounts to the requirement that all eigenvalues of the 2×2 matrix $M_{\alpha\beta} = \sqrt{k_{\alpha\beta}} w_\beta$ have negative real parts. Fulfilling this condition by choosing appropriate values for $k_{\alpha\beta}$ (see example in Fig 6A and Section C.3 in S1 Text) guarantees that the outlier(s) due to the nonzero means are to the left of the bulk covariance spectrum so that the largest eigenvalue is $x_+(g)$, Eq (6). For $K = O(1)$, the results in [43] show that the above condition is sufficient but not necessary for stability. For example, when all $k_{\alpha\beta}$ are equal to k, which corresponds to a balance of excitation and inhibition [45], all eigenvalues of M are 0 and the dynamics is stable for $g < 1$ for large N. At the same time, there can be two outlying eigenvalues on the two sides of the bulk covariance spectrum (Fig 6B), whose expected location can be predicted (Section 3.4 and Section C.3 in S1 Text). Several additional examples including all inhibitory networks are considered in the Section F in S1 Text.

3.6 Frequency dependent covariance

We have so far focused on the long time window covariance matrix. This would be especially suitable for neural activity recordings with limited temporal resolution such as calcium
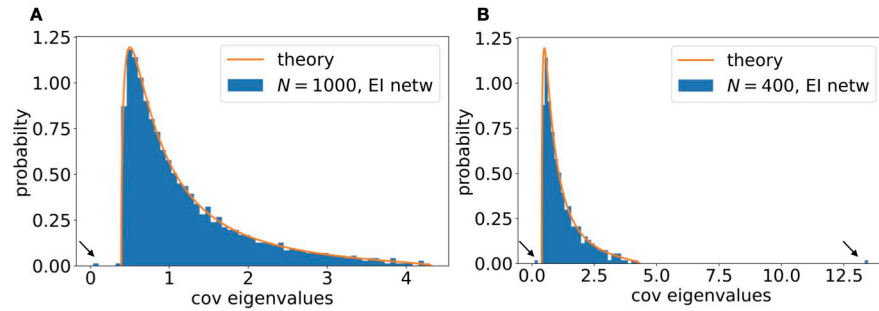
Fig 6. EI networks. A. One realization of the covariance eigenvalues by Eq (2) with an EI network satisfying the stability condition (see text). The bulk spectrum is well described by the Gaussian random connectivity theory (solid line, Eq (5)). There is one small outlier to the left of the bulk (arrow). The parameters are $g = w_0 = 0.4$, $\sqrt{k_{ee}} = 0.5$, $\sqrt{k_{ei}} = 1.5$, $\sqrt{k_{ie}} = 1$, $\sqrt{k_{ii}} = 2$, $K = 60$, $N = 1000$. To improve the accuracy of the theory to finite K, N, here we use a slightly modified connection weight, $w_0/\sqrt{K_{\alpha\beta}(1 - K_{\alpha\beta}/N)}$, for all excitatory non-zero connections, and similarly $-w_0/\sqrt{K_{\alpha\beta}(1 - K_{\alpha\beta}/N)}$ for inhibitory connections, such that $\mathrm{var}(J_{ij}) = w_0^2/N$ holds exactly for finite N. **B**. Similar as A but for balanced EI network (see text) with $k_{\alpha\beta} = k = 1$, $g = 0.4$, $K = 40$, $N = 400$. Note there are two outliers on both sides of the bulk.

imaging [46]. Temporal structures of correlation beyond the slow time scale can be described by the frequency covariance matrix (or coherence matrix)

$$C_{ij}(\omega) = \lim_{\Delta T \to \infty} \langle \Delta s_i(\omega) \Delta s_j^\dagger(\omega) \rangle, \tag{19}$$

where $\Delta s_i(\omega) = \frac{1}{\sqrt{\Delta T}} \int_0^{\Delta T} \Delta e^{-i\omega} x_i(t) dt$ is the Fourier transform of the neural activity and $z^\dagger$ is the complex conjugate. $C_{ij}(\omega)$ can also be calculated by the Fourier transform of the time-lagged cross-correlation functions $C_{ij}(\tau) = \langle x_i(t) x_j(t - \tau) \rangle$ (Wiener-Khinchin theorem). Analogous to Eq (2) $C(\omega)$ obeys [25],

$$C(\omega) = \sigma^2 |a(\omega)|^2 (I - a(\omega)J)^{-1}(I - a^\dagger(\omega)J)^{-T}. \tag{20}$$

Here $z^\dagger$ is the complex conjugate and $|z|$ is the norm for a complex z. The transfer function $a(\omega) = (1 + i\tau\omega)^{-1}$ summarizes the dynamics of single neurons in the network and corresponds to a response filter of $e^{-t/\tau}/\tau$, $t > 0$ for the model of Eq (1) (see also Section 5.2). The long time window covariance we have studied corresponds to $C(\omega = 0)$ (Wiener-Khinchin theorem).

For the i.i.d. Gaussian random connectivity J, we can show that the spectrum of $C(\omega)$ is given by the same Eq (5) for $C(0)$ (up to a constant scaling) by replacing g with a frequency dependent $g(\omega)$ (compare with Eq (3))

$$g(\omega) = |a(\omega)| \sqrt{N \, \mathrm{var}(J_{ij})} = \frac{g}{\sqrt{1 + \tau^2 \omega^2}}. \tag{21}$$

We can use Eq (21) to study the scaling of frequency as g approaches the critical value of 1. In many cases, we can expect that the neuronal and synaptic dynamics lead to a smooth effective low-pass filtering of the recurrent input, such that for small frequency $g(0) - g(\omega) \propto \omega^2$. For the specific $g(\omega)$ in Eq (21), this can be directly verified. The low-pass filtering implies that the frequencies showing a critical covariance spectrum are those with $|\omega| = o\left((1 - g)^{\frac{1}{2}}\right)$.

Note, however, that the simple replacement by an effective g may not apply to a connectivity matrix that does not have i.i.d. entries. For example, for networks with non-zero reciprocal motifs, the covariance spectrum changes qualitatively with frequency (Section D.1 in S1 Text).

3.7 Sampling in time and space

The theoretical spectra we have discussed are based on the exact covariance matrix (Eq (2)). For neural data, this is equivalent to the limit of the sample covariance $\hat{C}$ (Eq (37)) when the number of time samples M is much larger than the number of neurons N. Note that if the activity data is first averaged or summed over a time window/bin (ΔT in Eq (37)) before calculating the sample covariance, then M is the number of bins. However, many large-scale neural recordings are in the so-called *high dimensional* regime, where N and M are comparable, that is, the ratio $\alpha = N/M$ remains finite or even greater than 1 for large N, M. It is thus important to study this effect of temporal sampling on the covariance eigenvalues to better relate to experimental data [7].

We refer to $\hat{C}$ and $p_{\hat{C}}(x)$ as the *time-sampled* covariance and spectrum. The relation between the original spectrum $p_C(x)$ and the time-sampled spectrum $p_{\hat{C}}(x)$ for a finite $\alpha \geq 0$ has been studied in [47]. The authors derived a general relation between the generating function of the eigenvalue distribution $W(z) = \sum_{n=1}^{\infty} z^n \mu_n$, where μ_n is the n-th moments of the eigenvalue distribution, and the counterpart $\hat{W}(z)$ for the sampled distribution,

$$\hat{W}(z \cdot (1 + \alpha W(z))) = W(z), \quad \text{and conversely} \quad W\left(\frac{z}{1 + \alpha \hat{W}(z)}\right) = \hat{W}(z). \qquad (22)$$

We give an alternative derivation of this result using free probability [48–50] (Section H in S1 Text), which allows us to also generalize to the spatial sampling case (see below). For simplicity, here we describe the results for $0 \leq \alpha \leq 1$. For $\alpha > 1$ where time samples are severely limited, the spectrum of the $N - M$ nonzero eigenvalues can be calculated with small modifications (Section H.2 in S1 Text).

One corollary of Eq (22) is a simple formula for how the (relative) dimension changes under time sampling

$$D(\hat{C}) = D(C) \frac{N}{N + \alpha D(C)}, \quad \hat{D}(\hat{C}) = \frac{\hat{D}(C)}{1 + \alpha \hat{D}(C)}. \qquad (23)$$

These formulas show that both D and $\hat{D} = D/N$ decrease with α (fewer time samples).

The relations Eqs (22) to (23) apply to any covariance matrix spectrum. For example, it reproduces the time-sampled dimension derived in [7] for a different model of covariance C. We now apply Eq 22 to the case of i.i.d. Gaussian connectivity to derive specific results of the time-sampled spectrum $p_{\hat{C}}(x) =: p_{g,\alpha}(x)$. Here Eq (22) becomes a cubic equation and can be solved analytically (Section H.3 in S1 Text). Consistent with the dimension, when α increases from 0 to 1, the support of the time-sampled distribution expands from both sides (Fig 7A). In particular, for any fixed $\alpha < 1$ (so $\hat{C}$ is positive definite), the left edge of the support x_- decreases with g but is always bounded away from 0 even as $g \to 1$, where $x_- \to \frac{2}{27}\left((1 + 3\alpha)^{\frac{3}{2}} + 1 - 9\alpha\right)$ (see also Fig L in S1 Text). Interestingly, the approximate power law of $p_C(x)$ (Eq 7) still holds under time sampling for any fixed α as $g \to 1$ (Section H.3 in S1 Text).

Another challenge for fitting to neural data is that often only a subset of neurons are observed instead of the entire local recurrent network. The unobserved neurons have an impact on the dynamics and affect the eigenvalues of the observed covariance matrix. We study this by considering randomly selecting $N_s = fN$, $0 < f \leq 1$ neurons and define their covariance $\tilde{C}$ as the *space-sampled* covariance. Using the free probability approach, we derive similar results as Eqs (22) and (23) (Section I in S1 Text) and apply them to derive the
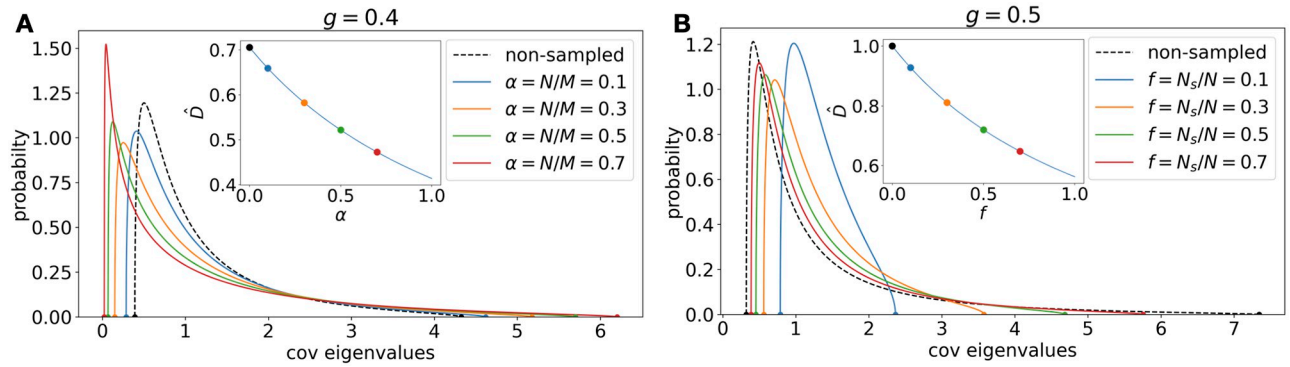
Fig 7. Effects of sampling in time and space on the covariance spectrum. A. For the i.i.d. Gaussian random connectivity, how different levels of time samples α change the spectrum (Eq (22)). The non-sampled case corresponds to $\alpha = 0$. g is fixed at 0.4. Inset: The relative dimension $\hat{D}$ vs. α (Eq 23). The dots correspond to the pdfs with matched colors. **B**. Same as A. but for the spatial subsampling (Section I.2 in S1 Text), at $g = 0.5$. The non-sampled case corresponds to $f = 1$. The relative dimension in the inset is based on Section I.1 in S1 Text.

https://doi.org/10.1371/journal.pcbi.1010327.g007

spectrum and dimension for the i.i.d. Gaussian connectivity under spatial sampling. In particular, the relative dimension $\hat{D}(\tilde{C}) = D(\tilde{C})/(fN)$ increases with spatial sampling (i.e., decrease f) which is consistent with the shape of the spectrum where its support narrows in (Fig 7). Lastly, the power-law feature is also preserved under spatial sampling. For any fixed $0 < f \leq 1$, we show that as $g \to 1$ and $x \to \infty$ (see examples with g close to 1 in Fig L in S1 Text)

$$p_{g,f}(x) \approx \frac{\sqrt{3}}{2\pi} f^{-\frac{1}{3}} x^{-\frac{5}{3}}. \tag{24}$$

3.8 Fitting the theoretical spectrum to data

Our theory for the bulk covariance spectrum can be fitted to neural activity whenever the covariance eigenvalues can be calculated. The best-fitting theoretical spectrum can be found by minimizing the L^2 or L^∞ error between the empirical and theoretical cumulative distributions (Methods) with respect to parameters such as g. We note that the availability of closed-form or analytic solutions of the theoretical distributions makes this optimization highly efficient.

In many settings, the value of the baseline neuronal variance σ^2 in Eq (2) is not known. But this can be easily addressed by scaling both the observed and the predicted eigenvalues to have a mean equal to 1. After fitting the connectivity parameter g for the normalized eigenvalues, σ^2 can then be estimated using the original means of data and theory. For our theoretical spectra, the mean μ of covariance eigenvalues is available in closed-form (Eqs (9) and (18)), 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 activity to z-scores). In this case, we need to analyze the eigenvalues of a *correlation matrix* whose entries are normalized as $C_{ij}/\sqrt{C_{ii}C_{jj}}$. Interestingly, we found that the correlation eigenvalue distribution for our random connectivity models in the large N limit is the same as the rescaled $p_R(x)$ above. This is because the diagonal entries of C become uniform (thus converge to μ) for large N (Section J in S1 Text).

The fitting of the spectrum is also robust to outliers in the covariance eigenvalues (Section 3.4). In Section K in S1 Text we demonstrate an example where a rank-2 component is added to the covariance C. Since in practice the rank of the perturbation is unknown *a priori*, we use
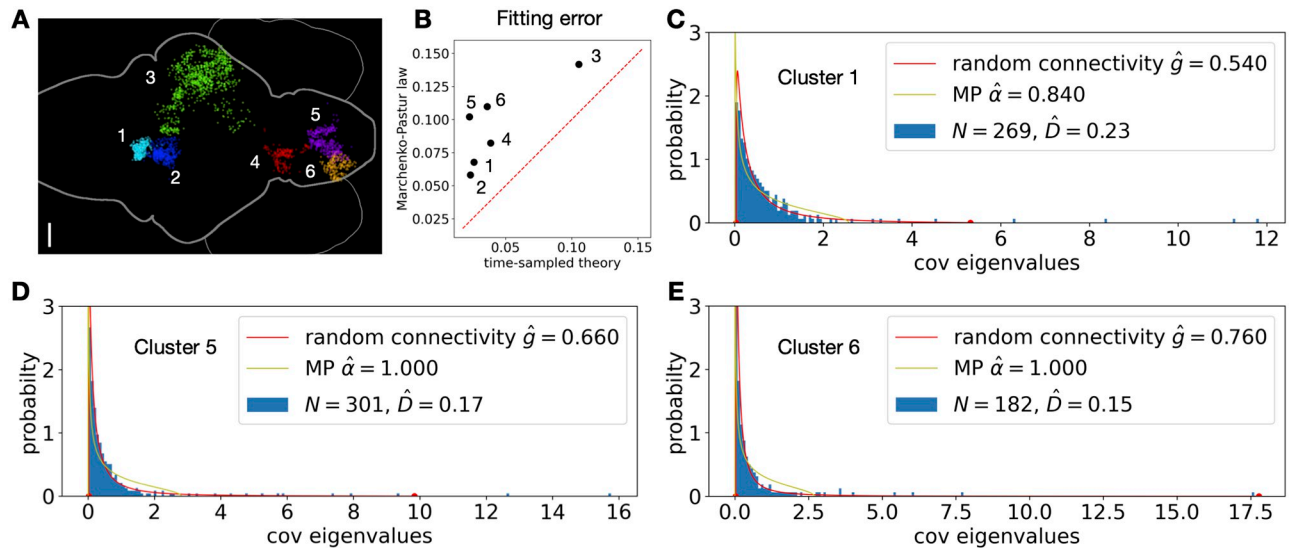
Fig 8. Fitting the theoretical spectrum to data. A. The anatomical map of neurons (dots) in the example functional clusters (different colors) across a larval zebrafish brain (scale bar is 50 μm, see text and [52]). **B**. Comparing the fitting error of the time-sampled random connectivity theory (Section 3.7) and the Marchenko–Pastur law. The errors are measured by Eq (38). The red dashed line is the diagonal. $\hat{D} = D/N$ labeled on each plot is the relative dimensionality (Eq 4). The calcium activity is recorded at a frame rate of 2 Hz and a total of 600 frames of spontaneous activity [52] are used in here. See more details in Methods. Fitting results for all other clusters are in Fig O in S1 Text.

all eigenvalues in the fitting, and the fitted *g* is highly accurate despite the presence of outliers. We can also use the fitted *g* to help identify the outliers by separating them out based on the upper edge of $p_C(x)$ support [32, 51].

We conclude with a preliminary application to whole-brain calcium imaging data in larval zebrafish. In [52], activities of the majority of the neurons in the larval zebrafish brain were imaged simultaneously during presentations of various visual stimuli and grouped into *functional clusters* based on their response similarity. These clusters reveal potential neural circuits and, in some cases, reveal a good match with known brain nuclei. Here we select a few clusters that contain a large number of neurons and are anatomically localized (Fig 8B). For each cluster, we calculate its sample correlation matrix during the spontaneous condition (no stimulus was presented) and then fitted the eigenvalues to the time-sampled spectrum with i.i.d. Gaussian connectivity (Section 3.7). Here the calcium activity is used but the correlation matrix is expected to be approximately the same if firing rates or long time window spike counts were used (Section K in S1 Text). Despite the simplicity of the model with only one parameter to tune, the results show good agreement with data and is significantly better than fitting using the Marchenko–Pastur law (Fig 8C), which models spatially independent noise (Section 5.4). Therefore, our theory provides a quantitative mechanistic explanation of how a long tail of covariance eigenvalues or equivalently low dimensional activity occurs in recurrent neural circuits.

4 Discussion

In this work, we studied the eigenvalue density distribution of the covariance matrix of a randomly connected neuronal network, whose activity is approximated as noise driven linear fluctuations around a steady state. We derived an explicit expression for eigenvalue distribution in the large-network limit analytically in terms of the statistics of the connectivity such as coupling strength and second-order motifs. Our results also include closed-form expressions for

the dimension measure generalizing known results [6]. Some of these dimensionality results are also derived in a recent parallel work [37], whereas the shape of the covariance spectrum was not studied in [37]. Knowing the exact shape and support of the bulk spectrum can facilitate separating outlying eigenvalues corresponding to low dimensional structure (coming from other unmodeled effects such as external input) from variability due to noise [51] (see an example in Fig N in S1 Text). The shape of the bulk spectrum reflects structured amplification of the neuronal noise by the random recurrent interactions. Since the bulk spectrum is not altered by low rank perturbations to the connectivity or to the activity (Section C in S1 Text), this allows for distinguishing different sources of variability in neural data. As the connection strength increases towards the critical edge of stability, the spectrum exhibits a power-law tail of large eigenvalues, with exponent $-5/3$ in pdf (or $-3/2$ in eigenvalue vs. rank plot). This power-law shape near the critical g provides concise theoretical characterizations of the spectra under various connectivities. Intriguingly, the same power law persists even when the shape of the spectrum is modified by connectivity motifs or due to finite temporal and spatial sample size. In contrast, when we move away from the asymmetric, random connectivity model, the exponent of the power law (if any) becomes different: $-7/4$ for symmetric random connectivity (Eq (16)), -2 for a normal connectivity J^n with matching eigenvalue distribution as i.i.d. Gaussian J (Section E.3 in S1 Text), and $-d/4 - 1$ for a d-dimensional ring network (see below). Based on these results, we conjecture that a power-law tail, whenever present for any covariance spectrum, reflects the qualitative nature of the connectivity and is a robust feature that will survive both temporal and spatial sampling with the same exponent (precise statement in Section I.3 in S1 Text).

Unlike the large eigenvalues [53], the interpretation of the bulk spectrum of PCA of neural activity data has received little attention. A notable exception is a recent work [54] which studied the power law of covariance spectrum of data near criticality based on the renormalization group method. By fitting experimental data to the theoretical spectrum, information from eigenvalues of all sizes can be used to estimate the effective connection strength (Fig 8). Our theory thus provides an important benchmark to compare with experimental data and advocates the bulk covariance spectrum as a powerful global description of collective dynamics in neuronal networks.

4.1 Nonlinear dynamics

One limitation of the work is the assumed dynamic regime where fluctuations of the neuronal activity are described by the linear response theory [22, 23] around a fixed point. While extending the theory to highly nonlinear activity such as chaotic dynamics [2] is left for future research, we provide some numerical examples of networks with nonlinear neurons to illustrate the applicability of our results.

We consider a model with nonlinear rate neurons driven by external noise by introducing an activation function $\phi(x) = \tanh(x)$ in Eq (1) [2] [55]. The nonlinearity transforms the currents $h_i(t)$ to firing rates $r_i(t) = \phi(h_i(t))$, and the recurrent interaction term is now $\sum_{j=1}^{N} J_{ij} r_i(t)$ (Eq (33) in Methods). When the connection strength $g = N\text{var}(J_{ij})$ and the noise magnitude σ are small, the neural activity will mainly fluctuate near 0 where $\phi(x)$ can be approximated by linearizing around 0 (Fig 9A, bottom). Indeed, the spectrum based on the linear theory (green dashed line in Fig 9A, top) approximates the numerical spectrum. For larger g and σ the deviation from the linear theory becomes significant as the firing rates are now strongly shaped by the nonlinearity (Fig 9B, 9C and 9D, bottom). Interestingly, these spectra can be well fitted by the linear-theory spectrum if g is replaced by a smaller effective $\hat{g}$. This reduction in the effective g can be qualitatively understood as the average slope of $\phi(x)$ over the distribution of $h_i(t)$
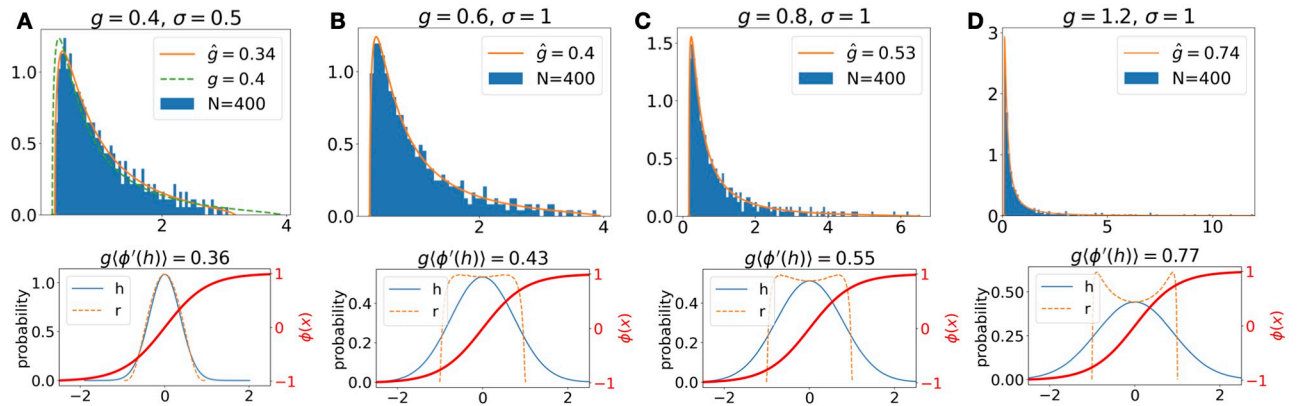
Fig 9. Covariance spectrum in nonlinear dynamics. A. Top: A histogram of covariance eigenvalues calculated from firing rate activities $r_i(t)$ of simulating a $N = 400$ network model according to Eq 33. The eigenvalues are normalized to have a mean equal to 1 for easy comparison of the shape (or equivalently the eigenvalues of the correlation matrix Section 3.8). Here $g = 0.4$ and $\sigma = 0.5$ (see Methods for additional numerical details). The green dashed line is the time-sampled theoretical spectrum (Section 3.7) using actual g and α, only shown in A for clarity of the plots. The length of the simulated data corresponds to $\alpha = N/T = 0.1$. The orange curve is also the time-sampled theoretical spectrum except for using an effective $\hat{g}$ that is fitted numerically to best match the simulated eigenvalues. Bottom: The blue curve is the histogram of $h_i(t)$ (aggregated across i and t) and the orange dashed curve is the histogram of $r_i(t)$. The overlaying red curve shows the nonlinearity $\phi(x)$ as a reference. The $\langle \cdot \rangle$ in the title of each plot represents averaging $\phi'(h_i(t))$ over all i and t. **B-D**. Same as A except for $\sigma = 1$ and $g = 0.6, 0.8, 1.2$, respectively. Only the fitted theoretical spectrum (orange curve) is shown for clarity.

(Fig 9 bottom). Note that, unlike the noiseless model in [2], the cases studied here are not chaotic even with $g > 1$. As shown in [55], the presence of external noise shifts the transition to chaotic dynamics from $g = 1$ to larger values. A theoretical characterization of effective g and other strongly nonlinear dynamics such as chaotic dynamics is left for future research. Future work could also consider cases with transient activity which are common in nonlinear systems and more general network architectures, such as multiple populations of EI networks [21] and incorporating distant dependent connectivity patterns based on known cortical microcircuit architectures [56–58].

4.2 Ordered vs. disorder connectivity

We have studied the covariance spectrum under random connectivity, which is used as a model for complex recurrent networks. Here we ask whether features of these spectra are distinct results of the connectivity being random. To address this, we briefly explored the covariance spectra from several widely used examples of ordered connectivity for comparison.

First, consider a ring network [56] with translation invariant long-range connections, where the connection strength between neurons depends smoothly on their distance (Fig 10A inset, see Methods). In the large-network limit, the covariance spectrum becomes a delta distribution at 1 with a few discretely located large eigenvalues (Fig 10A). Next, we consider the ring network with short-range, in particular, Nearest-Neighbor (NN) connections. The covariance spectrum is now continuous (no outliers) and supported on an interval, but the pdf diverges at both edges as $(\Delta x)^{-\frac{1}{2}}$ (Fig 10B).

To seek further examples of ordered connectivity leading to a qualitatively similar covariance spectrum as the random connectivity, we consider the d-dimensional generalization of the NN ring (Methods). As dimension d increases, the smoothness of the pdf within and at the edges of the support increases, and the covariance spectrum becomes qualitatively similar to the random case [59] (Fig 10). Interestingly, as the connection strength approaches its critical value for stability, the covariance spectra also exhibit a power-law tail $p_C(x) \propto x^{-\frac{d}{4}-1}$ (Section
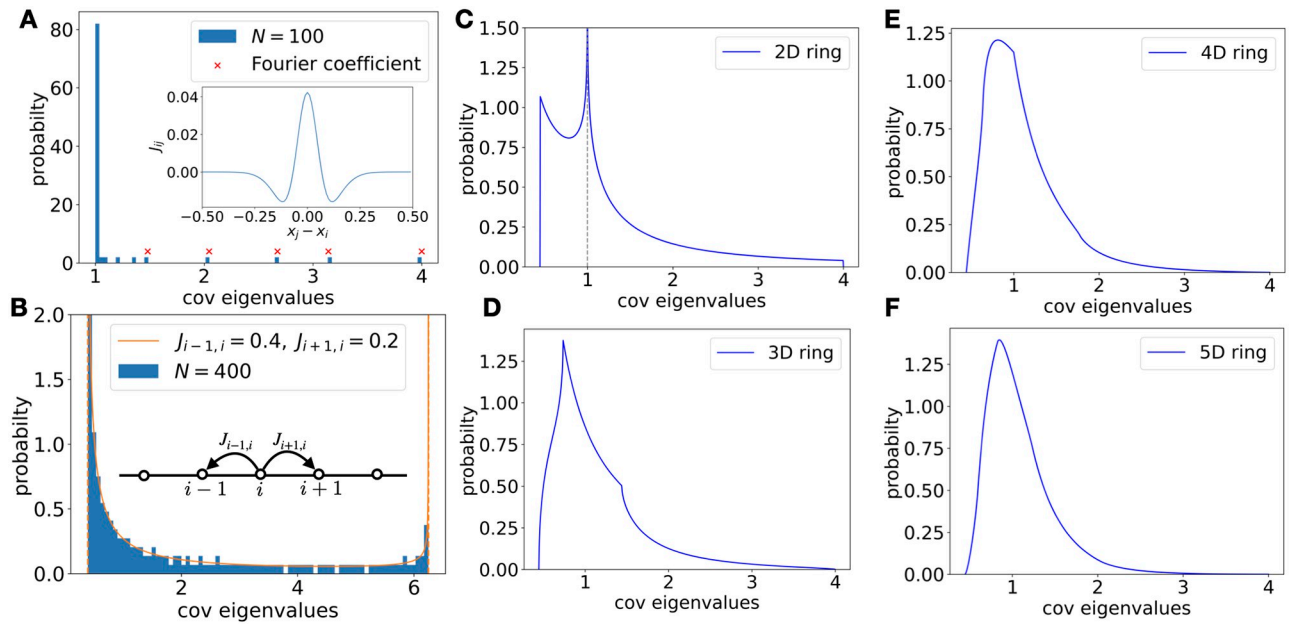
Fig 10. Covariance spectra under some deterministic connectivity models. A. Histogram of the covariance eigenvalues of a ring network with a long-range connection profile (inset, $N = 100$). Most eigenvalues are close to 1 and the rest of eigenvalues converge to discrete locations predicted by top Fourier coefficients (crosses) of the connection profile (Eq (36)). **B.** Same as A. but for a ring network with Nearest-Neighbor connections: $J_{i-1} = 0.4$, $J_{i+1,i} = 0.2$. The solid line is theoretical spectrum in large N limit which has two diverging singularities at both support edges. The effect of such singularities is also evident in the finite-size network at $N = 400$ (a single realization). **C-F.** Higher dimensional Nearest-Neighbor ring network ($ad = 0.6$, see Methods). As the dimension increases, the singularities in the pdf become milder and less evident, and the overall shape becomes qualitatively similar to the random connectivity case (Fig 1).

L.3.1 in S1 Text; $x^{-\frac{d}{2}-1}$ is also possible under other cases using [60]). To match the exponent of the random network d would be $8/3 \approx 2.67$. These comparisons suggest that the covariance spectrum's overall smooth density and long tail shape is a shared property in highly connected networks with high rank connectivity matrices, including random networks and high dimensional short-range spatially invariant networks.

5 Methods

5.1 Models of random connectivity

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

- i.i.d. Gaussian random connectivity $J_{ij} \sim \mathcal{N}(0, g^2/N)$: closed-form pdf and endpoints (Eq (5)), including the frequency-dependent covariance spectrum (Section 3.6), and a power-law tail approximation (Eq (7)).

- Gaussian random connectivity with reciprocal motifs/asymmetry $\kappa = \rho(J_{ij}, J_{ji})$ (Section 3.3): analytic solution and endpoints (quartic root, Section E in S1 Text) and a power-law tail approximation (Eq 17). For special case of symmetric and ant-symmetric connectivity, closed-form pdf Eqs (14) and (15), including a frequency dependent covariance spectrum (Section G in S1 Text).

- Erdős-Rényi and certain EI network Section 3.5: same bulk spectrum as the i.i.d. Gaussian case.

For all cases, the mean μ and the dimension D are derived in closed-form (Eqs 9, (10) and (18)).

For simplicity, we do not require J_{ii} to be zero (i.e., no self-coupling), but allow it, for example in the i.i.d. Gaussian model, to be distributed in the same way as other entries J_{ij}. In the 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 (Section A.3 in S1 Text).

5.2 Applications to alternative neuronal models and signal covariance

Although the relation between C and J (Eq (2)) is derived here in a linear rate neuron network Eq (1), it also arises in other models of networked systems.

Linearly interacting Poisson neurons. This is also called a multivariate Hawkes model [27]. This is a simple model for spiking neuron networks, but is versatile enough to capture for example the temporal spiking correlations seen in other more sophisticated nonlinear spiking neuron networks [26, 28]. A time-dependent Poisson firing rate is calculated as a filtered input spike train $s_j(t)$ (sum of delta functions), and spikes are then drawn as a Poisson process given $y_i(t)$,

$$y_i(t) = y_0 + \int_0^\infty A(t - \tau)\left(\sum_j W_{ij}s_j(t - \tau)\right)d\tau. \tag{25}$$

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 [27]

$$C = (I - aW)^{-1}C^0(I - aW)^{-T}, \quad C^0 = \text{diag}\{Y_1, Y_2, , \ldots, Y_N\}, \quad Y = (I - aW)^{-1}Y^0, \tag{26}$$

which is valid if the time varying $y_i(t)$ does not often become 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 (26) becomes

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

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

Another condition that ensures a uniform Y and does not restrict connections to be weak is a *row balance* condition of W sometimes assumed in EI networks [45],

$$\sum_{j=1}^N J_{i_1 j} = \sum_{j=1}^N J_{i_2 j}. \tag{27}$$

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

Integrate-and-Fire neurons. As shown in [23, 26] the linear response theory [22] can be used to approximately describe the covariance structure of a network of generalized integrate-

and-fire (IF) neurons

$$\tau \frac{dV_i}{dt} = -(V_i - E_{L,i}) + \psi(V_i) + E_i + \sqrt{\sigma_i^2 \tau_i} \xi_i(t) + \sum_j W_{ij} F_{ij}(t) * y_j(t). \tag{28}$$

Here V_i is the membrane potential and a spike is generated when V_i reaches a threshold. $y_i(t) = \Sigma_k \delta(t - t_{i,k})$ is the spike train, and $y_i^0(t)$ in Eq (29) is the "unperturbed" spike train in absence of recurrent connections W. Different choices of $\psi(V)$ realize types of IF neurons, such as $\psi(V) = \Delta_T \exp((V - V_T)/\Delta_T)$ for the exponential IF neurons. During the asynchronous firing of neurons (no strong synchronized firing across the network), Eq (28) can be well approximated by

$$\Delta y_i(t) = A_i(t) * \left(\sum_{j=1}^{N} W_{ij} F_{ij}(t) * \Delta y_j(t) \right) + \Delta y_i^0(t), \quad i = 1, 2, \ldots, N. \tag{29}$$

Here $a(t) * b(t) = (a * b)(t) = \int_0^t a(s)b(t - s)ds$ denotes convolution. $W = \{W_{ij}\}$ is the matrix of recurrent connection weights. $A_i(t)$ is the linear response kernel for neuron i (e.g., an exponential decay) $F_{ij}(t)$ is the temporal kernel of the synapse. For simplicity, we assume that both A and F are uniform across the network.

It it shown [23, 26] that the long time window spike count covariance matrix C (in fact also the frequency covariance Eq (19)) is well approximated by

$$C = a^2 \langle (\Delta y_i^0(t))^2 \rangle (I - aW)^{-1} (I - aW)^{-T}.$$

Here the scalar $a = (\int_0^\infty A(t)dt)(\int_0^\infty F(t)dt)$ summarizes the cellular and synaptic dynamics. $\langle (\Delta y_i^0(t))^2 \rangle$ can be thought of as the baseline neuronal variance in the absence of recurrent connections ($A = 0$ in Eq (29). This expression of the covariance matrix again matches with Eq 2.

Fixed points 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 the *noise covariance* [61]. We can also consider the spectrum for a *signal covariance*. This perspective is needed to use our results to interpret experimental data where the neural activity is largely driven by stimuli for example [10].

Consider a network of linear firing rate neurons,

$$\tau \frac{dr_i}{dt} = -r_i + \sum_{j=1}^{N} J_{ij} r_j + u_i, \quad i = 1, 2, \ldots, N. \tag{30}$$

Here u_i is the external input to neuron i. Assume the network settles to a steady state, so the fixed point firing rates are given by

$$\vec{r} = (I - J)^{-1} \vec{u}. \tag{31}$$

Now consider the network activity across an ensemble of input patterns, which has whitened statistics [62],

$$\text{var}(u_i) = \sigma^2, \quad \text{cov}(u_i, u_j) = 0. \tag{32}$$

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

We note that Eq (31) 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) [29].

Nonlinear rate neurons.   The following is a simple and classic nonlinear rate neuron model similar to that used in [2],

$$\tau \dot{h}_i(t) = -h_i(t) + \sum_{j=1}^{N} J_{ij} \phi(h_j(t)) + \xi_i(t), \quad i = 1, \ldots, N. \tag{33}$$

Here $\phi(x) = \tanh(x)$ converts currents $h_i(t)$ into firing rates $r_i(t) = \phi(h_i(t))$. The white noise ξ_i and i.i.d. Gaussian random J are the same as in Eqs (1) and (3). The presence of noise allows for nontrivial activity for $g < 1$. Due to symmetry of $\phi(x)$ and J distribution, the average activity over time $\langle h_i(t) \rangle = 0$ for large N. Note that the largest slope of $\phi(x)$ is 1 at $x = 0$, our main model Eq (1) can be viewed as a linear approximation to Eq (33) and the definition of connection strength g is consistent.

In Fig 9, $N = 400$, $\tau = 1$ and Eq (33) is simulated using the Euler-Maruyama method with a time step of 0.01 for a duration of $T_0 = 40000$. To get the long time window covariance, the firing rates $r_i(t)$ are binned by a time window of 10, which is sufficient based on examining the decay of its autocorrelation function. After binning, the simulated data correspond to a time-sample parameter $\alpha = N/T = 0.1$ (3.7) and this α is used in calculating the theoretical spectra. The empirical covariance matrix is then calculated from the binned firing rates according to Eq (37). As shown in Fig 9, the theoretical spectrum based on the linear dynamics can still be used to fit the simulated spectrum in this nonlinear network when g is replaced by a smaller effective $\hat{g}$.

5.3 Power-law approximation of the eigenvalue distribution

The power-law property of $p_C(x)$ for i.i.d. J_{ij} under critical g is probably known in random matrix theory (private communication), by results from the equivalent distribution studied in [30], although we do not know of a specific reference. We include a derivation based on the explicit expression of $p_C(x)$ (Eq (5)) that is outlined below.

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 the Taylor expansion of $(1 - g^2)$ 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,

$$x \to \infty, \quad x(1 - g^2)^3 \to 0. \tag{34}$$

Note that since $x_+/x_- \sim (1 - g^2)^{-3}$, there is plenty range of x to satisfy the above for strong connections when g is close to 1. Under these limits, Eq (5) greatly simplifies as various terms vanish leading to

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

This explains the validity of the 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* $x \to \infty$), then there is a simpler alternative derivation that can we also apply to other connectivity models (see Section A.2 in S1 Text).

5.4 Comparison with the Marchenko–Pastur distribution

The Marchenko–Pastur distribution is widely used for modeling covariance eigenvalues arising from noise [32, 51, 63]. It is also the limit of the time-sampled spectrum $p_{g,\alpha}(x)$ (Fig 7 and Section 3.7) at weak connections $g = 0$. The Marchenko–Pastur law has one shape parameter α. We focus on the case when the covariance is positive definite which restricts $0 < \alpha < 1$ (otherwise there is a delta distribution at 0) and the pdf is

$$p_{MP}(x) = \frac{\sqrt{(\alpha_+ - x)(x - \alpha_-)}}{2\pi\alpha x}, \quad \alpha_\pm = (1 \pm \sqrt{\alpha})^2, \tag{35}$$

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.

5.5 Deterministic connectivity

5.5.1 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 $x_i - x_j$ is therefore translation invariant.

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

$$J_{ij} = \frac{1}{N}f(x_i - x_j) = \frac{1}{N}f\left(\frac{i - j}{N}\right). \tag{36}$$

In the large-network limit, the covariance eigenvalues have an approximate delta distribution at 1 except for a finite number of discretely located larger eigenvalues (Fig 10A). A precise statement of this result is described in Section L.1 in S1 Text. The outlying eigenvalues correspond to the leading 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 $N \to \infty$.

5.5.2 Multi-dimensional ring network. For a 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 10).

To characterize the shape of the covariance spectrum (Fig 10), we further simplify by setting $J^k_{i_k-1,i_k} = J^k_{i_k+1,i_k} = a$ (see also Section L.3 in S1 Text for motivations based on 1D ring) and analytically derived $p_C(x)$ (Section L.3.2 in S1 Text). 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 identify in empirical eigenvalue distributions from a finite-size network (not shown).

5.6 Fitting the theoretical spectrum to data

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

$$C_{ij} = \frac{1}{\Delta T}\frac{1}{M-1}\sum_{t=1}^{M}(s_i(t) - \bar{s}_i)(s_j(t) - \bar{s}_j), \quad \bar{s}_i = \frac{1}{M}\sum_{t=1}^{M}s_i(t). \tag{37}$$

For calcium imaging data, the fluorescence is approximately integrating the spikes over the indicator time constant. So we can still apply Eq (37) 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 the theory, Section 3.8).

We fit the theoretical spectrum to empirical eigenvalues by finding the connectivity parameter g that minimizes 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 the theoretical pdf (Eq 5) to get its cdf. As seen below, the theoretical cdf only needs to be calculated at the empirical eigenvalues.

Motivated by methods of hypothesis testing on distributions, we measure the L^2 norm cdf error using the Cramer-von Mises statistic

$$D_{CvM}^2 = \int (F(x) - F_n(x))^2 dF_n(x) = \frac{1}{12n^2} + \frac{1}{n}\sum_{i=1}^{n}\left(F(x_i) - \frac{2i-1}{2n}\right)^2 \tag{38}$$

Here n is the number of samples and x_i are the i-th empirical eigenvalues. Alternatively, the error can also be measured under L^∞ norm based on the Kolmogorov-Smirnov statistic

$$D_{KS} = \sup_x |F_n(x) - F(x)|. \tag{39}$$

where x_i are samples. Our code implements both measures.

In Fig 8B–8E, we fit the time-sampled theoretical spectrum with i.i.d. Gaussian connectivity (Section 3.7) to calcium imaging data in larval zebrafish [52]. The theoretical spectrum (once normalized by the mean, see Section 3.8) depends on two parameters g and α, but the latter is fixed to be N/M based on the data. Here N is the number of neurons in a cluster, and M is the number of time frames used in calculating the sample correlation matrix (Eq (37)). The calcium fluorescence $\Delta F/F$ traces of each neuron are normalized to z-scores [52], which is consistent with calculating the eigenvalues of the correlation matrix (Section 3.8). g is then optimized to minimize the Cramer-von Mises error (Eq 37) between the data. The largest eigenvalue for each cluster is often much larger than the rest and is thus removed before the fitting. For comparison, we fit the same data to the Marchenko–Pastur law (Section 5.4) whose shape depends on the parameter α. Here we allow α to vary so that both models (random connectivity and MP law) have one parameter to be optimized to fit data.

Supporting information

S1 Text. Supplementary material. Analytical derivations and additional figures.
(PDF)

References

1. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. Nature. 1998; 393(6684):440–2. https://doi.org/10.1038/30918 PMID: 9623998

2. Sompolinsky H, Crisanti a, Sommers HJ. Chaos in random networks. Phys Rev Lett. 1988; 61(3):259–262. https://doi.org/10.1103/PhysRevLett.61.259 PMID: 10039285

3. Ginzburg I, Sompolinsky H. Theory of correlations in stochastic neural networks. Physical Review E. 1994; 50(4):3171–3191. https://doi.org/10.1103/PhysRevE.50.3171 PMID: 9962363

4. van Vreeswijk C, Sompolinsky H. Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity. Science. 1996; 274(5293):1724 LP–1726. https://doi.org/10.1126/science.274.5293.1724 PMID: 8939866

5. Hu Y, Trousdale J, Josić K, Shea-Brown E. Motif statistics and spike correlations in neuronal networks. Journal of Statistical Mechanics: Theory and Experiment. 2013; 2013(03):P03012. https://doi.org/10.1088/1742-5468/2013/03/P03012

6. Dahmen D, Grün S, Diesmann M, Helias M. 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. 2019; 116(26):13051–13060. https://doi.org/10.1073/pnas.1818972116 PMID: 31189590

7. Mazzucato L, Fontanini A, La Camera G. Stimuli reduce the dimensionality of cortical activity. Frontiers in Systems Neuroscience. 2016; 10(FEB). https://doi.org/10.3389/fnsys.2016.00011 PMID: 26924968

8. Litwin-Kumar A, Harris KD, Axel R, Sompolinsky H, Abbott LF. Optimal Degrees of Synaptic Connectivity. Neuron. 2017; 93(5):1153–1164.e7. https://doi.org/10.1016/j.neuron.2017.01.030 PMID: 28215558

9. Semedo JD, Zandvakili A, Machens CK, Yu BM, Kohn A. Cortical Areas Interact through a Communication Subspace. Neuron. 2019; 102(1):249–259.e4. https://doi.org/10.1016/j.neuron.2019.01.026 PMID: 30770252

10. Stringer C, Pachitariu M, Steinmetz N, Carandini M, Harris KD. High-dimensional geometry of population responses in visual cortex. Nature. 2019. https://doi.org/10.1038/s41586-019-1346-5 PMID: 31243367

11. Recanatesi S, Ocker GK, Buice MA, Shea-Brown E. Dimensionality in recurrent spiking networks: Global trends in activity and local origins in connectivity. PLoS Computational Biology. 2019; 15(7):1–29. https://doi.org/10.1371/journal.pcbi.1006446 PMID: 31299044

12. Sadtler PT, Quick KM, Golub MD, Chase SM, Ryu SI, Tyler-Kabara EC, et al. Neural constraints on learning. Nature. 2014; 512(7515):423–426. https://doi.org/10.1038/nature13665 PMID: 25164754

13. Gallego JA, Perich MG, Naufel SN, Ethier C, Solla SA, Miller LE. Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. Nature Communications. 2018; 9(1):1–13. https://doi.org/10.1038/s41467-018-06560-z PMID: 30315158

14. Fusi S, Miller EK, Rigotti M. Why neurons mix: High dimensionality for higher cognition. Current Opinion in Neurobiology. 2016; 37:66–74. https://doi.org/10.1016/j.conb.2016.01.010 PMID: 26851755

15. Farrell M, Recanatesi S, Moore T, Lajoie G, Shea-Brown E. Recurrent neural networks learn robust representations by dynamically balancing compression and expansion. bioRxiv. 2019; p. 1–20.

16. Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB. Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol. 2005; 3(3):e68. https://doi.org/10.1371/journal.pbio.0030068 PMID: 15737062

17. Perin R, Berger TK, Markram H. A synaptic organizing principle for cortical neuronal groups. P Natl Acad Sci USA. 2011; 108(13):5419–5424. https://doi.org/10.1073/pnas.1016051108 PMID: 21383177

18. Sporns O, Ko R. Motifs in Brain Networks. PLoS Biol. 2004; 2. https://doi.org/10.1371/journal.pbio.0020369 PMID: 15510229

19. Zhao L, Beverlin B, Netoff T, Nykamp DQ. Synchronization from Second Order Network Connectivity Statistics. Front Comput Neurosci. 2011; 5:1–16. https://doi.org/10.3389/fncom.2011.00028 PMID: 21779239

20. Hu Y, Trousdale J, Josić K, Shea-Brown E. Local paths to global coherence: Cutting networks down to size. Physical Review E—Statistical, Nonlinear, and Soft Matter Physics. 2014; 89(3):1–16.

21. Kadmon J, Sompolinsky H. Transition to chaos in random neuronal networks. Physical Review X. 2015; 5(4):1–28. https://doi.org/10.1103/PhysRevX.5.041030

22. Lindner B, Doiron B, Longtin A. Theory of oscillatory firing induced by spatially correlated noise and delayed inhibitory feedback. Physical Review E. 2005; 72(6):061919. https://doi.org/10.1103/PhysRevE.72.061919 PMID: 16485986

23. Trousdale J, Hu Y, Shea-Brown E, Josić K. Impact of Network Structure and Cellular Response on Spike Time Correlations. PLoS Comput Biol. 2012; 8(3):e1002408. https://doi.org/10.1371/journal.pcbi.1002408 PMID: 22457608

24. Bair W, Zohary E, Newsome WT. 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. 2001; 21(5):1676–1697. https://doi.org/10.1523/JNEUROSCI.21-05-01676.2001

25. Gardiner C. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. Berlin: Springer-Verlag; 2009.

26. Grytskyy D, Tetzlaff T, Diesmann M, Helias M. A unified view on weakly correlated recurrent networks. Frontiers in Computational Neuroscience. 2013; 7(October):1–19. https://doi.org/10.3389/fncom.2013.00131 PMID: 24151463

27. Hawkes AG. Spectra of some self-exciting and mutually exciting point processes. Biometrika. 1971; 58 (1):83–90. https://doi.org/10.1093/biomet/58.1.83

28. Pernice V, Staude B, Cardanobile S, Rotter S. How Structure Determines Correlations in Neuronal Networks. PLoS Computational Biology. 2011; 7(5):e1002059. https://doi.org/10.1371/journal.pcbi.1002059 PMID: 21625580

29. Améndola C, Dettling P, Drton M, Onori F, Wu J. Structure Learning for Cyclic Linear Causal Models. arXiv. 2020; p. 1–19.

30. Bai ZD. Circular law. Annals of Probability. 1997; 25(1):494–529. https://doi.org/10.1214/aop/1024404298

31. Sommers HJ,Crisanti a, Sompolinsky H, Stein Y. Spectrum of large random asymmetric matrices. Physical Review Letters. 1988; 60(19):1895–1898. https://doi.org/10.1103/PhysRevLett.60.1895 PMID: 10038170

32. Gavish M, Donoho DL. The Optimal Hard Threshold for Singular Values is 4/sqrt(3). IEEE Transactions on Information Theory. 2014; 60(8):5040–5053. https://doi.org/10.1109/TIT.2014.2323359

33. Rajan K, Abbott LF, Sompolinsky H. Interactions between Intrinsic and Stimulus-Evoked Activity in Recurrent Neural Networks. In: Ding M, Glanzman DL, editors. The Dynamic Brain: An Exploration of Neuronal Variability and Its Functional Significance. Oxford University Press; 2011. p. 65–82.

34. Girko VL. Circular law. Theory of probability and its applications. 1983;.

35. Götze F, Tikhomirov A. The circular law for random matrices. Annals of Probability. 2010; 38(4):1444–1491.

36. Meshulam L, Gauthier JL, Brody CD, Tank DW, Bialek W. Coarse graining, fixed points, and scaling in a large population of neurons. Physical Review Letters. 2019; 123(17):178103. https://doi.org/10.1103/PhysRevLett.123.178103 PMID: 31702278

37. Dahmen D, Recanatesi S, Ocker GK, Jia X, Helias M, Shea-Brown E. Strong coupling and local control of dimensionality across brain areas. bioRxiv. 2020.

38. Hu Y, Brunton SL, Cain N, Mihalas S, Kutz JN, Shea-brown E. Feedback through graph motifs relates structure and function in complex networks. Physical Review E. 2018; 062312:1–25.

39. Mastrogiuseppe F, Ostojic S. Linking Connectivity, Dynamics, and Computations in Low-Rank Recurrent Neural Networks. Neuron. 2018; 99(3):609–623.e29. https://doi.org/10.1016/j.neuron.2018.07.003 PMID: 30057201

40. Rivkind A, Barak O. Local Dynamics in Trained Recurrent Neural Networks. Physical Review Letters. 2017; 118(25):1–5. https://doi.org/10.1103/PhysRevLett.118.258101 PMID: 28696758

41. Amit DJ, Gutfreund H, Sompolinsky H. Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks. Phys Rev Lett. 1985;(September):1530–1533. https://doi.org/10.1103/PhysRevLett. 55.1530 PMID: 10031847

42. Schuessler F, Mastrogiuseppe F, Dubreuil A, Ostojic S, Barak O. The interplay between randomness and structure during learning in RNNs. In: Larochelle H, Ranzato M, Hadsell R, Balcan MF, Lin H, editors. Advances in Neural Information Processing Systems. vol. 33. Curran Associates, Inc.; 2020. p. 13352–13362. Available from: https://proceedings.neurips.cc/paper/2020/file/ 9ac1382fd8fc4b631594aa135d16ad75-Paper.pdf.

43. Tao T. Outliers in the spectrum of iid matrices with bounded rank perturbations. Probability Theory and Related Fields. 2013; 155(1-2):231–263. https://doi.org/10.1007/s00440-011-0397-9

44. Horn RA, Johnson CR. Matrix Analysis. Cambridge University Press; 1990.

45. Rajan K, Abbott LF. Eigenvalue spectra of random matrices for neural networks. Phys Rev Lett. 2006; 97:188104. https://doi.org/10.1103/PhysRevLett.97.188104 PMID: 17155583

46. Vladimirov N, Mu Y, Kawashima T, Bennett DV, Yang CT, Looger LL, et al. Light-sheet functional imaging in fictively behaving zebrafish. Nature Methods. 2014; 11(9):883–884. https://doi.org/10.1038/ nmeth.3040 PMID: 25068735

47. Burda Z, Görlich A, Jarosz A, Jurkiewicz J. Signal and noise in correlation matrix. Physica A: Statistical Mechanics and its Applications. 2004; 343(1-4):295–310. https://doi.org/10.1016/j.physa.2004.05.048

48. Voiculescu D. Multiplication of certain non-commuting random variables. Journal of Operator Theory. 1987; 18:2223–2235.

49. Mingo JA, Speicher R. Free Probability and Random Matrices; 2017.

50. Speicher R. Free Probability Theory and Random Matrices. In: Vershik AM, Yakubovich Y, editors. Asymptotic Combinatorics with Applications to Mathematical Physics. 1. Springer, Berlin, Heidelberg; 2003. p. 53–73.

51. Peyrache A, Benchenane K, Khamassi M, Wiener SI, Battaglia FP. Principal component analysis of ensemble recordings reveals cell assemblies at high temporal resolution. Journal of Computational Neuroscience. 2010; 29(1-2):309–325.

52. Chen X, Mu Y, Hu Y, Kuan AT, Nikitchenko M, Randlett O, et al. Brain-wide Organization of Neuronal Activity and Convergent Sensorimotor Transformations in Larval Zebrafish. Neuron. 2018; 100(4):876–890.e5. https://doi.org/10.1016/j.neuron.2018.09.042 PMID: 30473013

53. Okun M, Steinmetz Na, Cossell L, Iacaruso MF, Ko H, Barthó P, et al. Diverse coupling of neurons to populations in sensory cortex. Nature. 2015; 521(7553):511–515. https://doi.org/10.1038/nature14273 PMID: 25849776

54. Bradde S, Bialek W. PCA Meets RG. Journal of Statistical Physics. 2017; 167(3-4):462–475. https://doi. org/10.1007/s10955-017-1770-6 PMID: 30034029

55. Schuecker J, Goedeke S, Helias M. Optimal Sequence Memory in Driven Random Networks. Physical Review X. 2018; 8(4):041029. https://doi.org/10.1103/PhysRevX.8.041029

56. Ben-Yishai R, Lev Bar-Or R, Sompolinsky H. Theory of orientation tuning in visual cortex. Proceedings of the National Academy of Sciences of the United States of America. 1995; 92(9):3844–3848. https:// doi.org/10.1073/pnas.92.9.3844 PMID: 7731993

57. Burak Y, Fiete IR. Accurate path integration in continuous attractor network models of grid cells. PLoS Computational Biology. 2009; 5(2). https://doi.org/10.1371/journal.pcbi.1000291 PMID: 19229307

58. Levy RB, Reyes AD. Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. Journal of Neuroscience. 2012; 32(16):5609–5619. https://doi.org/10.1523/ JNEUROSCI.5158-11.2012 PMID: 22514322

59. Goldberg J, Rokni U, Sompolinsky H. Patterns of Ongoing Activity and the Functional Architecture of the Primary Visual Cortex. Neuron. 2004; 42:489–500. https://doi.org/10.1016/S0896-6273(04)00197-7 PMID: 15134644

60. Wang X. Volumes of Generalized Unit Balls. Mathematics Magazine. 2005; 78(5):390–395. https://doi. org/10.2307/30044198

61. Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. Nature reviews Neuroscience. 2006; 7(5):358–366. https://doi.org/10.1038/nrn1888 PMID: 16760916

62. Field DJ. Relations between the statistics of natural images and the response properties of cortical cells. Journal of the Optical Society of America A. 1987; 4(12):2379. https://doi.org/10.1364/JOSAA.4. 002379 PMID: 3430225

63. Marchenko VA, Pastur LA. Distribution of eigenvalues for some sets of random matrices. Mathematics of the USSR-Sbornik. 1967; 1(4):457. https://doi.org/10.1070/SM1967v001n04ABEH001994