Connectivity Structure and Dynamics of Nonlinear Recurrent Neural Networks

David G. Clark,[1,2,*,‡] Owen Marschall,[1] Alexander van Meegen,[3,§] and Ashok Litwin-Kumar[1,2,†]

[1]*Zuckerman Institute, Columbia University, New York, New York 10027, USA*
[2]*Kavli Institute for Brain Science, Columbia University, New York, New York 10027, USA*
[3]*Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA*

(Received 6 September 2024; accepted 29 September 2025; published 3 November 2025)

Studies of the dynamics of nonlinear recurrent neural networks often assume independent and identically distributed couplings, but large-scale connectomics data indicate that biological neural circuits exhibit markedly different connectivity properties. These include rapidly decaying singular-value spectra and structured singular-vector overlaps. Here, we develop a theory to analyze how these forms of structure shape high-dimensional collective activity in nonlinear recurrent neural networks. We first introduce the random-mode model, a random-matrix ensemble related to the singular-value decomposition that enables control over the spectrum and right-left mode overlaps. Then, using a novel path-integral calculation, we derive analytical expressions that reveal how connectivity structure affects features of collective dynamics: the dimension of activity, which quantifies the number of high-variance collective-activity fluctuations, and the temporal correlations that characterize the timescales of these fluctuations. We show that connectivity structure can be invisible in single-neuron activities, while dramatically shaping collective activity. Furthermore, despite the nonlinear, high-dimensional nature of these networks, the dimension of activity depends on just two connectivity parameters—the variance of the couplings and the effective rank of the coupling matrix, which quantifies the number of dominant rank-one connectivity components. We contrast the effects of single-neuron heterogeneity and low-dimensional connectivity, making predictions about how z-scoring data affects the dimension of activity. Finally, we demonstrate the presence of structured overlaps between left and right modes in the *Drosophila* connectome, incorporate them into the theory, and show how they further shape collective dynamics.

DOI: 10.1103/2jt7-c8cq

Subject Areas: Biological Physics, Nonlinear Dynamics, Statistical Physics

I. INTRODUCTION

The collective activity of a high-dimensional nonlinear system is determined by the structure of the interactions between its elements, but the mapping from structure to activity is generally analytically tractable only for limited forms of structure. In systems neuroscience, structure in synaptic connectivity controls the function of a neural system by determining what patterns of activity are produced by populations of neurons. This paper studies this mapping in the context of nonlinear recurrent neural networks, used widely in neuroscience as models of neural-circuit dynamics [1–3] and in machine learning as systems for sequence processing [4–7]. While these models omit many aspects of actual neural circuits, they capture several of their fundamental features, including their large scale, nonlinear units, and recurrent interactions. Functionally, a key property of such systems is that, like real neural circuits, they can generate rich time-varying activity in the absence of external input [8,9].

Theoretical studies of nonlinear recurrent neural networks, including the pioneering work of Sompolinsky *et al.* (Ref. [10]), often focus on the case of independent and identically distributed (i.i.d.) couplings or simple variants thereof [11]. Such disordered networks produce high-dimensional chaotic activity, reminiscent of asynchronous cortical activity observed *in vivo* [10]. Initializing recurrent neural networks so that they exhibit chaotic dynamics has also been shown to facilitate subsequent learning [12–16].

*Contact author: dgclark@fas.harvard.edu
†Contact author: a.litwin-kumar@columbia.edu
‡Present address: Kempner Institute for the Study of Natural and Artificial Intelligence, Harvard University, Cambridge, Massachusetts 02138, USA.
§Present address: School of Life Sciences and School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Experimental access to large-scale connectivity structure creates an opportunity and imperative to develop theories that explain models beyond those with i.i.d. couplings. Recently, large-scale synaptic reconstructions of neural circuits, or "connectomes," have become available in multiple species, including the full brain of *Drosophila melanogaster* [17] and increasingly large portions of mammalian circuits [18–21]. These large-scale connectomes reveal that neural circuits deviate substantially from the i.i.d. assumption, as previously demonstrated via the presence of small-scale motifs [22]. In particular, in accordance with real-world networks more generally [23], neural circuits appear to exhibit low-rank structure in which the connectivity is well described by fewer rank-one components than the network size (we demonstrate this explicitly for the *Drosophila* connectome in Sec. II A). In neural-network models, recurrent or otherwise, such structure emerges naturally in several contexts, including when connections depend on distance in physical or feature spaces [24–26] or when networks are trained on tasks with low-dimensional structure [27–29].

In parallel with the development of large-scale connectivity maps, recent advances in recording technologies, including silicon probes [30] and large-scale calcium imaging [31,32], now enable the activities of many neurons to be monitored simultaneously, with the latest datasets including hundreds of thousands of neurons or more [33]. This technological progress has enabled researchers to characterize collective properties of neural activity rather than focusing solely on single-neuron responses. These collective properties capture how neurons interact and reveal distributed computational processes that are visible only with high-yield recording technologies [34]. Such population-level features are typically analyzed using dimensionality reduction techniques like principal components analysis [35–38].

A. Relating connectivity and collective-activity structure analytically

We currently lack analytical tools to predict relationships between connectivity structure, described by connectomics or other datasets, and collective activity, described by large-scale neural recordings. Dynamical mean-field theory (DMFT) is a theoretical tool from statistical physics widely used to analytically characterize activity in large nonlinear recurrent neural networks. Although recent advances in DMFT have extended these techniques to describe collective features in networks producing high-dimensional chaotic activity [39], these calculations have been limited to i.i.d. connectivity, or connectivity with simple correlations between reciprocal couplings. Extending such calculations to more complicated connectivity structures rapidly becomes unwieldy. Consequently, such analyses have not been performed for networks whose connections are constrained by the statistics of connectomic or other datasets, limiting our ability to leverage these datasets for theoretical insight.

One tractable approach has been to study networks with *very* low-dimensional connectivity and thus activity (specifically, in the limit where the rank of the coupling matrix remains finite while the network becomes large; see Sec. II C). However, connectomes, like that of *Drosophila* considered in Sec. II A, do not support this finite-rank assumption. Furthermore, while neural activity recorded in experiments is often low dimensional, this may be inherited from the low dimensionality of experimental tasks; in more complicated tasks or spontaneous states, activity is typically higher dimensional [31–33].

To address these limitations and challenges, this paper makes two contributions.

(1) A tractable model of neural-network coupling matrices that we call the random-mode model. Motivated by the *Drosophila* connectome considered in Sec. II A, we introduce the random-mode model in Sec. II B, in which the couplings are generated as a sum of rank-one outer products of random left and right modes with component-specific strengths. Unlike well-studied low-rank recurrent neural-network models, we scale the rank with the network size (Sec. II C). We show in Sec. II D that parametrizing these strengths provides control over the singular-value spectrum of the coupling matrix.

(2) A new path-integral calculation of a specific four-point function of network activity, originally defined by Clark *et al.* (Ref. [39]), that captures features of collective activity including its dimensionality. We review these and other summary statistics in Sec. III. In Secs. IV A–IV C, we describe the calculation, based on fluctuations around the saddle point in a path integral, and apply it to the random-mode model. In Sec. IV D, we use the solution to study how features of the coupling matrix control collective activity.

Additionally, in Sec. V, we analyze notable limiting cases of the random-mode model. In Sec. VI, we study a generalization of the model featuring heterogeneity among single-neuron properties and contrast these effects with those of low-rank structure in connectivity. Finally, in Sec. VII, we incorporate structured overlaps between the left and right modes into the model, demonstrating their presence in the *Drosophila* connectome and showing how they further shape collective dynamics.

II. RANDOM-MODE MODEL

A. Spectral structure of neural-circuit connectivity

Let $\boldsymbol{J}$ denote an $N \times N$ weighted, directed coupling matrix among N neurons. Many theoretical studies have examined single-neuron [10] and collective [39] properties of activity in nonlinear recurrent neural networks with i.i.d.

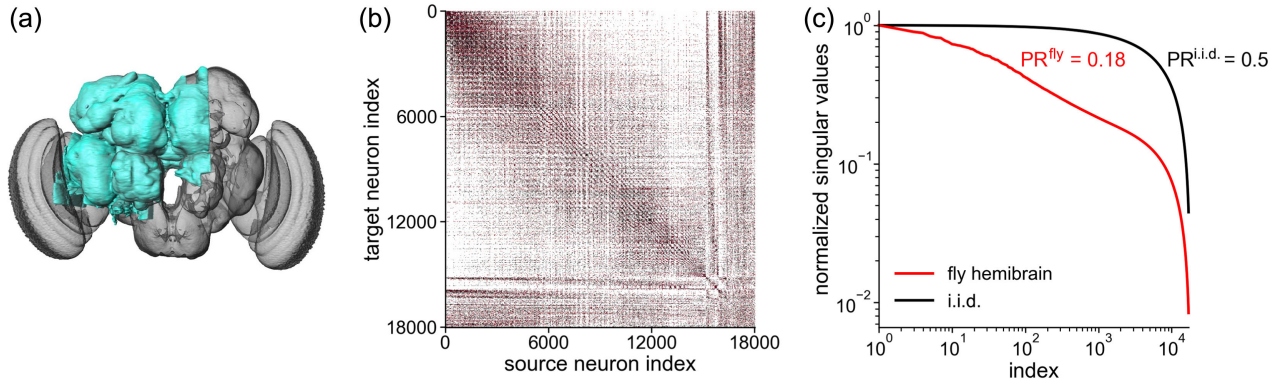

FIG. 1. Analysis of *Drosophila* central-brain connectome. (a) Volume of fly brain for which reconstruction was performed (blue; reproduced from Ref. [17]). (b) Normalized coupling matrix (elements summed within 10×10 blocks to aid visualization). (c) Singular-value spectra of the normalized coupling matrix (red) and an i.i.d. random matrix (gray). The fly connectome exhibits a smooth spectrum that decays quickly, corresponding to a reduced participation ratio. $N = 18028$ neurons.

J_{ij}. However, the coupling matrices of real-world networks, including neural circuits, often exhibit approximate low-rank structure that deviates dramatically from the i.i.d. assumption. To illustrate this, we analyzed a central-brain connectome of the fruit fly *Drosophila melanogaster* using the singular-value decomposition (SVD), which decomposes any matrix as a sum of rank-one components, with each component's strength given by the corresponding singular value [Figs. 1(a) and 1(b); details in Appendix B]. Specifically, the SVD decomposes $\boldsymbol{J}$ in the form

$$\boldsymbol{J} = \sum_{a=1}^{M} S_a \boldsymbol{u}_a \boldsymbol{v}_a^T, \tag{1}$$

where, for $a = 1, \ldots, M$, $S_a > 0$ are the singular values; $\boldsymbol{u}_a$ and $\boldsymbol{v}_a$ are N-dimensional left and right singular vectors, respectively; and M is the rank of $\boldsymbol{J}$.

The ranked singular values of the *Drosophila* connectome decay much more rapidly than those of an i.i.d. matrix of the same size, which follow a universal distribution independent of single-element statistics [Fig. 1(c); similar analysis appears in Thibeault *et al.* (Ref. [23])]. We quantify this rapid decay using the participation ratio (PR) of the squared singular-value spectrum, $\mathrm{PR}^S = (\sum_{a=1}^{M} S_a^2)^2 / \sum_{a=1}^{M} S_a^4$. The connectome's participation ratio is 0.18, substantially lower than the value 0.5 characteristic of i.i.d. matrices, confirming that this biological network exhibits a much more concentrated low-rank structure than would arise from i.i.d. couplings.

B. Definition of the random-mode model

Both the singular-value spectrum and, as we will show later, the overlaps of the left and right singular vectors of the *Drosophila* connectome deviate from the predictions of an i.i.d. coupling matrix. We therefore introduce the random-mode model, a generative model for coupling matrices that allows us to capture these properties within a random-matrix ensemble (Fig. 2).

The random-mode model has the same mathematical form as an SVD, but is a statistical generative process for $\boldsymbol{J}$ rather than a matrix factorization. The model generates a coupling matrix of the form

$$\boldsymbol{J} = \sum_{a=1}^{M} D_a \boldsymbol{\ell}_a \boldsymbol{r}_a^T, \tag{2}$$

where, for each $a = 1, \ldots, M$, $D_a > 0$ are the component strengths; $\boldsymbol{r}_a$ and $\boldsymbol{\ell}_a$ are N-dimensional right and left modes, respectively; and M is the number of components, where $\mathrm{rank} \leq \min(N, M)$. We now describe the properties of these quantities. See Appendix A for a glossary of terms.

1. Component strengths

The component strengths D_a are analogous to the singular values S_a of $\boldsymbol{J}$. We model these quantities deterministically with the requirement that they are defined for arbitrarily large M and that empirical averages $M^{-1} \sum_{a=1}^{M} f(D_a)$ converge to limiting values, denoted $\langle f(D) \rangle_D$, as $M \to \infty$.

2. Left and right modes

The left and right modes $\boldsymbol{\ell}_a$ and $\boldsymbol{r}_a$ are analogous to the left and right singular vectors $\boldsymbol{u}_a$ and $\boldsymbol{v}_a$ of $\boldsymbol{J}$. We assume they are sampled i.i.d. across the neuron index i,

$$P(\{\boldsymbol{\ell}_a, \boldsymbol{r}_a\}_{a=1}^{M}) = \prod_{i=1}^{N} P(\{\ell_{ai}, r_{ai}\}_{a=1}^{M}), \tag{3}$$

where $P(\{\ell_a, r_a\}_{a=1}^{M})$ is the joint distribution over the $2M$ mode components, which must be specified. We assume that this distribution has the following moments fixed:

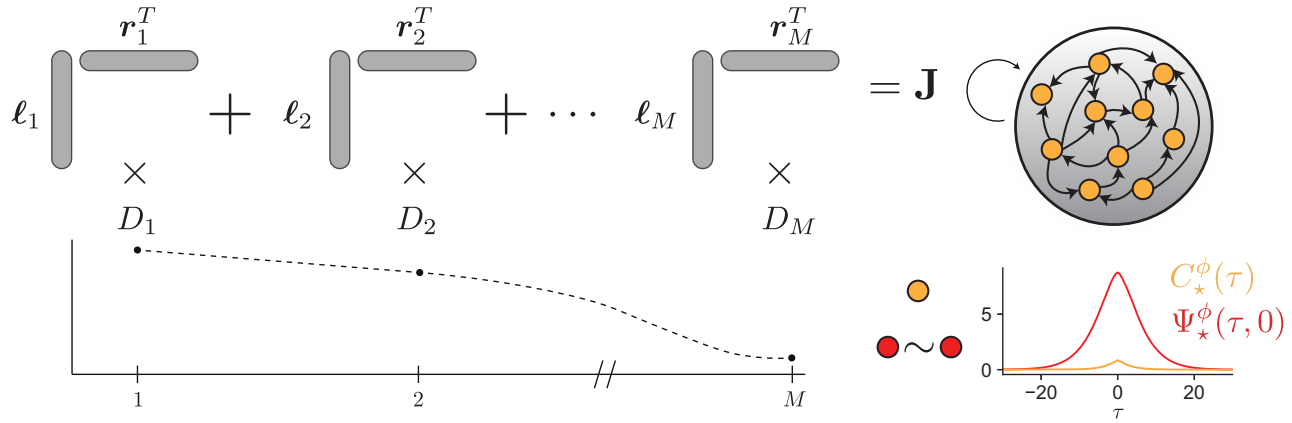

FIG. 2. Schematic of the random-mode model. Upper: couplings $\boldsymbol{J}$ are generated as a sum of outer products, $\boldsymbol{\ell}_a \boldsymbol{r}_a^T$, with component strengths D_a. Lower: the two-point function $C_\star^\phi(\tau)$ and four-point function $\Psi_\star^\phi(\tau)$ are calculated in terms of the statistics of D_a. The two-point function depends only on the effective gain g_{eff}, while the four-point function depends on both g_{eff} and PR^D, the effective dimension of the connectivity determined by the D_a distribution.

$$\langle \ell_a \rangle = \langle r_a \rangle = 0, \tag{4}$$

$$\langle \ell_a \ell_b \rangle = \langle r_a r_b \rangle = \frac{1}{N} \delta_{ab}. \tag{5}$$

This yields orthonormality of modes in expectation,

$$\begin{aligned} \langle \boldsymbol{\ell}_a^T \boldsymbol{\ell}_b \rangle &= \delta_{ab}, \\ \langle \boldsymbol{r}_a^T \boldsymbol{r}_b \rangle &= \delta_{ab}, \end{aligned} \tag{6}$$

in analogy to singular vectors, for which orthonormality holds exactly [40]. Individual mode components are $\mathcal{O}(1/\sqrt{N})$. Note that we have not constrained second moments that mix left and right mode components.

C. Rank scaling and comparison to low-rank recurrent neural networks

A crucial modeling choice is how the number of modes M scales relative to the network size N. The two main possibilities are $M = \mathcal{O}(N)$ or $M = \mathcal{O}(1)$, corresponding to extensive and intensive rank scaling, respectively. The random-mode model is characterized by extensive rank scaling, where both M and N approach infinity while maintaining a fixed ratio,

$$\alpha = \frac{M}{N} = \mathcal{O}(1). \tag{7}$$

By choosing α to be small, or by choosing the spectrum of ranked component strengths to decay rapidly, we can construct connectivity that is low rank or well described by a small number of rank-one components, compared to the number of neurons.

The alternative approach of intensive rank scaling, where M remains finite while $N \to \infty$, leads to what has been termed "low-rank recurrent neural networks" in the literature [42–45]. In some instances, the intensive-rank component is added to an i.i.d. matrix modeling unstructured "background" connectivity [42,43]. Note that, in this paper, when we use "low rank" or "low dimensional," we refer to quantities that are small as a fraction of N, even if they are extensive; this differs from some prior works (such as those on low-rank recurrent neural networks) where these terms refer to intensive quantities.

The distinction between extensive- and intensive-rank scaling has important implications for specifying the distribution over the $2M$ mode components $P(\{\ell_a, r_a\}_{a=1}^M)$. For finite M, one can directly specify this distribution. At large N, such a distribution encodes the relative geometry of the left and right modes, since the inner products between modes converge to (N times) the second moments of $P(\{\ell_a, r_a\}_{a=1}^M)$, with negligible fluctuations. A common choice is a $2M$-dimensional multivariate Gaussian distribution. An important insight from prior work on the intensive-rank case is that these overlaps are crucial for shaping the dynamics and computations implemented by the network, as overlaps between connectivity modes dictate how activity modes interact [42–45].

Our goal is to go beyond the intensive-rank setup, which produces correspondingly low-dimensional activity, to model more sophisticated, higher-dimensional dynamics. However, specifying the full distribution $P(\{\ell_a, r_a\}_{a=1}^M)$ in the random-mode model, where $M \to \infty$, is more complicated. First, it must be well defined for all M to allow taking the limit. More fundamentally, we would like to use the random-mode model to construct generative models that capture certain essential spectral features of coupling matrices, over which we have systematic control. Using a joint Gaussian distribution would require specifying M^2 parameters which, since $M = \mathcal{O}(N)$, is comparable to defining the full $N \times N$ coupling matrix.

For this reason, this paper specializes to a case where $P(\{\ell_a, r_a\}_{a=1}^M)$ factorizes across the mode index a,

$$P(\{\ell_a, r_a\}_{a=1}^M) = \prod_{a=1}^M P(\ell_a, r_a), \tag{8}$$

where $P(\ell_a, r_a)$ must be specified. Moreover, until Sec. VII, we further specialize to the case where the left and right modes are fully independent: $P(\ell_a, r_a) = P(\ell_a)P(r_a)$, with each marginal factor $P(\ell_a)$ or $P(r_a)$ being a univariate Gaussian with mean zero and variance $1/N$ [46]. In this case, all $2MN$ components specifying the modes are i.i.d. with mean zero and variance $1/N$; we refer to this case as "i.i.d. modes." In Sec. VII we use a nonfactorized form of $P(\ell_a, r_a)$ to specify correlations between ℓ_a and r_a, introducing M deterministic parameters (in addition to the M component strengths) to specify the left-right correlation for each a.

D. Relationship to singular-value decomposition

In the SVD, the left and right singular vectors are orthonormal: $\boldsymbol{u}_a^T\boldsymbol{u}_b = \boldsymbol{v}_a^T\boldsymbol{v}_b = \delta_{ab}$. In contrast, the left and right modes are orthonormal only in expectation [Eq. (6)]; once sampled, there are random $\mathcal{O}(1/\sqrt{N})$ overlaps,

$$\begin{aligned}\boldsymbol{\ell}_a^T\boldsymbol{\ell}_b &= \delta_{ab} + \mathcal{O}(1/\sqrt{N}),\\ \boldsymbol{r}_a^T\boldsymbol{r}_b &= \delta_{ab} + \mathcal{O}(1/\sqrt{N}).\end{aligned} \tag{9}$$

When the effective rank (defined below) is small and N is large, the component strengths and modes approximate singular values and vectors, respectively, since the effects of the $\mathcal{O}(1/\sqrt{N})$ deviations from orthonormality in the sampled modes are negligible. Indeed, a well-known result in high-dimensional geometry is that finite numbers of random vectors with i.i.d. components approximate orthonormal bases at large N [47]. When the effective rank is not small, overlaps between different modes produce a discrepancy between the component strengths D_a and singular values S_a. We now characterize this discrepancy analytically.

We denote the nth moments and participation ratio of the component-strength distribution by

$$r_n = \langle D^n \rangle_D, \tag{10}$$

$$\mathrm{PR}^D = \frac{r_2^2}{r_4}, \tag{11}$$

from which we define the "effective rank" of the couplings,

$$\text{effective rank} = \alpha \mathrm{PR}^D. \tag{12}$$

The effective rank is related to the participation ratio of the squared singular-value spectrum at large N [48] through

$$\mathrm{PR}^S = \frac{\alpha \mathrm{PR}^D}{1 + 2\alpha \mathrm{PR}^D}. \tag{13}$$

Expanding in small $\alpha\mathrm{PR}^D$, $\mathrm{PR}^S = \alpha\mathrm{PR}^D + \mathcal{O}((\alpha\mathrm{PR}^D)^2)$, demonstrating that the component strengths are closely related to the singular values in the low-dimensional regime. While we express the analytic results in this paper in terms of $\alpha\mathrm{PR}^D$, the above equation can always be used to translate between $\alpha\mathrm{PR}^D$ and PR^S.

We further quantify this discrepancy in the case where $D_a = 1$ for all a. Using methods from free probability theory [49,50] (Appendix C; Fig. 8), the M singular values are distributed over a range with boundaries

$$S_\pm = \sqrt{1 + \frac{5\alpha}{2} - \frac{\alpha^2}{8} \pm \left(1 + \frac{\alpha}{8}\right)^{3/2}\sqrt{8\alpha}}. \tag{14}$$

For small α, $S_\pm = 1 \pm \sqrt{2\alpha} + \mathcal{O}(\alpha)$. Thus, while the random-mode model spreads the nonzero singular values over a range, this spread becomes negligible for small α. More generally, the distributions of D_a and S_a coincide for small $\alpha\mathrm{PR}^D$.

Having presented a generative model of connectivity, we now present the network model that transforms this connectivity into neural activity. We then define summary statistics that characterize the structure of this activity, allowing us to analyze its dependence on the connectivity.

III. NETWORK MODEL AND SUMMARY STATISTICS

A. Recurrent neural-network model

We study a recurrent neural network of N neurons. Each neuron $i \in \{1, \ldots, N\}$ is characterized by its preactivation $x_i(t)$ and activation $\phi_i(t) = \phi(x_i(t))$, where $\phi(\cdot)$ is a scalar nonlinearity; we use $\phi(x) = \mathrm{erf}(\sqrt{\pi}x/2)$, which is sigmoid-shaped and has $\phi'(0) = 1$ [51]. The network dynamics are governed by

$$T[x_i](t) = \sum_{j=1}^N J_{ij}\phi_j(t), \tag{15a}$$

$$\text{where } T[x](t) = (1 + \partial_t)x(t). \tag{15b}$$

Here, J_{ij} denotes the synaptic coupling from neuron j to neuron i, and $T[\cdot]$ is a causal functional that specifies the single-neuron dynamics for which Eq. (15b) is a canonical choice. We expect our results to be agnostic to the specific

choice of $T[\cdot]$. We next introduce correlation functions that capture single-neuron and collective activity properties.

B. Two-point functions and duality of neuronal and temporal covariances

Let $a \in \{x, \phi\}$ denote either preactivations or activations. We define two complementary covariance functions, averaging over either time or neurons,

$$C^a_{ij}(\tau) = \langle a_i(t) a_j(t+\tau) \rangle_t, \tag{16}$$

$$C^a(t_1, t_2) = \frac{1}{N} \sum_{i=1}^{N} a_i(t_1) a_i(t_2). \tag{17}$$

These functions exhibit a form of duality: for any lag τ, the matrix $\boldsymbol{C}^a(\tau)$ with elements $C^a_{ij}(\tau)$ indexed by (i, j) and the matrix with elements $C^a(t_1, t_2 + \tau)$ indexed by (t_1, t_2) have the same eigenvalue spectrum, up to scaling (to see this, note that both matrices can be computed from a time-by-neuron activity matrix $\boldsymbol{A}$ as $\boldsymbol{A}^T\boldsymbol{A}$ and $\boldsymbol{A}\boldsymbol{A}^T$, respectively, which have identical spectra) [52].

We also define the neuron-averaged response function, describing the propagation of infinitesimal perturbations,

$$S^a(t_1, t_2) = \frac{1}{N} \sum_{i=1}^{N} \left. \frac{\delta a_i(t_1)}{\delta I_i(t_2)} \right|_{\boldsymbol{I}=0}, \tag{18}$$

where $I_i(t)$ is a source term added to the right-hand side of the network dynamics: $T[x_i](t) = \sum_{j=1}^{N} J_{ij}\phi_j(t) + I_i(t)$. Finally, we define "stationarized" counterparts,

$$C^a(\tau) = \langle C^a(t, t \pm \tau) \rangle_t, \qquad S^a(\tau) = \langle S^a(t, t - \tau) \rangle_t. \tag{19}$$

C. Four-point functions, dimension of activity, and principal-component timescales

In contrast to the above two-point functions, which characterize single-neuron activity, four-point functions describe features of collective activity. The same four-point functions were used in Ref. [39]. Here, we motivate these functions in more detail and express them in two ways via the duality between neuronal and temporal covariances.

To motivate their definition, consider the dimension of activity, which can be quantified as the participation ratio of the spectrum of the equal-time covariance matrix $\boldsymbol{C}^a(0)$ [9,36,53–59]. Given the eigenvectors $\boldsymbol{v}^a_k$ and eigenvalues λ^a_k for $k = 1, \ldots, N$, as one would compute as part of principal components analysis, the participation ratio is

$$\mathrm{PR}^a = \frac{1}{N} \frac{(\sum_{k=1}^{N} \lambda^a_k)^2}{\sum_{k=1}^{N} (\lambda^a_k)^2}. \tag{20}$$

To see why PR^a provides a meaningful measure of dimension, consider the case where D eigenvalues equal a positive constant and the remaining eigenvalues are zero. Then, $\mathrm{PR}^a = D/N$. More generally, when the spectrum exhibits a smooth decay rather than a hard cutoff, the participation ratio identifies the characteristic decay scale (divided by N) [55]. We compare the participation ratio to alternative measures of effective dimensionality in Appendix D. Without the normalization factor $1/N$, this quantity would vary between 1 and N; with normalization, it varies between $1/N$ and 1.

In the limit $N \to \infty$, this quantity could have three qualitatively different behaviors:

(i) Subextensive dimensionality: $\mathrm{PR}^a = 0$. The number of dimensions filled by activity grows sublinearly with N.
(ii) Nontrivial extensive dimensionality: $0 < \mathrm{PR}^a < 1$. The number of dimensions filled by activity grows linearly with N and not all dimensions are filled equally.
(iii) Trivial extensive dimensionality: $\mathrm{PR}^a = 1$. All dimensions are filled equally by activity.

An advantageous property of the participation ratio is that it can be expressed as the ratio of the squared trace and Frobenius norm of $\boldsymbol{C}^a(0)$, both of which can be further expressed in terms of matrix elements $C^a_{ij}(0)$,

$$\begin{aligned} \mathrm{PR}^a &= \frac{(\mathrm{tr}\boldsymbol{C}^a(0))^2}{\|\boldsymbol{C}^a(0)\|_F^2} \\ &= \frac{C^a(0)^2}{\frac{1}{N}\sum_{i=1}^{N} C^a_{ii}(0)^2 + \frac{1}{N}\sum_{i\neq j} C^a_{ij}(0)^2}, \end{aligned} \tag{21}$$

where $C^a(0)$ is given by Eq. (19). Assuming that diagonal elements are uniform with negligible fluctuations, i.e., $C^a_{ii}(0) = C^a(0)$ for all i (as occurs, for example, under i.i.d. connectivity [10]), this simplifies to

$$\mathrm{PR}^a = \frac{C^a(0)^2}{C^a(0)^2 + \psi^a(0, 0)}, \tag{22}$$

where, following Ref. [39], we define the four-point function

$$\psi^a(\boldsymbol{\tau}) = \frac{1}{N} \sum_{i\neq j} C^a_{ij}(\tau_1) C^a_{ij}(\tau_2), \tag{23}$$

where $\boldsymbol{\tau} = (\tau_1, \tau_2)$. The three scaling behaviors of dimensionality outlined above correspond in the limit $N \to \infty$ to $\psi^a(0, 0) \to \infty$, $\psi^a(0, 0) \to$ const., or $\psi^a(0, 0) \to 0$, respectively. Since $\psi^a(0, 0)$ is given by a sum over $\mathcal{O}(N^2)$ squared cross-covariances times $1/N$, its magnitude is N times that of an individual squared cross-covariance. Both i.i.d. matrices and the extensive-rank matrices generated by the random-mode model lead to $\mathcal{O}(1/\sqrt{N})$ cross-covariances. Thus, $\psi^a(0, 0) \to$ const. as $N \to \infty$, leading to nontrivial extensive dimensionality, $0 < \mathrm{PR}^a < 1$.

In this paper, we focus on a related function $\Psi^a(\boldsymbol{\tau})$ that includes the diagonal terms,

$$\Psi^a(\boldsymbol{\tau}) = \frac{1}{N}\sum_{i,j=1}^{N} C_{ij}^a(\tau_1)C_{ij}^a(\tau_2). \tag{24}$$

As per the duality between neuronal and temporal covariances, $\Psi^a(\boldsymbol{\tau})$ can also be expressed in terms of the time-by-time covariance $C^a(t_1,t_2)$ as

$$\Psi^a(\boldsymbol{\tau}) = N\langle C^a(t_1,t_2)C^a(t_1+\tau_1,t_2+\tau_2)\rangle_{t_1,t_2}. \tag{25}$$

The neuron-by-neuron [Eq. (24)] and time-by-time [Eq. (25)] definitions of $\Psi^a(\boldsymbol{\tau})$ form the basis of the cavity and path-integral calculations of this function, respectively. Finally, the dimension of activity is given by

$$\mathrm{PR}^a = \frac{C^a(0)^2}{\Psi^a(0,0)}, \tag{26}$$

which, unlike Eq. (22), holds even when single-neuron variances $C_{ii}^a(0)$ are nonuniform across neurons.

To study timescales of collective activity, we consider the principal components of activity,

$$p_k^a(t) = \frac{1}{\sqrt{\lambda_k^a}}(\boldsymbol{v}_k^a)^T\boldsymbol{a}(t). \tag{27}$$

These principal components are the basis of much of modern analysis of high-dimensional neural data [35–38]. They all have unit variance due to the $1/\sqrt{\lambda_k^a}$ normalization, but potentially very different characteristic timescales [Fig. 9(a)]. To extract the timescales of just the leading components, we weight them by the squares of their corresponding eigenvalues. This gives

$$\frac{1}{N}\sum_k (\lambda_k^a)^2\langle p_k^a(t)p_k^a(t+\tau)\rangle_t = \Psi^a(\tau,0), \tag{28}$$

showing that $\Psi^a(\boldsymbol{\tau})$ captures the temporal structure of the leading principal components of activity [Figs. 9(c) and 9(d)]. Note that weighting by the eigenvalues themselves, rather than their squares, gives $N^{-1}\sum_k \lambda_k^a\langle p_k^a(t)p_k^a(t+\tau)\rangle_t = C^a(\tau)$, recovering single-neuron information [Figs. 9(b) and 9(d)].

IV. PATH-INTEGRAL ANALYSIS OF THE RANDOM-MODE MODEL

We aim to compute the summary statistics $C^a(t_1,t_2)$, $S^a(t_1,t_2)$, and $\Psi^a(\boldsymbol{\tau})$ in the limit $N\to\infty$ under the disorder average. For a general function $\mathcal{F}$, we use the subscript $\star$ to denote such limiting values,

$$\mathcal{F}_\star = \lim_{N\to\infty}\langle\mathcal{F}\rangle_{\boldsymbol{J}}. \tag{29}$$

These values also correspond to the saddle point of a path integral. The summary statistics we have defined are self-averaging, meaning that the same values are obtained, up to $\mathcal{O}(1/\sqrt{N})$ fluctuations, when they are computed based on a single realization of a large network. In this paper, we are interested in statistically stationary states such that, for $N\to\infty$, the two-point functions depend on time differences only and thus can be replaced by their stationarized counterparts [Eq. (19)], i.e., for all t,

$$C_\star^a(t,t\pm\tau) = C_\star^a(\tau), \qquad S_\star^a(t,t-\tau) = S_\star^a(\tau). \tag{30}$$

Calculating two-point functions is well-established "standard DMFT" [60–62]. As we review in Appendix E, four-point functions were previously calculated for networks with i.i.d. couplings using a two-site version of the cavity method [39]. In this section, we introduce a new path-integral approach to calculating four-point functions that allows for the rapid solution of the i.i.d. model and generalizes to more complicated models, including the random-mode model.

A. Relating Ψ^a to temporal covariance fluctuations

To cast the problem in the language of path integrals, we turn to the time-by-time definition of $\Psi^a(\boldsymbol{\tau})$ [Eq. (25)]. If the time-by-time covariance $C^a(t_1,t_2)$ were equal to its limiting value, $C_\star^a(t_2-t_1)$, we would have

$$\Psi_\star^a(\boldsymbol{\tau}) \overset{?}{=} N\langle C_\star^a(t_2-t_1)C_\star^a(t_2-t_1+\tau_2-\tau_1)\rangle_{t_1,t_2} = 0$$

since, for large time differences, $C_\star^a(t_2-t_1)$ decays exponentially in $|t_2-t_1|$ [10]. However, based on the neuron-by-neuron definition [Eq. (24)], $\Psi_\star^a(\boldsymbol{\tau})$ is clearly a nonzero, order-one quantity. This apparent contradiction is resolved by considering the $\mathcal{O}(1/\sqrt{N})$ fluctuations around the limiting value. The necessity of these fluctuations becomes clear when considering the duality between neuronal and temporal covariances: the presence of nonzero cross-covariances $C_{ij}^a(0)$ indicates low-dimensional structure in the system; equivalently, this structure manifests as the system revisiting similar states over time, resulting in nonzero temporal covariance $C^a(t_1,t_2)$ even for large time differences $|t_2-t_1|\gg 1$ (Fig. 3). To capture this behavior, we define fluctuations around the saddle point as

$$\delta C^a(t_1,t_2) = C^a(t_1,t_2) - C_\star^a(t_2-t_1). \tag{31}$$

We express $\Psi^a(\boldsymbol{\tau})$ purely in terms of these fluctuations as

$$\Psi^a(\boldsymbol{\tau}) = N\langle\delta C^a(t_1,t_2)\delta C^a(t_1+\tau_1,t_2+\tau_2)\rangle_{t_1,t_2}, \tag{32}$$

noting that terms involving $C_\star^a(t_2-t_1)$ vanish at large time differences. Our task, therefore, is to compute the covariance of the fluctuations, $\delta C^a(t_1,t_2)$, captured asymptotically by

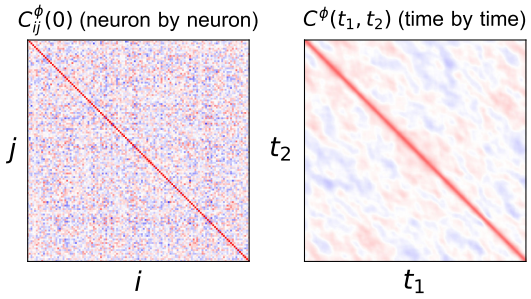

FIG. 3. Duality of neuron-by-neuron and time-by-time covariances and its relation to the dimension of activity. Both plots are based on the same simulation of a network of $N = 2500$ neurons with $g = 2.25$ and i.i.d. couplings. The dimension can be computed either by computing the statistics of the off diagonals of the neuron-by-neuron covariance $C^{\phi}_{ij}(0)$ (left) or by computing the fluctuations away from the translation-invariant mean-field form of the time-by-time covariance $C^{\phi}(t_1, t_2)$ (right).

$$\langle N\delta C^a(t_1,t_2)\delta C^a(t_3,t_4)\rangle_\star. \tag{33}$$

Then, to compute $\Psi^a_\star(\boldsymbol{\tau})$, we take in Eq. (33) the limits

$$\begin{aligned} t_3 &= t_1 + \tau_1, \qquad |t_2 - t_1| \to \infty, \\ t_4 &= t_2 + \tau_2, \qquad \tau_1, \tau_2 \sim \mathcal{O}(1). \end{aligned} \tag{34}$$

B. Calculation of two- and four-point functions

Details of the calculation are given in Appendix F. Briefly, we compute the covariance of fluctuations [Eq. (33)] using the Martin-Siggia-Rose-Janssen–de Dominicis (MSRJD) path-integral formalism [60–65]. For any fixed coupling matrix $\boldsymbol{J}$, the path integral is

$$Z[\boldsymbol{J}] = \int \mathcal{D}\boldsymbol{x} \int \mathcal{D}\hat{\boldsymbol{x}} \exp\left\{ i \sum_{i=1}^{N} \int dt \hat{x}_i(t) \left(T[x_i](t) - \sum_{j=1}^{N} J_{ij}\phi_j(t) \right) \right\}. \tag{35}$$

The auxiliary field $\hat{x}_i(t)$ is conjugate to $x_i(t)$ and enforces the network's equations of motion [Eq. (15a)] through the δ-function representation $(2\pi)^{-1}\int d\hat{x} e^{i\hat{x}x} = \delta(x)$. Upon introduction of source terms (that are not needed for our purposes), this path integral serves as a generating functional for correlation and response functions.

In principle, to compute $\Psi^a_\star(\boldsymbol{\tau})$ in this framework, after integrating out $\boldsymbol{J}$ from $Z[\boldsymbol{J}]$ to obtain a statistical field theory governing two-point functions, we (1) find the saddle-point solution for two-point functions by extremizing the action (i.e., standard DMFT); (2) compute the time^2-by-time^2 Hessian matrix of the action describing fluctuations around this saddle point; (3) invert this Hessian and extract the sub-block corresponding to the covariance of $\delta C^{\phi}(t_1, t_2)$ [Eq. (33)]; and (4) apply the temporal limits [Eq. (34)] to this sub-block to obtain $\Psi^{\phi}_\star(\boldsymbol{\tau})$.

Direct inversion of the time^2-by-time^2 Hessian is analytically intractable. However, we show that the temporal limits commute with Hessian inversion, allowing us to invert a low-dimensional frequency-dependent Hessian instead (Appendix G). This readily yields the Fourier-space function $\Psi^{\phi}_\star(\boldsymbol{\omega})$, where $\boldsymbol{\omega} = (\omega_1, \omega_2)$. A few additional steps yield $\Psi^{x}_\star(\boldsymbol{\omega})$. This approach allows for the efficient calculation of two- and four-point functions for a given model, requiring only a few lines of computation directly from the action of the theory.

As a demonstration of this approach, we apply it to a generalization of the i.i.d. model with correlated reciprocal couplings, recovering results of Ref. [39] (Appendix H).

C. Solution of the random-mode model

Applied to the random-mode model, the self-consistent equations that determine $C^{\phi}_\star(t_1, t_2)$ are (Appendix I 1)

$$T[x](t) = \eta^x(t), \tag{36a}$$

$$\eta^x \sim \mathcal{GP}(0, \alpha r_2 C^{\phi}_\star), \tag{36b}$$

$$C^{\phi}_\star(t_1,t_2) = \langle \phi(t_1)\phi(t_2)\rangle_\star, \tag{36c}$$

where Eq. (36b) indicates that $\eta^x(t)$ is a Gaussian field with zero mean and correlation function $\alpha r_2 C^{\phi}_\star(t,t')$. These equations are equivalent to those for an i.i.d. model [Eqs. (E2)–(E4)] with coupling strength

$$g_{\text{eff}} = \sqrt{\alpha r_2}. \tag{37}$$

The four-point function, our main object of interest, is (Appendix I 2)

$$\Psi^{\phi}_\star = \frac{1 + \frac{1}{\alpha \text{PR}^D} |g^2_{\text{eff}} S^{\phi}_{12}|^2}{|1 - g^2_{\text{eff}} S^{\phi}_{12}|^2} C^{\phi}_{12}, \tag{38}$$

where we have simplified notation by suppressing frequency variables via the shorthand $C^{\phi}_{12} = C^{\phi}_\star(\omega_1)C^{\phi}_\star(\omega_2)$ and $S^{\phi}_{12} = S^{\phi}_\star(\omega_1)S^{\phi}_\star(\omega_2)$. The same solution is obtained through a more involved, but intuitive, two-site cavity calculation (Appendix L; Fig. 11).

This solution for the random-mode model is one of our main results, which we now spend some time interpreting.

D. Interpreting the solution

Our analysis reveals two features of connectivity that characterize how it shapes activity, namely, the effective coupling strength g_{eff} and the effective rank αPR^D. The former characterizes local structure (the magnitudes of individual couplings), while the latter characterizes global structure (the number of large connectivity components).

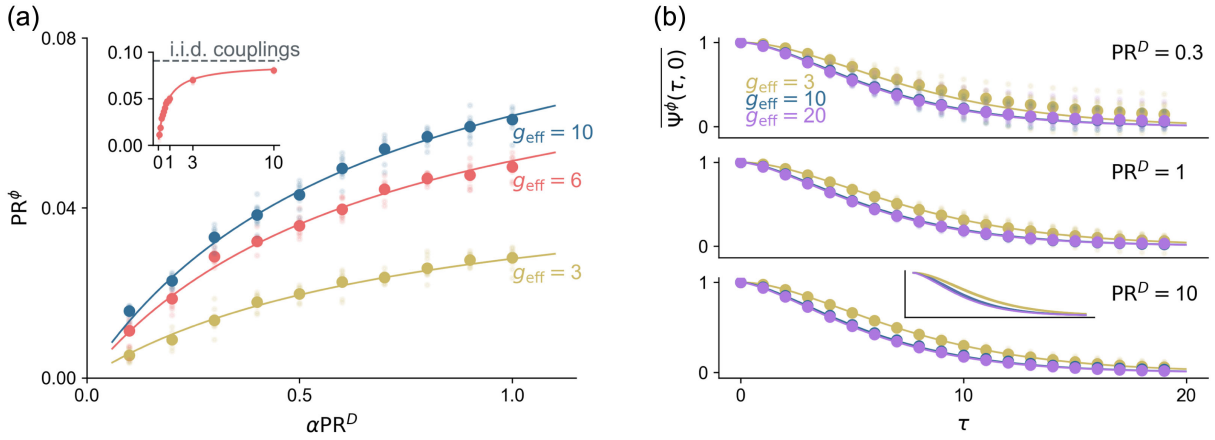

FIG. 4. Dimension of activity and collective timescales in the random-mode model. (a) Activity dimension PR^{ϕ} versus effective rank αPR^{D} for various coupling strengths g_{eff}. Thin dots, individual simulations; thick dots, means over 10 simulations; lines, theoretical predictions. Inset: extended $g_{\text{eff}} = 6$ case, showing convergence to i.i.d.-coupling behavior with growing α. (b) Normalized four-point function $\overline{\Psi^{\phi}(\tau, 0)} \equiv \Psi^{\phi}(\tau, 0)/\Psi^{\phi}(0, 0)$ for various g_{eff} and αPR^{D}. Inset: theory curves for all αPR^{D}, demonstrating relative invariance of collective timescales to the effective rank. Simulations use $N = 5000$ neurons.

The above equations determining $C^{\phi}_{\star}(\tau)$ and $\Psi^{\phi}_{\star}(\boldsymbol{\tau})$ immediately reveal that global structure in the connectivity is invisible at the single-neuron level but shapes collective activity. Specifically, the correlation function $C^{\phi}_{\star}(\tau)$ is fully determined by g_{eff} and is insensitive to the effective rank of connectivity. This has important implications for neural data analysis. In particular, analyses of single-neuron activity properties may miss signatures of structured connectivity that become apparent only when analyzing collective-activity properties among simultaneously recorded neurons.

What is the nature of the dependence of collective-activity properties, such as the dimension of activity, on the effective rank? To plot this relationship and validate the theory, we simulated networks and computed the participation ratio from the empirical equal-time covariance matrix of the activations (see Appendix M for numerical details). We used a component-strength spectrum $D_a \propto \exp(-\beta_D a/M)$, which yields $\text{PR}^{D} = \beta_D^{-1}\tanh(\beta_D)$. To attain a desired effective rank $\alpha\text{PR}^{D} < 1$, we set $\alpha = M/N = 1$ and solved for β; for $\alpha\text{PR}^{D} > 1$, we fixed $\text{PR}^{D} = 1$ and increased α. We adjusted g_{eff} by rescaling D_a uniformly. We find excellent theory-simulation agreement (Fig. 4). Figure 4(a) illustrates how the dimension of activity, PR^{ϕ}, varies with αPR^{D} for different values of g_{eff}. Consistent with i.i.d.-coupling networks, PR^{ϕ} increases monotonically with g_{eff} [39]. It also increases monotonically with αPR^{D}.

Taking $\alpha\text{PR}^{D} \to \infty$ for fixed g_{eff} in Eq. (38) recovers the formula for i.i.d. couplings [Eq. (E5)] of Ref. [39]. Indeed, in this limit, $\boldsymbol{J}$ approaches an i.i.d. matrix. Correspondingly, PR^{ϕ} approaches the activity dimension of a network with i.i.d. couplings [Fig. 4(a) inset]. Comparing Eq. (E5) to the corresponding i.i.d. formula ($1/\alpha\text{PR}^{D} = 0$) and noting that single-neuron properties are unaffected by αPR^{D}, one sees that any structure beyond i.i.d. connectivity ($1/\alpha\text{PR}^{D}$ finite) strictly decreases the dimension of activity.

The dimension of activity depends on $\Psi^{\phi}_{\star}(\boldsymbol{\tau})$ at $\boldsymbol{\tau} = (0, 0)$; we now examine the temporal profile of this function. Figure 4(b) shows a normalized version of $\Psi^{\phi}(\tau, 0)$, which is related to the correlation functions of leading principal components, for various values of g_{eff} and αPR^{D}. The decay timescale of $\Psi^{\phi}_{\star}(\tau, 0)$ decreases with increasing g_{eff}, approaching a limiting behavior at $g_{\text{eff}} \approx 10$. While αPR^{D} affects the overall scale of this function and, consequently, the dimension of activity, it has little effect on the decay timescale and, hence, little effect on the timescales of leading principal components [Fig. 4(b), inset]. Thus, the effective rank primarily influences the dimension of activity rather than the collective temporal structure.

A remarkable feature of these results is that only two connectivity statistics, g_{eff} and αPR^{D}, are sufficient to determine the two- and four-point functions. Because of the nonlinear, high-dimensional nature of the network, this reduction is unexpected. This reduction breaks down when structured mode overlaps are introduced (Sec. VII), in which case, the formula for the four-point function requires knowledge of the full joint distribution of component strengths and overlaps.

V. LIMITING BEHAVIOR OF THE RANDOM-MODE MODEL

Here, we examine the behavior of the random-mode model in two limiting cases: as the system approaches the transition to chaos from above ($g_{\text{eff}} \to 1^{+}$) and in the limit of low effective rank ($\alpha\text{PR}^{D} \to 0^{+}$).

We first consider the limit $g_{\text{eff}} \to 1^{+}$ for arbitrary αPR^{D}. The relevant expansion parameter is $0 < g_{\text{eff}} - 1 \ll 1$. For

networks with i.i.d. couplings with coupling strength g, it is known that, as $g \to 1^+$, $\Psi^a_{\star,\text{i.i.d.}}(\boldsymbol{\tau}) \sim 1/(g-1)$ and $\text{PR}^a_{\text{i.i.d.}} \sim (g-1)^3$ [39]. For the random-mode model, to leading order in $1/(g_{\text{eff}}-1)$,

$$\Psi^a_\star(\boldsymbol{\tau}) = \left(1 + \frac{1}{\alpha \text{PR}^D}\right) \underbrace{\Psi^a_{\star,\text{i.i.d.}}(\boldsymbol{\tau})}_{\sim 1/(g_{\text{eff}}-1)}. \tag{39}$$

Consequently, to leading order in $g_{\text{eff}} - 1$, the dimension of activity behaves as

$$\text{PR}^a = \left(1 + \frac{1}{\alpha \text{PR}^D}\right)^{-1} \underbrace{\text{PR}^a_{\text{i.i.d.}}}_{\sim (g_{\text{eff}}-1)^3}. \tag{40}$$

Thus, as the network approaches the transition to chaos from above, structured connectivity reduces the dimension by a factor of $(1 + 1/\alpha \text{PR}^D)^{-1}$ compared to a network with i.i.d. couplings. Taking the further limit of $\alpha \text{PR}^D \to 0^+$ yields the simple relation

$$\text{PR}^a = \alpha \text{PR}^D \text{PR}^a_{\text{i.i.d.}}. \tag{41}$$

The difference between αPR^D and PR^S is a higher-order correction, so we also have $\text{PR}^a = \text{PR}^S \text{PR}^a_{\text{i.i.d.}}$; that is, in the limits of both low effective rank and small effective coupling strength, the dimension of activity is equal to the dimension of the singular-value spectrum multiplied by the dimension of activity for an i.i.d. network.

We next consider the limit $\alpha \text{PR}^D \to 0^+$, for arbitrary g_{eff}. In this limit, the component strengths closely approximate the singular values. To leading order in αPR^D,

$$\text{PR}^a = K^a(g_{\text{eff}}) \alpha \text{PR}^D, \tag{42}$$

$$K^a(g_{\text{eff}}) = C^a_\star(0)^2 \left[\int d\boldsymbol{\omega} \left| \frac{g^2_{\text{eff}} S^a_{12}}{1 - g^2_{\text{eff}} S^\phi_{12}} \right|^2 C^\phi_{12} \right]^{-1}. \tag{43}$$

We verify this limiting behavior in simulations using $\alpha = 1$ and a step function for the component strengths: $D_a = \text{const.}$ for $a/N \le \text{PR}^D$ and $D_a = 0$ otherwise. We varied PR^D, observing linear relationships with PR^ϕ and PR^x in the $\text{PR}^D \to 0^+$ limit [Fig. 10(a)].

Focusing on $K^\phi(g_{\text{eff}})$, note that its expression [Eq. (43)] resembles the formula for the dimension of activity in an i.i.d. network with coupling strength g_{eff} [Eq. (E5)], but includes an additional factor $|g^2_{\text{eff}} S^\phi_{12}|^2$ in the integral. Numerical evaluation confirms that, as expected from Eq. (41), $K^\phi(g_{\text{eff}}) \to \text{PR}^\phi_{\text{i.i.d.}}$ as $g_{\text{eff}} \to 1^+$. As g_{eff} increases, $K^\phi(g_{\text{eff}})$ increases monotonically to approximately $1.53\text{PR}^\phi_{\text{i.i.d.}}$ [Fig. 10(a), inset]. Thus, for networks with low effective rank, $\text{PR}^\phi \sim \alpha \text{PR}^D \text{PR}^\phi_{\text{i.i.d.}}$ up to an order-one fudge factor that becomes exactly one as $g_{\text{eff}} \to 1^+$.

Finally, we consider the dependence of the full eigenvalue spectrum λ^a_i of $\boldsymbol{C}^a(0)$ on the spectrum of component strengths D_a. The latter leads to the former via a highly nonlinear relationship. Indeed, using the same step-function component-strength spectrum, the ranked eigenvalues λ^ϕ and λ^x of the covariance matrix decay smoothly rather than exhibiting a cutoff [Fig. 10(b)]. Nevertheless, certain properties of these connectivity and activity spectra are linked. In particular, in the limit $\alpha \text{PR}^D \to 0^+$, αPR^D and PR^a exhibit a linear relationship [with proportionality factor $K^a(g_{\text{eff}})$], implying that the ratios of the squared second moments and fourth moments of these spectra are linked (despite, even in this limit, the full spectra not coinciding).

VI. SINGLE-NEURON HETEROGENEITY

Cortical neurons display broad distributions of firing rates [66]. This heterogeneity among neurons is analogous to the heterogeneity of component strengths in the random-mode model, but operates in the neuron basis rather than the component basis. We now compare how low-dimensional connectivity structure and heterogeneous neuronal properties affect collective activity by extending our framework to include single-neuron heterogeneity.

In a general formulation described in Appendix J, we extend the nonlinearity of neuron i to $\Phi_{\boldsymbol{\theta}_i}(x)$, which depends on a vector of neuron-specific parameters $\boldsymbol{\theta}_i$, and solve for the resulting two- and four-point functions. Here, to model firing-rate heterogeneity specifically, we take $\boldsymbol{\theta}_i$ to consist of a single gain parameter G_i, with $\Phi_{\boldsymbol{\theta}_i}(x) = G_i \phi(x)$. In analogy with the component-strength distribution, we define

$$q_n = \langle G^n \rangle_G, \tag{44}$$

$$\text{PR}^G = \frac{q_2^2}{q_4}. \tag{45}$$

The single-site picture that determines the two-point functions is largely similar to Eq. (36c).

We analyze collective activity (four-point functions) for (1) unnormalized activations $\Phi_{\boldsymbol{\theta}_i}(t) = G_i \phi(x_i(t))$ and (2) normalized activations $\phi_i(t) = \phi(x_i(t))$, the latter being analogous to z-scored firing rates in neural recordings, where z scoring removes single-neuron heterogeneity while preserving structure from cross-neuron and temporal correlations. The four-point functions for the unnormalized and normalized variables are found to be

$$\Psi^\Phi_\star = \frac{\frac{1}{\text{PR}^G} + \frac{1}{\alpha \text{PR}^D} |g^2_{\text{eff}} S^\phi_{12}|^2}{|1 - g^2_{\text{eff}} S^\phi_{12}|^2} C^\Phi_{12}, \tag{46a}$$

$$\Psi^\phi_\star = \frac{1 + (\frac{1}{\text{PR}^G} + \frac{1}{\alpha \text{PR}^D} - 1)|g^2_{\text{eff}} S^\phi_{12}|^2}{|1 - g^2_{\text{eff}} S^\phi_{12}|^2} C^\phi_{12}, \tag{46b}$$

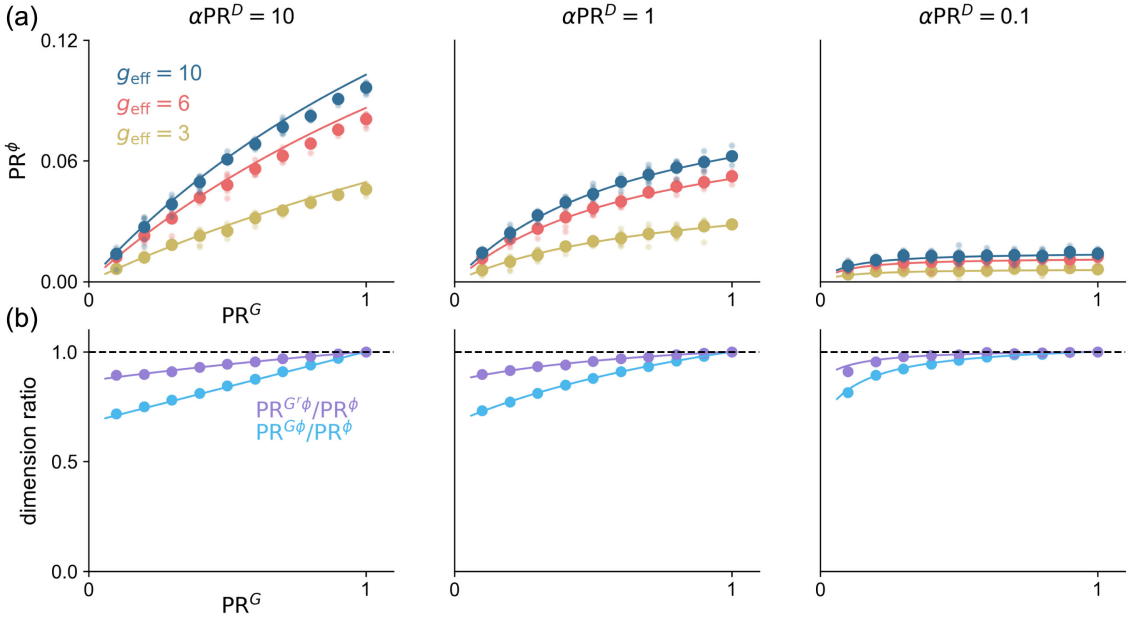

FIG. 5. Effect of single-neuron heterogeneity on dimension of activity in the random-mode model. (a) Dimension of normalized activity PR^{ϕ} versus participation ratio of gain distribution PR^{G} for various coupling strengths g_{eff} and effective ranks $\alpha\mathrm{PR}^{D}$. Thin dots, individual simulations; thick dots, means over ten simulations; lines, theoretical predictions. (b) Fractional reduction in dimension for weighted activations relative to PR^{ϕ}. Blue, $G_i\phi_i$; purple, $G_i^{\mathrm{readout}}\phi_i$. $g_{\mathrm{eff}} = 10$. Thick dots, means over ten simulations, averaging before taking the ratio; lines, theoretical predictions. All simulations use $N = 5000$ neurons.

respectively, where $C_{12}^{\Phi} = q_2^2 C_{12}^{\phi}$ and we have updated the definition of the effective coupling strength to $g_{\mathrm{eff}} = \sqrt{\alpha r_2 q_2}$. Equation (46b) demonstrates that reductions in effective rank, described by $\alpha\mathrm{PR}^{D}$, and reductions in the heterogeneity of gains, described by PR^{G}, have symmetric effects on the dimension of normalized activity, with their combined effect captured by the harmonic mean $[(1/\alpha\mathrm{PR}^{D}) + (1/\mathrm{PR}^{G})]^{-1}$ [67]. One might expect that normalizing the activations before computing the dimension of activity would remove the effect of heterogeneous gains on the dimension. This is not the case, however, since neuronal gains affect the recurrent dynamics, not merely the observed firing rates.

To further explore these effects, we compare the activity dimensions for three sets of activations: unnormalized activations and normalized activations (1 and 2 above) and (3) normalized activations multiplied by an independent set of "readout" gains, $\Phi_i^{\mathrm{readout}}(t) = G_i^{\mathrm{readout}}\phi_i(t)$, where the readout gains have the same distribution as the actual gains. Case (3), whose four-point function is given by Eq. (J18), considers variables with the same distribution of firing rates as the unnormalized activations, but where this heterogeneity is unrelated to the recurrent dynamics.

To validate our theoretical predictions [Eqs. (46a), (46b), and (J18)], we simulated networks with component strengths and gains given by $D_a \propto \exp(-\beta_D a/M)$ and $G_i \propto \exp(-\beta_G i/N)$ (Fig. 5). The dimension of activity is determined by the coupling strength g_{eff}, the effective rank $\alpha\mathrm{PR}^{D}$, and the participation ratio of the gain distribution PR^{G}. Increases in each of these parameters lead to a higher activity dimension for all three sets of activations.

Normalized activations exhibit the highest dimension. While scaling normalized activations by heterogeneous factors reduces dimension as expected, the magnitude of this reduction depends on whether G_i or G_i^{readout} is used, even when their distributions are identical. This occurs because neurons with the largest gains preferentially participate in the leading modes of the normalized activations. Further scaling these already-dominant neurons by their gains results in overrepresentation of these neurons compared to using random gains. Thus, the dimension of unnormalized activations is lowest, with the dimension of random-readout activations falling between those of the normalized and unnormalized activations.

In experimental data analysis, it is common to normalize single-neuron activities to control for differences in firing rates across neurons. Our analysis implies that such normalization may not eliminate the effects of single-neuron heterogeneity on collective-activity properties. The persistence of these effects through normalization provides an experimental signature for distinguishing between heterogeneity that affects recurrent dynamics and heterogeneity that affects only the observed firing rates, for example, in a circuit that is not recurrent.

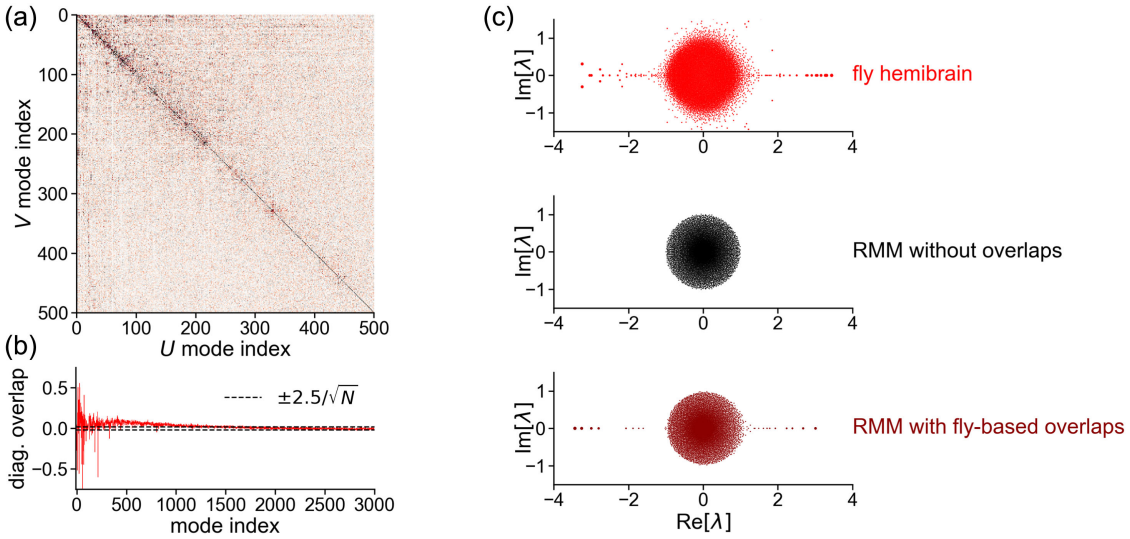

FIG. 6. Right-left mode overlaps and eigenvalue spectra of the fly-brain connectome. (a) Overlap matrix between right and left singular vectors of the connectome. First 500 modes are shown. (b) First 3000 diagonal elements of the overlap matrix from (a). Horizontal lines indicate an approximate cutoff for $\mathcal{O}(1/\sqrt{N})$ random overlaps. (c) Eigenvalue spectra of various connectome-based matrices. Top: actual hemibrain spectrum, exhibiting large outliers along the real axis. Middle: eigenvalues from a realization of the random-mode model (RMM), using singular values from the connectome as component strengths D_a, resulting in a circularly symmetric spectrum. Bottom: random-mode model realization with right-left mode overlaps incorporated, where overlaps are given by the inner products of left and right singular vectors from the fly connectome SVD. Data from Scheffer *et al.* [17], $N = 18028$ neurons.

VII. MODE OVERLAPS

Thus far, we have assumed i.i.d. modes, with the factorization $P(\ell_a, r_a) = P(\ell_a)P(r_a)$ implying that left-right overlaps $\boldsymbol{\ell}_a^T \boldsymbol{r}_a$ are random and $\mathcal{O}(1/\sqrt{N})$. The same holds for $\boldsymbol{\ell}_a^T \boldsymbol{r}_b$ with $a \neq b$ due to the more fundamental assumption that $P(\{\ell_a, r_a\}_{a=1}^M)$ factorizes across a. However, computations in biological circuits require structured interactions between modes, suggesting stronger, nonrandom overlaps between left and right modes. To test this, we return to the *Drosophila* connectome and compute the overlap matrix between left and right singular vectors, $O_{ab} = \boldsymbol{u}_a^T \boldsymbol{v}_b$ [Fig. 6(a)]. Contrary to the small, unstructured overlaps expected under the i.i.d. assumption, the connectome exhibits large, structured overlaps, particularly along the diagonal [Fig. 6(b)].

Motivated by the presence of these large diagonal overlaps, we extend the random-mode model to include correlations between corresponding left and right modes. Specifically, we take $P(\ell_a, r_a)$ to be a zero-mean bivariate Gaussian with marginal variance $1/N$ and covariance ρ_a/N, where ρ_a is the correlation between the ath mode pair. Both D_a and ρ_a are treated deterministically, and we assume that $M^{-1}\sum_a f(D_a, \rho_a)$ converges to a limiting value $\langle f(D,\rho)\rangle_{D,\rho}$ as $N \to \infty$. In this setting, $\boldsymbol{\ell}_a^T \boldsymbol{r}_a = \rho_a$, with negligible $\mathcal{O}(1/\sqrt{N})$ fluctuations around this structured overlap, while $\boldsymbol{\ell}_a^T \boldsymbol{r}_b$ for $a \neq b$ remains unstructured and $\mathcal{O}(1/\sqrt{N})$. This additional structure produces eigenvalue spectra resembling that of the connectome, including large real outliers [Fig. 6(c)] [68].

This purely diagonal overlap structure, however, can generate *only* real eigenvalue outliers. The presence of complex outliers in the connectome, together with the full overlap matrix itself [Fig. 6(a)], implies the presence of richer correlations. Capturing such features would require relaxing the full factorization of $P(\{\ell_a, r_a\})$ across a while retaining $\mathcal{O}(M)$ parameter scaling—for example, via blockwise factorization or Markovian correlations across a. We leave such generalizations for future work. We also note that, by setting $\rho_a = 1$ for all a, the model becomes a Hopfield network storing an extensive set of Gaussian patterns. In such networks, each mode can be "condensed" [$\mathcal{O}(1)$ overlap with activity] or "uncondensed" [$\mathcal{O}(1/\sqrt{N})$ overlap with activity]. Our analysis assumes no condensed modes, but sufficiently strong positive overlaps break this assumption. The Gaussian case, however, is unable to generate multiple multistable attractors as in a memory system [44]. Note that the form of the single-element density is relevant for the condensed patterns, but not for the uncondensed patterns. Allowing for condensed patterns within our analysis would require specifying this density. See Ref. [69] for a DMFT analysis of generalized Hopfield dynamics in which these ideas are implemented.

In Appendix K, we compute the two- and four-point functions. As in previous cases, the dynamics reduce to single-site processes,

$$T[x](t) = \eta^x(t) + \int^t dt' R_\star(t,t')\phi(t'), \tag{47a}$$

$$\eta^x \sim \mathcal{GP}(0, Q_\star), \tag{47b}$$

$$C^\phi_\star(t_1,t_2) = \langle \phi(t_1)\phi(t_2)\rangle_\star, \tag{47c}$$

$$S^\phi_\star(t_1,t_2) = \left\langle \frac{\delta\phi(t_1)}{\delta I(t_2)} \right\rangle_\star. \tag{47d}$$

Assuming temporal stationarity, the kernels are given in Fourier space by

$$Q_\star(\omega) = \alpha \left\langle \frac{D^2}{|1 - D\rho S^\phi_\star(\omega)|^2} \right\rangle_{D,\rho} C^\phi_\star(\omega), \tag{48}$$

$$R_\star(\omega) = \alpha \left\langle \frac{D\rho}{1 - D\rho S^\phi_\star(\omega)} \right\rangle_{D,\rho}. \tag{49}$$

Unlike the i.i.d. modes case, the single-site picture is not equivalent to that of an i.i.d. network for some g_{eff}.

To better understand the self-coupling kernel $R_\star(\omega)$, we expand in powers of ρ. The coefficients encode connectivity motifs of increasing orders,

$$R_\star(\omega) = \langle J_{ii}\rangle_\star + \langle NJ_{ij}J_{ji}\rangle_\star \, S^\phi_\star(\omega) + \cdots, \quad \text{where}$$
$$\langle J_{ii}\rangle_\star = \alpha\langle D\rho\rangle_{D,\rho}, \qquad \langle NJ_{ij}J_{ji}\rangle_\star = \alpha\langle (D\rho)^2\rangle_{D,\rho}. \tag{50}$$

Here, $\langle J_{ii}\rangle_\star$ corresponds to deterministic self-couplings or autapses. To avoid such biologically implausible $\mathcal{O}(1)$ self-connections, we set $\langle D\rho\rangle_{D,\rho} = 0$, yielding $J_{ii} = \mathcal{O}(1/\sqrt{N})$ [70]. The second term in the expansion (note that $i \neq j$) reflects reciprocal correlations in an otherwise i.i.d. network (Appendix H). Nonzero higher-order terms indicate that this structure cannot be reduced to either of these simpler cases. Notably, if both $\langle NJ_{ii}\rangle_\star$ and $\langle NJ_{ij}J_{ji}\rangle_\star$ vanish, all higher-order terms vanish as well, making reciprocal covariance necessary and sufficient for a nonzero self-coupling kernel in the single-site description.

The four-point function, meanwhile, is given by

$$\Psi^\phi_\star = \frac{V - 1 + \{\frac{(1+X_{12})(1+X^*_{21})}{1-W^*} + \text{H.c.}\}}{|1-U|^2} C^\phi_{12},$$
$$U = \alpha\langle D^2\Sigma_1\Sigma_2\rangle_{D,\rho}, \qquad V = \alpha\langle D^4|\Sigma_1\Sigma_2|^2\rangle_{D,\rho},$$
$$W = \alpha\langle D^2\rho^2\Sigma^*_1\Sigma_2\rangle_{D,\rho}, \qquad X_{12} = \alpha\langle D^3\rho\Sigma_1|\Sigma_2|^2\rangle_{D,\rho},$$
$$\Sigma_k = \frac{S^\phi_k}{1 - D\rho S^\phi_k} \qquad (k = 1,2). \tag{51}$$

We validate this theory using $\alpha = 1$ and $D_a \propto \exp(-\beta_D a/N)$, as before. For overlaps, we use a sigmoidal spectrum for ρ_a qualitatively inspired by Fig. 6(b), with multiplicative prefactor γ_ρ and other parameters chosen such that $\langle D\rho\rangle_{D,\rho} = 0$ (Appendix K 3). Across values of g_{eff} and PR^D, sufficiently strong overlaps, of either sign, reduce the dimension of activity (Fig. 7). The maximum dimension occurs for $\gamma_\rho < 0$, where dynamics most strongly suppress activity aligned with dominant connectivity modes. This effect weakens as g_{eff} increases.

In summary, our analysis reveals how connectivity structure shapes collective activity and its dimensionality. When mode overlaps are unstructured, connectivity structure reduces to two parameters: the effective rank αPR^D and effective coupling strength g_{eff}. Structured left-right mode overlaps break this simplification: Collective dynamics then depend on the full joint distribution of component strengths and overlaps, providing an additional mechanism for controlling activity dimensionality.

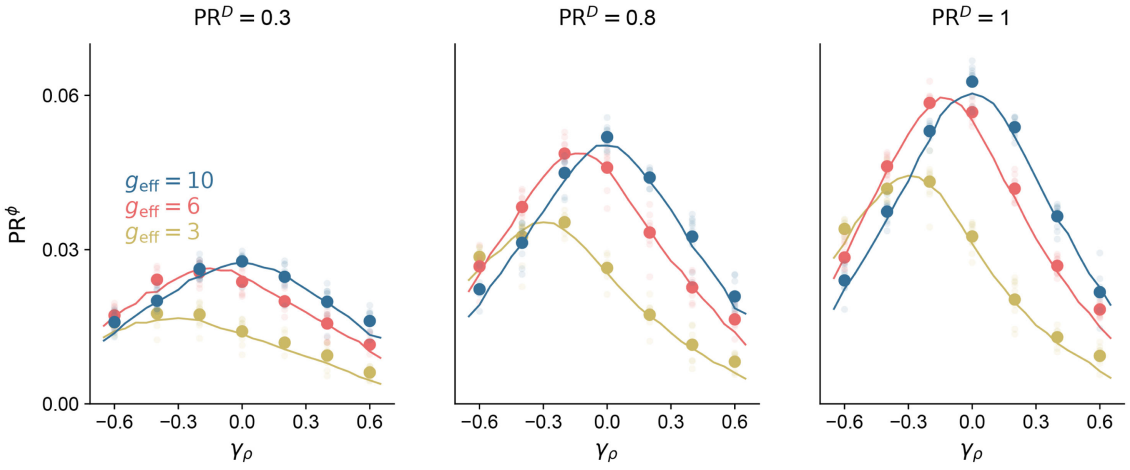

FIG. 7. Effect of right-left mode correlations on dimension of activity in the random-mode model. Dimension of activity PR^ϕ versus correlation parameter γ_ρ in the (D,ρ) joint distribution, for various coupling strengths g_{eff} and effective ranks αPR^D. Thin dots, individual simulations; thick dots, means over ten simulations; lines, theoretical predictions. All simulations use $N = 5000$ neurons.

VIII. DISCUSSION

A. Motifs versus global structure

Neuronal connectivity is often described using statistics of motifs: local patterns like chains or cycles among small groups of neurons [22,73–75]. These can be estimated from partial observations of a weight matrix which, until recently, set the limits of experimental measurement. Advances in whole-brain connectome reconstruction now allow access to global connectivity, including spectral features. Our parametrization of coupling matrices represents this global structure, with control over both the spectrum and mode overlaps. These spectral features determine how activity modes interact dynamically and thus may more directly dictate neural computations compared to motifs.

That said, in some studies the primary goal is to capture motifs or graph-theoretic features (e.g., degree distributions) over which the random-mode model does not provide control (although some of these features have spectral signatures that could be incorporated). When such features are the focus, other connectivity models are better suited—for instance, maximum-entropy random graph models constrained by fixed structural features (e.g., the configuration model for matching degree distributions).

B. Linking connectivity and activity datasets

The random-mode model, together with its dynamical theory, could aid in the interpretation of large-scale neural activity recordings in the context of anatomical connectivity, such as in *Drosophila*, where both the connectome and whole-brain activity data are available [76]. In the current work, we used this connectome to characterize the spectral structure of an actual large neural circuit, stopping short of considering the activity that it generates. Future work could progressively incorporate richer connectome constraints into the model to identify which features most strongly shape collective activity. One possible sequence would be to add the spectrum of component strengths, then diagonal mode overlaps, and finally off-diagonal overlaps, assessing at each step how much the correspondence between modeled and observed activity improves. This is similar to approaches in statistical physics that aim to recover higher-order correlations by constraining lower-order statistics [38,77,78].

C. Learning

Learning alters collective dynamics by modifying connectivity, the core principle of machine learning. Our analysis focuses on structured but untrained networks that produce chaotic activity. Nevertheless, the extensive low-rank connectivity structures we study provide a starting point for understanding trained networks.

Large networks trained on simple tasks often show intensive low-rank modifications to their initial couplings, resulting in a finite number of task-relevant eigenvalues [27,43]. Existing theoretical tools can already describe such systems by taking the trained weights, performing a low-rank decomposition, modeling mode overlaps statistically, and studying the resulting low-rank DMFT [45]. By contrast, empirical studies of large networks trained on real-world tasks suggest the emergence of extensive low-rank structure, with smooth spectral profiles with many significant components [28]. This structure resembles that modeled by the random-mode model. Whether mode overlaps extracted from such networks can be incorporated into the random-mode model framework, and thereby yield a working theory for the extensive-rank case, is an open question.

A more ambitious direction is to develop a DMFT for networks whose weights are learned from data, rather than first training networks and then retrofitting a statistical model. Such a theory would describe not only how data shape the weights, as addressed by classical works [79], but also how data interact with recurrent network dynamics. The path-integral formalism is well suited to this, as it allows derivation of DMFT equations for coupling matrices drawn from a Gibbs ensemble with an energy function given by a task loss. While this approach has been applied to low-dimensional tasks [80], applying it to high-dimensional tasks could yield a theory of how such tasks generate extensive-rank weights that, in turn, produce extensive-dimensional activity whose single-neuron and collective properties the theory would describe.

Another route to extensive-rank structure is via networks that perform many tasks, which could be modeled, for example, by representing mode overlaps through blockwise factorization of $P(\{\ell_a, r_a\}_{a=1}^{M})$ across a. This "multitask" scenario is increasingly relevant in both neuroscience and machine learning [81,82].

Even in standard recurrent neural network training, initializing couplings with low-rank structure could improve procedures such as FORCE learning, which rely on suppressing chaos [12,13]. Lower-dimensional attractors may be easier to control, potentially leading to faster convergence.

D. Modeling functional consequences of single-neuron heterogeneity

A central question in neuroscience is how neuronal diversity shapes collective dynamics and computation [83]. Appendix J presents a general framework that assigns each neuron a parameter vector $\boldsymbol{\theta}$ and enables the study of how different forms of heterogeneity affect properties such as dimensionality. Section VI focuses on gain heterogeneity to model firing-rate diversity, but this framework could also incorporate more intricate cell-specific features, such as electrophysiological characteristics [84], linking single-neuron heterogeneity to functional consequences for population activity.

E. Transient dynamics and excitation-inhibition

Our analysis has focused on stationary states with translation-invariant temporal correlations. Many neural computations, however, rely on transient dynamics that break stationarity and depend on absolute time. Extending the DMFT accordingly is numerically tractable for two-point functions but would become unwieldy for four-point functions, $\Psi^a_\star(t_1, t_2, t_3, t_4)$. Representing and manipulating such a large object may require approximations or low-rank decomposition.

One notable transient phenomenon is transient amplification, linked to non-normal coupling matrices that naturally arise from excitatory-inhibitory segregation (Dale's law) [85,86]. Although the random-mode model does not enforce Dale's law, the resulting matrices are nevertheless non-normal and may exhibit transient amplification; such effects may be amplified through structured off-diagonal overlaps between left and right modes. Non-normality also decouples the highest-variance dimensions from the slowest modes. Relatedly, extending the random-mode model to respect Dale's law and incorporating a nonsaturating $\phi(x)$—achieving stability via excitation-inhibition balance rather than saturation [87,88]—could enable more direct comparisons to both connectivity and activity data.

F. Random basis property

The random-mode model assumes that the $2M$ mode components associated with each neuron are sampled i.i.d. for each neuron [Eq. (3)]. This leads to neuronal permutation symmetry of the distribution over $\boldsymbol{J}$. Furthermore, when the marginal distributions over left and right mode components are isotropic Gaussians, as we have assumed, the embedding into neuronal space is, by definition, random and Gaussian. This "random basis property" may underlie heterogeneous tuning observed in neural circuits [29], but in our case it imposes modeling limitations: While the random-mode model has configurable spectral properties, it precludes configuration of spatial or anatomical properties.

This limitation could be addressed by conditioning the distribution over the $2M$ mode components on neuron-local properties such as brain region or spatial location, provided the number of effectively distinct possibilities is intensive. This would yield a regionally or spatially embedded random-mode model, apt for connectome studies.

Because we assume zero-mean mode components [Eq. (4)], $\langle J_{ij}\rangle = 0$ for $i \neq j$. Thus, the low-rank structure is *not* inherited from the expected weight matrix. This differs from the mechanism described by Thibeault *et al.* [23], who analyze cases where low-rank structure arises because $\langle \boldsymbol{J}\rangle$ is low rank, with the random part treated as noise. By Weyl's inequality, the singular values of $\boldsymbol{J}$ differ from those of $\langle \boldsymbol{J}\rangle$ by, at most, the spectral norm of the noise. This mechanism is fundamentally different from the random-mode model's zero-mean, randomly embedded structure. Which of these mechanisms is more prevalent in real-world networks is an open question.

G. Alternative generative models

Tiberi *et al.* [89] proposed a generative model for coupling matrices that parametrizes the eigendecomposition of $\boldsymbol{J}$ rather than its SVD. Their formulation does not capture relationships between eigenvalues and eigenvectors. Our SVD-like approach allows simultaneous control over component strengths and their corresponding mode overlaps. Whereas Tiberi *et al.* [89] studied noise-driven linear networks within a random-matrix-theory framework, our DMFT formulation applies to nonlinear dynamics.

Connectivity parametrizations resembling the random-mode model have been used to study the correspondence between spiking and rate-based recurrent neural networks [90] and to expedite training spiking networks [91].

IX. CONCLUSION

The random-mode model and associated analytical techniques provide tools for exploring the relationship between connectivity and collective dynamics. Our parametrization of connectivity is a middle ground between intensive-rank models and i.i.d. random networks. Building on this framework stands to further bridge neuronal connectivity, activity, and function.

ACKNOWLEDGMENTS

We thank L. F. Abbott for feedback. A. v. M. was supported by the Swartz Foundation and the Kempner Institute for the Study of Natural and Artificial Intelligence. D. G. C, O. M., and A. L.-K. were supported by the Kavli Foundation and Gatsby Charitable Foundation (Grant No. GAT3708). D. G. C. was also supported by the National Institutes of Health (Grant No. T32NS064929). A. L.-K. was also supported by the National Science Foundation (Grant No. 2443158) and the National Institutes of Health (Grant No. RF1DA060772).

APPENDIX A: GLOSSARY OF TERMS

Number of neurons (N): Total number of neurons in the network.

Number of modes (M): Number of components in the random-mode model.

(Effective) coupling strength (g, g_{eff}): Standard deviation of the couplings times $\sqrt{N}$ in the i.i.d. connectivity model, random-mode model [Eq. (37)], or variants thereof [e.g., see just below Eqs. (46a) and (46b)].

Random-mode model: Network connectivity model: $\boldsymbol{J} = \sum_{a=1}^{M} D_a \boldsymbol{\ell}_a \boldsymbol{r}_a^T$ [Eq. (2)].

Connectivity component ($\boldsymbol{\ell}_a \boldsymbol{r}_a^T$): Outer product of output and right modes.

TABLE I. Mapping of neurotransmitters to synaptic signs.

Neurotransmitter	Synaptic sign
GABA	−1 (Inhibitory)
Acetylcholine	+1 (Excitatory)
Glutamate	−1 (Inhibitory)
Serotonin	Ignored
Octopamine	Ignored
Dopamine	Ignored
Neither	Ignored

Right mode ($\boldsymbol{r}_a$): Neuronal pattern in the ath component onto which activity patterns are projected.
Left mode ($\boldsymbol{\ell}_a$): Neuronal pattern in the ath component along which the projected activity pattern is expanded.
Component strength (D_a): Scaling factor for the ath component.
Ratio of modes to neurons (α): Given by $\alpha = M/N$ [Eq. (7)].
Participation ratio of component strengths (PR^D): Given by $\mathrm{PR}^D = r_2^2/r_4$ [Eq. (11)], where $r_n = \langle D^n \rangle_D$ [Eq. (10)].
Effective rank ($\alpha\mathrm{PR}^D$): Connectivity dimensionality measure [Eq. (12)].
Two-point function: Correlation function given by $C^{\phi}(t_1, t_2) = (1/N)\sum_{i=1}^{N} \phi_i(t_1)\phi_i(t_2)$ [Eq. (17)].
Four-point function: Higher-order correlation function given by $\Psi^{\phi}(\tau_1, \tau_2) = (1/N)\sum_{i,j=1}^{N} C_{ij}^{\phi}(\tau_1)C_{ij}^{\phi}(\tau_2)$ [Eq. (24)]. See also time-by-time definition [Eq. (25)].

APPENDIX B: PREPROCESSING OF FLY HEMIBRAIN CONNECTOME

We use the dataset from Scheffer *et al.* [17], in which synaptic counts are reported for each connection. These counts determine the magnitudes of the coupling matrix elements. Neurotransmitter probabilities for each neuron are obtained from a machine-learning analysis of the electron microscopy data [92]. Each neuron is assigned its most probable neurotransmitter, which is then mapped to a synaptic sign (excitatory or inhibitory) according to Table I [93,94].

Synaptic counts do not directly correspond to effective synaptic strengths, so it is reasonable to choose "fudge factors" to relate the two. Our primary goal in choosing these factors is to standardize the spectrum of the coupling matrix so that it is roughly confined to the unit disk. At the same time, we want to preserve key structural features of the original connectivity.

Specifically, the preprocessing is designed to preserve:
(i) the sparsity pattern (for which the connections are zero versus nonzero);
(ii) the sign of each connection (excitatory versus inhibitory);
(iii) approximately, the relative magnitudes of same-sign inputs to each neuron; and
(iv) approximately, the relative magnitudes of same-sign outputs from each neuron.

The last two properties are preserved only approximately because they are enforced simultaneously by the iterative normalization procedure described below.

We normalize the coupling matrix using an iterative scaling method analogous to the Sinkhorn-Knopp algorithm. First, for each row, we rescale the positive elements by a row-specific factor so that their $L2$ norm equals $1/\sqrt{2}$. We then rescale the negative elements, using a separate factor, so that their $L2$ norm also equals $1/\sqrt{2}$. This yields rows with a total $L2$ norm of 1 and balanced excitatory and inhibitory contributions. A small number of neurons with only excitatory or only inhibitory inputs retain an $L2$ norm of $1/\sqrt{2}$.

Next, we apply the same rescaling procedure to each column, which disrupts the previous row normalization. We alternate the row and column scaling steps until convergence. The result can be interpreted as the projection of the original matrix onto the set of matrices that are properly normalized along both rows and columns, while keeping the same nonzero elements and signs.

This iterative normalization provides a principled way to choose the required scale factors while achieving our goal of a standardized spectrum.

APPENDIX C: FREE PROBABILITY CALCULATION

We wish to determine the limiting density of the singular values of

$$\boldsymbol{J} = \boldsymbol{L}\boldsymbol{R}^T, \tag{C1}$$

where $\boldsymbol{L}$ and $\boldsymbol{R}$ are independent $N \times M$ random matrices with $\alpha = M/N$ and entries of variance $1/N$. Equivalently, we can compute the eigenvalues of

$$\boldsymbol{J}^T\boldsymbol{J} = \boldsymbol{R}\boldsymbol{L}^T\boldsymbol{L}\boldsymbol{R}^T \tag{C2}$$

and then take the square root to obtain the singular values. These eigenvalues are the same as those of

$$\boldsymbol{L}^T\boldsymbol{L}\boldsymbol{R}^T\boldsymbol{R}, \tag{C3}$$

which is the product of two independent Wishart matrices,

$$\boldsymbol{L}^T\boldsymbol{L}\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{W}_L\boldsymbol{W}_R. \tag{C4}$$

The Stieltjes transform $G(z)$ of a spectral density $\rho(\lambda)$ is defined as

$$G(z) = \int_{-\infty}^{\infty} \frac{d\lambda \rho(\lambda)}{z - \lambda}, \qquad z \in \mathbb{C} \backslash \mathbb{R}, \tag{C5}$$

$$\rho(\lambda) = -\frac{1}{\pi} \lim_{\epsilon \to 0^+} \operatorname{Im} G(\lambda + i\epsilon). \tag{C6}$$

In a two-dimensional electrostatics analogy, the real and imaginary parts of $G(z)$ correspond to the Cartesian components of the electric field generated by a line charge density $\rho(\lambda)$ on the real axis. The inversion formula is then analogous to Gauss's law, relating the discontinuity of the normal field component to the local charge density.

The moment-generating function and S transform are defined by

$$M(z) = zG(z) - 1, \tag{C7}$$

$$S(z) = \frac{1 + z}{zM^{-1}(z)}, \tag{C8}$$

where $M^{-1}(z)$ is the functional inverse of $M(z)$. For a Marchenko-Pastur distributed Wishart matrix $\boldsymbol{W}$ with aspect ratio α,

$$S(z) = \frac{1}{1 + \alpha z}. \tag{C9}$$

A key fact in free probability is that the S transforms of free variables multiply. Thus, for $\boldsymbol{W}_1 \boldsymbol{W}_2$, the S transform is

$$S(z) = \frac{1}{(1 + \alpha z)^2}. \tag{C10}$$

From the definition of the S transform, the inverse-moment-generating function is

$$M^{-1}(z) = \frac{1 + z}{zS(z)} = \frac{(1 + z)(1 + \alpha z)^2}{z}. \tag{C11}$$

Therefore, $M(z)$ satisfies

$$z = \frac{[1 + M(z)][1 + \alpha M(z)]^2}{M(z)}. \tag{C12}$$

Using $M(z) = zG(z) - 1$, we obtain a cubic equation for $G(z)$,

$$\begin{aligned} &-\alpha^2 z^3 G(z)^3 + 2z^2(\alpha^2 - \alpha) G(z)^2 \\ &\quad + z(z + 2\alpha - \alpha^2 - 1) G(z) - z = 0. \end{aligned} \tag{C13}$$

In principle, this can be solved using the cubic formula, but the explicit form is cumbersome. For the support boundaries of $\rho(\lambda)$, we need only the points where $\operatorname{Im} G(z)$ changes from zero to nonzero. This occurs when the discriminant of the cubic vanishes. The discriminant is

$$\begin{aligned} \Delta = \alpha^2 \lambda^7 [&4\lambda^2 + (\alpha^2 - 20\alpha - 8)\lambda \\ &+ (-4\alpha^3 + 12\alpha^2 - 12\alpha + 4)], \end{aligned} \tag{C14}$$

where we replaced z with λ to indicate restriction to the real axis. Setting $\Delta = 0$ gives $\lambda = 0$ or the nontrivial quadratic equation

$$4\lambda^2 + (\alpha^2 - 20\alpha - 8)\lambda + (-4\alpha^3 + 12\alpha^2 - 12\alpha + 4) = 0. \tag{C15}$$

The roots of this quadratic are

$$\lambda_\pm = 1 + \frac{5\alpha}{2} - \frac{\alpha^2}{8} \pm \left(1 + \frac{\alpha}{8}\right)^{3/2} \sqrt{8\alpha}. \tag{C16}$$

The singular values correspond to the square roots of the eigenvalues, given by Eq. (14). We confirm this formula numerically in Fig. 8.

APPENDIX D: ALTERNATIVE MEASURES OF EFFECTIVE DIMENSIONALITY

Alternative measures of effective dimensionality, which could be applied to both connectivity and activity in place of the participation ratio, have been proposed in the literature. Roy and Vetterli [95] define effective rank as the exponential of the entropy of a categorical probability distribution obtained by normalizing the eigenvalues to sum to 1, i.e., $p_i = \lambda_i / \sum_{j=1}^{N} \lambda_j$. Written in terms of this distribution, the participation ratio is $\mathrm{PR} = (\sum_{i=1}^{N} p_i^2)^{-1}$, while the exponentiated entropy is $\exp(H) = \exp(-\sum_{i=1}^{N} p_i \log p_i)$. By Jensen's inequality, $\mathrm{PR} \leq \exp(H)$. Exact equality holds in certain cases, including when exactly $K \leq N$ eigenvalues are equal and the rest are zero, yielding $\mathrm{PR} = \exp(H) = K$.

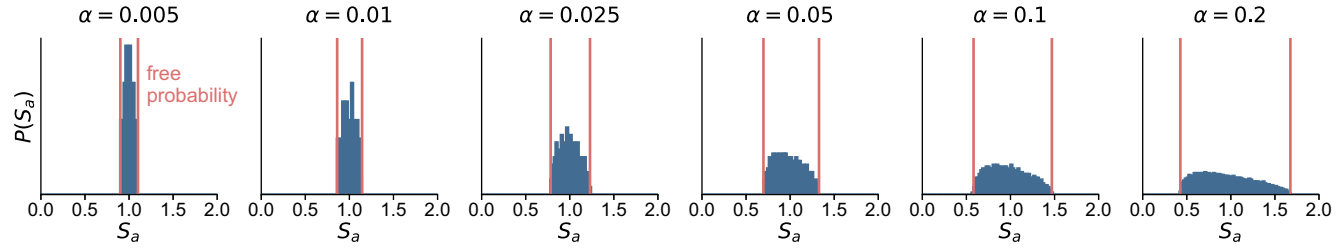

FIG. 8. Distribution of singular values in the random-mode model. Histograms show singular values of $\boldsymbol{LDR}^T$ with $D_a = 1$ for various α. For $\alpha \ll 1$, the singular values concentrate at 1. Vertical lines, prediction of Eq. (14). $N = 5000$.

Furthermore, for smoothly decaying spectra as discussed above [for which $p_i = f(i/Nw)/\sum_{j=1}^{N} f(j/Nw)$, with $w \ll 1$], both measures have $\mathrm{PR} = c_1\, Nw$ and $\exp(H) = c_2\, Nw$, where c_1 and c_2 are N- and w-independent constants depending on the specific form of f. Thus, while the measures differ in detail, they provide similar characterizations of effective dimensionality. This entropy-based definition has been used in neuroscience applications, for example, in Ref. [96] to analyze transient amplification in neural networks. Both measures capture the number of dominant modes in the system. Thibeault *et al.* (Ref. [23]) call (a metric closely related to) the participation ratio-based metric the "srank" (the difference being whether the maximum eigenvalue or sum over all eigenvalues is used as a normalizing factor) and call the entropylike metric the "erank."

APPENDIX E: REVIEW: CALCULATING TWO- AND FOUR-POINT FUNCTIONS FOR I.I.D. COUPLINGS

We now review how to compute these two- and four-point functions in a classic network model with i.i.d. couplings $\boldsymbol{J}$. The first- and second-order coupling statistics are

$$\langle J_{ij} \rangle = 0, \qquad \langle J_{ij}^2 \rangle = \frac{g^2}{N}. \tag{E1}$$

The magnitude of a typical coupling is $g/\sqrt{N}$. We refer to g as the coupling strength. For the canonical dynamics $T[x](t) = (1 + \partial_t)x(t)$, the network is quiescent for $g < 1$ and chaotic for $g > 1$, with a sharp phase transition as $N \to \infty$. In this paper, we assume that the network is nonquiescent.

The two-point function $C_\star^a(t_1, t_2)$ can be computed through a single-site picture that describes the dynamics of a typical neuron embedded in the rest of the network, with preactivation $x(t)$ and activation $\phi(t)$. The single-site dynamics are given by

$$T[x](t) = \eta^x(t), \tag{E2}$$

where $\eta^x(t)$ is a Gaussian field with mean zero and covariance $g^2 C_\star^\phi(t_1, t_2)$. We denote this by

$$\eta^x \sim \mathcal{GP}(0, g^2 C_\star^\phi). \tag{E3}$$

$C_\star^\phi(t_1, t_2)$ is determined self-consistently by enforcing

$$C_\star^\phi(t_1, t_2) = \langle \phi(t_1)\phi(t_2) \rangle_\star, \tag{E4}$$

where $\langle \cdots \rangle_\star$ denotes an average within this single-site process, i.e., with respect to the Gaussian distribution of $\eta^x(t)$. Once $C_\star^\phi(t_1, t_2)$ has been determined, $C_\star^x(t_1, t_2)$ follows easily. This single-site problem can be derived through either a single-site cavity calculation (see Ref. [97]) or a saddle-point condition in a path integral (see Refs. [60,61]). For the i.i.d. couplings considered here, there is a simpler heuristic derivation: In the neuronal input $\sum_{j=1}^{N} J_{ij}\phi_j(t)$, the correlations between the couplings J_{ij} and dynamic variables $\phi_j(t)$ can be safely neglected to leading order in $1/N$, yielding both Gaussianity of $\eta^x(t)$ by the central limit theorem and the second-order statistics of Eq. (E3).

Clark *et al.* [39] first computed $\Psi_\star^a(\boldsymbol{\tau})$ using a dynamic, two-site version of the cavity method, based on the neuron-by-neuron definition [Eq. (24)]. This method finds $\psi_\star^a(\boldsymbol{\tau})$, the off-diagonal contribution to $\Psi_\star^a(\boldsymbol{\tau})$, noting that the on-diagonal contribution is simply $C_\star^a(\tau_1)C_\star^a(\tau_2)$. A cavity is first created by removing two neurons from the network and allowing the rest of the network, the reservoir, to generate dynamic activity. The cavity neurons are then introduced, and their effect on the reservoir is treated

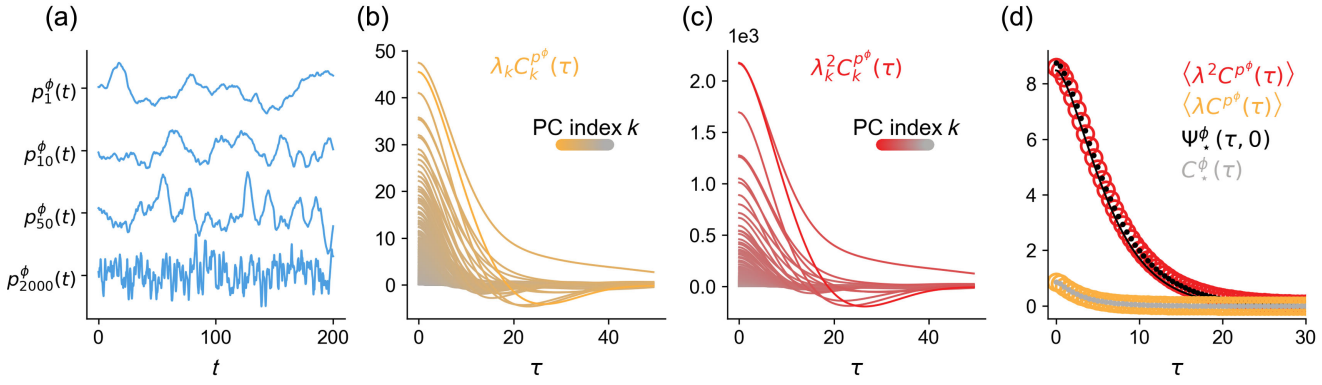

FIG. 9. Relationship between principal component timescales and two- and four-point correlation functions. (a) Example activity traces of principal components $p_k^\phi(t)$. (b) Empirical correlation functions for each principal component (PC), weighted by their eigenvalues. (c) Same as (b) but weighted by the squared eigenvalue. (d) Averages over curves in (b) and (c), with comparisons to $C_\star^\phi(\tau)$ and $\Psi_\star^\phi(\tau, 0)$. Dots, simulation results; lines, theoretical predictions. All panels describe the same network of $N = 4000$ neurons with i.i.d. couplings with variance g^2/N with coupling strength $g = 6$.

perturbatively. This yields a pair of coupled mean-field equations for the cavity units, generalizing the single-site picture discussed above to a two-site picture. Finally, self-consistency conditions are constructed by recognizing that the cavity pair is statistically equivalent to any reservoir pair. This calculation results in expressions for the four-point function in Fourier space, given by

$$\Psi^{\phi}_{\star}(\boldsymbol{\omega}) = \frac{C^{\phi}_{\star}(\omega_1)C^{\phi}_{\star}(\omega_2)}{|1 - g^2 S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2)|^2} \tag{E5}$$

for the activations, and

$$\Psi^{x}_{\star}(\boldsymbol{\omega}) = C^{x}_{\star}(\omega_1)C^{x}_{\star}(\omega_2) + |U(\boldsymbol{\omega})|^2 C^{\phi}_{\star}(\omega_1)C^{\phi}_{\star}(\omega_2) + U(\boldsymbol{\omega})C^{x\phi}_{\star}(\omega_1)C^{x\phi}_{\star}(\omega_2) + \text{H.c.}, \tag{E6a}$$

$$U(\boldsymbol{\omega}) = \frac{g^2 S^{x}_{\star}(\omega_1)S^{x}_{\star}(\omega_2)}{1 - g^2 S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2)} \tag{E6b}$$

for the preactivations, where $\boldsymbol{\omega} = (\omega_1, \omega_2)$ and $C^{x\phi}_{\star}(\omega)$ is a cross-covariance between the preactivation and activation. Note that, if the joint distribution of preactivations were Gaussian, $\Psi^{x}_{\star}(\boldsymbol{\tau})$ and $\Psi^{\phi}_{\star}(\boldsymbol{\tau})$ would differ only by a proportionality constant due to Price's theorem [98]. The more complex relationship observed here reflects the non-Gaussian joint statistics across different neurons, which is relevant because the network is nonlinear.

APPENDIX F: PATH-INTEGRAL CALCULATION OF TWO- AND FOUR-POINT FUNCTIONS FOR I.I.D. COUPLINGS

The general program, for i.i.d. or structured couplings, is as follows:

(1) Formulate the field theory. Begin with the path integral $Z[\boldsymbol{J}]$ and average over the connectivity disorder. Introduce auxiliary fields representing two-point functions [such as $C^{\phi}(t_1, t_2)$] along with their conjugate partners. Use integral representations of δ functions to enforce field definitions, factorizing the exponential over extensive dimensions. This yields a statistical field theory with action N times an order-one function of the auxiliary fields.
(2) Solve the saddle-point equations. Find the saddle point by setting action derivatives to zero. This produces self-consistent equations for the two-point functions, completing the standard DMFT analysis.
(3) Compute the Hessian. Calculate the time^2–by–time^2 Hessian matrix, which depends on four time variables (t_1, t_2, t_3, t_4) and characterizes fluctuations around the saddle point.
(4) Apply temporal separation limits. Impose the temporal limits [Eq. (34)], making the Hessian blocks translation-invariant functions of time differences (τ_1, τ_2). Transform to Fourier space to obtain a low-dimensional, frequency-dependent Hessian whose elements are functions of (ω_1, ω_2).
(5) Invert the frequency-space Hessian. Perform matrix inversion to directly obtain $\Psi^{a}_{\star}(\boldsymbol{\omega})$. If needed, apply inverse Fourier transform to recover $\Psi^{a}_{\star}(\boldsymbol{\tau})$.

Here, we consider i.i.d. $\boldsymbol{J}$ with mean zero and variance g^2/N. After performing the Gaussian integration over $\boldsymbol{J}$, we introduce an auxiliary field $C^{\phi}(t_1, t_2)$, defined by Eq. (17), and its conjugate $\hat{C}^{\phi}(t_1, t_2)$ to factorize the action across the neuron index i. This leads to a partition function for a statistical field theory involving $C^{\phi}(t_1, t_2)$ and $\hat{C}^{\phi}(t_1, t_2)$,

$$Z = \int \mathcal{D}C^{\phi} \int \mathcal{D}\hat{C}^{\phi} \exp\left(-N\mathcal{S}[C^{\phi}, \hat{C}^{\phi}]\right), \tag{F1}$$

where the intensive action is

$$\mathcal{S}[C^{\phi}, \hat{C}^{\phi}] = -\frac{1}{2} C^{\phi}\hat{C}^{\phi} - \log W[C^{\phi}, \hat{C}^{\phi}], \tag{F2}$$

and the single-site path integral is

$$W[C^{\phi}, \hat{C}^{\phi}] = \int \mathcal{D}x \int \mathcal{D}\hat{x} \exp\left(i\hat{x}T[x] - \frac{g^2}{2}\hat{x}C^{\phi}\hat{x} - \frac{1}{2}\phi\hat{C}^{\phi}\phi\right). \tag{F3}$$

1. Two-point functions

In the limit $N \to \infty$, the saddle point dominates this integral. The derivatives of the action, which are zero at the saddle point, are

$$\frac{\delta\mathcal{S}}{\delta C^{\phi}(t_1, t_2)} = -\hat{C}^{\phi}(t_1, t_2) + g^2\langle\hat{x}(t_1)\hat{x}(t_2)\rangle_W, \tag{F4}$$

$$\frac{\delta\mathcal{S}}{\delta \hat{C}^{\phi}(t_1, t_2)} = -C^{\phi}(t_1, t_2) + \langle\phi(t_1)\phi(t_2)\rangle_W, \tag{F5}$$

where $\langle\cdots\rangle_W$ denotes an average within the dynamic process described by $W[C^{\phi}, \hat{C}^{\phi}]$. In evaluating derivatives of the action, we follow the rule that derivatives with respect to $C^{\phi}(t_1, t_2)$ also affect $C^{\phi}(t_2, t_1)$, and likewise for $\hat{C}^{\phi}(t_1, t_2)$, due to the symmetry in the action. Using the vanishing of correlation functions involving only the conjugate field $\hat{x}(t)$, the saddle-point conditions yield

$$\hat{C}^{\phi}_{\star}(t_1, t_2) = 0, \tag{F6}$$

$$C^{\phi}_{\star}(t_1,t_2) = \langle \phi(t_1)\phi(t_2)\rangle_{\star}, \tag{F7}$$

where $\langle\cdots\rangle_{\star}$ denotes an average within the dynamic process described by $W[C^{\phi}_{\star},0]$. This recovers the same single-site process [Eqs. (E2) and (E3)] and self-consistency condition [Eq. (E4)] described in Appendix E.

2. Four-point functions

We now compute $\Psi^{a}_{\star}(\boldsymbol{\tau})$ within the path-integral formalism. Fluctuations around the saddle point derived above are governed by the Hessian of the action [60,99]. This Hessian has blocks given by

$$\begin{aligned} H_{C^{\phi}C^{\phi}}(t_1,t_2,t_3,t_4) &= \frac{\delta^2 S}{\delta C^{\phi}(t_1,t_2)\delta C^{\phi}(t_3,t_4)} \\ &= -g^4\langle\langle \hat{x}(t_1)\hat{x}(t_2),\hat{x}(t_3)\hat{x}(t_4)\rangle\rangle_{\star}, \end{aligned} \tag{F8a}$$

$$\begin{aligned} H_{C^{\phi}\hat{C}^{\phi}}(t_1,t_2,t_3,t_4) &= \frac{\delta^2 S}{\delta C^{\phi}(t_1,t_2)\delta \hat{C}^{\phi}(t_3,t_4)} \\ &= -I(t_1,t_2,t_3,t_4) \\ &\quad - g^2\langle\langle \hat{x}(t_1)\hat{x}(t_2),\phi(t_3)\phi(t_4)\rangle\rangle_{\star}, \end{aligned} \tag{F8b}$$

$$\begin{aligned} H_{\hat{C}^{\phi}\hat{C}^{\phi}}(t_1,t_2,t_3,t_4) &= \frac{\delta^2 S}{\delta \hat{C}^{\phi}(t_1,t_2)\delta \hat{C}^{\phi}(t_3,t_4)} \\ &= -\langle\langle \phi(t_1)\phi(t_2),\phi(t_3)\phi(t_4)\rangle\rangle_{\star}, \end{aligned} \tag{F8c}$$

where $\langle\langle A,B\rangle\rangle_{\star} = \langle AB\rangle_{\star} - \langle A\rangle_{\star}\langle B\rangle_{\star}$, and $I(t_1,t_2,t_3,t_4) = \delta(t_1-t_3)\delta(t_2-t_4) + \delta(t_1-t_4)\delta(t_2-t_3)$. Expanding the action to second order around the saddle point, the path integral becomes

$$Z = \exp\left(-N\mathcal{S}[C^{\phi}_{\star},0]\right)\int \mathcal{D}\delta C^{\phi}\mathcal{D}\delta\hat{C}^{\phi}\exp\left(-\frac{N}{2}\begin{pmatrix}\delta C^{\phi}\\ \delta\hat{C}^{\phi}\end{pmatrix}^{T}\begin{pmatrix}H_{C^{\phi}C^{\phi}} & H_{C^{\phi}\hat{C}^{\phi}}\\ H^{T}_{\hat{C}^{\phi}C^{\phi}} & H_{\hat{C}^{\phi}\hat{C}^{\phi}}\end{pmatrix}\begin{pmatrix}\delta C^{\phi}\\ \delta\hat{C}^{\phi}\end{pmatrix}\right),$$

revealing the structure of the Gaussian fluctuations around the saddle point. The covariance matrix among the fluctuation variables $\delta C^{\phi}(t_1,t_2)$ and $\delta\hat{C}^{\phi}(t_1,t_2)$ is $1/N$ times the inverse Hessian, yielding

$$N\langle \delta C^{\phi}(t_1,t_2)\delta C^{\phi}(t_3,t_4)\rangle_{\star} = [H_{C^{\phi}C^{\phi}} - H_{C^{\phi}\hat{C}^{\phi}}H^{-1}_{\hat{C}^{\phi}\hat{C}^{\phi}}H_{\hat{C}^{\phi}C^{\phi}}]^{-1}(t_1,t_2,t_3,t_4).$$

In principle, we have all the necessary components to evaluate this expression: The single-site dynamics are known via the saddle-point condition, and the Hessian blocks are given in terms of connected correlation functions in this single-site process. We could then take the temporal separation limits of Eq. (34) to obtain $\Psi^{\phi}_{\star}(\boldsymbol{\tau})$. The challenge is that each Hessian block is a complicated time^2–by–time^2 matrix that, for general values of the time variables, has no analytic inverse. To circumvent this problem, we show in Appendix G that the temporal limits of Eq. (34) can be taken *before* taking the inverses. Under these temporal limits, the Hessian blocks are

$$H_{\hat{C}^{\phi}\hat{C}^{\phi}}(t_1,t_2,t_3,t_4) = -C^{\phi}_{\star}(\tau_1)C^{\phi}_{\star}(\tau_2), \tag{F9}$$

$$H_{C^{\phi}\hat{C}^{\phi}}(t_1,t_2,t_3,t_4) = -\delta(\tau_1)\delta(\tau_2) + g^2 S^{\phi}_{\star}(\tau_1)S^{\phi}_{\star}(\tau_2), \tag{F10}$$

$$H_{C^{\phi}C^{\phi}}(t_1,t_2,t_3,t_4) = 0, \tag{F11}$$

where we used the fact that multiplication by $-i\hat{x}(t)$ is equivalent to a functional derivative with respect to a source at time t. Since the relevant quantities are now translation invariant, i.e., depend only on τ_1 and τ_2, we can transform to Fourier space. In this representation, each Hessian block becomes a frequency-dependent scalar,

$$H_{\hat{C}^{\phi}\hat{C}^{\phi}}(\boldsymbol{\omega}) = -C^{\phi}_{\star}(\omega_1)C^{\phi}_{\star}(\omega_2), \tag{F12}$$

$$H_{C^{\phi}\hat{C}^{\phi}}(\boldsymbol{\omega}) = -1 + g^2 S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2), \tag{F13}$$

$$H_{C^{\phi}C^{\phi}}(\boldsymbol{\omega}) = 0. \tag{F14}$$

In summary, the full Hessian can be replaced by a 2×2, frequency-dependent Hessian given by

$$\boldsymbol{H} = \begin{pmatrix} 0 & -1 + g^2 (S^{\phi}_{12})^{*} \\ -1 + g^2 S^{\phi}_{12} & -C^{\phi}_{12}\end{pmatrix}, \tag{F15}$$

where we simplified notation by suppressing frequency arguments via the shorthand $C^{\phi}_{12} = C^{\phi}_{\star}(\omega_1)C^{\phi}_{\star}(\omega_2)$ and $S^{\phi}_{12} = S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2)$. The Fourier-space function $\Psi^{\phi}_{\star}(\boldsymbol{\omega})$ is the upper-left element of the inverse of this matrix. Performing the 2×2 matrix inversion gives

$$\Psi^{\phi}_{\star} = \frac{C^{\phi}_{12}}{|1 - g^2 S^{\phi}_{12}|^2}, \tag{F16}$$

which agrees with the two-site cavity result [Eq. (E5)].

3. Adding sources

Above, computing $\Psi^{\phi}_{\star}(\boldsymbol{\tau})$ was simplified by $C^{\phi}(t_1, t_2)$ appearing naturally in the statistical field theory resulting from integrating out $\boldsymbol{J}$. To compute $\Psi^{x}_{\star}(\boldsymbol{\tau})$, we need to introduce a source-field term in the action.

Consider a general intensive action $\mathcal{S}[\boldsymbol{C}; \mathcal{J}]$, where $\boldsymbol{C}$ is a collection of fields (e.g., $\boldsymbol{C} = \{C^{\phi}, \hat{C}^{\phi}\}$), and $\mathcal{J}$ is a source. Our goal is to compute

$$\frac{\delta^2}{\delta \mathcal{J}^2} \log Z \bigg|_{\mathcal{J}=0}, \tag{F17}$$

which gives the fluctuations around the mean of the quantity multiplying $\mathcal{J}$ in the action. To facilitate this calculation, we introduce a new field U and set it equal to $\mathcal{J}$ using the conjugate $\hat{U}$,

$$Z[\mathcal{J}] = \int \mathcal{D}\boldsymbol{C} \int \mathcal{D}U \int \mathcal{D}\hat{U} \times \exp\left(-N\hat{U}U + N\hat{U}\mathcal{J} - N\mathcal{S}[\boldsymbol{C}; U]\right). \tag{F18}$$

Instead of taking the second derivative with respect to $\mathcal{J}$, we set $\mathcal{J} = 0$ and compute the fluctuations of $\hat{U}$, measured by Eq. (F17), in the augmented theory whose path integral is

$$Z = \int \mathcal{D}\boldsymbol{C} \int \mathcal{D}U \int \mathcal{D}\hat{U} \exp(-N\tilde{\mathcal{S}}[\boldsymbol{C}, U, \hat{U}]), \tag{F19}$$

where

$$\tilde{\mathcal{S}}[\boldsymbol{C}, U, \hat{U}] = \hat{U}U + \mathcal{S}[\boldsymbol{C}; U]. \tag{F20}$$

The Hessian of the augmented action, in terms of the Hessian of the original action, is

$$\left(\begin{array}{c|cc} \boldsymbol{H} & \frac{\partial^2 S}{\partial \mathcal{J} \partial \boldsymbol{C}} & \boldsymbol{0} \\ \hline \left(\frac{\partial^2 S}{\partial \mathcal{J} \partial \boldsymbol{C}}\right)^T & \frac{\partial^2 S}{\partial \mathcal{J}^2} & 1 \\ \boldsymbol{0}^T & 1 & 0 \end{array}\right), \tag{F21}$$

where

$$H_{ab} = \frac{\delta^2 S}{\delta C_a \delta C_b}, \qquad \left[\frac{\delta^2 S}{\delta \mathcal{J} \delta \boldsymbol{C}}\right]_a = \frac{\delta^2 S}{\delta \mathcal{J} \delta C_a}. \tag{F22}$$

The fluctuations of $\hat{U}$ are given by the bottom-right element of the inverse Hessian evaluated at the saddle point. Using a Schur complement to compute this element, we obtain

$$\frac{\delta^2}{\delta \mathcal{J}^2} \log Z \bigg|_{\mathcal{J}=0} = \left[-\frac{\delta^2 S}{\delta \mathcal{J}^2} + \left(\frac{\delta^2 S}{\delta \mathcal{J} \delta \boldsymbol{C}}\right)^T \boldsymbol{H}^{-1} \frac{\delta^2 S}{\delta \mathcal{J} \delta \boldsymbol{C}}\right]_{\mathcal{J}=0}, \tag{F23}$$

with all quantities evaluated at the saddle point. By applying this formula to the action of the i.i.d. model with the source term $-\frac{1}{2} x \mathcal{J} x$ added to the single-site path integral [Eq. (F3)], we recover the expression for $\Psi^{x}_{\star}(\boldsymbol{\tau})$ from the two-site cavity method [Clark *et al.* [39]; Eq. (F3)]. For actions of the form $\mathcal{S}[\boldsymbol{C}; \mathcal{J}] = -\mathcal{J} C_a + \mathcal{J}$-independent terms, this formula reduces to $(\delta^2/\delta \mathcal{J}^2) \log Z|_{\mathcal{J}=0} = [\boldsymbol{H}^{-1}]_{aa}$, as expected.

We now apply Eq. (F23) to the action of the i.i.d. connectivity model with a source term for correlations among preactivations. The action is

$$\mathcal{S}[C^{\phi}, \hat{C}^{\phi}] = -\frac{1}{2} C^{\phi} \hat{C}^{\phi} - \log \int \mathcal{D}x \int \mathcal{D}\hat{x} \times \exp\left(i\hat{x}T[x] - \frac{g^2}{2} \hat{x} C^{\phi} \hat{x} - \frac{1}{2} \phi \hat{C}^{\phi} \phi - \frac{1}{2} x \mathcal{J} x\right). \tag{F24}$$

Computing the necessary quantities and taking the temporal limits of Eq. (34), we obtain, in Fourier space,

$$\frac{\delta^2 S}{\delta \mathcal{J}^2} = -C^{x}_{12}, \qquad \frac{\delta^2 S}{\delta \mathcal{J} \delta \boldsymbol{C}} = \begin{pmatrix} g^2 (S^{x}_{12})^* \\ -C^{x\phi}_{12} \end{pmatrix}, \tag{F25}$$

where $C^{x\phi}_{12} = C^{x\phi}_{\star}(\omega_1) C^{x\phi}_{\star}(\omega_2)$ and $S^{x}_{12} = S^{x}_{\star}(\omega_1) S^{x}_{\star}(\omega_2)$. Substituting these into Eq. (F23) and using the frequency-dependent Hessian from Eq. (F15), we obtain

$$\Psi^{x}_{\star} = C^{x}_{12} + |U|^2 C^{\phi}_{12} + U C^{x\phi}_{12} + \text{H.c.}, \quad \text{where } U = \frac{g^2 S^{x}_{12}}{1 - g^2 S^{\phi}_{12}}, \tag{F26}$$

in agreement with the two-site cavity result of Clark *et al.* (Ref. [39]), given in Eq. (E6).

4. Other kinds of path integrals

Different path-integral formalisms have been applied to neural-network models depending on the underlying dynamics. The Doi-Peliti formalism [100] derives path integrals from master equations for systems with non-negative integer degrees of freedom (such as spiking neural networks) that evolve according to master equations, using an operator formalism as an intermediate step. In contrast, we use the MSRJD formalism for continuous-time stochastic differential equations, which naturally describes our rate-based network dynamics. For a comprehensive discussion of different path-integral approaches in neural-network theory, see Ref. [69].

APPENDIX G: TEMPORAL LIMITS AND INVERSES

Here, we demonstrate, for the i.i.d. model, that the temporal limits in Eq. (34) can be taken before the inverses. This allows us to work with a frequency-dependent Hessian. In this section, we use Einstein notation for integrals over time variables (i.e., repeated time indices are integrated over). We would like to evaluate

$$\Psi^{\phi}(\tau_1,\tau_2) = U^{-1}(t_1,t_2,s_1,s_2)K(s_1,s_2,s_3,s_4)U^{-1}(t_1+\tau_1,t_2+\tau_2,s_3,s_4) \tag{G1}$$

in the limits of Eq. (34), where $U(t_1,t_2,t_3,t_4) = H_{\hat{C}^{\phi}C^{\phi}}(t_1,t_2,t_3,t_4)$ and $K(t_1,t_2,t_3,t_4) = H_{\hat{C}^{\phi}\hat{C}^{\phi}}(t_1,t_2,t_3,t_4)$ [note that $H_{C^{\phi}C^{\phi}}(t_1,t_2,t_3,t_4) = 0$ due to the vanishing of correlation functions involving only the conjugate field]. Restating Eq. (F8), we have

$$U(t_1,t_2,t_3,t_4) = -[\delta(t_1-t_3)\delta(t_2-t_4)+\delta(t_1-t_4)\delta(t_2-t_3)] - g^2\langle\!\langle \phi(t_1)\phi(t_2),\hat{x}(t_3)\hat{x}(t_4)\rangle\!\rangle_{\star}, \tag{G2}$$

$$K(t_1,t_2,t_3,t_4) = \langle\!\langle \phi(t_1)\phi(t_2),\phi(t_3)\phi(t_4)\rangle\!\rangle_{\star}. \tag{G3}$$

We assume temporal stationarity throughout this section and consider only $a = \phi$.

We start by examining the properties of $U(t_1,t_2,t_3,t_4)$. U vanishes if one of the four time points is far [measured on the timescale of $C^{\phi}_{\star}(\tau)$] from the other three. For t_1 or t_2, this is a consequence of the perturbation associated with $\hat{x}(t_1)$ or $\hat{x}(t_2)$ being far away; for t_3 or t_4, this is a consequence of $\langle\phi(t_i)\rangle_{\star} = 0$ separating from the expectation. U also vanishes if there is more than one time point which is far from the remaining ones by the same arguments. Thus, the time points need to be close in pairs. However, if t_1, t_2 are close and t_3, t_4 are close, but the pairs are distant, U still vanishes since the perturbations associated with $\hat{x}(t_i)$ are both far. This leaves three possibilities for a nonvanishing U: t_1, t_3 close and t_2, t_4 close, but the pairs are distant; t_1, t_4 close and t_2, t_3 close, but the pairs are distant; or all time points are close.

Separating the three possibilities into two subdomains where all time points are close (c) and where the time points belong to either of the two pairwise separate (s) possibilities, we write $U = U_s + U_c$, where U_s and U_c are defined to vanish on the respective other domain. On the separated domain, U simplifies to

$$\begin{aligned} U_s(t_1,t_2,t_3,t_4) = &-[\delta(t_1-t_3)\delta(t_2-t_4)+\delta(t_1-t_4)\delta(t_2-t_3)] \\ &+ g^2[S^{\phi}_{\star}(t_1-t_3)S^{\phi}_{\star}(t_2-t_4)+S^{\phi}_{\star}(t_1-t_4)S^{\phi}_{\star}(t_2-t_3)], \end{aligned} \tag{G4}$$

where we used that the two distant contributions are nonoverlapping to add them.

We assume that we can also decompose the inverse into a contribution for separate pairs and a contribution where all time points are close, $U^{-1} = U_s^{-1} + U_c^{-1}$, where again U_s^{-1} and U_c^{-1} vanish on the respective other domain. We demonstrate that U_s^{-1} is the inverse of U_s, and that U_c^{-1} is the inverse of U_c, by noting that $U^{-1}(t_1,t_2,s_1,s_2)$ $U(s_1,s_2,t_3,t_4)$ reduces to the contraction of U_c^{-1} and U_c if t_3 and t_4, and hence s_1 and s_2, are close; or to the contraction of U_s^{-1} and U_s if t_3 and t_4, and hence s_1 and s_2, are distant.

For Eq. (G1), we need U_s^{-1} due to the limits of Eq. (34). The fact that U reduces to U_s in the limits of Eq. (34) justifies taking the limits before inverting U. Furthermore, these limits imply that, in Eq. (G1), s_1 and s_2, as well as s_3 and s_4, are far apart. In this case, K simplifies to

$$\begin{aligned} K_s(s_1,s_2,s_3,s_4) = &C^{\phi}_{\star}(s_1-s_3)C^{\phi}_{\star}(s_2-s_4) \\ &+ C^{\phi}_{\star}(s_1-s_4)C^{\phi}_{\star}(s_2-s_3). \end{aligned} \tag{G5}$$

This justifies taking the limits of Eq. (34) for K.

APPENDIX H: PARTIALLY SYMMETRIC DISORDER

To demonstrate the flexibility of the path-integral approach to computing the four-point function, we now consider a generalization of the classic i.i.d. model in which we introduce correlations between reciprocal couplings, characterized by the following statistics:

$$\langle J_{ij}J_{k\ell}\rangle = \frac{g^2}{N}\delta_{ik}\delta_{j\ell} + \frac{g^2\rho}{N}\delta_{i\ell}\delta_{jk}. \tag{H1}$$

Here, $\rho \in [-1,1]$ is a parameter that controls the degree of symmetry in the connectivity. When $\rho = 1$, the connectivity is fully symmetric, while $\rho = -1$ corresponds to fully antisymmetric connectivity. The case $\rho = 0$ recovers the i.i.d. model. We express this partially symmetric connectivity as a linear combination of two matrices, $\boldsymbol{X}$ and $\boldsymbol{Y}$, that are i.i.d. random with mean zero and variance $1/N$,

$$J_{ij} = \sqrt{g^2-|\sigma|}X_{ij} + \sqrt{\frac{|\sigma|}{2}}[Y_{ij}+\mathrm{sgn}(\sigma)Y_{ji}], \tag{H2}$$

where $\sigma = \rho g^2$. In terms of $\boldsymbol{X}$ and $\boldsymbol{Y}$, the path integral for this partially symmetric model is

$$Z[\boldsymbol{X},\boldsymbol{Y}] = \int \mathcal{D}\boldsymbol{x} \int \mathcal{D}\hat{\boldsymbol{x}} \exp\left(i \sum_{i=1}^{N} \hat{x}_i T[x_i] - i\sqrt{g^2 - |\sigma|} \sum_{i,j=1}^{N} X_{ij} \hat{x}_i \phi_j - i\sqrt{\frac{|\sigma|}{2}} \sum_{i,j=1}^{N} Y_{ij} [\hat{x}_i \phi_j + \mathrm{sgn}(\sigma) \phi_i \hat{x}_j] \right). \quad \text{(H3)}$$

We integrate out $\boldsymbol{X}$ and $\boldsymbol{Y}$. To factorize the action over indices spanning extensive dimensions, we introduce

$$C^{\phi}(t_1,t_2) = \frac{1}{N}\sum_{i=1}^{N} \phi_i(t_1)\phi_i(t_2), \quad \text{conjugate: } \hat{C}^{\phi}(t_1,t_2); \quad \text{(H4)}$$

$$S^{\phi}(t_1,t_2) = \frac{1}{N}\sum_{i=1}^{N} \frac{\delta \phi_i(t_1)}{\delta I_i(t_2)}, \qquad \text{conjugate: } \hat{S}^{\phi}(t_1,t_2). \quad \text{(H5)}$$

We obtain, for the disorder-averaged path integral,

$$Z = \int \mathcal{D}C^{\phi} \int \mathcal{D}\hat{C}^{\phi} \int \mathcal{D}R \int \mathcal{D}\hat{R} \exp(-N\mathcal{S}[C^{\phi}, \hat{C}^{\phi}, S^{\phi}, \hat{S}^{\phi}]), \quad \text{(H6)}$$

$$\mathcal{S}[C^{\phi}, \hat{C}^{\phi}, S^{\phi}, \hat{S}^{\phi}] = -\frac{1}{2} C^{\phi} \hat{C}^{\phi} + \frac{\sigma}{2} S^{\phi} \hat{S}^{\phi} - \log W[C^{\phi}, \hat{C}^{\phi}, S^{\phi}, \hat{S}^{\phi}], \quad \text{(H7)}$$

$$W[C^{\phi}, \hat{C}^{\phi}, S^{\phi}, \hat{S}^{\phi}] = \int \mathcal{D}\hat{x} \int \mathcal{D}x \exp\left(i\hat{x} T[x] - \frac{1}{2}\phi \hat{C}^{\phi} \phi - \frac{g^2}{2} \hat{x} C^{\phi} \hat{x} - \frac{i\sigma}{2} \phi \hat{S}^{\phi} \hat{x} - \frac{i\sigma}{2} \hat{x} S^{\phi} \phi \right). \quad \text{(H8)}$$

1. Two-point functions

To obtain the saddle-point solution, we compute the derivatives of the action, under the rule that derivatives with respect to $S^{\phi}(t_1,t_2)$ also affect $\hat{S}^{\phi}(t_2,t_1)$, and vice versa, due to the symmetry in the action. This gives

$$\frac{\delta \mathcal{S}}{\delta C^{\phi}(t_1,t_2)} = -\hat{C}^{\phi}(t_1,t_2) + g^2 \langle \hat{x}(t_1)\hat{x}(t_2) \rangle_W, \quad \text{(H9)}$$

$$\frac{\delta \mathcal{S}}{\delta \hat{C}^{\phi}(t_1,t_2)} = -C^{\phi}(t_1,t_2) + \langle \phi(t_1)\phi(t_2) \rangle_W, \quad \text{(H10)}$$

$$\frac{\delta \mathcal{S}}{\delta S^{\phi}(t_1,t_2)} = \sigma \hat{S}^{\phi}(t_1,t_2) + i\sigma \langle \hat{x}(t_1)\phi(t_2) \rangle_W, \quad \text{(H11)}$$

$$\frac{\delta \mathcal{S}}{\delta \hat{S}^{\phi}(t_1,t_2)} = \sigma S^{\phi}(t_1,t_2) + i\sigma \langle \phi(t_1)\hat{x}(t_2) \rangle_W. \quad \text{(H12)}$$

Setting these derivatives to zero yields the saddle-point conditions,

$$\hat{C}^{\phi}_{\star}(t_1,t_2) = 0, \quad \text{(H13)}$$

$$C^{\phi}_{\star}(t_1,t_2) = \langle \phi(t_1)\phi(t_2) \rangle_{\star}, \quad \text{(H14)}$$

$$S^{\phi}_{\star}(t_1,t_2) = \hat{S}^{\phi}_{\star}(t_2,t_1) = \left\langle \frac{\delta \phi(t_1)}{\delta I(t_2)} \right\rangle_{\star}, \quad \text{(H15)}$$

where $\langle \cdots \rangle_{\star}$ denotes an average within the dynamic process described by $W[C^{\phi}_{\star}, 0, S^{\phi}_{\star}, \hat{S}^{\phi}_{\star}]$. The single-site process at the saddle point is described by

$$T[x](t) = \eta^x(t) + \sigma[S^{\phi}_{\star} \circ \phi](t), \quad \text{(H16)}$$

$$\eta^x \sim \mathcal{GP}(0, g^2 C^{\phi}_{\star}), \quad \text{(H17)}$$

where $\circ$ denotes convolution. The symmetric structure provides a convolutional, nonlinear self-coupling $\sigma[S^{\phi}_{\star} \circ \phi](t)$ in the single-site dynamics. The two-point correlation and response functions $C^{\phi}_{\star}(\tau)$ and $S^{\phi}_{\star}(\tau)$ must be determined self-consistently within this single-site picture.

2. Four-point functions

Introducing the notation $C^{\phi}_k = C^{\phi}_{\star}(\omega_k)$, $S^{\phi}_k = S^{\phi}_{\star}(\omega_k)$ for $k = 1, 2$, the frequency-dependent Hessian at the saddle point is

$$\boldsymbol{H} = \begin{pmatrix} 0 & -1 + g^2 S^{\phi}_{12} & 0 & 0 \\ -1 + g^2 (S^{\phi}_{12})^* & -C^{\phi}_{12} & \sigma (S^{\phi}_1)^* C^{\phi}_2 & \sigma (S^{\phi}_2)^* C^{\phi}_1 \\ 0 & \sigma S^{\phi}_1 C^{\phi}_2 & 0 & \sigma - \sigma^2 S^{\phi}_1 (S^{\phi}_2)^* \\ 0 & \sigma S^{\phi}_2 C^{\phi}_1 & \sigma - \sigma^2 (S^{\phi}_1)^* S^{\phi}_2 & 0 \end{pmatrix}. \quad \text{(H18)}$$

Inverting this matrix and isolating the upper-left element, we obtain, in agreement with Clark *et al.* (Ref. [39]),

$$\Psi^{\phi}_{\star} = \frac{1}{|1 - g^2 S^{\phi}_{12}|^2} \frac{1 - |\sigma(S^{\phi}_1)^* S^{\phi}_2|^2}{|1 - \sigma(S^{\phi}_1)^* S^{\phi}_2|^2} C^{\phi}_{12}. \tag{H19}$$

APPENDIX I: PATH-INTEGRAL CALCULATION OF TWO- AND FOUR-POINT FUNCTIONS FOR THE RANDOM-MODE MODEL

Having demonstrated the path-integral approach to computing two- and four-point functions for i.i.d. couplings, we now apply this formalism to the random-mode model. In terms of the mode matrices $\boldsymbol{L}$ and $\boldsymbol{R}$, the path integral is

$$Z[\boldsymbol{L}, \boldsymbol{R}] = \int \mathcal{D}\boldsymbol{x} \int \mathcal{D}\hat{\boldsymbol{x}} \exp\left(i \sum_{i=1}^{N} \hat{x}_i T[x_i] - i \sum_{i,j=1}^{N} [\boldsymbol{LDR}]_{ij} \hat{x}_i \phi_j \right). \tag{I1}$$

We would like to integrate out $\boldsymbol{L}$ and $\boldsymbol{R}$, but this is complicated by them appearing together in a quadratic term in the action. To simplify the integration, we introduce a set of M latent variables,

$$z_a(t) = D_a \sum_{i=1}^{N} R_{ia} \phi_i(t), \tag{I2}$$

and their conjugates $\hat{z}_a(t)$. We use indices i, j for neurons and a, b for latent variables. The introduction of these latent variables makes the action linear in $\boldsymbol{L}$ and $\boldsymbol{R}$,

$$Z[\boldsymbol{L}, \boldsymbol{R}] = \int \mathcal{D}\boldsymbol{x} \int \mathcal{D}\hat{\boldsymbol{x}} \int \mathcal{D}\boldsymbol{z} \int \mathcal{D}\hat{\boldsymbol{z}} \exp\left(i \sum_{i=1}^{N} \hat{x}_i T[x_i] + i \sum_{a=1}^{M} z_a \hat{z}_a - i \sum_{i,a} L_{ia} \hat{x}_i z_a - i \sum_{i,a} D_a R_{ia} \phi_i \hat{z}_a \right). \tag{I3}$$

We can now easily integrate out $\boldsymbol{L}$ and $\boldsymbol{R}$. To factorize the action over indices spanning extensive dimensions, we introduce the field $C^{\phi}(t_1, t_2)$ [Eq. (17)] with conjugate $\hat{Q}(t_1, t_2)$ as well as

$$Q(t_1, t_2) = \frac{1}{N} \sum_{a=1}^{M} z_a(t_1) z_a(t_2) \tag{I4}$$

with conjugate $\hat{C}^{\phi}(t_1, t_2)$. The disorder-averaged path integral is

$$Z = \int \mathcal{D}C^{\phi} \int \mathcal{D}\hat{C}^{\phi} \int \mathcal{D}Q \int \mathcal{D}\hat{Q} \exp\left(-N \mathcal{S}[C^{\phi}, \hat{C}^{\phi}, Q, \hat{Q}]\right), \tag{I5}$$

$$\mathcal{S}[C^{\phi}, \hat{C}^{\phi}, Q, \hat{Q}] = -\frac{1}{2} C^{\phi} \hat{Q} - \frac{1}{2} Q \hat{C}^{\phi} - \log W^{x}[Q, \hat{Q}] - \alpha \langle \log W^{z}_{D}[C^{\phi}, \hat{C}^{\phi}] \rangle_{D}, \tag{I6}$$

$$W^{x}[Q, \hat{Q}] = \int \mathcal{D}x \int \mathcal{D}\hat{x} \exp\left(i\hat{x} T[x] - \frac{1}{2} \phi \hat{Q} \phi - \frac{1}{2} \hat{x} Q \hat{x} \right), \tag{I7}$$

$$W^{z}_{D}[C^{\phi}, \hat{C}^{\phi}] = \int \mathcal{D}z \int \mathcal{D}\hat{z} \exp\left(i\hat{z} z - \frac{1}{2} z \hat{C}^{\phi} z - \frac{1}{2} D^2 \hat{z} C^{\phi} \hat{z} \right). \tag{I8}$$

Thus, analysis of the path integral is made tractable by recasting the problem as interacting neurons and latent variables. The neuronal single-site picture is described by $W^{x}[Q, \hat{Q}]$, and the latent-variable single-site picture is described by $W^{z}_{D}[C^{\phi}, \hat{C}^{\phi}]$, where the subscript D reflects that each latent variable has its own component strength. This formulation introduces an extensive set of latent variables $z_a(t)$, in contrast to the finite set of latent variables, conventionally denoted by $\kappa_a(t)$, in low-rank neural networks [42,44]. The statistics of the Gaussian input to each single-site process (neurons or latent variables) are determined by the statistics of the other process, creating a bipartite, mutually referential structure that also arises in the cavity treatment of the problem.

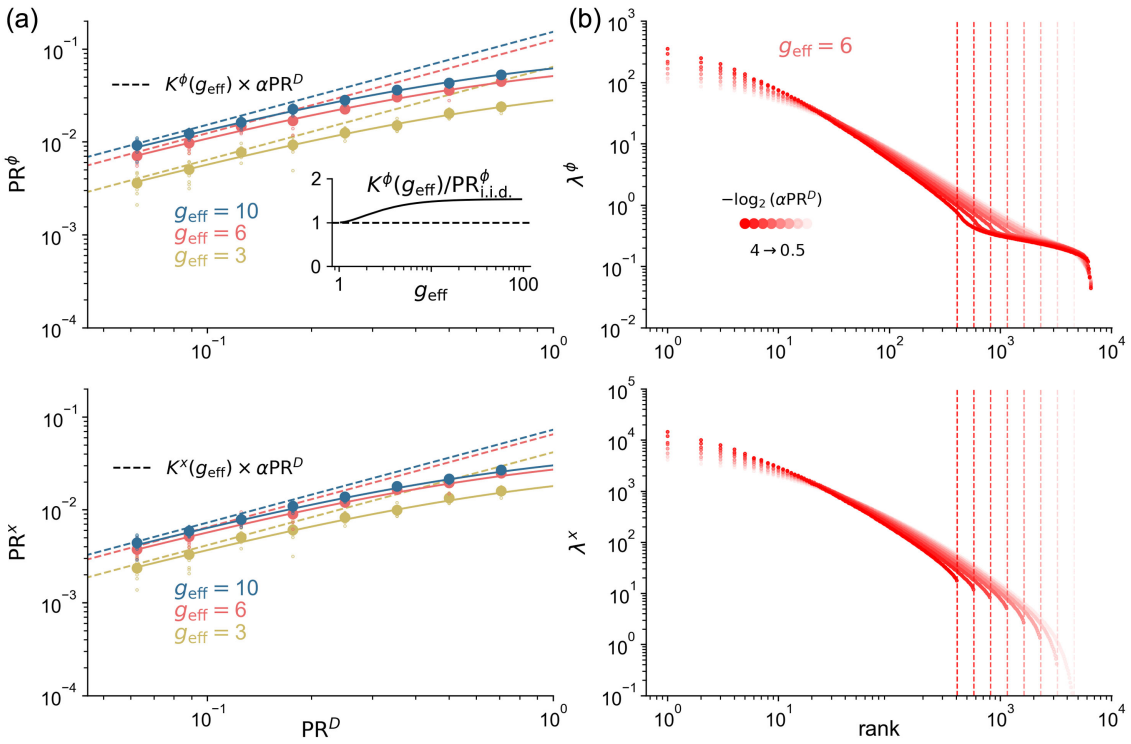

FIG. 10. Analysis of dimension of activity and covariance spectrum in the low-effective-rank limit. (a) Top: dimension of activity PR^ϕ versus effective rank $\alpha \mathrm{PR}^D$ ($\alpha = 1$) for various coupling strengths $g_{\rm eff}$. Component strengths D_a are constant for $a/N \le \mathrm{PR}^D$ and 0 otherwise. Small dots, individual simulations; large dots, means over ten simulations; solid lines, theoretical predictions; dotted lines, small-PR^D expansion, $K^\phi(g_{\rm eff})\mathrm{PR}^D$. Inset: ratio of proportionality factor $K^\phi(g_{\rm eff})$ to $\mathrm{PR}^\phi_{\rm i.i.d.}$ versus $g_{\rm eff}$. Bottom: same analysis for preactivations. (b) Top: rank-ordered eigenvalue plots for activation covariance matrix $C^\phi_{ij}(0)$. Dotted vertical lines indicate $N\mathrm{PR}^D$. Bottom: same for preactivations. All simulations use $N = 6500$ neurons.

1. Two-point functions

Following the steps outlined at the end of Sec. F 2, we first compute the DMFT by calculating the saddle point. The derivatives of the action are

$$\frac{\delta \mathcal{S}}{\delta C^\phi(t_1,t_2)} = -\hat{Q}^\phi(t_1,t_2) + \alpha \langle D^2 \langle \hat{z}(t_1)\hat{z}(t_2)\rangle_{W^z_D}\rangle_D, \tag{I9}$$

$$\frac{\delta \mathcal{S}}{\delta \hat{C}^\phi(t_1,t_2)} = -Q(t_1,t_2) + \alpha \langle \langle z(t_1)z(t_2)\rangle_{W^z_D}\rangle_D, \tag{I10}$$

$$\frac{\delta \mathcal{S}}{\delta Q(t_1,t_2)} = -\hat{C}^\phi(t_1,t_2) + \langle \hat{x}(t_1)\hat{x}(t_2)\rangle_{W^x}, \tag{I11}$$

$$\frac{\delta \mathcal{S}}{\delta \hat{Q}(t_1,t_2)} = -C^\phi(t_1,t_2) + \langle \phi(t_1)\phi(t_2)\rangle_{W^x}, \tag{I12}$$

where $\langle\cdots\rangle_{W^x}$ and $\langle\cdots\rangle_{W^z_D}$ are averages within the dynamic processes described by $W^x[Q,\hat{Q}]$ and $W^z_D[C^\phi,\hat{C}^\phi]$, respectively. Setting these derivatives to zero yields

$$\hat{C}^\phi_\star(t_1,t_2) = \hat{Q}_\star(t_1,t_2) = 0, \tag{I13}$$

$$C^\phi_\star(t_1,t_2) = \langle \phi(t_1)\phi(t_2)\rangle_\star, \tag{I14}$$

$$Q_\star(t_1,t_2) = \alpha \langle \langle z(t_1)z(t_2)\rangle_{\star D}\rangle_D, \tag{I15}$$

where $\langle\cdots\rangle_\star$ and $\langle\cdots\rangle_{\star D}$ are averages within the dynamic processes described by $W^x[Q_\star,0]$ and $W^z_D[C^\phi_\star,0]$, respectively. These single-site processes are

$$T[x](t) = \eta^x(t), \tag{I16}$$

$$z(t) = \eta^z(t), \tag{I17}$$

where $\eta^x(t)$ and $\eta^z(t)$ are Gaussian fields,

$$\eta^x \sim \mathcal{GP}(0, Q_\star), \tag{I18}$$

$$\eta^z \sim \mathcal{GP}(0, D^2 C^\phi_\star), \tag{I19}$$

showing that the neuronal and latent-variable single-site processes have mutually referential statistics [101].

Solving for $Q_\star(t_1,t_2)$ by combining Eqs. (I15), (I17), and (I19), we obtain $Q_\star(t_1,t_2)=\alpha r_2 C^\phi_\star(t_1,t_2)$, with r_n defined by Eq. (10). Consolidating these results yields Eq. (36c).

2. Four-point functions

The frequency-dependent saddle-point Hessian is

$$\boldsymbol{H}=\left(\begin{array}{cc|cc} 0 & \alpha r_2 & 0 & -1 \\ \alpha r_2 & -\alpha r_4 C^\phi_{12} & -1 & 0 \\ \hline 0 & -1 & 0 & S^\phi_{12} \\ -1 & 0 & (S^\phi_{12})^* & -C^\phi_{12} \end{array}\right). \tag{I20}$$

The upper-left block, corresponding to the latent variables, is related to the Hessian of an i.i.d. network with linear dynamics and a heterogeneous distribution of single-neuron scale factors D_a; the lower-right block, corresponding to the neurons, is related to the Hessian of an i.i.d. network of nonlinear neurons [Eq. (F15)]. They are coupled through the simple off-diagonal blocks, reflecting the interaction between neurons and latent variables.

This yields the expression for $\Psi^\phi_\star$ [Eq. (38)]. To obtain $\Psi^x_\star$, we add $-\frac{1}{2}x\mathcal{J}x$ to the neuronal single-site path integral [Eq. (I7)] and compute

$$\frac{\delta^2 S}{\delta\mathcal{J}^2}=-C^x_{12},\qquad \frac{\delta^2 S}{\delta\mathcal{J}\delta\boldsymbol{C}}=\begin{pmatrix}0\\0\\(S^x_{12})^*\\-C^{x\phi}_{12}\end{pmatrix}. \tag{I21}$$

Inserting these into Eq. (F23) we obtain

$$\Psi^x_\star=C^x_{12}+\left(1+\frac{1}{\alpha \mathrm{PR}^D}\right)|U|^2C^\phi_{12}+UC^{x\phi}_{12}+\mathrm{H.c.}, \tag{I22}$$

where $U=g^2_{\mathrm{eff}}S^x_{12}/(1-g^2_{\mathrm{eff}}S^\phi_{12})$ [Eq. (E6b) with $g=g_{\mathrm{eff}}$]. Here, we have used the following formula to invert the Hessian:

$$\begin{pmatrix}0&a&0&-1\\a^*&b&-1&0\\0&-1&0&c\\-1&0&c^*&d\end{pmatrix}^{-1}=\begin{pmatrix}-\frac{b+|c|^2d}{|1-ac|^2}&-\frac{c^*}{1-(ac)^*}&-\frac{ad+bc^*}{|1-ac|^2}&-\frac{1}{1-(ac)^*}\\-\frac{c}{1-ac}&0&-\frac{1}{1-ac}&0\\-\frac{a^*d+bc}{|1-ac|^2}&-\frac{1}{1-(ac)^*}&-\frac{|a|^2d+b}{|1-ac|^2}&-\frac{a^*}{1-(ac)^*}\\-\frac{1}{1-ac}&0&-\frac{a}{1-ac}&0\end{pmatrix}. \tag{I23}$$

APPENDIX J: PATH-INTEGRAL CALCULATION OF TWO- AND FOUR-POINT FUNCTIONS FOR THE RANDOM-MODE MODEL WITH SINGLE-UNIT HETEROGENEITY

We incorporate the generalized nonlinearity $\Phi_{\boldsymbol{\theta}_i}(x)$ into the path integral, as well as a source term for the normalized variables, and integrate out $\boldsymbol{L}$ and $\boldsymbol{R}$. To factorize the action over indices spanning extensive dimensions, we introduce the order parameters

$$C^\Phi(t_1,t_2)=\frac{1}{N}\sum_{i=1}^N\Phi_{\boldsymbol{\theta}_i}(t_1)\Phi_{\boldsymbol{\theta}_i}(t_2),\qquad \text{conjugate: } \hat{Q}(t_1,t_2); \tag{J1}$$

$$Q(t_1,t_2)=\frac{1}{N}\sum_{a=1}^M z_a(t_1)z_a(t_2),\qquad \text{conjugate: } \hat{C}^\phi(t_1,t_2), \tag{J2}$$

where $\Phi_{\boldsymbol{\theta}_i}(t)=\Phi_{\boldsymbol{\theta}_i}[x_i(t)]$. The resulting action and single-site path integrals for this system are

$$\mathcal{S}[C^\Phi,\hat{C}^\Phi,Q,\hat{Q}]=-\frac{1}{2}C^\Phi\hat{Q}-\frac{1}{2}Q\hat{C}^\phi-\langle\log W_{\boldsymbol{\theta}}[Q,\hat{Q};\mathcal{J}]\rangle_{\boldsymbol{\theta}}-\alpha\langle\log W^z_D[C^\Phi,\hat{C}^\Phi]\rangle_D, \tag{J3}$$

$$W^x_{\boldsymbol{\theta}}[Q,\hat{Q};\mathcal{J}]=\int\mathcal{D}x\int\mathcal{D}\hat{x}\exp\left(i\hat{x}T[x]-\frac{1}{2}\Phi_{\boldsymbol{\theta}}\hat{Q}\Phi_{\boldsymbol{\theta}}-\frac{1}{2}\hat{x}Q\hat{x}-\frac{1}{2}\phi\mathcal{J}\phi\right), \tag{J4}$$

$$W^z_D[C^\Phi,\hat{C}^\Phi]=\int\mathcal{D}z\int\mathcal{D}\hat{z}\exp\left(i\hat{z}z-\frac{1}{2}z\hat{C}^\Phi z-\frac{1}{2}D^2\hat{z}C^\Phi\hat{z}\right). \tag{J5}$$

1. Two-point functions

The saddle-point calculation proceeds similarly to the nonheterogeneous case of the random-mode model. We obtain

$$\hat{C}^{\Phi}_{\star}(t_1,t_2)=\hat{Q}_{\star}(t_1,t_2)=0, \tag{J6}$$

$$C^{\Phi}_{\star}(t_1,t_2)=\langle\langle\Phi_{\boldsymbol{\theta}}(t_1)\Phi_{\boldsymbol{\theta}}(t_2)\rangle_{\star}\rangle_{\boldsymbol{\theta}}, \tag{J7}$$

$$Q_{\star}(t_1,t_2)=\alpha\langle\langle z(t_1)z(t_2)\rangle_{\star D}\rangle_{D}, \tag{J8}$$

where $\langle\cdots\rangle_{\star}$ and $\langle\cdots\rangle_{\star D}$ are averages with respect to $W^{x}_{\boldsymbol{\theta}}[Q_{\star},0;0]$ and $W^{z}_{D}[C^{\Phi}_{\star},0]$, respectively (note that $W^{x}_{\boldsymbol{\theta}}[Q_{\star},0;0]$ does not depend on $\boldsymbol{\theta}$). Using $\langle z(t_1)z(t_2)\rangle_{\star D}=D^2C^{\Phi}_{\star}(t_1,t_2)$ and, consequently, $Q_{\star}(t_1,t_2)=\alpha r_2C^{\Phi}_{\star}(t_1,t_2)$, we obtain the self-consistent single-site dynamic process

$$T[x](t)=\eta^{x}(t), \tag{J9a}$$

$$\eta^{x}\sim\mathcal{GP}(0,\alpha r_2C^{\Phi}_{\star}), \tag{J9b}$$

$$C^{\Phi}_{\star}(t_1,t_2)=\langle\langle\Phi_{\boldsymbol{\theta}}(t_1)\Phi_{\boldsymbol{\theta}}(t_2)\rangle_{\star}\rangle_{\boldsymbol{\theta}}. \tag{J9c}$$

The self-consistency condition [Eq. (J9c)] is identical to that of an i.i.d. network, but with an average over $\boldsymbol{\theta}\sim P(\boldsymbol{\theta})$, in addition to the usual Gaussian average over $\eta^{x}(t)$, to account for single-neuron heterogeneity.

2. Four-point functions

To write down the solution for the four-point function, we define new two-frequency correlation functions with an outer average over $\boldsymbol{\theta}\sim P(\boldsymbol{\theta})$,

$$\mathfrak{C}^{\Phi}_{12}=\langle\langle\Phi_{\boldsymbol{\theta}}\Phi_{\boldsymbol{\theta}}\rangle_{\star}(\omega_1)\langle\Phi_{\boldsymbol{\theta}}\Phi_{\boldsymbol{\theta}}\rangle_{\star}(\omega_2)\rangle_{\boldsymbol{\theta}}, \tag{J10a}$$

$$\mathfrak{s}^{\Phi}_{12}=\left\langle\left\langle\frac{\delta\Phi_{\boldsymbol{\theta}}}{\delta I}\right\rangle_{\star}(\omega_1)\left\langle\frac{\delta\Phi_{\boldsymbol{\theta}}}{\delta I}\right\rangle_{\star}(\omega_2)\right\rangle_{\boldsymbol{\theta}}, \tag{J10b}$$

$$\mathfrak{C}^{\Phi\phi}_{12}=\langle\langle\Phi_{\boldsymbol{\theta}}\phi\rangle_{\star}(\omega_1)\langle\Phi_{\boldsymbol{\theta}}\phi\rangle_{\star}(\omega_2)\rangle_{\boldsymbol{\theta}}, \tag{J10c}$$

along with the usual shorthand $C^{\Phi}_{12}=C^{\Phi}_{\star}(\omega_1)C^{\Phi}_{\star}(\omega_2)$.

The frequency-dependent Hessian at the saddle point is

$$\boldsymbol{H}=\left(\begin{array}{cc|cc}0&\alpha r_2&0&-1\\ \alpha r_2&-\alpha r_4C^{\Phi}_{12}&-1&0\\ \hline 0&-1&0&\mathfrak{s}^{\Phi}_{12}\\ -1&0&(\mathfrak{s}^{\Phi}_{12})^{*}&-\mathfrak{C}^{\Phi}_{12}\end{array}\right), \tag{J11}$$

where the new Fraktur variables are defined in Eq. (J10). We invert the Hessian using Eq. (I23) and identify the upper-left element as $\Phi_{\boldsymbol{\theta}_i}(\boldsymbol{\omega})$,

$$\Psi^{\Phi}_{\star}=\frac{\frac{\mathfrak{C}^{\Phi}_{12}}{C^{\Phi}_{12}}+\frac{1}{\alpha\mathrm{PR}^{D}}|\alpha r_2\mathfrak{s}^{\Phi}_{12}|^2}{|1-\alpha r_2\mathfrak{s}^{\Phi}_{12}|^2}C^{\Phi}_{12}. \tag{J12}$$

Combining the inverted Hessian with the relevant quantities for computing fluctuations of normalized variables,

$$\frac{\delta^2S}{\delta\mathcal{J}^2}=-C^{\phi}_{12},\qquad \frac{\delta^2S}{\delta\mathcal{J}\delta\boldsymbol{C}}=\begin{pmatrix}0\\0\\(S^{\phi}_{12})^{*}\\-\mathfrak{C}^{\Phi\phi}_{12}\end{pmatrix}, \tag{J13}$$

where

$$\mathfrak{C}^{\Phi\phi}_{12}=\langle\langle\Phi_{\boldsymbol{\theta}}\phi\rangle_{\star}(\omega_1)\langle\Phi_{\boldsymbol{\theta}}\phi\rangle_{\star}(\omega_2)\rangle_{\boldsymbol{\theta}}, \tag{J14}$$

$$S^{\phi}_{12}=\left\langle\frac{\delta\phi}{\delta I}\right\rangle_{\star}(\omega_1)\left\langle\frac{\delta\phi}{\delta I}\right\rangle_{\star}(\omega_2), \tag{J15}$$

we obtain for the normalized variables $\phi_i(t)$,

$$\Psi^{\phi}_{\star}=\left[|A|^2\left(\frac{\mathfrak{C}^{\Phi}_{12}}{C^{\phi}_{12}}+\frac{1}{\alpha\mathrm{PR}^{D}}\frac{C^{\Phi}_{12}}{C^{\phi}_{12}}\right)+\left(A\frac{\mathfrak{C}^{\Phi\phi}_{12}}{C^{\phi}_{12}}+\mathrm{H.c.}\right)+1\right]C^{\phi}_{12}, \tag{J16}$$

where

$$A=\frac{\alpha r_2S^{\phi}_{12}}{1-\alpha r_2\mathfrak{s}^{\phi}_{12}}. \tag{J17}$$

3. Four-point functions for firing-rate heterogeneity

Upon specializing the above general results to the case where $\boldsymbol{\theta}_i$ contains a single gain parameter G_i, we obtain the four-point functions for unnormalized [Eq. (46a)] and normalized [Eq. (46b)] variables given in the main text.

Note that, like the distribution over component strengths D_a, two-point functions of activity depend only on the second moment of G_i, and four-point functions depend only on the second and fourth moments of G_i.

For "random-readout" gains with the same distribution as the recurrent gains (see main text), the four-point function is

$$\Psi^{\Phi^{\mathrm{readout}}}=\left[\left(\frac{1}{\mathrm{PR}^{G}}-1\right)(|1-g^2_{\mathrm{eff}}S^{\phi}_{12}|^2+|g^2_{\mathrm{eff}}S^{\phi}_{12}|^2)+1+\frac{1}{\alpha\mathrm{PR}^{D}}|g^2_{\mathrm{eff}}S^{\phi}_{12}|^2\right]\frac{C^{\Phi^{\mathrm{readout}}}_{12}}{|1-g^2_{\mathrm{eff}}S^{\phi}_{12}|^2}, \tag{J18}$$

where $C^{\Phi^{\mathrm{readout}}}_{12}=C^{\Phi}_{12}=q_2^2C^{\phi}_{12}$.

We conclude with a note on gain modulation in recurrent neural networks. It has been shown that modulating single-neuron gains while keeping synaptic weights fixed has great expressive power over the dynamics of recurrent neural networks [102]. Here, we consider random gains rather than judiciously chosen or learned gains. In particular, they are sampled independent of the connectivity. However, our framework could be extended to model gains that are related to the connectivity and, more ambitiously, could be extended to model learned gain parameters (or, as explored in the Discussion, learned connectivity).

APPENDIX K: PATH-INTEGRAL CALCULATION OF TWO- AND FOUR-POINT FUNCTIONS FOR THE RANDOM-MODE MODEL WITH DIAGONALLY STRUCTURED OVERLAPS

To generate matrices $\boldsymbol{L}$ and $\boldsymbol{R}$ with appropriate correlation structure, we express them in terms of independent Gaussian random matrices $\boldsymbol{X}^1$, $\boldsymbol{X}^2$, and $\boldsymbol{Y}$, each with variance $1/N$,

$$L_{ia} = \sqrt{1-|\rho_a|}X^1_{ia} + \sqrt{|\rho_a|}Y_{ia}, \tag{K1}$$

$$R_{ia} = \sqrt{1-|\rho_a|}X^2_{ia} + \mathrm{sgn}(\rho_a)\sqrt{|\rho_a|}Y_{ia}. \tag{K2}$$

We begin with the path integral containing latent variables, insert this parametrization of $\boldsymbol{L}$ and $\boldsymbol{R}$, and then integrate out $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, and $\boldsymbol{Y}$. To factorize the action over indices spanning extensive dimensions, we introduce

$$C^\phi(t_1,t_2) = \frac{1}{N}\sum_{i=1}^N \phi_i(t_1)\phi_i(t_2), \qquad \text{conjugate: } \hat{Q}(t_1,t_2); \tag{K3}$$

$$S^\phi(t_1,t_2) = \frac{1}{N}\sum_{i=1}^N \frac{\delta\phi_i(t_1)}{\delta I_i(t_2)}, \qquad \text{conjugate: } \hat{R}(t_1,t_2); \tag{K4}$$

$$Q(t_1,t_2) = \frac{1}{N}\sum_{a=1}^M z_a(t_1)z_a(t_2), \qquad \text{conjugate: } \hat{C}^\phi(t_1,t_2); \tag{K5}$$

$$R(t_1,t_2) = \frac{1}{N}\sum_{a=1}^M D_a\rho_a\frac{\delta z_a(t_1)}{\delta I_a(t_2)}, \qquad \text{conjugate: } \hat{S}^\phi(t_1,t_2). \tag{K6}$$

The resulting path integral, action, and single-site path integrals are given by

$$Z = \int \mathcal{D}C^\phi \int \mathcal{D}\hat{C}^\phi \int \mathcal{D}S^\phi \int \mathcal{D}\hat{S}^\phi \int \mathcal{D}Q \int \mathcal{D}\hat{Q} \int \mathcal{D}R \int \mathcal{D}\hat{R}\, \exp\left(-N\mathcal{S}[C^\phi,\hat{C}^\phi,S^\phi,\hat{S}^\phi,Q,\hat{Q},R,\hat{R}]\right), \tag{K7}$$

$$\mathcal{S}[C^\phi,\hat{C}^\phi,S^\phi,\hat{S}^\phi,Q,\hat{Q},R,\hat{R}] = -\frac{1}{2}C^\phi\hat{Q} - \frac{1}{2}Q\hat{C}^\phi + \frac{1}{2}S^\phi\hat{R} + \frac{1}{2}R\hat{S}^\phi \tag{K8}$$

$$- \log W^x[Q,\hat{Q},R,\hat{R}] - \alpha\langle \log W^z_{D,\rho}[C^\phi,\hat{C}^\phi,S^\phi,\hat{S}^\phi]\rangle_{D,\rho}, \tag{K9}$$

$$W^x[Q,\hat{Q},R,\hat{R}] = \int \mathcal{D}x \int \mathcal{D}\hat{x}\, \exp\left(i\hat{x}T[x] - \frac{1}{2}\hat{x}Q\hat{x} - \frac{1}{2}\phi\hat{Q}\phi - \frac{i}{2}\hat{x}R\phi - \frac{i}{2}\phi\hat{R}\,\hat{x}\right), \tag{K10}$$

$$W^z_{D,\rho}[C^\phi,\hat{C}^\phi,S^\phi,\hat{S}^\phi] = \int \mathcal{D}z \int \mathcal{D}\hat{z}\, \exp\left(i\hat{z}z - \frac{1}{2}D^2\hat{z}C^\phi\hat{z} - \frac{1}{2}z\hat{C}^\phi z - \frac{i}{2}D\rho\hat{z}S^\phi z - \frac{i}{2}D\rho z\hat{S}^\phi\hat{z}\right). \tag{K11}$$

1. Two-point functions

We first calculate the saddle-point conditions by taking the derivatives of the action,

$$\frac{\delta\mathcal{S}}{\delta C^\phi(t_1,t_2)} = -\hat{Q}(t_1,t_2) + \alpha\left\langle D^2\langle\hat{z}(t_1)\hat{z}(t_2)\rangle_{W^z_{D,\rho}}\right\rangle_{D,\rho}, \tag{K12}$$

$$\frac{\delta\mathcal{S}}{\delta\hat{C}^\phi(t_1,t_2)} = -Q(t_1,t_2) + \alpha\left\langle\langle z(t_1)z(t_2)\rangle_{W^z_{D,\rho}}\right\rangle_{D,\rho}, \tag{K13}$$

$$\frac{\delta\mathcal{S}}{\delta S^\phi(t_1,t_2)} = \hat{R}(t_1,t_2) + i\alpha\left\langle D\rho\langle\hat{z}(t_1)z(t_2)\rangle_{W^z_{D,\rho}}\right\rangle_{D,\rho}, \tag{K14}$$

$$\frac{\delta\mathcal{S}}{\delta\hat{S}^\phi(t_1,t_2)} = R(t_1,t_2) + i\alpha\left\langle D\rho\langle z(t_1)\hat{z}(t_2)\rangle_{W^z_{D,\rho}}\right\rangle_{D,\rho}, \tag{K15}$$

$$\frac{\delta\mathcal{S}}{\delta Q(t_1,t_2)} = -\hat{C}^\phi(t_1,t_2) + \langle\hat{x}(t_1)\hat{x}(t_2)\rangle_{W^x}, \tag{K16}$$

$$\frac{\delta \mathcal{S}}{\delta \hat{Q}(t_1,t_2)} = -C^{\phi}(t_1,t_2) + \langle \phi(t_1)\phi(t_2)\rangle_{W^x}, \tag{K17}$$

$$\frac{\delta \mathcal{S}}{\delta R(t_1,t_2)} = \hat{S}^{\phi}(t_1,t_2) + i\langle \hat{x}(t_1)\phi(t_2)\rangle_{W^x}, \tag{K18}$$

$$\frac{\delta \mathcal{S}}{\delta \hat{R}(t_1,t_2)} = S^{\phi}(t_1,t_2) + i\langle \phi(t_1)\hat{x}(t_2)\rangle_{W^x}. \tag{K19}$$

The saddle-point conditions yield

$$\hat{C}^{\phi}(t_1,t_2) = \hat{Q}_{\star}(t_1,t_2) = 0, \tag{K20}$$

$$C^{\phi}_{\star}(t_1,t_2) = \langle \phi(t_1)\phi(t_2)\rangle_{\star}, \tag{K21}$$

$$S^{\phi}_{\star}(t_1,t_2) = \hat{S}^{\phi}_{\star}(t_2,t_1) = \left\langle \frac{\delta \phi(t_1)}{\delta I(t_2)} \right\rangle_{\star}, \tag{K22}$$

$$Q_{\star}(t_1,t_2) = \alpha \langle \langle z(t_1) z(t_2)\rangle_{\star D,\rho}\rangle_{D,\rho}, \tag{K23}$$

$$R_{\star}(t_1,t_2) = \hat{R}_{\star}(t_2,t_1) = \alpha \left\langle D\rho \left\langle \frac{\delta z(t_1)}{\delta I(t_2)} \right\rangle_{\star D,\rho} \right\rangle_{D,\rho}, \tag{K24}$$

where $\langle\cdots\rangle_{\star D,\rho}$ and $\langle\cdots\rangle_{\star}$ are averages within the dynamic processes described by $W^z_{D,\rho}[C^{\phi}_{\star}, 0, S^{\phi}_{\star}, \hat{S}^{\phi}_{\star}]$ and $W^x[Q_{\star}, 0, R_{\star}, \hat{R}_{\star}]$, respectively. This yields neuronal and latent-variable single-site processes described by

$$T[x](t) = \eta^x(t) + [R_{\star} \circ \phi](t) \tag{K25}$$

$$z(t) = \eta^z(t) + D\rho[S^{\phi}_{\star} \circ z](t), \tag{K26}$$

where the Gaussian fields have statistics

$$\eta^x \sim \mathcal{GP}(0, Q_{\star}), \tag{K27}$$

$$\eta^z \sim \mathcal{GP}(0, D^2 C^{\phi}_{\star}). \tag{K28}$$

This gives rise to the single-site problem stated in Sec. VII.

2. Four-point functions

The frequency-dependent Hessian at the saddle point is

$$\boldsymbol{H} = \left(\begin{array}{cccc|cccc} 0 & u & 0 & 0 & 0 & -1 & 0 & 0 \\ u^* & -vC^{\phi}_1 C^{\phi}_2 & x^*_{12} C^{\phi}_2 & x^*_{21} C^{\phi}_1 & -1 & 0 & 0 & 0 \\ 0 & x_{12} C^{\phi}_2 & 0 & -w^* & 0 & 0 & 0 & 1 \\ 0 & x_{21} C^{\phi}_1 & -w & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & -1 & 0 & 0 & 0 & S^{\phi}_1 S^{\phi}_2 & 0 & 0 \\ -1 & 0 & 0 & 0 & (S^{\phi}_1 S^{\phi}_2)^* & -C^{\phi}_1 C^{\phi}_2 & (S^{\phi}_1)^* C^{\phi}_2 & (S^{\phi}_2)^* C^{\phi}_1 \\ 0 & 0 & 0 & 1 & 0 & S^{\phi}_1 C^{\phi}_2 & 0 & -S^{\phi}_1 (S^{\phi}_2)^* \\ 0 & 0 & 1 & 0 & 0 & S^{\phi}_2 C^{\phi}_1 & -(S^{\phi}_1)^* S^{\phi}_2 & 0 \end{array}\right) \tag{K29}$$

where

$$u = \alpha \langle D^2 \sigma_1 \sigma_2 \rangle_{D,\rho}, \tag{K30}$$

$$v = \alpha \langle D^4 |\sigma_1|^2 |\sigma_2|^2 \rangle_{D,\rho}, \tag{K31}$$

$$w = \alpha \langle D^2 \rho^2 \sigma_1^* \sigma_2 \rangle_{D,\rho}, \tag{K32}$$

$$x_{12} = \alpha \langle D^3 \rho \sigma_1 |\sigma_2|^2 \rangle_{D,\rho}, \tag{K33}$$

$$\sigma_k = \frac{1}{1 - D\rho S^{\phi}_k} \quad \text{for } k = 1, 2. \tag{K34}$$

Isolating the upper-left element of the inverse of this matrix yields Eq. (51).

3. Sigmoidal parametrization of overlaps

To model structured left-right mode overlaps, we used the following sigmoidal form:

$$\rho_a = \gamma_\rho \left(1 - \frac{2}{1 + \exp[-\beta_\rho (a/N - u_{\text{flip}})]}\right), \tag{K35}$$

where $\rho_a = 0$ at $a/N = u_{\text{flip}}$. The parameter γ_ρ sets the overall magnitude and sign of the overlaps, while β_ρ controls the steepness of the transition. We fixed $\beta_\rho = 16$, varied γ_ρ, and set u_{flip} such that $\langle D\rho \rangle_{D,\rho} = 0$, thereby eliminating strong self-couplings.

APPENDIX L: TWO-SITE CAVITY CALCULATION OF FOUR-POINT FUNCTIONS FOR THE RANDOM-MODE MODEL

While the path-integral approach provides a systematic framework for analyzing neural-network dynamics, the cavity method offers an intuitive, complementary perspective. This section provides an overview of the cavity analysis for the random-mode model. While these approaches yield equivalent results, they differ in their methodology: The path-integral approach focuses on fluctuations around the saddle point in a field theory, whereas the cavity approach examines perturbations induced by a subset of held-out variables. Here, we compute the four-point function using a two-site cavity approach, noting that the two-point function (which, if desired, could be computed using a single-site cavity method) is the same as that of an i.i.d. network with appropriate g_{eff}.

Leveraging Eq. (I2), we reformulate the network as a bipartite system of neurons and latent variables and extend the cavity calculation of Clark *et al.* (Ref. [39]) to both groups. This leads to separate, mutually referential cavity pictures for each group, mirroring the two single-site path integrals (Sec. IV). As in the path-integral formalism, self-consistent expressions in one picture are defined using averages from the other. Unlike the path-integral approach, which derives $\Psi^{\phi}(\boldsymbol{\tau})$ using its time-by-time definition [Eq. (25)], the cavity method derives this function using its neuron-by-neuron definition [Eq. (24)].

The cavity method we employ, schematized in Fig. 11, has two key features that distinguish it from simpler cavity calculations. First, it uses a two-site structure: We consider the simultaneous removal of two neurons, or two latent variables, from the network, rather than just one. This allows us to study cross-correlations between neurons. This approach was previously used in Clark *et al.* (Ref. [39]). Second, it incorporates a bipartite structure: We treat both neurons and latent variables as dynamic objects, creating separate but mutually referential cavity pictures for each group. Such bipartite structure, without the two-site structure, has been used in cavity calculations in the Hopfield model [69,103,104]. Note that these two features are distinct: The two-site aspect refers to the number of neurons removed in each cavity, while the bipartite aspect refers to the types of variables considered (neurons and latent variables).

To apply cavity techniques, we reformulate the network as a bipartite system of neurons and latent variables,

$$T[x_i](t) = \sum_{a=1}^{M} L_{ia} z_a(t), \tag{L1}$$

$$z_a(t) = D_a \sum_{j=1}^{N} R_{ja} \phi_j(t), \tag{L2}$$

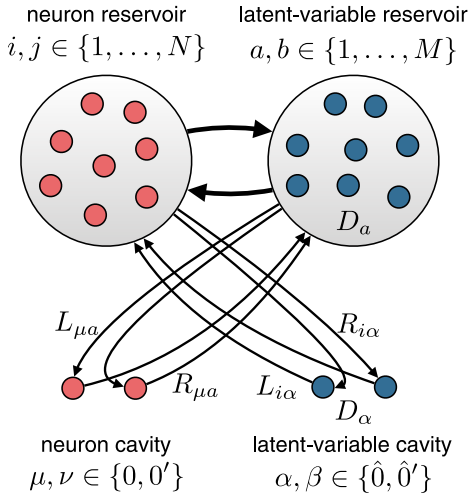

FIG. 11. Schematic of the two-site cavity approach for the bipartite representation of the random-mode model. Left: neuron cavity, considering the introduction of neurons $x_0(t)$ and $x_{0'}(t)$. Right: latent-variable cavity, considering the introduction of latent variables $z_{\hat{0}}(t)$ and $z_{\hat{0}'}(t)$. The cavity pictures are coupled, with averages from one appearing in self-consistent equations for the other.

as done in the path-integral approach. In the two-site cavity calculation, we keep track of various intermediate quantities to order $1/\sqrt{N}$, leading to an expression for $\Psi^{\phi}(\boldsymbol{\tau})$ accurate to leading order (order one), which we identify as $\Psi^{\phi}_{\star}(\boldsymbol{\tau})$.

To distinguish between different types of variables and their indices, we use the following notation. Neurons are indexed by $i, j \in \{1, \ldots, N\}$ and latent variables by $a, b \in \{1, \ldots, M\}$, as usual. For the cavity variables, we introduce special indices: Neuronal cavity variables are indexed by $\mu, \nu \in \{0, 0'\}$, while latent cavity variables are indexed by $\alpha, \beta \in \{\hat{0}, \hat{0}'\}$.

1. Neuronal cavity

We begin by introducing two neurons, $x_0(t)$ and $x_{0'}(t)$. The leading-order $(1/\sqrt{N})$ effect on the latent variables is

$$\delta z_a(t) = \int^t dt' \sum_{b=1}^{M} S^z_{ab}(t_1, t_2) D_b \sum_{\mu \in \{0,0'\}} R_{\mu b} \phi_\mu(t'), \tag{L3}$$

where $S^z_{ab}(t_1, t_2)$ is the response function of the latent variables. The dynamic equations for the cavity neurons are

$$T[x_\mu](t) = \eta_\mu(t) + \sum_{\nu \in \{0,0'\}} \int^t dt' F_{\mu\nu}(t_1, t_2) \phi_\nu(t'), \tag{L4}$$

where

$$\eta_\mu(t) = \sum_{a=1}^{M} L_{\mu a} z_a(t), \tag{L5}$$

$$F_{\mu\nu}(t_1,t_2)=\sqrt{N}\sum_{a,b=1}^{M}L_{\mu a}D_bR_{\nu b}S^z_{ab}(t_1,t_2). \quad \text{(L6)}$$

Defining the cavity-field time-average cross-covariance as

$$C^{\eta}_{00'}(\omega)=\sum_{a,b}L_{0a}L_{0'b}C^z_{ab}(\omega), \quad \text{(L7)}$$

we obtain the following expression for the cross-covariance of the cavity units, up to order $1/\sqrt{N}$:

$$C^{\phi}_{00'}(\omega)=|S^{\phi}_{\star}(\omega)|^2C^{\eta}_{00'}(\omega)+\frac{1}{\sqrt{N}}[(S^{\phi}_{\star}(\omega)F_{00'}(\omega))^*+S^{\phi}_{\star}(\omega)F_{0'0}(\omega)]C^{\phi}_{\star}(\omega). \quad \text{(L8)}$$

Our goal is to compute the parameter

$$\psi^{\phi}(\boldsymbol{\omega})=N\langle C^{\phi}_{00'}(\omega_1)C^{\phi}_{00'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}, \quad \text{(L9)}$$

where $\langle\cdot\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}$ denotes an average over $\boldsymbol{L}$, $\boldsymbol{D}$, and $\boldsymbol{R}$. To evaluate this, we need to square and disorder average Eq. (L8). This requires us to consider several two-frequency correlation functions, which we define as

$$\Gamma_{F_{00'}F_{00'}}(\boldsymbol{\omega})=\langle F_{00'}(\omega_1)F_{00'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}, \quad \text{(L10)}$$

$$\Gamma_{F^*_{00'}F_{0'0}}(\boldsymbol{\omega})=\langle F^*_{00'}(\omega_1)F_{0'0}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}, \quad \text{(L11)}$$

$$\Gamma_{F^*_{00'}C^{\eta}_{00'}}(\boldsymbol{\omega})=\sqrt{N}\langle F^*_{00'}(\omega_1)C^{\eta}_{00'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}, \quad \text{(L12)}$$

$$\Gamma_{C^{\eta}_{00'}C^{\eta}_{00'}}(\boldsymbol{\omega})=N\langle C^{\eta}_{00'}(\omega_1)C^{\eta}_{00'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}. \quad \text{(L13)}$$

These have all been scaled to be order one. We can evaluate these $\Gamma_{\ldots}(\boldsymbol{\omega})$ functions due to the independence of the couplings and dynamic variables in the expressions defining them, which is a consequence of the cavity construction. Of these, the only nonvanishing ones are

$$\Gamma_{C^{\eta}_{00'}C^{\eta}_{00'}}(\boldsymbol{\omega})=\alpha\langle C^z_{\hat{0}\hat{0}}(\omega_1)C^z_{\hat{0}\hat{0}}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}+\alpha^2N\langle C^z_{\hat{0}\hat{0}'}(\omega_1)C^z_{\hat{0}\hat{0}'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}, \quad \text{(L14)}$$

$$\Gamma_{F_{00'}F_{00'}}(\boldsymbol{\omega})=\alpha r_2+\alpha^2N\langle D^2_{\hat{0}'}S^z_{\hat{0}\hat{0}'}(\omega_1)S^z_{\hat{0}\hat{0}'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}, \quad \text{(L15)}$$

where $\hat{0},\hat{0}'$ denote the indices in the latent-variable cavity picture.

2. Latent-variable cavity

We now introduce two new latent variables, $z_{\hat{0}}(t)$ and $z_{\hat{0}'}(t)$. The leading-order $(1/\sqrt{N})$ effect on the neurons is

$$\delta\phi_i(t)=\int^t dt'\sum_{j=1}^{N}S^{\phi}_{ij}(t_1,t_2)\sum_{\alpha\in\{\hat{0},\hat{0}'\}}L_{j\alpha}z_{\alpha}(t'), \quad \text{(L16)}$$

where $S^{\phi}_{ij}(t_1,t_2)$ is the response function of the neurons. The dynamic equations for the cavity latent variables are

$$z_{\alpha}(t)=D_{\alpha}\left[\gamma_{\alpha}(t)+\frac{1}{\sqrt{N}}\sum_{\beta\in\{\hat{0},\hat{0}'\}}\int^t dt'G_{\alpha\beta}(t_1,t_2)z_{\beta}(t')\right], \quad \text{(L17)}$$

where

$$\gamma_{\alpha}(t)=\sum_{i=1}^{N}R_{i\alpha}\phi_i(t), \quad \text{(L18)}$$

$$G_{\alpha\beta}(t_1,t_2)=\sqrt{N}\sum_{i,j=1}^{N}R_{i\alpha}L_{j\beta}S^{\phi}_{ij}(t_1,t_2). \quad \text{(L19)}$$

Defining the time-average cavity-field cross-covariance as

$$C^{\gamma}_{\hat{0}\hat{0}'}(\omega)=\sum_{i,j=1}^{N}R_{i\hat{0}}R_{j\hat{0}'}C^{\phi}_{ij}(\omega), \quad \text{(L20)}$$

we obtain the following expression for the cross-covariance of the cavity latent variables up to order $1/\sqrt{N}$:

$$C^z_{\hat{0}\hat{0}'}(\omega)=D_{\hat{0}}D_{\hat{0}'}C^{\gamma}_{\hat{0}\hat{0}'}(\omega)+\frac{1}{\sqrt{N}}[D_{\hat{0}}G^*_{\hat{0}\hat{0}'}(\omega)C^z_{\hat{0}'\hat{0}'}(\omega)+D_{\hat{0}'}G_{\hat{0}'\hat{0}}(\omega)C^z_{\hat{0}\hat{0}}(\omega)]. \quad \text{(L21)}$$

We aim to compute the parameter

$$\psi^z(\boldsymbol{\omega})=N\langle C^z_{\hat{0}\hat{0}'}(\omega_1)C^z_{\hat{0}\hat{0}'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}} \quad \text{(L22)}$$

by squaring and disorder averaging Eq. (L21). This requires us to consider several two-frequency correlation functions, which we define as

$$\Gamma_{G_{\hat{0}\hat{0}'}G_{\hat{0}\hat{0}'}}(\boldsymbol{\omega})=\langle G_{\hat{0}\hat{0}'}(\omega_1)G_{\hat{0}\hat{0}'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{R},\ \boldsymbol{D}\setminus\{D_{\hat{0}},D_{\hat{0}'}\}}, \quad \text{(L23)}$$

$$\Gamma_{G^*_{\hat{0}\hat{0}'}G_{\hat{0}'\hat{0}}}(\boldsymbol{\omega})=\langle G^*_{\hat{0}\hat{0}'}(\omega_1)G_{\hat{0}'\hat{0}}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{R},\ \boldsymbol{D}\setminus\{D_{\hat{0}},D_{\hat{0}'}\}}, \quad \text{(L24)}$$

$$\Gamma_{G^*_{\hat{0}\hat{0}'}C^{\gamma}_{\hat{0}\hat{0}'}}(\boldsymbol{\omega})=\sqrt{N}\langle G^*_{\hat{0}\hat{0}'}(\omega_1)C^{\gamma}_{\hat{0}\hat{0}'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{R},\ \boldsymbol{D}\setminus\{D_{\hat{0}},D_{\hat{0}'}\}}, \quad \text{(L25)}$$

$$\Gamma_{C^{\gamma}_{\hat{0}\hat{0}'}C^{\gamma}_{\hat{0}\hat{0}'}}(\boldsymbol{\omega})=N\langle C^{\gamma}_{\hat{0}\hat{0}'}(\omega_1)C^{\gamma}_{\hat{0}\hat{0}'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{R},\ \boldsymbol{D}\setminus\{D_{\hat{0}},D_{\hat{0}'}\}}. \quad \text{(L26)}$$

Again, we can evaluate these $\Gamma_{\ldots}(\boldsymbol{\omega})$ functions due to the independence of the couplings and dynamic variables in the

expressions defining them. The nonvanishing ones are

$$\Gamma_{C^{\gamma}_{\hat{0}0'}C^{\gamma}_{\hat{0}0'}}(\boldsymbol{\omega}) = C^{\phi}_{\star}(\omega_1)C^{\phi}_{\star}(\omega_2) + N\langle C^{\phi}_{00'}(\omega_1)C^{\phi}_{00'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}, \tag{L27}$$

$$\Gamma_{G_{\hat{0}0'}G_{\hat{0}0'}}(\boldsymbol{\omega}) = S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2) + N\langle S^{\phi}_{00'}(\omega_1)S^{\phi}_{00'}(\omega_2)\rangle_{\boldsymbol{L},\boldsymbol{D},\boldsymbol{R}}, \tag{L28}$$

where $0, 0'$ denote the indices in the neuronal cavity picture.

3. Combining the two two-site cavity pictures

To combine the results from both two-site cavity pictures, we use the following relations:

$$C^{z}_{\hat{0}\hat{0}}(\omega) = D^2_{\hat{0}}C^{\phi}_{\star}(\omega), \tag{L29}$$

$$S^{z}_{\hat{0}0'}(\omega) = \frac{D_{\hat{0}}}{\sqrt{N}}G_{\hat{0}0'}(\omega), \tag{L30}$$

$$S^{\phi}_{00'}(\omega) = \frac{1}{\sqrt{N}}[S^{\phi}_{\star}(\omega)]^2 F_{00'}(\omega). \tag{L31}$$

Using these relations and the definitions of the $\Gamma_{\ldots}(\boldsymbol{\omega})$ functions [Eqs. (L10)–(L13) and (L23)–(L26)], we obtain

$$\Gamma_{C^{\eta}_{00'}C^{\eta}_{00'}}(\boldsymbol{\omega}) = \alpha r_4 C^{\phi}_{\star}(\omega_1)C^{\phi}_{\star}(\omega_2) + \alpha^2\psi^{z}(\boldsymbol{\omega}), \tag{L32}$$

$$\Gamma_{F_{00'}F_{00'}}(\boldsymbol{\omega}) = \alpha r_2[1 + \alpha r_2\Gamma_{G_{\hat{0}\hat{0}'}G_{\hat{0}\hat{0}'}}(\boldsymbol{\omega})], \tag{L33}$$

$$\Gamma_{C^{\gamma}_{\hat{0}\hat{0}'}C^{\gamma}_{\hat{0}\hat{0}'}}(\boldsymbol{\omega}) = C^{\phi}_{\star}(\omega_1)C^{\phi}_{\star}(\omega_2) + \psi^{\phi}(\boldsymbol{\omega}), \tag{L34}$$

$$\Gamma_{G_{\hat{0}\hat{0}'}G_{\hat{0}\hat{0}'}}(\boldsymbol{\omega}) = S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2) \times [1 + S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2)\Gamma_{F_{00'}F_{00'}}(\boldsymbol{\omega})]. \tag{L35}$$

We solve Eqs. (L33) and (L35) simultaneously, yielding

$$\Gamma_{F_{00'}F_{00'}}(\boldsymbol{\omega}) = \frac{\alpha r_2}{1 - \alpha r_2 S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2)}, \tag{L36}$$

$$\Gamma_{G_{\hat{0}\hat{0}'}G_{\hat{0}\hat{0}'}}(\boldsymbol{\omega}) = \frac{S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2)}{1 - \alpha r_2 S^{\phi}_{\star}(\omega_1)S^{\phi}_{\star}(\omega_2)}. \tag{L37}$$

With these solutions, we express $\psi^{z}_{\star}(\boldsymbol{\omega})$ and $\psi^{\phi}_{\star}(\boldsymbol{\omega})$ in terms of the $\Gamma_{\ldots}(\boldsymbol{\omega})$ functions. Switching to frequency-suppressed notation,

$$\psi^{z}_{\star} = r_2^2\Gamma_{C^{\gamma}_{\hat{0}\hat{0}'}C^{\gamma}_{\hat{0}\hat{0}'}} + r_2 r_4(\Gamma_{G_{\hat{0}\hat{0}'}G_{\hat{0}\hat{0}'}} + \text{H.c.})C^{\phi}_{12}, \tag{L38}$$

$$\psi^{\phi}_{\star} = |S^{\phi}_{12}|^2\Gamma_{C^{\eta}_{00'}C^{\eta}_{00'}} + (\Gamma_{F_{00'}F_{00'}}S^{\phi}_{12} + \text{H.c.})C^{\phi}_{12}. \tag{L39}$$

Now, we substitute Eq. (L38) into Eq. (L39). Using $\Psi^{\phi}_{\star} = \psi^{\phi}_{\star} + C^{\phi}_{12}$, we obtain

$$\Psi^{\phi}_{\star} = |g^2_{\text{eff}}S^{\phi}_{12}|^2\left[\Psi^{\phi}_{\star} + \frac{1}{\alpha \text{PR}^{D}}\frac{1 - |g^2_{\text{eff}}S^{\phi}_{12}|^2}{|1 - g^2_{\text{eff}}S^{\phi}_{12}|^2}C^{\phi}_{12}\right] + \frac{1 - |g^2_{\text{eff}}S^{\phi}_{12}|^2}{|1 - g^2_{\text{eff}}S^{\phi}_{12}|^2}C^{\phi}_{12}, \tag{L40}$$

where we used the previously defined quantities $g^2_{\text{eff}} = \alpha r_2$ and $\text{PR}^{D} = r_2^2/r_4$. Equation (L40) can be solved for $\Psi^{\phi}_{\star}$, yielding

$$\Psi^{\phi}_{\star} = \frac{1 + \frac{1}{\alpha \text{PR}^{D}}|g^2_{\text{eff}}S^{\phi}_{12}|^2}{|1 - g^2_{\text{eff}}S^{\phi}_{12}|^2}C^{\phi}_{12}. \tag{L41}$$

This result is identical to Eq. (38) derived using the path-integral approach, demonstrating the consistency between the two methods.

APPENDIX M: NUMERICAL DETAILS

1. Numerics for theory

To validate the theory, we used networks with dynamics defined by $T[x](t) = (1 + \partial_t)x(t)$ and with nonlinearity $\phi(x) = \text{erf}(\sqrt{\pi}x/2)$. For all connectivity models, except those with structured $\boldsymbol{L} - \boldsymbol{R}$ overlaps (Sec. VII), the single-site two-point functions $C^{\phi}_{\star}(\tau)$ and $S^{\phi}_{\star}(\tau)$ depend only on g_{eff} and can be numerically calculated as in i.i.d. networks (see, e.g., Refs. [10,105]). In summary:

(1) Since $\eta(t)$ is Gaussian, so is $x(t)$ for this linear form of $T[\cdot]$, allowing us to write $C^{\phi}_{\star}(\tau)$ in terms of Gaussian integrals over $x(t)$ and $x(t+\tau)$, which have marginal variance $C^{\phi}_{\star}(0)$ and covariance $C^{\phi}_{\star}(\tau)$. Because of our choice of $\phi(x)$, this expression simplifies to

$$C^{\phi}_{\star}(\tau) = \frac{2}{\pi}\tan^{-1}\left[\frac{C^{x}_{\star}(\tau)}{\sqrt{(C^{x}_{\star}(0) + \frac{2}{\pi})^2 - C^{x}_{\star}(\tau)^2}}\right]. \tag{M1}$$

(2) Squaring the single-site picture gives a second-order ordinary differential equation (ODE) $\partial^2_{\tau}C^{x}_{\star}(\tau) = C^{x}_{\star}(\tau) - g^2_{\text{eff}}C^{\phi}_{\star}(\tau)$. Since the rhs depends only on $C^{x}_{\star}(\tau)$ and $C^{x}_{\star}(0)$, we can consider it to be the negative derivative $-\partial V/\partial C^{x}_{\star}(\tau)$ of a potential $V(C^{x}_{\star}(\tau), C^{x}_{\star}(0))$ with an explicit dependence on the initial condition $C^{x}_{\star}(0)$.

(3) The rhs is integrated with respect to $C^{x}_{\star}(\tau)$ to provide an expression for $V(C^{x}_{\star}(\tau), C^{x}_{\star}(0))$, which due to our choice of $\phi(x)$ simplifies to

$$V(C^x_\star(\tau), C^x_\star(0)) = -\frac{1}{2}C^x_\star(\tau)^2 + g^2_{\text{eff}}\left[\frac{2}{\pi}\sqrt{\left(C^x_\star(0) + \frac{2}{\pi}\right)^2 - C^x_\star(\tau)^2} + C^x_\star(\tau)C^\phi_\star(\tau)\right]. \tag{M2}$$

(4) Restricting to solutions with $C^x_\star(\tau) \to 0$ as $\tau \to \infty$ and $\partial_\tau C^x_\star(\tau)|_{\tau=0} = 0$, enforcing conservation of energy gives $V(C^x_\star(0), C^x_\star(0)) = V(0, C^x_\star(0))$, which can be solved for $C^x_\star(0)$ by numerical root finding.

(5) Finally, Euler integration of the $[C^x_\star(\tau), \partial_\tau C^x_\star(\tau)]$ dynamics with these initial conditions gives its full time course, along with $C^\phi_\star(\tau)$ via Eq. (M1).

We integrated the Newtonian ODE from $\tau = 0$ to $\tau_{\max} = 200$ with step size $d\tau = 0.025$. We then use an FFT to obtain the frequency-space representation. The frequency-space representation of the response function $S^\phi_\star(\omega)$ was computed directly as

$$S^\phi_\star(\omega) = \frac{\langle \phi' \rangle_\star}{1 + i\omega}, \tag{M3}$$

with $\langle \phi' \rangle_\star = 1/\sqrt{1 + (\pi/2)C^x_\star(0)}$. Various expressions for $\Psi^a_\star(\omega_1, \omega_2)$ were then computed in frequency space before transforming back to the time domain by a two-dimensional inverse FFT.

For structured $\boldsymbol{L}$-$\boldsymbol{R}$ overlaps, the $\mathcal{O}(1)$ self-coupling in the single-site process makes $x(t)$ non-Gaussian, preventing the use of Gaussian integrals in the DMFT solution. We therefore obtain the solution via standard numerical methods, enforcing self-consistency among the $\eta^x(t)$ correlation function $Q_\star(\tau)$, the self-coupling kernel $R_\star(\tau)$, the response function $S^\phi_\star(\tau)$, and the $\phi(t)$ correlation function $C^\phi_\star(\tau)$. In summary:

(1) Seed values for kernels are set as $C^\phi_0(\tau) = \exp[-(\tau/15)^2] + 10^{-4}\delta(\tau)$ and $S^\phi_0(\tau) = 0$.

(2) $Q_0(\tau)$ and $R_0(\tau)$ are computed in Fourier space using the current values of $C^\phi_0(\tau)$ and $S^\phi_0(\tau)$ via Eq. (49), with the average taken over (D, ρ).

(3) A set of $N_{\text{sample}} = 10^4$ instances of η^x are sampled according to the Gaussian process kernel $Q_0(\tau)$ by standard spectral methods.

(4) Each of N_{sample} single-site processes $x_s(t)$ is Euler integrated according to Eq. (47a) with the input $\eta^x_s(t)$ and self-feedback by the convolution between $R_0(\tau)$ and $\phi(x_s(t))$.

(5) These time series $\phi(x_s(t))$ and $\eta^x_s(t)$ are used to compute empirical values $(C^\phi_1(\tau), S^\phi_1(\tau))$ for the autocovariance and response. The latter is computed by the Furutsu-Novikov theorem: $S^\phi(\omega) = C^{\phi\eta}(\omega)/C^\eta(\omega)$. The numerator is estimated using these time series, while $C^\eta(\omega)$ is simply $Q_0(\omega)$.

(6) After $N_{\text{iter}} = 300$ repetitions of steps (1)–(5) with smoothed updates

$$C^\phi_m(\tau) \leftarrow 0.2C^\phi_{m-1}(\tau) + 0.8C^\phi_m(\tau), \tag{M4}$$

$$S^\phi_m(\tau) \leftarrow 0.2S^\phi_{m-1}(\tau) + 0.8S^\phi_m(\tau) \tag{M5}$$

as new seed values for the upcoming mth iteration, the final numerical estimates for $C^\phi_\star(\tau)$ and $S^\phi_\star(\tau)$ are computed as the averages over the last 50 iterations.

We used a temporal discretization $\tau_{\max} = 120$, $d\tau = 0.04$. With these quantities, the various components of Eq. (51) and the expression for $\Psi^\phi_\star$ are straightforward to compute. We took $\langle\cdot\rangle_{D,\rho}$ averages in this case by averaging over every other discrete value of D_a and ρ_a.

2. Simulation details

Each network was integrated via Runge-Kutta with $dt = 0.05$. The empirical lagged covariance matrix was computed by averaging over N_{loop} individual covariance estimates, each of which is based on $T = 2000$ discrete, evenly spaced time samples $t_k^{i_{\text{loop}}}$, with $t_k^{i_{\text{loop}}} - t_{k-1}^{i_{\text{loop}}} = 1$,

$$\widehat{C^\phi_{ij}(\tau)} = \left\langle \frac{1}{T-1}\sum_{k=1}^{T} \phi_i(t_k^{i_{\text{loop}}})\phi_j(t_k^{i_{\text{loop}}} - \tau) \right\rangle_{i_{\text{loop}}=1}^{N_{\text{loop}}}. \tag{M6}$$

The value of N_{loop} varied but was generally ≥ 50 and always chosen to ensure saturated dimension values, i.e., $TN_{\text{loop}} \gg N$. Specific quantities of interest, such as PR^ϕ or $\Psi^\phi(\tau, 0)$, were computed using this estimate.

[1] O. Barak, D. Sussillo, R. Romo, M. Tsodyks, and L. F. Abbott, *From fixed points to chaos: Three models of delayed discrimination*, Prog. Neurobiol. **103**, 214 (2013).

[2] V. Mante, D. Sussillo, K. V. Shenoy, and W. T. Newsome, *Context-dependent computation by recurrent dynamics in prefrontal cortex*, Nature (London) **503**, 78 (2013).

[3] D. Sussillo, M. M. Churchland, M. T. Kaufman, and K. V. Shenoy, *A neural network that finds a naturalistic solution for the production of muscle activity*, Nat. Neurosci. **18**, 1025 (2015).

[4] S. Hochreiter and J. Schmidhuber, *Long short-term memory*, Neural Comput. **9**, 1735 (1997).

[5] I. Sutskever, O. Vinyals, and Q. V. Le, *Sequence to sequence learning with neural networks*, Adv. Neural Inf. Process. Syst. **27**, 3104 (2014).

[6] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, *Learning phrase representations using RNN encoder–decoder for statistical machine translation*, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, edited by A. Moschitti, B. Pang, and W. Daelemans (Association for Computational Linguistics, Doha, Qatar, 2014), pp. 1724–1734.

[7] K. Krishnamurthy, T. Can, and D. J. Schwab, *Theory of gating in recurrent neural networks*, Phys. Rev. X **12**, 011011 (2022).

[8] K. Rajan, L. F. Abbott, and H. Sompolinsky, *Stimulus-dependent suppression of chaos in recurrent neural networks*, Phys. Rev. E **82**, 011903 (2010).

[9] R. Engelken, F. Wolf, and L. F. Abbott, *Lyapunov spectra of chaotic recurrent neural networks*, Phys. Rev. Res. **5**, 043044 (2023).

[10] H. Sompolinsky, A. Crisanti, and H.-J. Sommers, *Chaos in random neural networks*, Phys. Rev. Lett. **61**, 259 (1988).

[11] D. Martí, N. Brunel, and S. Ostojic, *Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks*, Phys. Rev. E **97**, 062314 (2018).

[12] D. Sussillo and L. F. Abbott, *Generating coherent patterns of activity from chaotic neural networks*, Neuron **63**, 544 (2009).

[13] R. Laje and D. V. Buonomano, *Robust timing and motor patterns by taming chaos in recurrent neural networks*, Nat. Neurosci. **16**, 925 (2013).

[14] J. Aljadeff, M. Stern, and T. Sharpee, *Transition to chaos in random networks with cell-type-specific connectivity*, Phys. Rev. Lett. **114**, 088101 (2015).

[15] B. DePasquale, C. J. Cueva, K. Rajan, G. S. Escola, and L. F. Abbott, *full-force: A target-based method for training recurrent networks*, PLoS One **13**, e0191527 (2018).

[16] T. Asabuki and C. Clopath, *Taming the chaos gently: A predictive alignment learning rule in recurrent neural networks*, Nat. Commun. **16**, 6784 (2025).

[17] L. K. Scheffer, C. S. Xu, M. Januszewski, Z. Lu, S.-y. Takemura, K. J. Hayworth, G. B. Huang, K. Shinomiya, J. Maitlin-Shepard, S. Berg *et al.*, *A connectome and analysis of the adult drosophila central brain*, eLife **9**, e57443 (2020).

[18] S. Loomba, J. Straehle, V. Gangadharan, N. Heike, A. Khalifa, A. Motta, N. Ju, M. Sievers, J. Gempt, H. S. Meyer *et al.*, *Connectomic comparison of mouse and human cortex*, Science **377**, eabo0924 (2022).

[19] A. Shapson-Coe, M. Januszewski, D. R. Berger, A. Pope, Y. Wu, T. Blakely, R. L. Schalek, P. H. Li, S. Wang, J. Maitin-Shepard *et al.*, *A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution*, Science **384**, eadk4858 (2024).

[20] *The MICrONS Consortium, Functional connectomics spanning multiple areas of mouse visual cortex*, Nature (London) **640**, 435 (2025).

[21] M. R. Tavakoli, J. Lyudchik, M. Januszewski, V. Vistunou, N. Agudelo Dueñas, J. Vorlaufer, C. Sommer, C. Kreuzinger, B. Oliveira, A. Cenameri *et al.*, *Light-microscopy-based connectomic reconstruction of mammalian brain tissue*, Nature (London) **642**, 398 (2025).

[22] S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, *Highly nonrandom features of synaptic connectivity in local cortical circuits*, PLOS Biol. **3**, e68 (2005).

[23] V. Thibeault, A. Allard, and P. Desrosiers, *The low-rank hypothesis of complex systems*, Nat. Phys. **20**, 294 (2024).

[24] I. Akjouj, M. Barbier, M. Clenet, W. Hachem, M. Maïda, F. Massol, J. Najim, and V. C. Tran, *Complex systems in ecology: A guided tour with large Lotka–Volterra models and random matrices*, Proc. R. Soc. A **480**, 20230284 (2024).

[25] Z. Wang, W. Mai, Y. Chai, K. Qi, H. Ren, C. Shen, S. Zhang, G. Tan, Y. Hu, and Q. Wen, *The geometry and dimensionality of brain-wide activity*, eLife **14**, RP100666 (2025).

[26] L. Pezon, V. Schmutz, and W. Gerstner, *Linking neural manifolds to circuit structure in recurrent networks*, bioRxiv 2024 (2024).

[27] F. Schuessler, F. Mastrogiuseppe, A. Dubreuil, S. Ostojic, and O. Barak, *The interplay between randomness and structure during learning in RNNs*, Adv. Neural Inf. Process. Syst. **33**, 13352 (2020).

[28] C. H. Martin and M. W. Mahoney, *Implicit self-regularization in deep neural networks: Evidence from random matrix theory and implications for learning*, J. Mach. Learn. Res. **22**, 1 (2021).

[29] D. G. Clark, L. F. Abbott, and H. Sompolinsky, *Symmetries and continuous attractors in disordered neural circuits*, bioRxiv 2025 (2025).

[30] E. M. Trautmann, J. K. Hesse, G. M. Stine, R. Xia, S. Zhu, D. J. O'Shea, B. Karsh, J. Colonell, F. F. Lanfranchi, S. Vyas *et al.*, *Large-scale high-density brain-wide neural recording in nonhuman primates*, Nat. Neurosci. **28**, 1562 (2025).

[31] C. Stringer, M. Pachitariu, N. Steinmetz, M. Carandini, and K. D. Harris, *High-dimensional geometry of population responses in visual cortex*, Nature (London) **571**, 361 (2019).

[32] C. Stringer, M. Pachitariu, N. Steinmetz, C. B. Reddy, M. Carandini, and K. D. Harris, *Spontaneous behaviors drive multidimensional, brainwide activity*, Science **364**, eaav7893 (2019).

[33] J. Manley, S. Lu, K. Barber, J. Demas, H. Kim, D. Meyer, F. M. Traub, and A. Vaziri, *Simultaneous, cortex-wide dynamics of up to 1 million neurons reveal unbounded scaling of dimensionality with neuron number*, Neuron **112**, 1694 (2024).

[34] S. Chung and L. F. Abbott, *Neural population geometry: An approach for understanding biological and artificial neural networks*, Curr. Opin. Neurobiol. **70**, 137 (2021).

[35] J. P. Cunningham and M. Y. Byron, *Dimensionality reduction for large-scale neural recordings*, Nat. Neurosci. **17**, 1500 (2014).

[36] P. Gao and S. Ganguli, *On simplicity and complexity in the brave new world of large-scale neuroscience*, Curr. Opin. Neurobiol. **32**, 148 (2015).

[37] E. M. Trautmann, S. D. Stavisky, S. Lahiri, K. C. Ames, M. T. Kaufman, D. J. O'Shea, S. Vyas, X. Sun, S. I. Ryu, S. Ganguli *et al.*, *Accurate estimation of neural population dynamics without spike sorting*, Neuron **103**, 292 (2019).

[38] L. Meshulam and W. Bialek, *Statistical mechanics for networks of real neurons*, Rev. Mod. Phys. (to be published), 10.1103/jcrn-3nrc.

[39] D. G. Clark, L. F. Abbott, and A. Litwin-Kumar, *Dimension of activity in random neural networks*, Phys. Rev. Lett. **131**, 118401 (2023).

[40] This random structure, rather than exact, deterministic orthonormality, offers analytical tractability by eliminating the need to enforce orthonormality—a complicated global constraint; but see also Ref. [41], which handles random orthonormal matrices in a regression setting.

[41] A. Ingrosso, *Optimal learning with excitatory and inhibitory synapses*, PLoS Comput. Biol. **16**, e1008536 (2020).

[42] F. Mastrogiuseppe and S. Ostojic, *Linking connectivity, dynamics, and computations in low-rank recurrent neural networks*, Neuron **99**, 609 (2018).

[43] F. Schuessler, A. Dubreuil, F. Mastrogiuseppe, S. Ostojic, and O. Barak, *Dynamics of random recurrent networks with correlated low-rank structure*, Phys. Rev. Res. **2**, 013111 (2020).

[44] M. Beiran, A. Dubreuil, A. Valente, F. Mastrogiuseppe, and S. Ostojic, *Shaping dynamics with multiple populations in low-rank recurrent networks*, Neural Comput. **33**, 1572 (2021).

[45] A. Dubreuil, A. Valente, M. Beiran, F. Mastrogiuseppe, and S. Ostojic, *The role of population structure in computations through neural dynamics*, Nat. Neurosci. **25**, 783 (2022).

[46] The Gaussian assumption for individual components is not crucial—our results depend on only first and second moments for large N due to the central limit theorem; higher-order cumulants (scaling appropriately with $1/N$) are subdominant in the path-integral saddle-point approximation.

[47] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science* (Cambridge University Press, Cambridge, England, 2018), Vol. 47.

[48] This follows from using Wick's theorem to evaluate the expectations of $\mathrm{tr}(\boldsymbol{LDR}^T\boldsymbol{RDL}^T)$ and $\mathrm{tr}(\boldsymbol{LDR}^T\boldsymbol{RDL}^T\boldsymbol{LDR}^T\boldsymbol{RDL}^T)$ for the numerator and denominator, respectively, of PR^S.

[49] J. Pennington, S. Schoenholz, and S. Ganguli, *Resurrecting the sigmoid in deep learning through dynamical isometry: theory and practice*, in *Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA* (2017), https://papers.nips.cc/paper_files/paper/2017/hash/d9fc0cdb67638d50f411432d0d41d0ba-Abstract.html.

[50] J. Pennington, S. Schoenholz, and S. Ganguli, *The emergence of spectral universality in deep networks*, in *International Conference on Artificial Intelligence and Statistics* (PMLR, 2018), pp. 1924–1932.

[51] This behaves similarly to the more conventional $\phi(x) = \tanh(x)$ but allows for analytical evaluation of Gaussian integrals.

[52] This duality is reminiscent of ergodicity theorems, which relate temporal and spatial averages, but differs in a variety of ways, including that it holds even for finite-time intervals, whereas ergodicity concerns infinite-time limits.

[53] K. Rajan, L. F. Abbott, and H. Sompolinsky, *Inferring stimulus selectivity from the spatial structure of neural network dynamics*, Adv. Neural Inf. Process. Syst. **23** (2010).

[54] L. F. Abbott, K. Rajan, and H. Sompolinsky, *Interactions between intrinsic and stimulus-evoked activity in recurrent neural networks*, in *The Dynamic Brain: An Exploration of Neuronal Variability and its Functional Significance*, edited by M. Ding and D. Glanzman (Oxford University Press, New York, 2011), pp. 65–82.

[55] P. Gao, E. Trautmann, B. Yu, G. Santhanam, S. Ryu, K. Shenoy, and S. Ganguli, *A theory of multineuronal dimensionality, dynamics and measurement*, bioRxiv 214262 (2017).

[56] A. Litwin-Kumar, K. D. Harris, R. Axel, H. Sompolinsky, and L. F. Abbott, *Optimal degrees of synaptic connectivity*, Neuron **93**, 1153 (2017).

[57] S. Recanatesi, G. K. Ocker, M. A. Buice, and E. Shea-Brown, *Dimensionality in recurrent spiking networks: Global trends in activity and local origins in connectivity*, PLoS Comput. Biol. **15**, 1 (2019).

[58] S. Recanatesi, S. Bradde, V. Balasubramanian, N. A. Steinmetz, and E. Shea-Brown, *A scale-dependent measure of system dimensionality*, Patterns **3**, 100555 (2022).

[59] Y. Hu and H. Sompolinsky, *The spectrum of covariance matrices of randomly connected recurrent neuronal networks with linear dynamics*, PLoS Comput. Biol. **18**, 1 (2022).

[60] A. Crisanti and H. Sompolinsky, *Path integral approach to random neural networks*, Phys. Rev. E **98**, 062120 (2018).

[61] M. Helias and D. Dahmen, *Statistical Field Theory for Neural Networks*, Lecture Notes in Physics (Springer, Cham, 2020), 10.1007/978-3-030-46444-8.

[62] W. Zou and H. Huang, *Introduction to dynamical mean-field theory of randomly connected neural networks with bidirectionally correlated couplings*, SciPost Phys. Lect. Notes 079 (2024).

[63] P. C. Martin, E. D. Siggia, and H. A. Rose, *Statistical dynamics of classical systems*, Phys. Rev. A **8**, 423 (1973).

[64] C. De Dominicis, *Dynamics as a substitute for replicas in systems with quenched random impurities*, Phys. Rev. B **18**, 4913 (1978).

[65] J. A. Hertz, Y. Roudi, and P. Sollich, *Path integral methods for the dynamics of stochastic and disordered systems*, J. Phys. A **50**, 033001 (2016).

[66] T. Hromádka, M. R. DeWeese, and A. M. Zador, *Sparse representation of sounds in the unanesthetized auditory cortex*, PLOS Biol. **6**, e16 (2008).

[67] This symmetry arises because the dynamics of the network are determined by an effective coupling matrix $\boldsymbol{LDR}^T\boldsymbol{G}$, where $\boldsymbol{D} = \mathrm{diag}(D_1, \ldots, D_M)$ and $\boldsymbol{G} = \mathrm{diag}(G_1, \ldots, G_N)$. Swapping the distributions of D_a and G_i does not change the spectrum when $\boldsymbol{L}$ and $\boldsymbol{R}$ are i.i.d., since $\boldsymbol{LDR}^T\boldsymbol{G}$ and $\boldsymbol{R}^T\boldsymbol{GLD}$ have the same nonzero eigenvalues.

[68] B. Rider and C. D. Sinclair, *Extremal laws for the real ginibre ensemble*, Ann. Appl. Probab. **24**, 1621 (2014).

[69] D. G. Clark, *Transient dynamics of associative memory models*, arXiv:2506.05303.

[70] If model neurons represent clusters of biological neurons, strong self-connections may instead be interpreted as within-cluster coupling [71,72].

[71] M. Stern, H. Sompolinsky, and L. F. Abbott, *Dynamics of random neural networks with bistable units*, Phys. Rev. E **90**, 062710 (2014).

[72] A. Litwin-Kumar and B. Doiron, *Slow dynamics and high variability in balanced cortical networks with clustered connections*, Nat. Neurosci. **15**, 1498 (2012).

[73] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, *Network motifs: simple building blocks of complex networks*, Science **298**, 824 (2002).

[74] Y. Hu, J. Trousdale, K. Josić, and E. Shea-Brown, *Motif statistics and spike correlations in neuronal networks*, J. Stat. Mech. (2013) P03012.

[75] P. V. Aceituno, T. Rogers, and H. Schomerus, *Universal hypotrochoidic law for random matrices with cyclic correlations*, Phys. Rev. E **100**, 010302(R) (2019).

[76] E. S. Schaffer, N. Mishra, M. R. Whiteway, W. Li, M. B. Vancura, J. Freedman, K. B. Patel, V. Voleti, L. Paninski, E. M. Hillman *et al.*, *The spatial and temporal structure of neural activity across the fly brain*, Nat. Commun. **14**, 5572 (2023).

[77] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek, *Weak pairwise correlations imply strongly correlated network states in a neural population*, Nature (London) **440**, 1007 (2006).

[78] O. Maoz, G. Tkačik, M. S. Esteki, R. Kiani, and E. Schneidman, *Learning probabilistic neural representations with randomly connected circuits*, Proc. Natl. Acad. Sci. U.S.A. **117**, 25066 (2020).

[79] E. Gardner, *The space of interactions in neural network models*, J. Phys. A **21**, 257 (1988).

[80] B. Bordelon, J. Zavatone-Veth, and C. Pehlevan, *Mean field theory of representation learning in large rnns*, in *Cosyne Abstracts 2024* (2024) abstract 3-038.

[81] G. R. Yang, M. R. Joglekar, H. F. Song, W. T. Newsome, and X.-J. Wang, *Task representations in neural networks trained to perform many cognitive tasks*, Nat. Neurosci. **22**, 297 (2019).

[82] L. N. Driscoll, K. Shenoy, and D. Sussillo, *Flexible multitask computation in recurrent networks utilizes shared dynamical motifs*, Nat. Neurosci. **27**, 1349 (2024).

[83] E. Marder and J.-M. Goaillard, *Variability, compensation and homeostasis in neuron and network function*, Nat. Rev. Neurosci. **7**, 563 (2006).

[84] C. Teeter, R. Iyer, V. Menon, N. Gouwens, D. Feng, J. Berg, A. Szafer, N. Cain, H. Zeng, M. Hawrylycz *et al.*, *Generalized leaky integrate-and-fire models classify multiple neuron types*, Nat. Commun. **9**, 709 (2018).

[85] B. K. Murphy and K. D. Miller, *Balanced amplification: A new mechanism of selective amplification of neural activity patterns*, Neuron **61**, 635 (2009).

[86] G. Hennequin, T. P. Vogels, and W. Gerstner, *Optimal control of transient dynamics in balanced networks supports generation of complex movements*, Neuron **82**, 1394 (2014).

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

[88] J. Kadmon and H. Sompolinsky, *Transition to chaos in random neuronal networks*, Phys. Rev. X **5**, 041030 (2015).

[89] L. Tiberi, D. Dahmen, and M. Helias, *Hidden connectivity structures control collective network dynamics*, arXiv:2303.02476.

[90] V. Schmutz, J. Brea, and W. Gerstner, *Emergent rate-based dynamics in duplicate-free populations of spiking neurons*, Phys. Rev. Lett. **134**, 018401 (2025).

[91] Z. Lin and H. Huang, *Spiking mode-based neural networks*, Phys. Rev. E **110**, 024306 (2024).

[92] N. Eckstein, A. S. Bates, A. Champion, M. Du, Y. Yin, P. Schlegel, A. K.-Y. Lu, T. Rymer, S. Finley-May, T. Paterson *et al.*, *Neurotransmitter classification from electron microscopy images at synaptic sites in drosophila melanogaster*, Cell **187**, 2574 (2024).

[93] W. W. Liu and R. I. Wilson, *Glutamate is an inhibitory neurotransmitter in the drosophila olfactory system*, Proc. Natl. Acad. Sci. U.S.A. **110**, 10294 (2013).

[94] S. Molina-Obando, J. F. Vargas-Fique, M. Henning, B. Gür, T. M. Schladt, J. Akhtar, T. K. Berger, and M. Silies, *On selectivity in the drosophila visual system is a multisynaptic process involving both glutamatergic and gabaergic inhibition*, eLife **8**, e49373 (2019).

[95] O. Roy and M. Vetterli, *The effective rank: A measure of effective dimensionality*, in *2007 15th European Signal Processing Conference, Poznan, Poland* (IEEE, 2007), pp. 606–610.

[96] G. Christodoulou, T. P. Vogels, and E. J. Agnes, *Regimes and mechanisms of transient amplification in abstract and biological neural networks*, PLoS Comput. Biol. **18**, e1010365 (2022).

[97] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications* (World Scientific Publishing Company, Singapore, 1987), Vol. 9.

[98] R. Price, *A useful theorem for nonlinear devices having Gaussian inputs*, IRE Trans. Inf. Theory **4**, 69 (1958).

[99] A. van Meegen, T. Kühn, and M. Helias, *Large-deviation approach to random recurrent neuronal networks: Parameter inference and fluctuation-induced transitions*, Phys. Rev. Lett. **127**, 158302 (2021).

[100] M. A. Buice and J. D. Cowan, *Field-theoretic approach to fluctuation effects in neural networks*, Phys. Rev. E **75**, 051919 (2007).

[101] Note that the variance of $\eta^z(t)$ is scaled by D^2; equivalently, we could factor out D^2 from the variance and include D as a multiplicative factor in the latent-variable single-site dynamics [Eq. (I17)].

[102] J. P. Stroud, M. A. Porter, G. Hennequin, and T. P. Vogels, *Motor primitives in space and time via targeted gain modulation in cortical networks*, Nat. Neurosci. **21**, 1774 (2018).

[103] M. Mézard, *The space of interactions in neural networks: Gardner's computation with the cavity method*, J. Phys. A **22**, 2181 (1989).

[104] M. Mézard, *Mean-field message-passing equations in the Hopfield model and its generalizations*, Phys. Rev. E **95**, 022117 (2017).

[105] F. Mastrogiuseppe, *From dynamics to computations in recurrent neural networks*, Ph.D. thesis, Université Paris sciences et lettres, 2017.