Relating local connectivity and global dynamics in recurrent excitatory-inhibitory networks

Yuxiu Shao, Srdjan Ostojic

Laboratoire de Neurosciences Cognitives et Computationnelles, INSERM U960, Ecole Normale Superieure - PSL Research University, 75005 Paris, France

* yuxiu.shao@ens.psl.eu, * srdjan.ostojic@ens.fr

Abstract

How the connectivity of cortical networks determines the neural dynamics and the resulting computations is one of the key questions in neuroscience. Previous works have pursued two complementary strategies to quantify the structure in connectivity, by specifying either the local statistics of connectivity motifs between small groups of neurons, or by defining network-wide low-rank patterns of connectivity that determine the resulting low-dimensional dynamics. A direct relationship between these two approaches is however currently missing, and in particular it remains to be clarified how local connectivity statistics are related to the global connectivity structure and shape the low-dimensional activity. To bridge this gap, here we develop a method for mapping local connectivity statistics onto an approximate global low-rank structure. Our method rests on approximating the global connectivity matrix using dominant eigenvectors, which we compute using perturbation theory for random matrices. This approach demonstrates that multi-population networks defined from local connectivity properties can in general be approximated by low-rank connectivity with Gaussian-mixture statistics. We specifically apply this method to excitatory-inhibitory networks, and show that it leads to accurate predictions for both the low-dimensional dynamics, and for the activity of individual neurons. Altogether, our approach allows us to disentangle the effects of mean connectivity and reciprocal motifs on the global recurrent feedback, and provides an intuitive picture of how local connectivity shapes global network dynamics.

Author summary

The structure of connections between neurons is believed to determine how cortical networks control behaviour. Current experimental methods typically measure connections between small numbers of simultaneously recorded neurons, and thereby provide information on statistics of local connectivity motifs. Collective network dynamics are however determined by network-wide patterns of connections. How these global patterns are related to local connectivity statistics and shape the dynamics is an open question that we address in this study. Starting from networks defined in terms of local statistics, we develop a method for approximating the resulting connectivity by global low-rank patterns. We apply this method to classical excitatory-inhibitory networks and show that it allows us to predict both collective and single-neuron activity. More generally, our approach provides a link between local connectivity statistics and global network dynamics.

Introduction

One of the central questions in neuroscience is how the connectivity structure of cortical networks determines the collective dynamics of neural activity and their function. Experimental assessments of connectivity are typically based on measurements of synaptic weights between small numbers of neurons recorded simultaneously [1–9]. The most common approach to quantify connectivity therefore focuses on *local statistics*, and starts by characterizing the connection probability between pairs of neurons based on their type, before considering progressively more complex connectivity motifs. Linking these local connectivity statistics to the emerging network dynamics has been an active topic of investigations [10–24]. A second approach, motivated by computational network models instead of experimental measurements [25–29], instead specifies the connectivity in terms of a low-rank structure defined by network-wide patterns of connectivity [30–39]. This global connectivity structure directly determines the low-dimensional dynamics and resulting computations [30, 31, 33], yet it remains unclear how it is related to local connectivity statistics that can be recorded experimentally. In this study, we aim to bridge this gap, by mapping local connectivity statistics onto a global, low-rank description of connectivity and comparing the resulting dynamics.

Starting from random networks with connectivity defined in terms of local, cell-type dependent statistics, we develop a low-rank approximation based on the dominant eigenmodes of the connectivity matrix. Using perturbation theory, we show that the obtained low-rank connectivity patterns universally obey Gaussian-mixture statistics and therefore lead to analytically tractable dynamics [31, 33]. We specifically apply this approach to excitatory-inhibitory networks with connections consisting of independent and reciprocal parts, and exploit the low-rank approximation to predict the emerging dynamics.

We first show that, although the dominant low-rank structure is set on average by the mean synaptic weights [40–42], a perturbative approach accurately predicts the components of individual neurons on the dominant eigenvectors for individual instances of the random connectivity. As a result, our low-rank approximation analytically predicts the activity of individual neurons in the original E-I network defined based on local statistics. The analytic description of the dynamics in the low-rank approximation moreover leads to the identification of two distinct sources of recurrent feedback corresponding respectively to the mean connectivity and reciprocal connections between neurons. In particular, the reciprocal motifs impact dynamics by modulating both the dominant eigenvalue and the corresponding eigenvectors, and can give rise to additional bistability in the network. Altogether, our analytical mapping of the local EI statistics to a low-rank description provides a quantitative and intuitive description of how local connectivity statistics determine global low-dimensional dynamics.

1 Results

1.1 Local vs. global representations of random recurrent connectivity.

We study networks of N rate units with random recurrent connectivity given by the connectivity matrix $\mathbf{J}$, where the entry J_{ij} corresponds to the strength of the synapse from neuron j to neuron i. A full statistical description of the random connectivity would require specifying the joint distribution $P(\{J_{ij}\})$ of the N^2 synaptic weights. Determining the dynamics from this high-dimensional distribution is however in general intractable. We therefore focus on connectivity models that make simplifying assumptions on the underlying statistics.

Our specific goal is to relate two different classes of such models, which we refer to as the *local* and the *global* representations of recurrent connectivity. Both representations assume that the network consists of P populations, and the statistics of connectivity depend only on the pre- and post-synaptic populations. The two representations however take as starting points different statistical features of the connectivity.

The local representation defines the connectivity statistics by starting from the marginal distributions $\text{Prob}(J_{ij} = J)$ of individual synaptic weights, and by including progressively higher-order correlations referred to as *connectivity motifs* [1, 10, 14]. In this work, we will consider only the first two orders, i. e. the distribution of individual weights and the pairwise correlations η_{ij} between reciprocal connections J_{ij} and J_{ji} that quantify pairwise motifs (Fig 1A). Our key assumption is that both the marginal distributions of J_{ij} and the correlations η_{ij} depend only on the populations p and q that the post- and pre-synaptic neurons belong to:

$$\text{Prob}(J_{ij} = J) = f^{pq}(J), \quad p, q = 1 \dots P,$$
$$\eta_{ij} = \eta_{pq}. \tag{1}$$

All synapses connecting the same two populations therefore have identical statistics, leading to a block-like statistical structure for the connectivity matrix $\mathbf{J}$ (Fig 1B left panel).

The global representation of connectivity instead refers to the situation where $\mathbf{J}$ is defined as a low-rank matrix [30, 31, 33]:

$$\mathbf{J} = \frac{1}{N} \sum_{r=1}^{R} \mathbf{m}^{(r)} \mathbf{n}^{(r)\mathsf{T}}. \tag{2}$$

Here $\mathbf{m}^{(r)} = \{m_i^{(r)}\}_{i=1\dots N}$ and $\mathbf{n}^{(r)} = \{n_i^{(r)}\}_{i=1\dots N}$ for $r = 1 \dots R$ are referred to as *connectivity vectors*, where R is the rank of $\mathbf{J}$. In this representation, the statistics of connectivity are defined by the distribution of vector elements, rather than directly by the distribution of synaptic weights as in the local representation. Specifically, each neuron i is characterized by its set of entries $(m_i^{(1)}, \dots, m_i^{(R)}, n_i^{(1)}, \dots, n_i^{(R)})$ over the connectivity vectors. For each neuron, these $2R$ entries are generated from a joint distribution, independently of the other neurons, and the parameters of this joint distribution depend on the population p the neuron belongs to. Here we focus on the broad class of Gaussian-mixture low-rank networks, in which for population p, the joint distribution of elements is a multi-variate Gaussian defined by the means and covariances of the $2R$ entries [31, 33] (Fig 1C).

To relate the local and the global representations of connectivity, a key observation is that any matrix $\mathbf{J}$ generated from the local statistics defined in Eqs. (1) can be expressed as
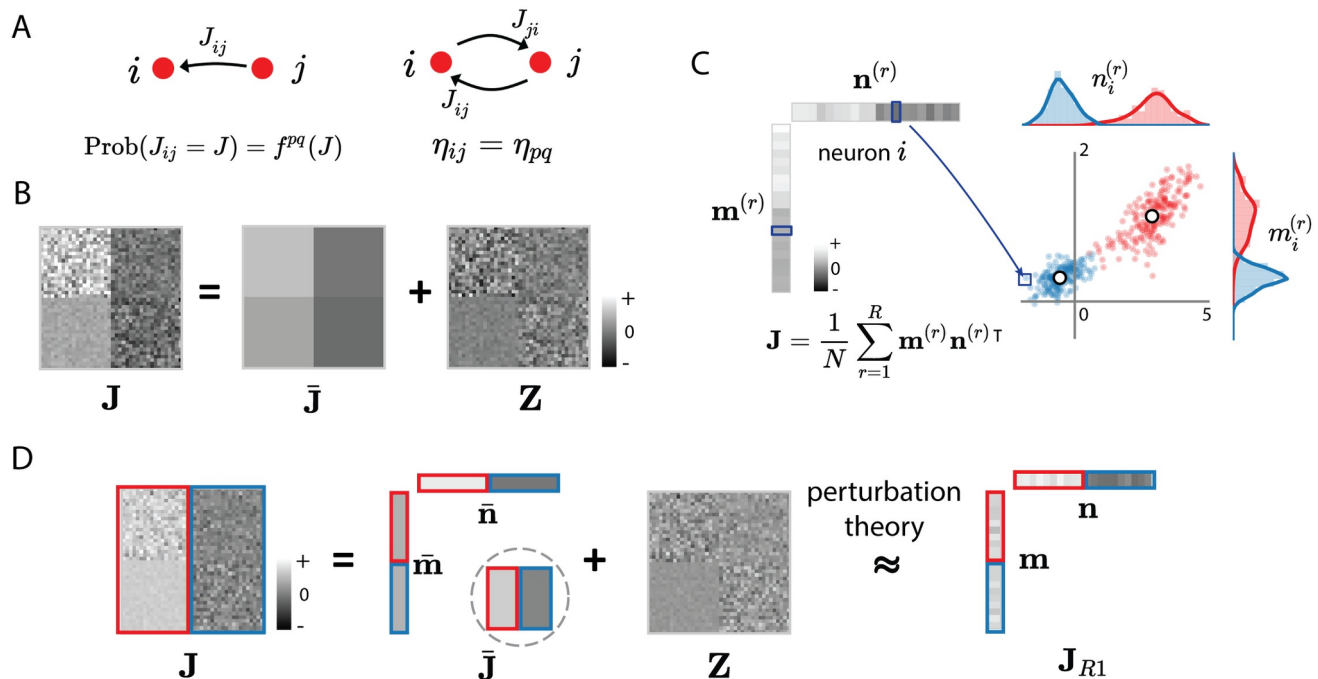
Fig 1. Local vs global representations of recurrent connectivity. (A) The local representation defines the statistics of synaptic weights J_{ij} by starting from the marginal probability distribution of individual synaptic weights (left) and then specifying reciprocal motifs in terms of correlations η_{ij} between reciprocal weights J_{ij} and J_{ji} connecting neurons i and j (right). Both the marginal distribution and the reciprocal correlations are assumed to depend only on the populations p and q that the neurons i and j belong to. (B) The resulting connectivity matrix $\mathbf{J}$ has block-structured statistics, where different blocks correspond to connections between the P different populations ($P = 2$ in this illustration). It can be decomposed into a superposition of a mean component $\bar{\mathbf{J}}$, and a remaining zero-mean random connectivity component $\mathbf{Z}$ that has block-structured variances. (C) The global, low-rank representation defines the connectivity matrix $\mathbf{J}$ as the sum of R outer products between connectivity vectors $\mathbf{m}^{(r)}$, $\mathbf{n}^{(r)}$ for $r = 1 \ldots R$. The statistics of connectivity are defined in terms of the joint probability distribution over neurons i of their entries $(m_i^{(1)}, \ldots, m_i^{(R)}, n_i^{(1)}, \ldots, n_i^{(R)})$ on connectivity vectors. We specifically consider the class of Gaussian-mixture low-rank models, where each neuron is first assigned to a population p, and within each population the entries on connectivity vectors are generated from a multivariate Gaussian distribution with fixed statistics. Here we illustrate this distribution for one pair of connectivity vectors ($R = 1$) and $P = 2$ populations. Each dot represents the connectivity parameters $(m_i^{(r)}, n_i^{(r)})$ of one neuron i, the red and blue colours denote the two populations, white dots and the rotations of the dot clouds indicate the mean and covariance of the distribution for each population. (D) Relating the local and global representations of recurrent connectivity for a simplified excitatory-inhibitory network. In this model, the mean connectivity depends only on the presynaptic population (indicated by red and blue colours). The mean connectivity $\bar{\mathbf{J}}$ is in this case rank-one, and can be written as an outer product of vectors $\bar{\mathbf{m}}$ and $\bar{\mathbf{n}}$. We approximate the full connectivity by a rank-one matrix, with connectivity vectors $\mathbf{m}$ and $\mathbf{n}$ obtained from $\bar{\mathbf{m}}$ and $\bar{\mathbf{n}}$ using perturbation theory.

$$\mathbf{J} = \bar{\mathbf{J}} + \mathbf{Z} \tag{3}$$

where $\bar{\mathbf{J}}$ contains the mean values of the connections, and $\mathbf{Z}$ contains the remaining, zero-mean random part [40]. Because of the underlying population structure (Eqs. (1)), $\bar{\mathbf{J}}$ consists of $P \times P$ blocks with identical values within each block (Fig 1B middle panel), and is therefore at most of rank P. The random part $\mathbf{Z}$ is instead in general of rank N, but obeys block-like statistics, with variance and normalized covariance parameters defined by $P \times P$ matrices (Methods Secs. 2.1.1-2.1.2, Eqs. (25), (30)).

For the sake of simplicity, in this study, we focus on a simplified excitatory-inhibitory model [41, 43]. This network consists of one excitatory and one inhibitory population, so $P = 2$ and in the following we use the population indices $p, q = E, I$. A central simplifying assumption in this model is that the mean synaptic weights depend only on the pre-synaptic population, so that $\bar{J}_{EE} = \bar{J}_{IE} > 0$ and $\bar{J}_{II} = \bar{J}_{EI} < 0$. The mean connectivity matrix $\bar{\mathbf{J}}$ therefore consists of only two blocks and is unit rank (Fig 1D). The statistics of the random part $\mathbf{Z}$ instead depend on both pre- and post-synaptic populations, and are therefore described by 2×2 matrices of variance and normalized covariance parameters (see Methods Sec. 2.1.3, Eq. (35)).

1.2 Approximating locally-defined connectivity with low-rank connectivity.

To relate the local and global representations of connectivity, we start from a connectivity matrix $\mathbf{J}$ generated from the local statistics (Eqs. (1)) and approximate it by a rank-R matrix of the form given in Eq. (2). As the locally-defined connectivity matrix $\mathbf{J}$ is of rank N, this is equivalent to the classical low-rank approximation problem, for which a variety of methods exist [30, 31, 33]. Here we use simple truncated eigen-decomposition as it preserves the dominant eigenvalues that determine non-linear dynamics.

Applying the standard eigenmode decomposition, $\mathbf{J}$ can be in general factored as

$$\mathbf{J} = \frac{1}{N} \sum_{r=1}^{N} \mathbf{m}^{(r)} \mathbf{n}^{(r)\mathsf{T}}, \tag{4}$$

where $\mathbf{m}^{(r)}$ and $\mathbf{n}^{(r)}$ are rescaled versions of the r-th right and left eigenvectors (Methods Sec. 2.3, Eqs (42)-(47)), ordered by the absolute value of their eigenvalue λ_r for $r = 1 \ldots N$. A rank-R approximation that preserves the top R eigenvalues can then be obtained by simply keeping the first R terms in the sum in Eq. (4). In this study, we focus on $R = 1$, corresponding to the dominant eigenvalue. Higher order approximations will be described elsewhere.

Eigenvalues and eigenvectors are in general complex non-linear functions of the entries of the matrix $\mathbf{J}$. To determine the dominant eigenvalues and the corresponding vectors of $\mathbf{J}$, we capitalize on the observation in Eq. (3) that a locally-defined connectivity matrix can in general be expressed as a sum of a low-rank matrix of mean values $\bar{\mathbf{J}}$ and the remaining random part $\mathbf{Z}$. Previous studies have found that the eigenspectra of matrices with such structure typically consist of two components in the complex plane: a continuously-distributed bulk determined by the random part, and discrete outliers controlled by the low-rank structure [30, 32, 39, 44–46]. In this study, we extend previous approaches to determine the influence of the block-like statistics of $\mathbf{Z}$ on the outliers that correspond to dominant eigenvalues. We then use perturbation theory to determine the corresponding left and right eigenvectors and their statistical

structure. Here we summarize the main steps of this analysis (full details are provided in Methods), and then apply it to specific cases in the following sections.

We focus on the simplified E-I network for which the mean part of the connectivity is unit rank and can therefore be written as $\bar{\mathbf{J}} = \bar{\mathbf{m}}\bar{\mathbf{n}}^\mathsf{T}/N$, so that the full connectivity matrix is

$$\mathbf{J} = \frac{1}{N}\bar{\mathbf{m}}\bar{\mathbf{n}}^\mathsf{T} + \mathbf{Z}. \tag{5}$$

The mean part $\bar{\mathbf{J}}$ of the connectivity has a unique non-trivial eigenvalue $\lambda_0 = \bar{\mathbf{n}}^\mathsf{T}\bar{\mathbf{m}}/N$ which can give rise to one or several outliers λ in the eigenspectrum of $\mathbf{J}$. To determine how the random part of the connectivity influences λ, we start from the characteristic equation for the eigenvalues of $\mathbf{J}$ and exploit Eq. (5) to apply the matrix determinant lemma (Eq. (57)). This leads to a non-linear equation for λ [32]:

$$\lambda = \sum_{k=0}^{\infty} \frac{\theta_k}{\lambda^k} \quad \text{with} \quad \theta_k = \frac{1}{N}\bar{\mathbf{n}}^\mathsf{T}\mathbf{Z}^k\bar{\mathbf{m}}. \tag{6}$$

Truncating the sum to second order yields an approximate third order polynomial for λ:

$$\lambda^3 = \lambda_0\lambda^2 + \theta_1\lambda + \theta_2, \quad \text{with} \quad \theta_1 = \bar{\mathbf{n}}^\mathsf{T}\mathbf{Z}\bar{\mathbf{m}}/N \quad \theta_2 = \bar{\mathbf{n}}^\mathsf{T}\mathbf{Z}^2\bar{\mathbf{m}}/N. \tag{7}$$

The statistics of the outlying eigenvalue can then be obtained by averaging over the random part of the connectivity $\mathbf{Z}$.

An approximate expression for the right and left connectivity vectors $\mathbf{m}$ and $\mathbf{n}$ of $\mathbf{J}$ corresponding to the outliers λ can be determined using first order perturbation theory [47]. We first note that $\bar{\mathbf{m}}$ and $\bar{\mathbf{n}}$ are the right- and left-eigenvectors of $\bar{\mathbf{J}}$ corresponding to the non-trivial eigenvalue λ_0. Interpreting the full connectivity matrix $\mathbf{J}$ as $\bar{\mathbf{J}}$ perturbed by a random matrix $\mathbf{Z}$, at first order $\mathbf{m}$ and $\mathbf{n}$ can be expressed as

$$\begin{aligned} \mathbf{m} &= \bar{\mathbf{m}} + \Delta\mathbf{m} \\ \mathbf{n} &= \bar{\mathbf{n}} + \Delta\mathbf{n}, \end{aligned} \tag{8}$$

with

$$\begin{aligned} \Delta\mathbf{m} &= \mathbf{Z}\bar{\mathbf{m}}/\lambda_0 \\ \Delta\mathbf{n} &= \mathbf{Z}^\mathsf{T}\bar{\mathbf{n}}/\lambda_0. \end{aligned} \tag{9}$$

A key observation is that each element of $\Delta\mathbf{m}$ and $\Delta\mathbf{n}$ is a sum of N random variables. The central limit theorem therefore predicts that, in the limit of large N, the statistics of Δm_i and Δn_i, and therefore m_i and n_i, follow a Gaussian distribution. In general, the mean and variance of m_i and n_i and their correlation are determined by the mean, variance and correlation of the elements of $\mathbf{Z}$, but not the specific form of the probability distribution. Since the matrix $\mathbf{Z}$ has block-like statistics determined by the population structure, the statistics of the resulting m_i and n_i depend on the population p the neuron i belongs to. Overall, the distribution of elements of $\mathbf{m}$ and $\mathbf{n}$ obtained from perturbation theory therefore follow a Gaussian-mixture distribution, so that our approach effectively approximates a locally-defined $\mathbf{J}$ by a Gaussian-mixture low-rank model specified by the means $\bar{m}^p$, $\bar{n}^p$, the variances $\sigma_{m^p}^2$, $\sigma_{n^p}^2$ and the covariances σ_{nm}^p of the entries on the connectivity vectors for $p = E, I$.

We next apply the perturbative approach described here to networks with independent random components, and then to networks with reciprocal motifs.
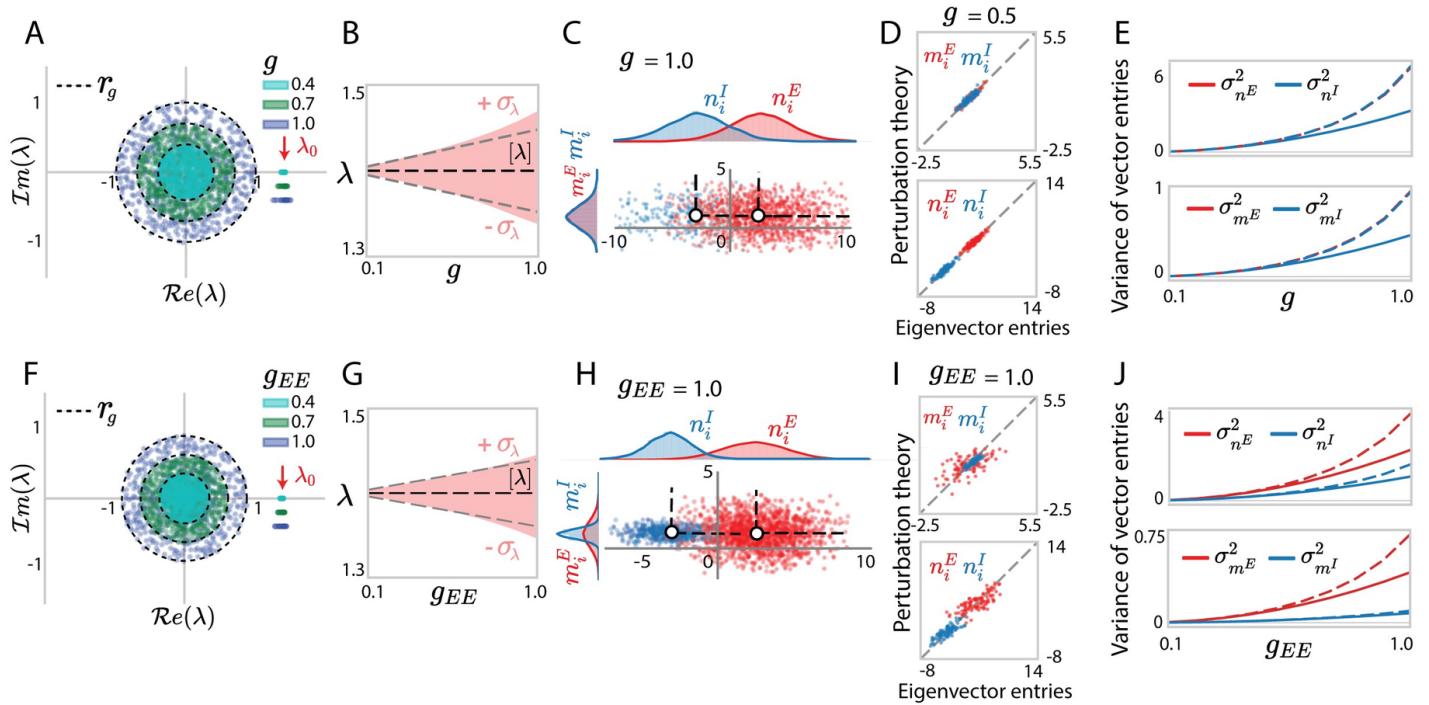
Fig 2. Eigenvalues and dominant eigenvectors for locally-defined Gaussian connectivity with independent synaptic weights. (A) Eigenvalue spectra of excitatory-inhibitory connectivity matrices $\mathbf{J}$ with elements generated from Gaussian distributions with identical variances g^2/N over neurons. The coloured dots in the circular bulk shows 600 eigenvalues for one realization of the random connectivity for each value of g. Different colours correspond to different values of g. Dashed envelopes indicate the theoretical predictions for the radius r_g of the circular bulk computed according to Eqs. (146), (147). Outlying eigenvalues are shown for 30 realizations of the random connectivity, and for different g their location on the y-axis is shifted to help visualization. The red arrow points to the eigenvalue λ_0 of the mean connectivity matrix $\bar{\mathbf{J}}$. (B) Statistics of outlying eigenvalues over realizations of random connectivity. Empirical distribution (the red area shows mean $\pm$ standard deviation), compared with the theoretical predictions for the mean (black dashed line) and standard deviation (gray dashed line) obtained using Eq. (78). (C) Scatter plot showing for each neuron i its entry n_i on the left eigenvector against its entry m_i on the right eigenvector. Red and blue colours represent respectively excitatory and inhibitory neurons. The white dots and the dashed lines respectively indicate the means and covariances for each population obtained from simulations. (D) Comparison between eigenvector entries obtained from direct eigen-decomposition of $\mathbf{J}$ with predictions of perturbation theory (Eqs. (8), (9)). (E) Comparison between simulations (full lines) and theory (dashed lines) for the variances $\sigma^2_{n^p}$, $\sigma^2_{m^p}$ of eigenvector entries corresponding to different populations (Eq. (88)). (F-J) Identical quantities for connectivity matrices in which the variance parameters are heterogeneous: $g_{EE} : g_{EI} : g_{IE} : g_{II} = 1.0 : 0.5 : 0.2 : 0.8$, g_{EE} increases from 0 to 1. Other network parameters $N_E = 4N_I = 1200$ and $J_E = 2.0, J_I = 0.6$ in all simulations.

1.3 Low-rank structure induced by independently generated synaptic connections

We first apply our approach for a low-rank approximation to the simplest version of the locally-defined excitatory-inhibitory network where each J_{ij} is generated independently from a Gaussian distribution with a mean that depends only on the pre-synaptic population, i. e. $[J_{ij}] = \bar{J}_{pq} = \bar{J}_q$ with $p, q \in E, I$. The entries of the eigenvectors $\bar{\mathbf{m}}$ and $\bar{\mathbf{n}}$ of the mean connectivity matrix $\bar{\mathbf{J}}$ are then given by:

$$
\begin{align}
\bar{m}_i &= 1, \quad i = 1 \ldots N \tag{10}\\
\bar{n}_i &= \bar{n}^E = \frac{N}{N_E} J_E \quad i \in N_E \tag{11}\\
\bar{n}_i &= \bar{n}^I = -\frac{N}{N_I} J_I \quad i \in N_I. \tag{12}
\end{align}
$$

We first consider the case where the variance of J_{ij} is uniform across all connections and given by g^2/N. In that situation, the entries of the random part of the connectivity $\mathbf{Z}$ are independent, identically distributed Gaussians, and the eigenvalue spectrum of $\mathbf{Z}$ converges to a uniform distribution on the complex plane within a circle of radius g [48–50]. Previous studies [30, 32, 44, 45] have shown that adding a unit rank matrix on top of an i. i. d. random matrix as in Eq. (3) leads to an eigenspectrum that is the superposition of the two spectra, and therefore consists of a circular bulk of radius g and an outlier on average located at the eigenvalue λ_0 of $\bar{\mathbf{J}}$. Our analyses of the eigenvalues of $\mathbf{J}$ confirm this result (Fig 2A). Indeed, averaging over $\mathbf{Z}$ in Eq. (6), $[\theta_k] = 0$ for all k [32], so that the outlier is on average given by $[\lambda] = \lambda_0$. Our approach moreover gives an expression for the standard deviation of the outlier which grows linearly with g (Fig 2B, Eq. (78)). Examining the entries of the left and right eigenvectors $\mathbf{n}$ and $\mathbf{m}$ of $\mathbf{J}$ corresponding to the outlier, we find that perturbation theory accurately predicts the individual entries of the eigenvectors as long as g is sufficiently below unity (Fig 2D). As expected, the distribution of (m_i, n_i) is well described by a mixture of two Gaussians centred at $\bar{m}^p$, $\bar{n}^p$. Perturbation theory provides a lower bound for the values of the corresponding variances (Fig 2E). For large values of g, the distributions remain Gaussian, but their variances increase above the predictions of perturbation theory. Importantly, the entries of the left and right eigenvectors are uncorrelated, and only their means, but not their variances, differ between the two populations (Figs 2C, E).

We next turned to the case where the variances of synaptic weights depend on the pre- and post-synaptic populations q, p, and are given by g_{pq}^2/N. In that case, the entries of the random part of the connectivity $\mathbf{Z}$ are independent, but not identically distributed Gaussians. Previous studies [11, 44] have shown that the spectrum of $\mathbf{Z}$ remains circularly symmetric, but its radius r_g is determined by a combination of variance parameters g_{pq} (S4 Appendix, Eqs. (146), (147)). Examining the resulting connectivity matrix $\mathbf{J}$, we found that the results for the uniform case directly extend to this heterogeneous situation. The eigenspectrum of $\mathbf{J}$ still consists of an independent superposition of the spectra of $\mathbf{Z}$ and $\bar{\mathbf{J}}$ (Fig 2F). In particular, the random part of the connectivity does not modify the average value of the outlier, but only impacts its variance, which now depends on a combination of the variances g_{pq} (Fig 2G, Eq. (88)). Similarly to the uniform case, the distribution of the entries of the left and right eigenvectors is well described by a mixture of two Gaussians, with variances predicted by perturbation theory (Fig 2I). The entries of the left and right eigenvectors are uncorrelated, but now both their means and variances depend on the population the neuron belongs to (Figs 2H, J).

In summary, when synaptic connections J_{ij} are generated independently across pairs of neurons, the equivalent global representation is a Gaussian-mixture low-rank model
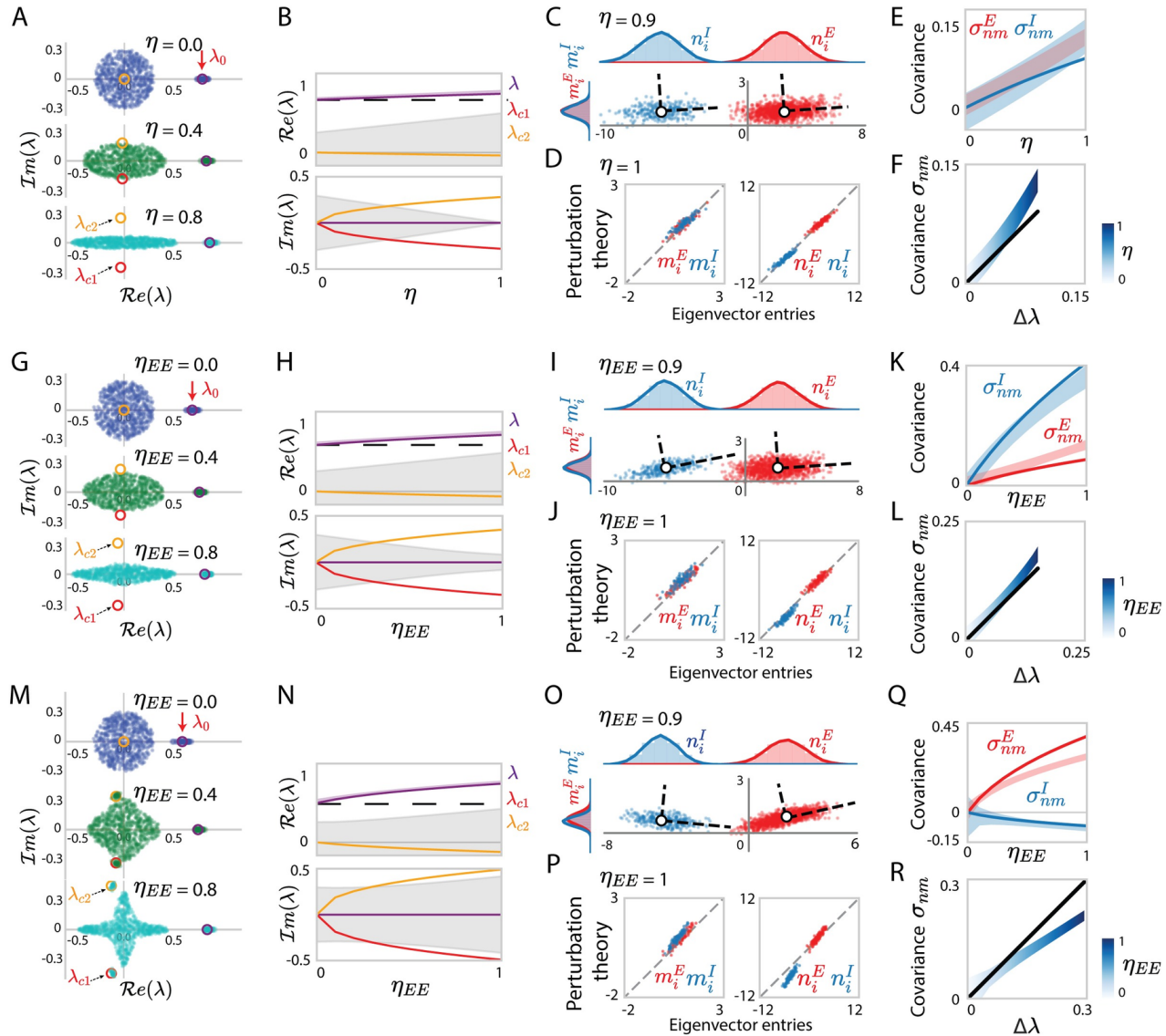
Fig 3. Eigenvalues and dominant eigenvectors for locally-defined Gaussian connectivity with reciprocal motifs. (A) Eigenvalue spectra of excitatory-inhibitory connectivity matrices $\mathbf{J}$, with homogeneous reciprocal correlations η. Different colours correspond to networks with different values of η. The dots in the elliptical bulk show 600 eigenvalues for one realization of the random connectivity. Outlying eigenvalues are shown for 30 realizations of the random connectivity. The red arrow on the top points to the eigenvalue λ_0 of the mean connectivity $\bar{\mathbf{J}}$. Coloured circles are the eigenvalues predicted using determinant lemma (Eq. (7)). (B) Comparison of the eigenvalues from the finite-size simulation with the predictions of the determinant lemma as the reciprocal correlation η is increased. The coloured solid lines show the roots of the third-order polynomial in Eq. (7). The light purple area indicates the empirical distribution of the dominant outlier, while the black dashed line is the unperturbed eigenvalue λ_0. The grey areas represent the areas covered by the eigenvalue bulk. (C) Scatter plot showing for each neuron i its entry n_i on the left eigenvector against its entry m_i on the right eigenvector. Red and blue colours represent respectively excitatory and inhibitory neurons. The white dots and the dashed lines respectively indicate the means and covariances for each population. (D) Comparison between eigenvector entries obtained from direct eigen-decomposition of $\mathbf{J}$ with predictions of perturbation theory (Eqs. (8), (9)). (E) Comparison between simulations (coloured areas) and theoretical predictions (coloured lines, Eq. (95)) for the population covariance σ_{nm}^p of the entries on the left and right connectivity eigenvectors to different populations. (F) Comparison of the overall covariance σ_{nm} (Eq. (70))

Fig 3. (Continued from previous page.) with the deviation $\Delta\lambda$ of the dominant outlying eigenvalue from the unperturbed value λ_0. Empirical covariance (gradient blue area, colour depth represents η) compared with the theoretical prediction (black line) obtained using Eqs. (95), (90). The x-axis uses the theoretical prediction of the deviation of the eigenvalue λ from λ_0. (G-L) Same as (A-F) for a connectivity matrix with heterogeneous reciprocal correlations: $\eta_{EE} = \eta_{EI} = -\eta_{II} > 0$, and η_{EE} increasing from 0 to 1. (M-R) Same as (A-F) for a connectivity matrix with heterogeneous reciprocal correlations: $\eta_{EE} = -\eta_{EI} = -\eta_{II} > 0$, and η_{EE} increasing from 0 to 1. Other network parameters: $N_E = 4N_I = 1200$ and homogeneous variance parameters $g_{pq} = g = 0.3$ in all simulations, $J_E = 2.0$, $J_I = 1.2$ for networks in (A-F), $J_E = 2.0$, $J_I = 1.3$ for networks in (G-L), $J_E = 2.0$, $J_I = 1.4$ for networks in (M-R).

where the entries of the structure vectors are independent with mean values determined by the low-rank structure of the mean connectivity matrix $\bar{\mathbf{J}}$. Importantly, in that situation, the dominant outlying eigenvalues of $\mathbf{J}$ are on average identical to those of $\bar{\mathbf{J}}$.

1.4 Low-rank structure induced by reciprocal motifs.

We next turn to locally-defined excitatory-inhibitory networks with reciprocal connectivity motifs quantified by the correlation η_{ij} between reciprocal synaptic weights J_{ij} and J_{ji}. We assumed that these reciprocal correlations are identical for any pair of neurons i and j belonging to a given pair of populations p and q, and used the corresponding parameters η_{pq} to generate the connectivity matrix $\mathbf{J}$ (Methods Sec. 2.1.2). Within the decomposition of $\mathbf{J}$ in a mean $\bar{\mathbf{J}}$ and random part $\mathbf{Z}$ (Eq. (22)), the additional reciprocal correlations affect only the statistics of $\mathbf{Z}$.

We first consider the homogeneous case where the reciprocal correlation is identical across all populations, i. e. $\eta_{pq} = \eta$. Previous studies have shown that a random matrix $\mathbf{Z}$ with zero mean and reciprocal correlations η has a continuous spectrum that is deformed from a circle into an ellipse as η is increased [18, 51]. Superpositions between correlated random matrices, and low-rank structure such as $\bar{\mathbf{J}}$ have to our knowledge not been previously studied. Inspecting the eigenspectrum of $\mathbf{J} = \bar{\mathbf{J}} + \mathbf{Z}$, we found that it still consists of a continuous bulk and discrete outliers (Fig 3A). The continuous bulk is contained in an ellipse in the complex plane identical to the spectrum of $\mathbf{Z}$, as in the uncorrelated case. In contrast, we found that the outliers deviated from the eigenvalues of $\bar{\mathbf{J}}$ as η was increased. These deviations were well captured by our analytic approach summarized in Eq. (7). Indeed, when averaging Eq. (7) over $\mathbf{Z}$, reciprocal correlations generate a non-zero $[\theta_2] = [\bar{\mathbf{n}}^\intercal \mathbf{Z}^2 \bar{\mathbf{m}}]/N$ due to $\mathbf{Z}^2$. This term leads to a cubic equation in Eq. (7) and therefore has two potential effects. First, the non-zero θ_2 induces deviations of the outliers from the eigenvalue λ_0 of $\bar{\mathbf{J}}$. The direction of these deviations is positive if excitation dominates ($\lambda_0 = J_E - J_I > 0$) and negative if inhibition dominates ($\lambda_0 = J_E - J_I < 0$, Figs 4F-H). Second, the cubic equation can have up to three solutions and therefore potentially generates additional outliers, and in particular complex conjugate ones (Fig 3M). Whether these additional outliers are observed depends on the accuracy of the third-order approximation to the determinant lemma, and on the norm of these outliers compared to the spectral radius (Fig 4E).

We next examine the right- and left-eigenvectors $\mathbf{m}$ and $\mathbf{n}$ corresponding to the dominant outlier. Analogous to the uncorrelated case in the networks with independent connections, the individual entries of these vectors are accurately predicted by perturbation theory and exhibit Gaussian-mixture statistics (Fig 3D). Unlike in the uncorrelated case, reciprocal correlations now induce correlations between Δm_i and Δn_i (Fig 3C). Indeed, perturbation theory predicts that the first-order effects $\Delta\mathbf{m}$ and $\Delta\mathbf{n}$ of the random connectivity on $\mathbf{m}$ and $\mathbf{n}$ are respectively determined by $\mathbf{Z}$ and its transpose $\mathbf{Z}^\intercal$ (Eq. (9)). Reciprocal correlations between z_{ij} and z_{ji} directly lead to correlations between $\mathbf{Z}$ and $\mathbf{Z}^\intercal$ and therefore a non-zero covariance σ_{nm} between elements of $\mathbf{m}$ and $\mathbf{n}$, that can be predicted by mean field theory (Fig 3E, Eq. (70)).
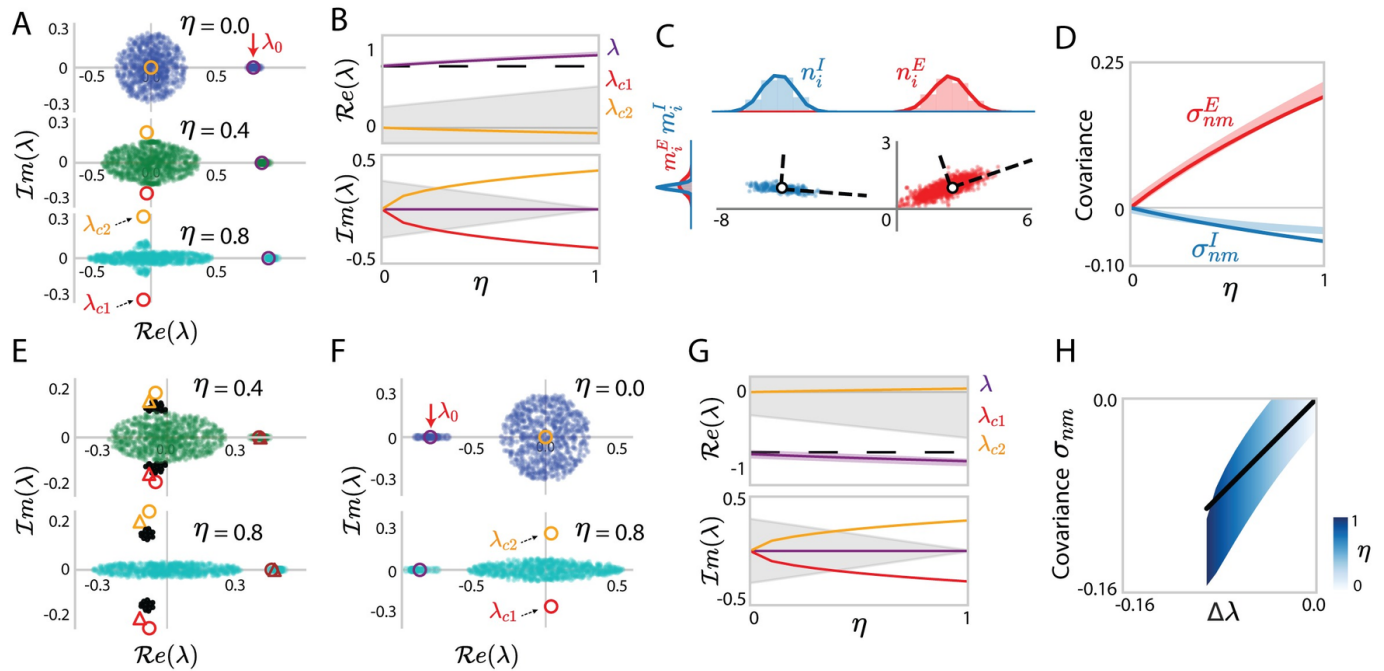
Fig 4. Eigenvalues and dominant eigenvectors for different network connectivity with reciprocal motifs. (A) Eigenvalue spectra of excitatory-inhibitory connectivity matrices $\mathbf{J}$, with homogeneous reciprocal correlations η but cell-type-dependent variance parameters $g_{EE} : g_{EI} : g_{IE} : g_{II} = 1.0 : 0.5 : 0.2 : 0.8$ and $g_{EE} = 0.3$. Different colours correspond to networks with different η. The elliptical bulk shows 600 randomly sampled eigenvalues for one realization of the random connectivity. Outlying eigenvalues are shown for 30 realizations of the random connectivity. The red arrow on the top points to the unperturbed eigenvalue λ_0 of the mean connectivity $\bar{\mathbf{J}}$. Coloured circles are the eigenvalues predicted using determinant lemma. (B) Comparison of the eigenvalues from the finite-size simulation with the predictions of the determinant lemma as the reciprocal correlation η is increased. The coloured solid lines show the roots of the third-order polynomial in Eq. (7). The light purple area indicates the empirical distribution of the dominant outlier, while the black dashed line is the unperturbed eigenvalue λ_0. The grey areas represent the eigenvalue bulk. (C) Scatter plot showing for each neuron i its entry n_i on the left eigenvector against its entry m_i on the right eigenvector, with the eigenvectors corresponding to eigenvalue outliers λ that deviate from λ_0. Red and blue colours represent respectively excitatory and inhibitory neurons. The white dots and the dashed lines respectively indicate the means and covariances for each population. (D) Comparison between the population covariance σ_{nm}^p of the entries on the left and right connectivity eigenvectors to different populations (coloured areas) and the predictions of perturbation theory (coloured lines, Eq. (95)). Other network parameters for (A-D): $J_E = 2.0$, $J_I = 1.2$, $N_E = 4N_I = 1200$. (E) Same as (A) for a connectivity matrix with homogeneous reciprocal correlations η, the coloured circles are solutions considering k up to 2 and the coloured triangles are solutions considering k up to 4. Other network parameters: $J_E = 1.5$, $J_I = 1.2$, $g_{EE} : g_{EI} : g_{IE} : g_{II} = 1.0, 0.5, 0.2, 0.8$, $g_{EE} = 0.2$ and $N_E = 4N_I = 1200$. (F, G) Same as (A, B) for an inhibition dominates connectivity matrix where $J_I = 2.0$, $J_E = 1.2$, with homogeneous reciprocity η and variance parameters $g = 0.3$. (H) Comparison of the overall covariance σ_{nm} with the deviation $\Delta\lambda$ of the dominant outlying eigenvalue from the unperturbed value λ_0.

When the network has both homogeneous variance parameters and correlation parameters, the excitatory and inhibitory populations have the same covariance $\sigma_{nm}^E = \sigma_{nm}^I = (J_E - J_I)g^2\eta/\lambda^2$ (Figs 3C, E, Eq. (95)). If the synaptic variances g_{pq} differ across populations, the covariances σ_{nm} are different for excitatory and inhibitory populations even if the reciprocal correlations are uniform (Figs 4C, D). The strength of the overall covariance reflects the strength of the additional feedback loop due to reciprocal correlations, and is therefore directly related to the deviations of the outlying eigenvalue from the uncorrelated value λ_0 (Fig 3F).

These results directly extend to networks with heterogeneous reciprocal correlations η_{pq}, $p, q = E, I$. In particular, in this case finite-size simulations confirm the presence of additional, complex conjugate outliers predicted by the cubic term in Eq. (7) (coloured circles containing outlier scatters at conjugate positions in Fig 3M). Moreover, the covariances σ_{nm}^p between the entries of low-rank connectivity vectors in this case differ between the excitatory and inhibitory population (Figs 3I, K, O, Q).

1.5 Approximating low-dimensional dynamics for locally-defined connectivity.

In previous sections, we developed a rank-one approximation of locally-defined excitatory-inhibitory connectivity. Here we use this approximation to describe the resulting low-dimensional dynamics. We consider networks of rate units, where the activation x_i of unit i obeys

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1}^{N} J_{ij}\phi(x_j(t)). \tag{13}$$

Here $\phi(x) = 1 + \tanh(x - \theta)$ is a positive transfer function, and for simplicity, we focus on autonomous dynamics without external inputs. We start from a locally-defined excitatory-inhibitory connectivity matrix, and compare the resulting activity with the theoretical predictions of our rank-one approximation, for which the dynamics are low-dimensional and analytically tractable. We first summarize the theoretical predictions for those dynamics, and then examine the specific cases of independent and reciprocally-correlated connectivity.

Recent works have showed that in networks with a rank R connectivity matrix, the trajectories $\mathbf{x}(t) = \{x_i(t)\}_{i=1...N}$ are confined to a low-dimensional subspace of the $N-$dimensional space describing the activity of all units [30–33]. In absence of external inputs, this subspace is R-dimensional and spanned by the set of connectivity eigenvectors $\mathbf{m}^{(r)}$ for $r = 1 \ldots R$, so that the trajectories can be parametrized as $\mathbf{x} = \sum_{r=1}^{R} \kappa_r \mathbf{m}^{(r)}$ where κ_r is a collective latent variable representing activity along $\mathbf{m}^{(r)}$. For a rank-one ($R = 1$) connectivity corresponding to an approximation of our locally-defined E-I network, the dynamics can therefore be represented by a single latent variable κ, so that the activation of unit i is given by

$$x_i(t) = \kappa m_i \tag{14}$$
$$= \kappa \bar{m}_i + \kappa \Delta m_i, \tag{15}$$

where we inserted the expression for m_i obtained from a first-order perturbation (Eq. (8), (9)). Note that since $\bar{m}_i = 1$, the first term in the r. h. s. of Eq. (15) corresponds to population-averaged activity, while the second term is the deviation of the activity of unit i from the population average. Inserting the values for Δm_i obtained using specific realizations of random connectivity in Eq. (9), the rank-one approximation provides predictions for the activity of single units in *individual instances* of locally-defined networks. Moreover, the rank-one theory predicts that both the

population-average and the standard deviation of activations in the network are proportional to κ.

The values taken by the latent variable κ can be determined by projecting Eq. (13) onto $\mathbf{m}$ and inserting Eq. (15) (Methods Secs. 2.6.1, 2.6.2). This leads to a closed equation for the dynamics of $\kappa(t)$:

$$\dot{\kappa} = -\kappa + \sum_{p=E,I} \alpha_p \Big(\bar{n}^p \langle \phi(\kappa \bar{m}^p, \kappa^2 \sigma_{m^p}^2) \rangle + \langle \phi'(\kappa \bar{m}^p, \kappa^2 \sigma_{m^p}^2) \rangle \sigma_{nm}^p \kappa \Big), \tag{16}$$

where the brackets denote a Gaussian average (see Eq. (110)). The steady state then obeys

$$\kappa = F_{mean}(\kappa) + F_{cov}(\kappa) \tag{17}$$

where

$$\begin{aligned} F_{mean}(\kappa) &= \sum_{p=E,I} \alpha_p \bar{n}^p \langle \phi(\kappa \bar{m}^p, \kappa^2 \sigma_{m^p}^2) \rangle, \\ F_{cov}(\kappa) &= \sum_{p=E,I} \alpha_p \langle \phi'(\kappa \bar{m}^p, \kappa^2 \sigma_{m^p}^2) \rangle \sigma_{nm}^p \kappa. \end{aligned} \tag{18}$$

The two terms in the r. h. s. of Eq. (17) show that the contributions of recurrent synaptic inputs to the latent dynamics κ come from two sources: (i) the population means of the left and right connectivity eigenvectors $\bar{n}^p$ and $\bar{m}^p$ that contribute to $F_{mean}(\kappa)$ (Eqs. (52)-(54)); (ii) the covariance σ_{nm}^p between the left and right connectivity eigenvectors that contributes only to $F_{cov}(\kappa)$. In the low-rank approximation of the locally-defined E-I connectivity, these two terms have distinct origins: the mean comes from the independent components of the connectivity (Eqs. (53), (54)); while the covariance comes from reciprocal correlations between connections (Eqs. (94), (95)). We therefore next examine separately the effects on dynamics of these two connectivity components.

1.5.1 Independently generated local connectivity

When synaptic connections are generated independently from a Gaussian distributions based on the identities of pre- and post-synaptic populations, the rank-one approximation of connectivity leads to uncorrelated left and right connectivity vectors $\mathbf{n}$ and $\mathbf{m}$, so that $\sigma_{nm}^p = 0$ for $p = E, I$. In consequence, only the first term is present in the r. h. s. of Eq. (17), and the fixed point of the latent dynamics is given by a difference between excitatory and an inhibitory feedback (Eq. (114)):

$$\kappa = J_E \langle \phi(\kappa, \kappa^2 \sigma_{m^E}^2) \rangle - J_I \langle \phi(\kappa, \kappa^2 \sigma_{m^I}^2) \rangle. \tag{19}$$

As long as the mean inhibition J_I is strong enough to balance the mean excitation J_E, Eq. (19) predicts a single fixed point. As J_E is increased, positive feedback begins to dominate and leads to a bifurcation to a bistable regime for the latent dynamic variable κ (Figs 5A-C, S2 Appendix).

This bistability due to positive feedback is expected on the basis of mean connectivity alone. Indeed replacing the connectivity matrix by its mean $\bar{\mathbf{J}}$ is equivalent to a rank-one approximation with $\mathbf{m} = \bar{\mathbf{m}}$ and $\mathbf{n} = \bar{\mathbf{n}}$ which lead to Eq. (113) with $\sigma_{m^p}^2 = 0$ for $p = E, I$. The additional first-order perturbation term in the rank-one approximation (Eqs. (8), (9)) additionally takes into account fluctuations in the connectivity, which leads to a non-zero $\sigma_{m^p}^2$, and modifies the fixed points predicted by
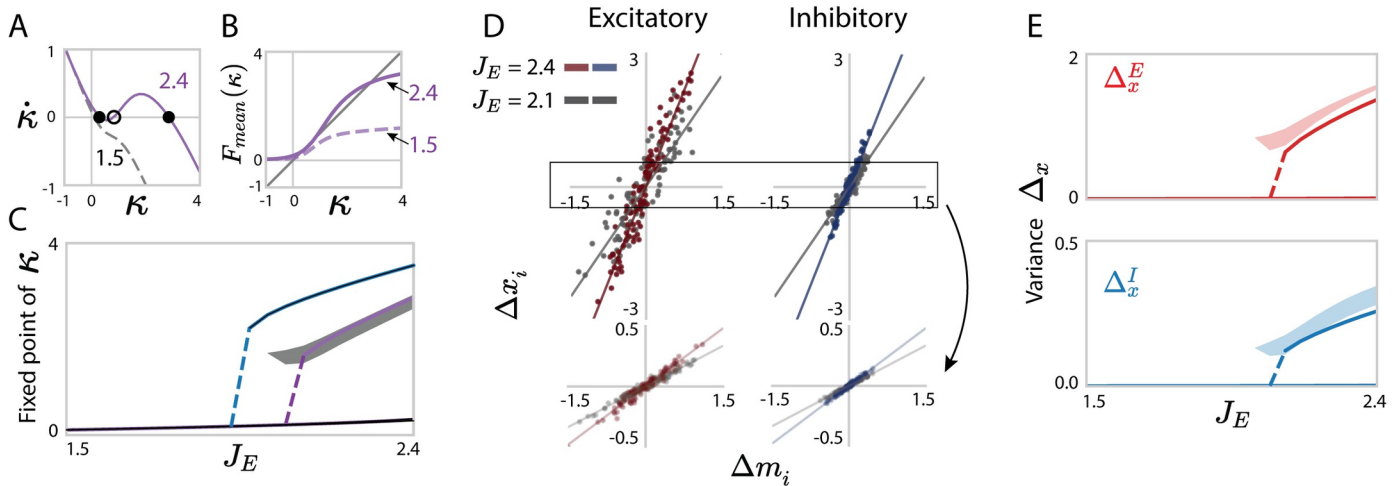
Fig 5. Predicting low-dimensional dynamics using a rank-one approximation of networks with independent Gaussian connectivity. (A) Fixed points of the latent variable κ in the rank-one approximation. The lines show the dynamics $\dot{\kappa}$ as function of κ, predicted by Eq. (16) (solid line: $J_E = 2.4$; dashed line: $J_E = 1.5$). The intersections with $y = 0$ correspond to fixed points (filled dots: stable; unfilled dot: unstable). (B) Contribution of mean connectivity to the latent dynamics, $F_{mean}(\kappa)$ in Eq. (18), for two values of J_E. (C) Bifurcation diagram for increasing J_E: analytical predictions of Eq. (19) compared with simulations of the full network with locally-defined connectivity. Blue line: analytical prediction including only the mean part of the connectivity ($\sigma^2_{m^p} = 0$ in Eq. (19)); purple line: analytical prediction including the first-order perturbation term in the rank-one approximation; gray: projection of simulated activity $\mathbf{x}$ onto the connectivity vector $\mathbf{m}$ computed by perturbation theory (Eq. (102)). (D) Comparison between predictions and simulations for the activity of individual units in a given realization of the random connectivity. For each unit i, a dot shows the deviation Δx_i of its steady-state activity from the population average, against its value Δm_i of the perturbed part of the connectivity vector $\mathbf{m}$ (Eq. (15)). The low-rank theory predicts $\Delta x_i = \kappa \Delta m_i$. Red and blue scatters show excitatory or inhibitory populations, each for several values of J_E. Lines represent $y = \kappa x$, where κ is obtained from Eq. (19). Upper panels show the result in a realization with a high fixed point, bottom panels show the result in a realization with a low fixed point. (E) Comparison between the predictions (solid lines) and simulations (shaded areas) for the population-averaged variances of Δx_i. Shaded areas show mean±std. Network parameters: $N_E = 4N_I = 1200$, $J_I = 0.6$, $g_{EE} : g_{EI} : g_{IE} : g_{II} = 1.0 : 0.5 : 0.2 : 0.8$ and $g_{EE} = 0.8$. The transfer function ϕ has parameter $\theta = 1.5$.

Eq. (19). In consequence, the bifurcation to bistability takes place at higher values of J_E than predicted from mean connectivity alone (purple lines compared to blue lines in Fig 5C).

More importantly, we find that the first-order perturbation in the rank-one approximation accurately predicts the firing rates of individual neurons for specific instances of the random, locally-defined connectivity (Fig 5D), and therefore the variance of the steady state of population dynamics Δ^p_x. In particular, cell-type dependent variances in the synaptic connectivity, lead to distinct variances Δ^E_x and Δ^I_x for excitatory and inhibitory populations (Eqs. (88), (129), Fig 5E).

Note that the independently generated local connectivity can be treated analytically without resorting to a rank-one approximation, by using a different variant of mean-field theory originally developed for randomly connected networks. [30, 52]. That theory is not perturbative, and takes into account an additional term in the variance (see Methods Sec. 2.6.3 and S3 Appendix for more details). However, in contrast to the rank-one approximation, it does not predict the activity of individual neurons, and is challenging to extend beyond independent random connectivity.
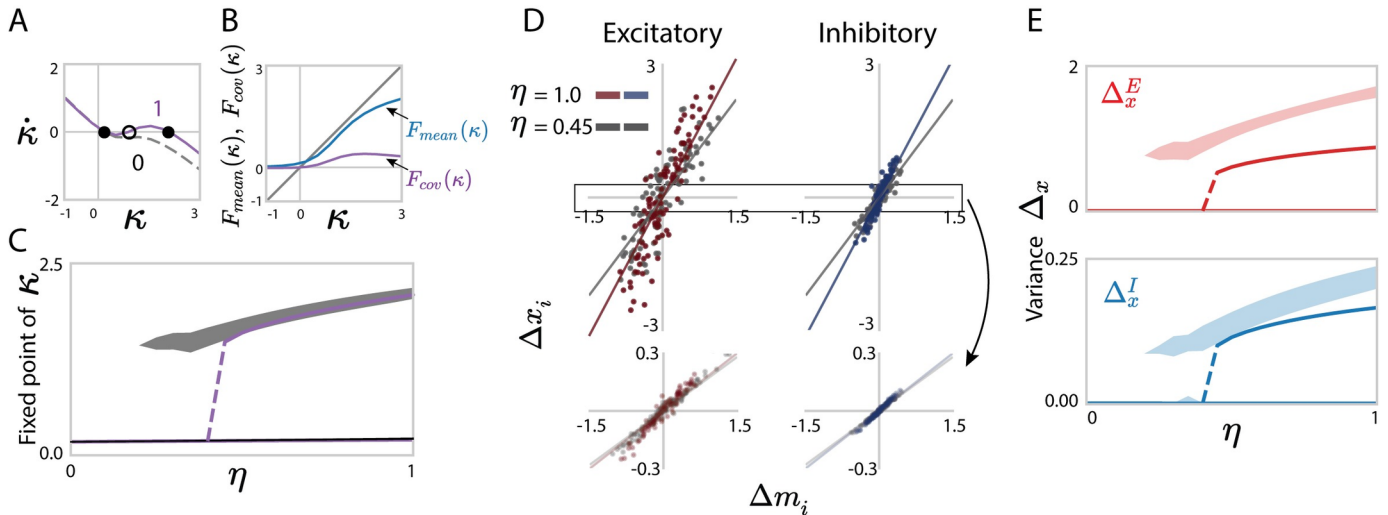
Fig 6. Predicting low-dimensional dynamics using a rank-one approximation of networks with homogeneous reciprocal motifs. (A) Influence of reciprocal correlations on fixed points of the latent variable κ in the rank-one approximation. The lines show the dynamics $\dot{\kappa}$ as function of κ, predicted by Eq. (16) (solid line: $\eta = 1$; dashed line: $\eta = 0$). The intersections with $y = 0$ correspond to fixed points (filled dots: stable; unfilled dot: unstable). (B) Comparison of the contributions of mean connectivity $F_{mean}(\kappa)$ and covariance $F_{cov}(\kappa)$ to the latent dynamics of κ (Eq. (18)) for $\eta = 1$. (C) Bifurcation diagram for increasing η at fixed J_E. Solid purple lines: analytical predictions of Eqs. (17), (18); gray areas: projection of simulated activity $\mathbf{x}$ onto the connectivity vector $\mathbf{m}$ computed by perturbation theory (Eq. (102)). (D) Comparison between predictions and simulations for the activity of individual units in a given realization of the random connectivity. For each unit i, a dot shows the deviation Δx_i of its steady-state activity from the population average, against its value Δm_i of the perturbed part of the connectivity vector $\mathbf{m}$ (Eq. (15)). The low-rank theory predicts $\Delta x_i = \kappa \Delta m_i$. Red and blue scatters show excitatory or inhibitory populations, each for two values of η. Lines represent $y = \kappa x$, where κ is obtained from Eqs. (17), (18). Upper panels show the result in a realization with a high fixed point, bottom panels show the result in a realization with a low fixed point. (E) Comparison between the predictions (solid lines, Eq. (133)) and simulations (shaded areas) for the population-averaged variances of Δx_i. Shaded areas show mean±std. Network parameters: $N_E = 4N_I = 1200$, $J_E = 1.9$, $J_I = 0.6$, $g_{EE} : g_{EI} : g_{IE} : g_{II} = 1.0 : 0.5 : 0.2 : 0.8$ and $g_{EE} = 0.8$. The transfer function ϕ has parameter $\theta = 1.5$.

1.5.2 Reciprocal motifs.

We next turn to the predictions of the rank-one approximation for dynamics resulting from locally-defined connectivity with reciprocal motifs. In this case, the additional reciprocal correlations in the random part of the connectivity lead to a non-zero covariance σ_{nm} between the connectivity vectors $\mathbf{n}$ and $\mathbf{m}$ in the rank-one approximation (Eqs. (90), (95)). This covariance in turn generates an additional feedback component in the dynamics of the latent variable, the second term in the r. h. s. of Eq. (17).

Specifically, positive reciprocal correlations combined with excitation-dominated connectivity enhance positive feedback with respect to mean connectivity alone. As a result, progressively increasing the reciprocal correlations can therefore induce a bifurcation to bistability, even if the mean excitation is not sufficient by itself to support two stable states (Figs 6A-C). This is a major novel effect of reciprocal motifs on collective dynamics. As in the case of independent connectivity, we moreover found that the perturbative term in the rank-one approximation correctly predicts the activity of individual neurons in specific realizations of the connectivity (Fig 6D), and therefore also the cell-type dependent variances of activity (Fig 6E).
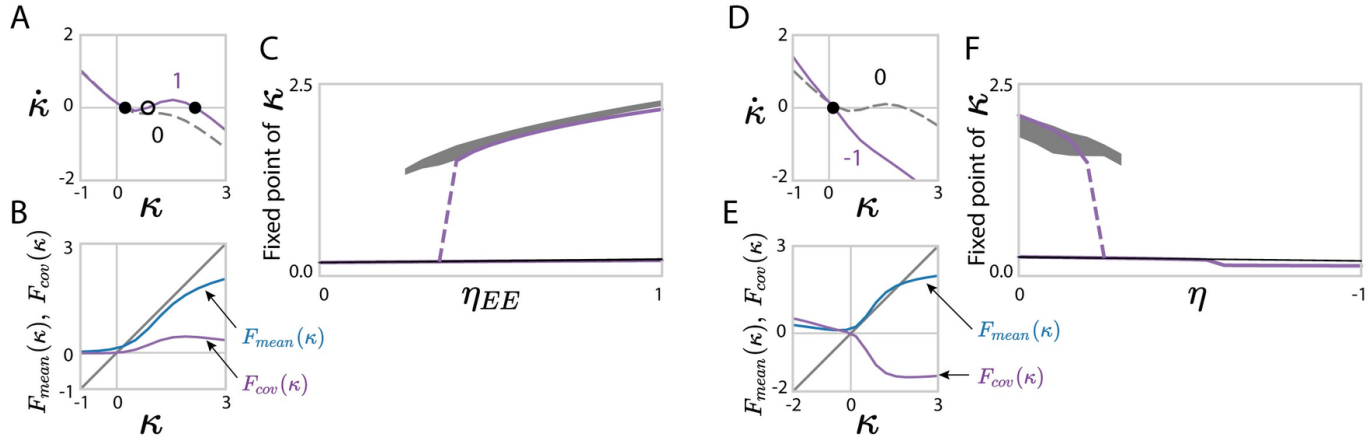
Fig 7. Predictions for low-dimensional dynamics using a rank-one approximation of networks with non-homogeneous and anti-symmetric reciprocal motifs. (A-D) Fixed points of latent dynamics in networks with heterogeneous, cell-type-dependent reciprocal correlations: $\eta_{EE} = \eta_{EI} = -\eta_{II}$. Network parameters: $N_E = 4N_I = 1200$, $g_{EE} : g_{EI} : g_{IE} : g_{II} = 1.0 : 0.5 : 0.2 : 0.8$, $g_{EE} = 0.8$, $J_E = 1.9$ and $J_I = 0.6$. (E-H) Fixed points of latent dynamics in networks with homogeneously anti-symmetric motifs, $\eta_{pq} = \eta \in [-1, 0]$. Network parameters: $N_E = 4N_I = 1200$, $g_{EE} : g_{EI} : g_{IE} : g_{II} = 1.0 : 0.5 : 0.2 : 0.8$, $g_{EE} = 0.8$, $J_E = 2.2$ and $J_I = 0.6$. In C and F, gray areas show projections of simulated activity $\mathbf{x}$ onto the connectivity vector $\mathbf{m}$ computed by perturbation theory (Eq. (102)), shaded areas show mean±std.

More generally, our rank-one approximation allows us to describe the latent dynamics when the degree of reciprocal correlation depends across the pre- and post-synaptic populations (Results Sec. 1.4). Such heterogeneity in reciprocal correlations can enhance different types of feedback. For example, antisymmetric connectivity within inhibitory populations ($\eta_{II} < 0$) disinhibits excitatory population and thus facilitates bistable transitions (Figs 7A-C) compared to networks with homogeneous reciprocal correlations. In contrast, excitation-dominated connectivity with homogeneous negative reciprocity ($\eta < 0$) generate negative feedback and therefore suppress the global dynamics from bistable state to quiescent (Figs 7D-F).

Importantly, describing the role of reciprocal correlations on latent dynamics relies on our global low-rank approximation of locally-defined connectivity. In particular, the effects of such correlations cannot be captured by considering only mean connectivity and population-averaged activity (first term in the r. h. s. of Eq. (17)). Moreover, including reciprocal correlations in classical mean-field approaches to randomly connected networks is technically challenging [18].

1.6 Extension: E-I networks with sparse connectivity.

In previous sections, we examined locally-defined connectivity generated using Gaussian distributions of individual synaptic weights (function f^{pq} in Eqs. (1)). Our results for the low-rank approximation of locally-defined connectivity are however independent of the precise form of the distribution f^{pq}. In particular, our finding that the resulting low-rank structure obeys Gaussian-mixture statistics is universal, in the sense that it is valid for any distribution f^{pq} for which the central limit theorem holds (see Discussion). To illustrate this universality, here we turn to networks with sparse connectivity, generated from Bernoulli distributions f^{pq} taking values 0 and A_q (where $A_q = A_E, A_I$ refer to strengths of excitatory and inhibitory connections), with a uniform fraction c of
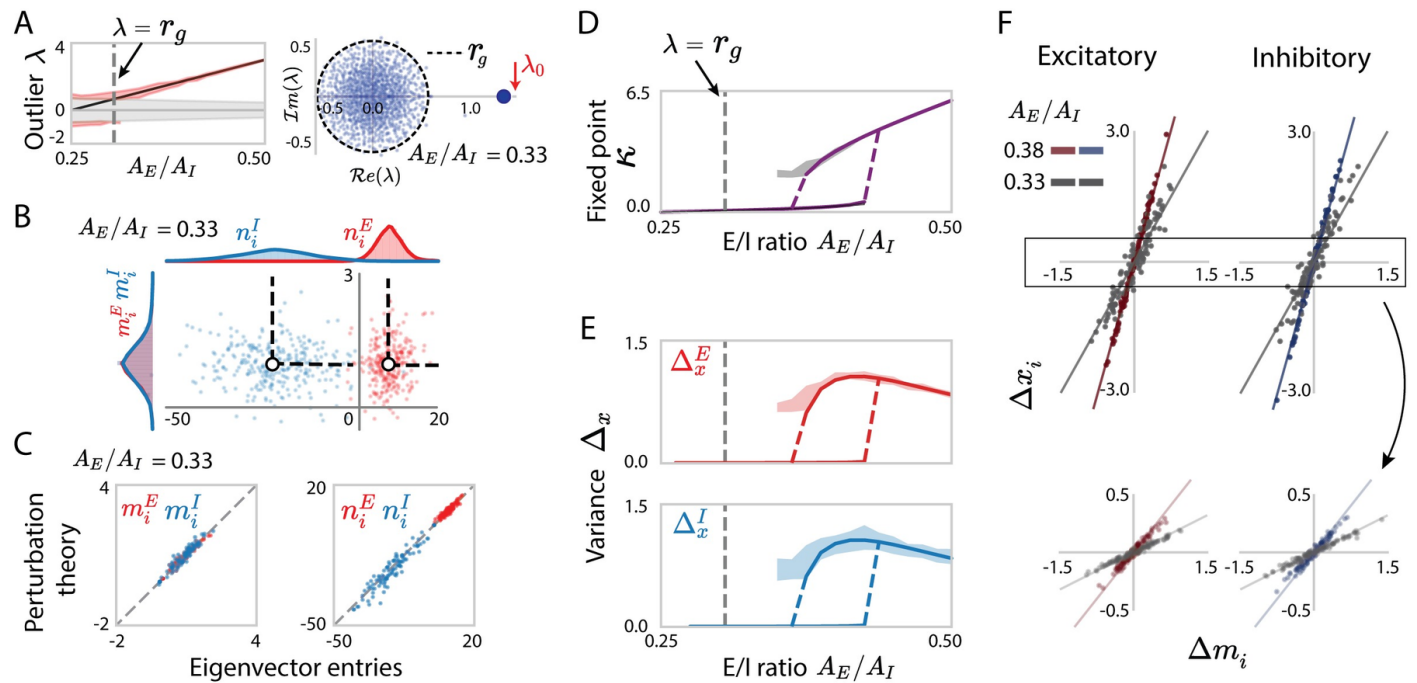
Fig 8. Rank-one approximation and predicted low-dimensional dynamics for sparse excitatory-inhibitory networks. (A) Left: Comparison of the predicted eigenvalue outlier $\lambda_0 = c(N_E A_E + N_I A_I)$ (black line) with finite-size simulations (red area shows mean±std for 30 realizations). The gray area represents the area covered by the eigenvalue bulk. Right: example spectrum of one realization of connectivity matrix with E/I ratio $A_E/A_I = 0.33$, and the radius r_g of the eigenvalue bulk computed from the statistically equivalent Gaussian connectivity (see S4 Appendix). (B) Scatter plot showing for each neuron i its entry n_i on the left eigenvector against its entry m_i on the right eigenvector. Red and blue colours represent respectively excitatory and inhibitory neurons. The white dots and the dash lines respectively indicate the means and covariances for each population obtained from simulations. For visualization purposes, the $x-$ and y-axis are scaled unequally. (C) Comparison between eigenvector entries obtained from direct eigen-decomposition of **J** with predictions of perturbation theory (Eqs. (8), (9)). (D) Bifurcation diagram for increasing the ratio A_E/A_I: analytical predictions of Eq. (115) compared with simulations of the full network with locally generated sparse connectivity. Purple line: analytical prediction including the first-order perturbation term in the rank-one approximation; gray: projection of simulated activity onto the connectivity vector **m** computed by perturbation theory Eq. (102). (E) Comparison between the predictions (solid lines) and simulations (shaded areas) for the population-averaged variances of Δx_i. Shaded areas show mean±std. (F) Comparison between predictions and simulations for the activity of individual units in a given realization of the sparse connectivity. For each unit i, a dot shows the deviation Δx_i of the steady-state activity from the population average, against the corresponding value Δm_i of the perturbed part of the connectivity vector **m** (Eq. (15)). The low-rank theory predicts $\Delta x_i = \kappa \Delta m_i$. Red and blue scatters show excitatory or inhibitory populations, each for two values of the ratio A_E/A_I. Lines represent $y = \kappa x$, where κ is obtained from Eqs. (17), (18). Upper panels show the result in a realization with a high fixed point, bottom panels show the result in a realization with a low fixed point. The gray vertical dashed line in A left, D, E correspond to the critical point c at which the absolute value of the outlier is equal to the radius of the eigenvalue bulk. Network parameters: $N_E = 4N_I = 800$, $c = 0.3$, $A_E = 0.025$. The transfer function ϕ has parameter $\theta = 1.5$.

non-zero connections (see Eq. (26)). The statistics of the corresponding rank-one approximation are fully determined by the mean, variance and covariance of synaptic weights in the excitatory and inhibitory populations. Here we compare the predictions of our perturbative approximation with direct simulations of full-rank networks with locally-defined connectivity. We first consider independently generated connectivity, and then turn to reciprocal motifs.

1.6.1 Independently generated sparse connectivity.

For networks with independently generated sparse connectivity, the mean and variance of individual synaptic weights are given by cA_q and $c(1-c)A_q^2$, where $q = E, I$ refers to the population of the presynaptic neuron. Previous works have shown that such sparse networks have a random, circularly distributed bulk of random eigenvalues with a radius r_g determined by the overall variance of synaptic weights (Eq. (148), Fig 8A) [53,54], as expected from the universality theorem for random matrices [49]. The overall mean of the synaptic weights instead determines the mean outlying eigenvalue (Fig 8A). In sparse networks, the main novelty with respect to the Gaussian case is that the mean and variance of synaptic weights are not independent parameters, but are instead both set by the synaptic couplings A_E and A_I as well as the network's sparsity c. In consequence, varying these couplings changes both the radius of the bulk and the outlying eigenvalue, and can lead to intersections where the outliers dip into the bulk (see S4 Appendix for details).

As long as the outlier lies outside of the eigenvalue bulk, our perturbative approximation Eq. (67) predicts well the individual entries of the right- and left-eigenvectors corresponding to the outlier from the fluctuation matrix $\mathbf{Z}$ (Fig 8C). As expected, the resulting statistics of eigenvectors entries are well described by a Gaussian-mixture distribution with parameters fully determined by the mean and variance of the synaptic weights (Fig 8B).

Our predictions for the low-dimensional dynamics based on the rank-one approximation therefore directly extend to sparse networks. Comparing with direct simulations, we found that Eq. (115) predicts well the global latent variable κ obtained by projecting the activity $\mathbf{x}$ onto the approximated rank-one eigenvector $\mathbf{m}$ (Eq. (102)). As the E/I ratio is increased, positive feedback increases, and the latent variable κ undergoes a transition from a single fixed point to two bistable states (Fig 8D).

For individual realizations of the sparse connectivity, the rank-one approximation $\mathbf{x} = \kappa\mathbf{m}$ predicts well the activation x_i of individual neurons in the simulations (Fig 8F). From the statistics of the right eigenvector $\mathbf{m}$ (Eq. (89)), our analysis moreover predicts the heterogeneity of activity in terms of population-averaged variances $\Delta_x^{E/I}$ (Fig 8E). This heterogeneity is identical in excitatory and inhibitory populations, as their right eigenvectors have identical fluctuations (Figs 8C, E).

1.6.2 Sparse EI networks with reciprocal motifs.

For sparse E-I networks, we generate reciprocal motifs by introducing a fraction ρ_{pq} of reciprocally connected pairs of neurons. Together with the sparsity c and the synaptic strengths, the parameter ρ_{pq} determines the cell-type dependent reciprocal correlation η_{pq} (Methods Sec. 2.1.2 Eqs. (32), (37), Fig 9A).

We first examine the effect of the reciprocal motifs on the statistical properties of the eigenvalues and eigenvectors. The spectrum still consists of continuous eigenvalues and a discrete outlier. Since in the independent case the outlier depends only on $A_{E/I}$ and sparsity c, here we fix these two variables but increase $\eta_{EE} = -\eta_{EI}$ while keeping η_{II} constant. As in the case of dense, Gaussian networks, the outlier increases with the increasing reciprocal correlation and deviates from the outlier of the corresponding
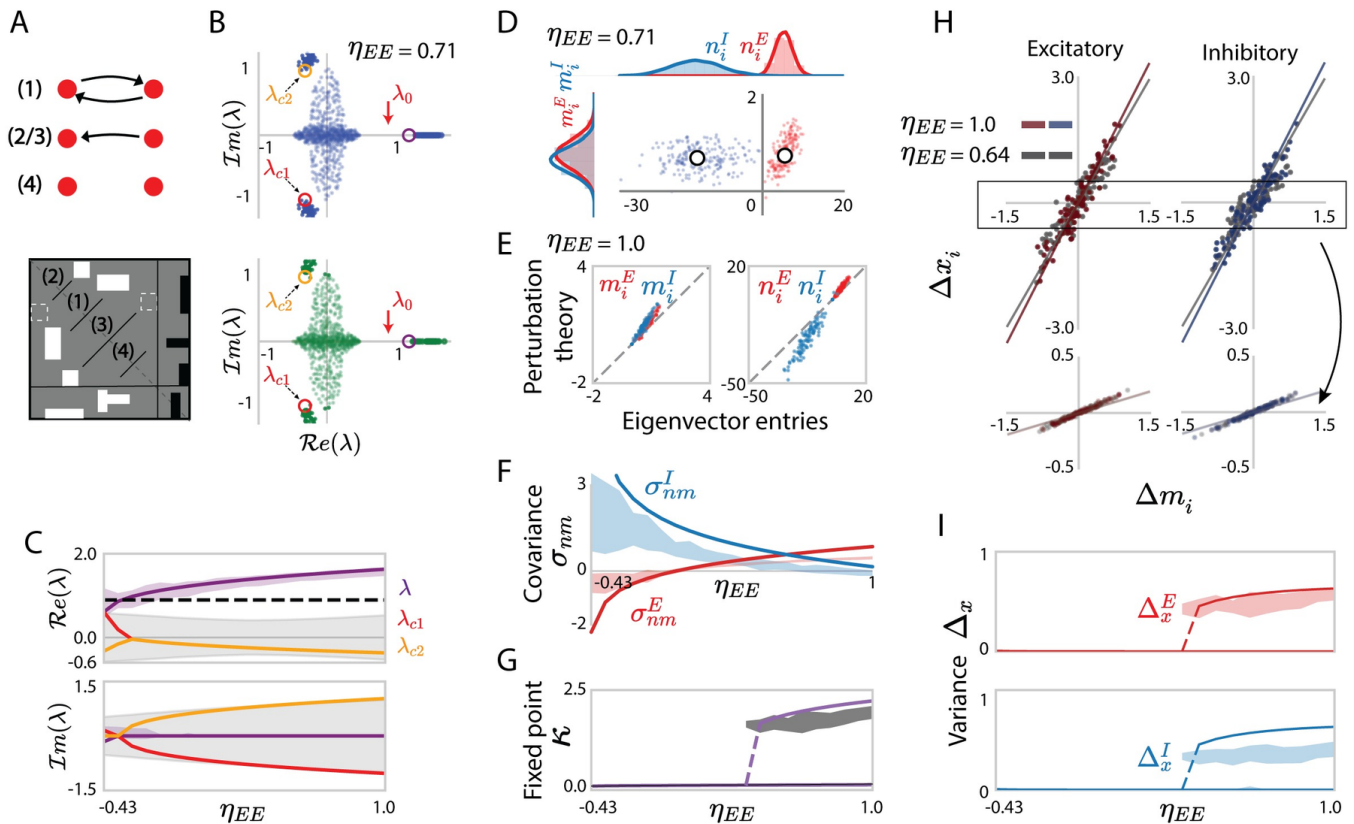
Fig 9. Characterizing connectivity statistical properties and low-dimensional dynamics for the sparse network with reciprocal motifs. (A) Schematics of a sparse EI network with four forms of paired connections. White and black rectangles represent the non-zero excitatory and inhibitory sparse connections. (B) Eigenvalue spectrum of the sparse connectivity (upper panel) and from the equivalent Gaussian connectivity (bottom panel) with reciprocal motifs. Cell-type-dependent reciprocal correlations are $\eta_{EE} = 0.71$, $\eta_{EI} = -0.71$, $\eta_{II} = -0.43$ in both connectivity matrices, continuous eigenvalue bulks show eigenvalues for one realization of the network connectivity. Red arrows point to the unperturbed eigenvalue λ_0. Outlying eigenvalues are shown for 30 realizations of the network connectivity. Coloured circles are the eigenvalues predicted using determinant lemma. (C) Comparison of the eigenvalues from the finite-size simulation of the sparse connectivity, with the predictions of the determinant lemma as progressively increasing the reciprocal correlation η_{EE} $(-\eta_{EI})$. The coloured solid lines show the roots of the third-order polynomial in Eq. (7). The purple area indicates the empirical distribution of the dominant outlier, while the black dashed line is the eigenvalue λ_0 of the corresponding independent sparse connectivity matrix (Eq. (76)). The gray areas correspond to the areas covered by the eigenvalue bulk. (D) Scatter plot showing for each neuron i its entry n_i on the left eigenvector against its entry m_i on the right eigenvector. Red and blue colours represent respectively excitatory and inhibitory neurons. The white dots indicate the means for each population obtained from simulations. For visualization purposes, the x- and y-axis are scaled unequally. (E) Comparison between eigenvector entries obtained from direct eigen-decomposition of $\mathbf{J}$ with predictions of perturbation theory (Eqs. (8), (9)). (F) Comparison between the population covariance σ_{nm}^p of the entries on the left and right connectivity eigenvectors to different populations (coloured areas) and the predictions of perturbation theory (coloured lines, Eq. (97)). (G) Bifurcation diagram for increasing the reciprocal correlation η_{EE} $(-\eta_{EI})$: analytical predictions of Eq. (117) compared with simulations of the full network with locally generated sparse connectivity and reciprocal motifs. Purple line: analytical prediction including the first-order perturbation term in the rank-one approximation; gray: projection of simulated activity onto the connectivity vector $\mathbf{m}$ computed by perturbation theory Eq. (102). (H) Comparison between predictions and simulations for the activity of individual units in a given realization of the sparse connectivity. For each unit i, a dot show the deviation Δx_i of the steady-state activity from the population average, against its value Δm_i of the perturbed part of the

Fig 9. (Continued from previous page.) connectivity vector $\mathbf{m}$ (Eq. (15)). The low-rank theory predicts $\Delta x_i = \kappa \Delta m_i$. Red and blue scatters show excitatory or inhibitory populations, eachfor two values of the reciprocal correlation η_{EE}. Lines represent $y = \kappa x$, where κ is obtained from Eqs. (17)-(18). Upper panels show the result in a realization with a high fixed point, bottom panels show the result in a realization with a low fixed point. (I) Comparison between the predictions (solid lines) and simulations (shaded areas) from the population-averaged variances of Δx_i, shaded areas show mean±std. In C, F, G, I, the reciprocal correlations $\eta_{EE} = -\eta_{EI}$ progressively increase from -0.43 to 1.0 while keeping $\eta_{II} = -0.43$ constant ($\rho_{EE} = \rho_{EI}$ increase from 0 to 1 and $\rho_{II} = 0$ is fixed). Network parameters: $N_E = 4N_I = 800$, $c = 0.3$, $A_E = 0.023$, $A_E/A_I = 0.3$. The transfer function ϕ has parameter $\theta = 1.5$.

independent sparse connectivity matrix (Figs 9B, C). Moreover, in the large network limit, we find that if means, variances, and reciprocal correlations are identical, dense Gaussian connectivity leads to the same eigenvalue spectrum as the sparse connectivity (Methods Secs. 2.1.2, 2.1.3 Eqs. (36), (37)). We furthermore mathematically predict two additional conjugate eigenvalue outliers generated by the reciprocal connections in the sparse case (Eqs. (80), (86), (87), Fig 9B).

As for uncorrelated connectivity, perturbation theory predicts the individual left and right eigenvector entries, which altogether follow Gaussian statistics as expected (Figs 9D, E). Importantly, reciprocal correlations induce a non-zero covariance σ_{nm} between the entries m_i and n_i of the right and left eigenvectors (Eq. (97), Fig 9F).

Finally, we examine the population dynamics in the sparse network with reciprocal motifs using the low-rank approximation derived above. The reciprocal motifs in the example network generate an overall positive feedback. Therefore, gradually increasing the reciprocal correlation η_{EE} ($-\eta_{EI}$) in the example network induces a bifurcation into bistability (Eq. (117), Fig 9G). Analogous to the predictions of individual eigenvector entries, the low-rank approximation gives analytical predictions for the activity of individual neurons in specific connectivity realizations (Fig 9H), and hence the cell-type dependent variances of neuronal activation $\Delta_x^{E/I}$ (Fig 9I) obtained from finite-size simulations of the original sparse networks.

Discussion

In this work, we unified two different descriptions of connectivity in multi-population networks and thereby connected two broad classes of models. Starting from local statistics of synaptic weights, we approximated the resulting connectivity matrix in terms of a low-rank structure. The obtained, approximate low-rank network model then allowed us to determine the influence of the local connectivity motifs on the global low-dimensional dynamics.

A key ingredient in our approach is a low-rank approximation of the locally-defined connectivity matrix. Approximating an arbitrary full-rank matrix by a rank-R one is a classical problem in numerical analysis, for which a number of different methods are available depending on the objective of approximation [55]. The most common method is to perform a singular value decomposition (SVD), and keep the top R terms [56]. This method minimizes the Frobenious norm of the difference between the original matrix and its low-rank approximation. Our goal in this study was however to obtain a low-rank approximation that preserves the dominant eigenvalues of the original matrix, as these eigenvalues determine the autonomous dynamics in the network. An SVD-based approximation preserves the top singular values, but in general not the top eigenvalues. We therefore opted for an approximation based on truncated eigen-decomposition. When studying input-driven and transient dynamics, different methods for low-rank approximation may be more appropriate, and are a topic of active research [57–61]

To perform the eigen-decomposition of excitatory-inhibitory connectivity matrices,

we leveraged the fact that they can be expressed as a sum of a block-like deterministic low-rank matrix and a full-rank random matrix with zero-mean [40]. The eigen-spectrum of such matrices in general consists of a continuously-distributed bulk and discrete outliers [44,45]. While a number of works have examined the bulk of the eigenvalue spectrum for random matrices [11,18,32,44,49,52,62,63], the outliers, and in particular the corresponding eigenvectors have to our knowledge received less attention. The main technical novelty in this work is the use of matrix perturbation theory [47,64] to approximate the eigenvectors corresponding to the outliers in the eigenspectrum of the locally-defined connectivity matrices. A key output of this approach is the finding that entries of the left- and right-eigenvectors follow multivariate Gaussian distributions, the statistics of which depend on the population the neurons belong to. This result provides a general theoretical mapping from locally-defined multi-population models to Gaussian-mixture low-rank networks [31,33]. It however holds only as long as the distribution of synaptic weights satisfies the assumptions of the central limit theorem. This in particular excludes heavy tailed distributions often found in experimental studies [1,65].

In the networks we considered, the non-random structure in connectivity comes only from the multi-population organization. More specifically, the low-rank skeleton of the locally-defined connectivity matrix is fully specified by the mean synaptic weights between different populations (Eq. (3)). This mean connectivity structure largely controls the outlying eigenvalue, and the average values of the corresponding eigenvector entries. The random part of the connectivity and reciprocal motifs can modify the outlying eigenvalue, and add heterogeneity as well as correlations to this underlying structure. Our perturbative theory allows us to quantify these effects and predict dynamics on a single-neuron basis. This approach can be directly extended to networks with additional structure, in which the low-rank skeleton is not solely determined by the mean connectivity but possibly by more general patterns.

A key insight from our study is a general relationship between reciprocal motifs in locally-defined connectivity and overlaps among connectivity vectors in low-rank networks. Indeed, we have shown that correlations between reciprocal synaptic weights generate overlaps beyond the mean in the corresponding low-rank approximation (Eq. (93)). Conversely, zero-mean overlaps between connectivity vectors in a low-rank model necessarily imply non-vanishing reciprocal correlations (Eq. (41) and S5 Appendix). Since overlaps between connectivity vectors determine the autonomous recurrent dynamics in low-rank networks, this relationship allowed us to quantify how reciprocal connectivity motifs contribute to network dynamics.

Local statistics of synaptic connectivity are believed to play an important role in the global network dynamics [1,17,32,62]. Our study, provides a mathematical theory that relates the local connectivity statistics to global recurrent dynamics through a low-rank approximation. This computational framework is not restricted to the reciprocal motifs that we have emphasized in this work, but can be extended to various forms of local connectivity motifs [10]. As a result, our framework can be applied more broadly to study experimentally-obtained connectivity databases and connectivity consisting of different types of motifs. [6,7,15].

Table 1. List of notations.

Notation	Description
$i,\ j$	Single neuron indices
$p,\ q$	Neuron population indices
α_p	Fraction of neurons belonging to population p
$\bar{J}_{pq}$	Mean synaptic weights between populations p, q
$J_{E/I}$	Re-scaled mean excitatory/inhibitory synaptic weights
$\sigma_{z_{pq}}$	Standard deviation of the synaptic weights between populations p, q
g_{pq}	Re-scaled standard deviation of the synaptic weights between populations p, q
η_{pq}	Reciprocal correlation of connectivity weights between populations p, q
$A_{E/I}$	Excitatory/inhibitory synaptic weights in the sparse network
c	Homogeneous sparsity of the sparse network, $c_{pq} = c$
$\sigma_{m^p}^2,\ \sigma_{n^p}^2$	Variances of components on connectivity vectors $\mathbf{m}^p$ and $\mathbf{n}^p$
σ_{nm}^p	Covariance between connectivity vectors $\mathbf{m}^p$ and $\mathbf{n}^p$
$\mu_x^p,\ \Delta_x^p$	Population-averaged mean and variance of activation

2   Materials and methods

Throughout this study, we consider recurrent networks of N neurons and denote by $\mathbf{J}$ the recurrent connectivity matrix, where J_{ij} is the synaptic strength of the connection from neuron j to neuron i.

2.1   Locally-defined multi-population connectivity

In this section, we introduce a first class of connectivity models, in which the synaptic couplings are generated based on *local* statistics determined by the identity of pre- and post-synaptic neurons. The N neurons in the network are organized in P populations, where population p has N_p neurons. Denoting by p and q the populations neurons i and j belong to, the value of the synaptic coupling J_{ij} is drawn randomly from a distribution in which statistics depend on the pre- and postsynaptic population q and p. The full connectivity matrix $\mathbf{J}$ therefore has a block structure, in the sense that all connections within the same block share *identical* statistics.

We examine two variants of this model class: (1) ***independent*** random connectivity [52]; (2) connectivity with ***reciprocal motifs*** [10,66]. In each case, we examine two specific examples of distributions of synaptic strengths, *Gaussian*, and *Sparse* distributions.

2.1.1   Independent random connectivity

For networks with independent random connectivity, the recurrent connections J_{ij} are sampled independently for each (i,j) pair from

$$\mathrm{Prob}(J_{ij} = J) = f^{pq}(\mathbf{J}), \tag{20}$$

where f^{pq} denotes a probability density function, and q, p are the pre- and post-synaptic populations. Separating the mean and random components, for an arbitrary distribution Eq. (20) can be re-expressed as

$$J_{ij} = \bar{J}_{pq} + z_{ij}. \tag{21}$$

Here $\bar{J}_{pq}$ is the mean value of the connections from population q to population p, and z_{ij} is the remaining zero-mean random part of each connection. Defining $\bar{\mathbf{J}}$ as the $N \times N$ deterministic matrix consisting of mean values, and $\mathbf{Z}$ as the noise matrix consisting of the random parts z_{ij}, the connectivity matrix $\mathbf{J}$ can be written as

$$\mathbf{J} = \bar{\mathbf{J}} + \mathbf{Z}. \tag{22}$$

The matrix $\bar{\mathbf{J}}$ is of size $N \times N$ and consists of P^2 blocks with identical values within each block. The rank of $\bar{\mathbf{J}}$ is therefore at most P [40]. In contrast, the noise matrix $\mathbf{Z}$ is in general of rank N. The full connectivity matrix $\mathbf{J}$ can then be interpreted as a rank-P deterministic matrix perturbed by the random matrix $\mathbf{Z}$ with block-dependent statistics.

In the case of Gaussian connectivity, connections from population q to population p are sampled independently from a Gaussian distribution.

$$f^{pq}(\mathbf{J}) = \mathcal{N}(\bar{J}_{pq}, \sigma^2_{z_{pq}}), \tag{23}$$

with variances

$$\sigma^2_{z_{pq}} = \frac{g^2_{pq}}{N}. \tag{24}$$

The noise matrix $\mathbf{Z}$ therefore has *block-structured variances* g^2_{pq}/N that we specify by a $P \times P$ matrix $\mathbf{G}_m$:

$$\mathbf{G}_m = \begin{bmatrix} g^2_{11}/N & \cdots & g^2_{1P}/N \\ \vdots & & \vdots \\ g^2_{P1}/N & \cdots & g^2_{PP}/N \end{bmatrix}. \tag{25}$$

In the case of sparse connectivity, J^{pq}_{ij} is a Bernoulli random variable. The connectivity weights J^{pq}_{ij} from population q to population p are non-zero with probability c_{pq} and zero otherwise. All non-zero connection weights within a block take the same value A_{pq}, so that analogously to Eq. (20), the sparse connectivity is defined as

$$f^{pq}(J_{ij} = J) = \begin{cases} c_{pq} & \text{for } J = A_{pq}, \\ 1 - c_{pq} & \text{for } J = 0. \end{cases} \tag{26}$$

The mean connectivity weight between populations p, q is then

$$\bar{J}_{pq} = c_{pq} A_{pq}. \tag{27}$$

and the variance of the remaining random part z_{ij} is

$$\sigma^2_{z_{pq}} = [z^2_{ij}]_{i \in N_p, j \in N_q} = (1 - c_{pq}) c_{pq} A^2_{pq}, \tag{28}$$

to simplify the parameters in sparse networks, we assume that A_{pq} depend only on presynaptic population q, and that the connection probability c_{pq} is a homogeneous network parameter independent of p, q that we denote by c.

2.1.2 Reciprocal connectivity motifs

To go beyond independent connectivity, we consider pairwise motifs, i. e. correlations between reciprocal pairs of weights J_{ij} and J_{ji}. We quantify this correlation using the normalized covariance η_{ij} defined as

$$\eta_{ij} = \left(\frac{[(J_{ij} - [J_{ij}])(J_{ji} - [J_{ji}])]}{\sqrt{[(J_{ij} - [J_{ij}])^2][(J_{ji} - [J_{ji}])^2]}} \right)_J, \tag{29}$$

where $[\cdot]$ denotes the average over the full connectivity distribution. Reciprocal connections are fully independent when $\eta_{ij} = 0$ for all i, j, fully symmetric when $\eta_{ij} = 1$ and fully anti-symmetric when $\eta_{ij} = -1$.

Our key assumption is that the statistics of connectivity are block-like, implying that all pairs of connections between populations p, q share the same correlation coefficient η_{pq}, so that the statistics are defined by a $P \times P$ reciprocal correlation matrix $\boldsymbol{\eta}_m$

$$
\boldsymbol{\eta}_m = \begin{bmatrix} \eta_{11} & \cdots & \eta_{1P} \\ \vdots & & \vdots \\ \eta_{P1} & \cdots & \eta_{PP} \end{bmatrix},
\tag{30}
$$

where, by definition $\eta_{pq} = \eta_{qp}$.

For Gaussian statistics, we generate connectivity matrices with a specified set of η_{pq} in the following manner. We first generate an $N \times N$ matrix $\mathbf{Y}'$ with entries independently sampled from the normal distribution $\mathcal{N}(0,1)$. We then form a linear combination $\mathbf{Y}'$ and its transpose $\mathbf{Y}'^{\mathsf{T}}$ to generate a matrix $\mathbf{Y}$ with reciprocal correlations η_{pq}. Specifically, we set

$$
y_{ij} = \gamma_{pq} y'_{ij} + \sqrt{1 - \gamma_{pq}^2} y'_{ji}
\tag{31}
$$

with $\gamma_{pq} = \sqrt{1 - \sqrt{1 - \eta_{pq}^2}/2}$ for $\eta_{pq} > 0$, and $-\sqrt{1 - \sqrt{1 - \eta_{pq}^2}/2}$ for $\eta_{pq} < 0$. Finally, we scale each block by $g_{pq}/\sqrt{N}$ to obtain the random connectivity component $\mathbf{Z}$, which is added to the mean connectivity component $\bar{\mathbf{J}}$ to finally obtain the full connectivity matrix $\mathbf{J}$.

For *sparse* networks, we first generate a connectivity matrix without reciprocal correlations. We then consider the upper triangle of this matrix, randomly select a fraction ρ_{pq} of the non-zero connections J_{ij} with value A_q and set their reciprocal connectivity weights J_{ji} to have a non-zero weight A_p. For the remaining $1 - \rho_{pq}$ fraction of non-zero connections in the upper triangle, we set the reciprocal connectivity weights to zero. The corresponding cell-type dependent reciprocal correlations for the multi-population sparse connectivity are then

$$
\eta_{pq} = \frac{A_p A_q (\rho_{pq} - c)}{|A_p A_q|(1 - c)}, \quad p, q = 1 \dots P,
\tag{32}
$$

where c is the homogeneous connection probability (Table. 1).

2.1.3 Excitatory-inhibitory networks

In this work, we specifically focus on excitatory-inhibitory networks composed of $P = 2$ populations, one excitatory and one inhibitory, with respectively N_E and N_I neurons. We denote the two populations by indices E and I, so that there are four types of connections: EE, EI, IE and II. Based on the usual anatomical estimates for neocortex, we choose $N_E = 0.8N$, $N_I = 0.2N$, and further define $\alpha_E = N_E/N$, $\alpha_I = N_I/N$, as the fractions of excitatory and inhibitory neurons.

For Gaussian networks, we enforce Dale's law only on the mean, i. e. we set $\bar{J}_{EE}$ and $\bar{J}_{IE}$ to be positive, while $\bar{J}_{EI}$ and $\bar{J}_{II}$ are negative. The $N \times N$ mean connectivity matrix $\bar{\mathbf{J}}$ is therefore in general rank-two. To further simplify the setting, we follow [43], and consider networks where the mean weights of all excitatory connections, and respectively all inhibitory connections, are equal and set by parameters J_E and J_I:

$$
\begin{aligned}
\bar{J}_{EE} &= \bar{J}_{IE} = J_E/N_E \\
\bar{J}_{EI} &= \bar{J}_{II} = -J_I/N_I.
\end{aligned}
\tag{33}
\tag{34}
$$

Under these additional assumptions, the entries in the first N_E columns of the mean connectivity matrix $\bar{\mathbf{J}}$ have the same positive weight J_E/N_E, and the entries in the following N_I columns have the same negative weight $-J_I/N_I$, so that $\bar{\mathbf{J}}$ becomes rank one.

We however allow the variances g_{pq}^2/N and reciprocal correlations η_{pq} to depend on both the pre- and post-synaptic population, so that the corresponding parameters form 2×2 matrices

$$\mathbf{G}_m = \left[\begin{array}{cc} g_{EE}^2/N & g_{EI}^2/N \\ g_{IE}^2/N & g_{II}^2/N \end{array} \right], \quad \boldsymbol{\eta}_m = \left[\begin{array}{cc} \eta_{EE} & \eta_{EI} \\ \eta_{IE} & \eta_{II} \end{array} \right], \tag{35}$$

where $\eta_{EI} = \eta_{IE}$.

For *sparse* excitatory-inhibitory networks, all non-zero excitatory (resp. inhibitory) synaptic weights are equal and positive, $A_E > 0$ (resp. $A_I < 0$). From Eqs. (27), (28) the mean and the variance of the synaptic weights in the sparse network can be matched to the parameters of the Gaussian model:

$$\begin{aligned} \frac{J_E}{N_E} &= cA_E, \quad g_{EE}^2/N = g_{IE}^2/N = A_E^2 c \left(1 - c\right), \\ -\frac{J_I}{N_I} &= cA_I, \quad g_{EI}^2/N = g_{II}^2/N = A_I^2 c \left(1 - c\right), \end{aligned} \tag{36}$$

In particular, for the sparse networks with pairwise reciprocal motifs, on top of the matching means and variances, the cell-type dependent reciprocal correlations satisfy (Eq. (32))

$$\eta_{EE} = \frac{\rho_{EE} - c}{1 - c}, \quad \eta_{EI} = -\frac{\rho_{EI} - c}{1 - c}, \quad \eta_{II} = \frac{\rho_{II} - c}{1 - c}. \tag{37}$$

2.2 Globally-defined connectivity: Gaussian-mixture low-rank networks

In this section, we introduce a second broad class of connectivity models, Gaussian-mixture low-rank networks [31, 33], in which the connectivity matrix is generated from a global statistics of vectors over neurons.

Low-rank networks are a class of recurrent neural networks in which the connectivity matrix $\mathbf{J}$ is restricted to be of rank R, assumed to be much smaller than the number of neurons N. Such a connectivity matrix can be expressed as a sum of R unit rank terms

$$\mathbf{J} = \frac{1}{N} \sum_{r=1}^{R} \mathbf{m}^{(r)} \mathbf{n}^{(r)\intercal}. \tag{38}$$

We refer to $\mathbf{n}^{(r)} = \{n_i^{(r)}\}_{i=1\ldots N}$ and $\mathbf{m}^{(r)} = \{m_i^{(r)}\}_{i=1\ldots N}$ as the r-th left and right *connectivity vectors*. The $2R$ connectivity vectors together fully specify the connectivity matrix. Each neuron i is then characterized by its set of $2R$ entries $(m_i^{(1)}, n_i^{(1)}, \ldots, m_i^{(R)}, n_i^{(R)})$ on these vectors. For unit-rank networks, the main focus of this study, the connectivity matrix is simply given by the outer product of a pair of connectivity vectors $\mathbf{m}$ and $\mathbf{n}$:

$$\mathbf{J}_{R1} = \frac{1}{N} \mathbf{m} \mathbf{n}^\intercal. \tag{39}$$

Gaussian-mixture low-rank networks are a subset of the class of low-rank networks, for which the entries of the connectivity vectors are drawn independently for each

neuron from a mixture of Gaussians distribution [31]. Specifically, a fraction α_p of neurons is assigned to a population p, and within each population, the entries on the connectivity vectors are generated from a given $2R$-dimensional Gaussian distribution. For a unit-rank network, for a neuron i in the population p, the connectivity parameters (m_i, n_i) are generated from a bi-variate Gaussian distribution with mean $(\bar{m}^p, \bar{n}^p)$, variance $(\sigma^2_{m^p}, \sigma^2_{n^p})$ and covariance σ^p_{nm}.

For any unit-rank matrix of the form in Eq. (39), the only potentially non-zero eigenvalue is given by $\lambda = \mathbf{n}^\intercal\mathbf{m}/N$, and the corresponding right (resp. left) eigenvector is $\mathbf{m}$ (resp. $\mathbf{n}$). For a Gaussian-mixture model, in the large N limit this eigenvalue becomes

$$\lambda = \sum_{p=1}^{P} \alpha_p \left(\bar{m}^p \bar{n}^p + \sigma^p_{nm} \right) \tag{40}$$

Starting from a Gaussian-mixture low-rank model in which the connectivity is globally defined, it is straightforward to compute the resulting local statistics of the connectivity, i. e. the mean $\bar{J}_{pq}$, variance $\sigma^2_{z_{pq}}$ (Methods Sec. 2.1.1) and reciprocal correlation η_{pq} (Methods Sec. 2.1.2) as:

$$
\begin{aligned}
\bar{J}_{pq} &= \frac{1}{N}\bar{m}^p \bar{n}^q, \\
\sigma^2_{z_{pq}} &= \frac{1}{N^2} \left(\sigma^2_{m^p}(\bar{n}^q)^2 + (\bar{m}^p)^2 \sigma^2_{n^q} + \sigma^2_{m^p}\sigma^2_{n^q} \right), \\
\eta_{pq} &= \frac{(\sigma^p_{nm}\bar{m}^q\bar{n}^q + \sigma^q_{nm}\bar{m}^p\bar{n}^p + \sigma^p_{nm}\sigma^q_{nm})}{\sqrt{\left(\sigma^2_{m^p}(\bar{n}^q)^2 + (\bar{m}^p)^2\sigma^2_{n^q} + \sigma^2_{m^p}\sigma^2_{n^q}\right)\left(\sigma^2_{m^q}(\bar{n}^p)^2 + (\bar{m}^q)^2\sigma^2_{n^p} + \sigma^2_{m^q}\sigma^2_{n^p}\right)}}.
\end{aligned}
\tag{41}
$$

The expression for the local statistics of network connectivity using rank-R connectivity is in S5 Appendix.

2.3 Approximating locally-defined connectivity with Gaussian-mixture low-rank models

In this section, we describe our general approach for approximating an arbitrary connectivity matrix $\mathbf{J}$ with a rank-R matrix $\mathbf{J}_R$. We then show that for $\mathbf{J}$ corresponding to locally-defined multi-population connectivity (Methods Sec. 2.1), the resulting approximation $\mathbf{J}_R$ in general obeys Gaussian-mixture low-rank statistics as defined in Methods Sec. 2.2.

To approximate a full rank matrix $\mathbf{J}$ with a rank-R matrix $\mathbf{J}_R$, we use truncated eigen-decomposition, which preserves the dominant eigenvalues. We start from the full eigen-decomposition of $\mathbf{J}$:

$$\mathbf{J} = \sum_{r=1}^{N} \lambda_r \mathbf{R}_r \mathbf{L}_r^T, \tag{42}$$

where λ_r is the r-th eigenvalue of $\mathbf{J}$ (ordered by decreasing absolute value), while $\mathbf{R}_r$ and $\mathbf{L}_r$ are the corresponding right- and left-eigenvectors that obey

$$\mathbf{J}\mathbf{R}_r = \mathbf{R}_r\lambda_r \tag{43}$$
$$\mathbf{L}_r^\intercal\mathbf{J} = \lambda_r\mathbf{L}_r^\intercal \tag{44}$$
$$\mathbf{L}_r^T\mathbf{R}_{r'} = \delta_{rr'}. \tag{45}$$

In the following, we constrain the right eigenvectors $\mathbf{R}_r$ to be of unit norm, while the normalization of the left eigenvector is determined by Eq. (45).

We obtain a rank-R approximation $\mathbf{J}_R$ of $\mathbf{J}$ by keeping the first R terms in Eq. (42): 669

$$\mathbf{J}_R = \sum_{r=1}^{R} \lambda_r \mathbf{R}_r \mathbf{L}_r^T. \tag{46}$$

The R non-trivial eigenvalues and eigenvectors of $\mathbf{J}_R$ therefore correspond to the first R 670
eigenvalues and eigenvectors of $\mathbf{J}$. We then set 671

$$\begin{aligned}
\mathbf{m}^{(r)} &= \sqrt{N}\mathbf{R}_r & (47) \\
\mathbf{n}^{(r)} &= \sqrt{N}\lambda_r \mathbf{L}_r & (48)
\end{aligned}$$

to have the same normalization for $\mathbf{J}_R$ as in Eq. (38). 672

    To obtain a low-rank approximation for a connectivity matrix $\mathbf{J}$ generated from 673
locally-defined statistics defined in Methods Sec. 2.1, we first determine its dominant 674
eigenvalues and eigenvectors. Starting from Eq. (22), this problem becomes equivalent 675
to finding the dominant eigenvalues and eigenvectors of a low-rank matrix $\bar{\mathbf{J}}$ perturbed 676
by a random matrix $\mathbf{Z}$ with block-like statistics. We compute the statistics of these 677
eigenvalues and eigenvectors by combining the Matrix's Determinant Lemma, the 678
Matrix Perturbation Theory and the Central Limit Theorem. Below we summarize this 679
general approach before applying it to different specific cases in Methods Secs. 2.4, 2.5. 680

    We focus on the case where $\bar{\mathbf{J}}$ is unit rank as in the simplified excitatory-inhibitory 681
network introduced in Methods Sec. 2.1.3. In that case, the unique non-zero eigenvalue 682
of $\bar{\mathbf{J}}$ is 683

$$\lambda_0 = J_E - J_I, \tag{49}$$

and the corresponding left and right eigenvectors are 684

$$\begin{aligned}
\bar{\mathbf{L}} &= \Big[\frac{\sqrt{N}J_E}{N_E(J_E - J_I)}, \dots, -\frac{\sqrt{N}J_I}{N_I(J_E - J_I)}, \dots \Big]^\intercal, \\
\bar{\mathbf{R}} &= \Big[\frac{1}{\sqrt{N}}, \dots, \frac{1}{\sqrt{N}}, \dots \Big]^\intercal.
\end{aligned} \tag{50}$$

$\bar{\mathbf{J}}$ can then be rewritten as 685

$$\bar{\mathbf{J}} = \frac{1}{N}\bar{\mathbf{m}}\bar{\mathbf{n}}^\intercal, \tag{51}$$

where the structure vectors $\bar{\mathbf{m}}$ and $\bar{\mathbf{n}}$ are uniquely defined by rescaling the left and right 686
eigenvectors $\bar{\mathbf{L}}$, $\bar{\mathbf{R}}$ of $\bar{\mathbf{J}}$ (Eq. (50)) as in Eq. (47), so that 687

$$\begin{aligned}
\bar{m}_i &= 1, \quad i = 1\dots N & (52) \\
\bar{n}_i &= \bar{n}^E = \frac{N}{N_E}J_E \quad i \in N_E & (53) \\
\bar{n}_i &= \bar{n}^I = -\frac{N}{N_I}J_I \quad i \in N_I, & (54)
\end{aligned}$$

The full connectivity matrix $\mathbf{J}$ can be then expressed as 688

$$\mathbf{J} = \bar{\mathbf{m}}\bar{\mathbf{n}}^\intercal/N + \mathbf{Z}. \tag{55}$$

Eigenvalues. For a random matrix $\mathbf{Z}$ with independently distributed elements, the 689
eigenvalues are distributed on a disk of radius r_g centred at the origin in the complex 690
plane [11, 44, 50]. Correlations between elements in general modify the shape of this 691
continuous spectrum [18, 67]. In contrast, adding a low-rank component typically 692
induces isolated eigenvalues outside the continuous part of the spectrum [44, 45]. To 693

obtain a low-rank approximation of the full matrix, we focus on determining these outliers when they exist.

All eigenvalues λ of $\mathbf{J}$ satisfy the characteristic equation

$$\det\left(\mathbf{J} - \mathbf{I}\lambda\right) = 0. \tag{56}$$

To determine the outlying eigenvalues of a random connectivity with low-rank structure, we apply the matrix determinant lemma to the l. h. s. of the characteristic equation [32]:

$$\det\left(\frac{1}{N}\bar{\mathbf{m}}\bar{\mathbf{n}}^{\mathsf{T}} + \mathbf{Z} - \mathbf{I}\lambda\right) = \left(1 + \frac{1}{N}\bar{\mathbf{n}}^{\mathsf{T}}\left(\mathbf{Z} - \mathbf{I}\lambda\right)^{-1}\bar{\mathbf{m}}\right)\det\left(\mathbf{Z} - \mathbf{I}\lambda\right). \tag{57}$$

As outliers by definition cannot be eigenvalues of $\mathbf{Z}$, they correspond to zeros of the first term in the r. h. s. , and therefore satisfy:

$$\lambda = \frac{1}{N}\bar{\mathbf{n}}^{\mathsf{T}}(1 - \mathbf{Z}/\lambda)^{-1}\bar{\mathbf{m}}. \tag{58}$$

Expanding $(1 - \mathbf{Z}/\lambda)^{-1}$ in series, we further get [32]

$$\lambda = \sum_{k=0}^{\infty}\frac{\theta_k}{\lambda^k}, \tag{59}$$

with

$$\theta_k = \frac{1}{N}\bar{\mathbf{n}}^{\mathsf{T}}\mathbf{Z}^k\bar{\mathbf{m}}. \tag{60}$$

Here θ_0 corresponds to the eigenvalue λ_0 of $\bar{\mathbf{J}}$ (Eq. (49)), and the higher order terms specify how this eigenvalue is modified by the random part of the connectivity. Truncating Eq. (59) at a given order, and averaging over $\mathbf{Z}$ yields a polynomial equation for the mean eigenvalues of $\mathbf{J}$. In Methods Sec. 2.4, we exploit this equation to determine the effects of different cell-type specific random connectivity $\mathbf{Z}$ on the outlying eigenvalues.

Note that within first-order perturbation theory, the eigenvalues are given by $\lambda = \lambda_0 + \Delta\lambda$ with

$$\Delta\lambda = \bar{\mathbf{L}}^{\mathsf{T}}\mathbf{Z}\bar{\mathbf{R}} = \frac{1}{N\lambda_0}\bar{\mathbf{n}}^{\mathsf{T}}\mathbf{Z}\bar{\mathbf{m}}. \tag{61}$$

Eigenvectors. To determine the eigenvectors corresponding to the outlying eigenvalue of $\mathbf{J}$, we treat it as $\bar{\mathbf{J}}$ perturbed by $\mathbf{Z}$ (Eq. (22)). Matrix perturbation theory then states that, at first order, the right- and left-eigenvectors $\mathbf{R}$ and $\mathbf{L}$ of $\mathbf{J}$ corresponding to the outlying eigenvalue λ are given by [47]:

$$\mathbf{R} = \bar{\mathbf{R}} + \Delta\mathbf{R} \tag{62}$$
$$\mathbf{L} = \bar{\mathbf{L}} + \Delta\mathbf{L} \tag{63}$$

where $\bar{\mathbf{R}}$ and $\bar{\mathbf{L}}$ are the right- and left-eigenvectors of $\bar{\mathbf{J}}$ defined in Eq. (50), and

$$\begin{aligned}\Delta\mathbf{R} &= \frac{1}{\lambda_0}\mathbf{Z}\bar{\mathbf{R}} \\ \Delta\mathbf{L}^{\mathsf{T}} &= \frac{1}{\lambda_0}\bar{\mathbf{L}}^{\mathsf{T}}\mathbf{Z}.\end{aligned} \tag{64}$$

Using the normalization in Eq. (47), we then get

$$\begin{aligned}\mathbf{m} &= \sqrt{N}\mathbf{R} = \bar{\mathbf{m}} + \Delta\mathbf{m}, \\ \mathbf{n}^{\mathsf{T}} &= \lambda\sqrt{N}\mathbf{L}^{\mathsf{T}} = \bar{\mathbf{n}}^{\mathsf{T}} + \Delta\mathbf{n}^{\mathsf{T}},\end{aligned} \tag{65}$$

with constant entries on $\bar{\mathbf{m}}$ and $\bar{\mathbf{n}}$ defined in Eqs. (52)-(54) and

$$
\begin{aligned}
\Delta\mathbf{m} &= \frac{1}{\lambda_0}\mathbf{Z}\bar{\mathbf{m}} \\
\Delta\mathbf{n}^{\mathsf{T}} &= \frac{1}{\lambda_0}\bar{\mathbf{n}}^{\mathsf{T}}\mathbf{Z}.
\end{aligned}
\tag{66}
$$

where we approximated λ at first order by λ_0.

Statistics of Eigenvector entries. While $\bar{\mathbf{m}}$ and $\bar{\mathbf{n}}$ are deterministic vectors, the perturbations $\Delta\bar{\mathbf{m}}$ and $\Delta\bar{\mathbf{n}}$ are random variables obtained by multiplying $\bar{\mathbf{m}}$ and $\bar{\mathbf{n}}$ with the random matrix $\mathbf{Z}$ (Eq. (66)). We therefore next consider the statistics of the elements m_i and n_i of $\mathbf{m}$ and $\mathbf{n}$ defined in Eq. (65).

Since the elements of $\mathbf{Z}$ have zero mean, the mean values of m_i and n_i are given by $\bar{m}_i$ and $\bar{n}_i$ defined in Eqs. (52)-(54). The mean value of n_i, but not m_i, therefore depends on whether the neuron i belongs to the excitatory or inhibitory population. Taking into account that $\mathbf{Z}$ has block-like statistics, we split the matrix product in Eq. (66) into the sum of items corresponding to excitatory and inhibitory pre-synaptic neurons. Using Eqs. (52)-(54), Δm_i and Δn_i can be written as

$$
\begin{aligned}
\Delta m_i &= \frac{1}{\lambda_0}\sum_{j\in N_E} z_{ij} + \frac{1}{\lambda_0}\sum_{j\in N_I} z_{ij} \\
\Delta n_i &= \frac{1}{\lambda_0}\sum_{j\in N_E} \bar{n}^E z_{ji} + \frac{1}{\lambda_0}\sum_{j\in N_I} \bar{n}^I z_{ji},
\end{aligned}
\tag{67}
$$

We next take the limit $N_E, N_I \to \infty$, and apply the central limit theorem, which states that each sum converges to a Gaussian random variable, so that we have

$$
\begin{aligned}
\Delta m_i^p &\sim \frac{1}{\lambda_0}\sum_{q=E,I} \sqrt{N_q}\mathcal{N}(0, \sigma_{z_{pq}}^2) \\
\Delta n_i^p &\sim \frac{1}{\lambda_0}\sum_{q=E,I} \sqrt{N_q}\bar{n}^q\mathcal{N}(0, \sigma_{z_{qp}}^2).
\end{aligned}
\tag{68}
$$

where $p \in E, I$ is the population the neuron i belongs to, and $\sigma_{z_{pq}}^2$, $\sigma_{z_{qp}}^2$ are the variance of z_{ij}, z_{ji} respectively, for i, j in populations p, q. The perturbations Δm_i and Δn_i. therefore converge to Gaussian random variables of zero mean and variances $\sigma_{m^p}^2$ and $\sigma_{n^p}^2$ given by:

$$
\begin{aligned}
\sigma_{m^p}^2 &= \sum_{q=E,I} \frac{N_q \sigma_{z_{pq}}^2}{\lambda_0^2} \\
\sigma_{n^p}^2 &= \sum_{q=E,I} \frac{N_q (\bar{n}^q)^2 \sigma_{z_{qp}}^2}{\lambda_0^2}.
\end{aligned}
\tag{69}
$$

The population covariance σ_{nm}^p between elements m_i and n_i with i belonging to population p can furthermore be written as

$$
\sigma_{nm}^p = \frac{1}{N_p}\Delta\mathbf{n}^{p\mathsf{T}}\Delta\mathbf{m}^p = \frac{1}{N_p\lambda_0^2}\sum_{s,q=E,I}\left(\bar{n}^s\sum_{i\in N_s}\sum_{j\in N_q}\sum_{k\in N_p} z_{ik}z_{kj}\bar{m}^q\right),
\tag{70}
$$

while the overall covariance σ_{nm} between all m_i and n_i reads

$$
\sigma_{nm} = \sum_{p=E,I} \alpha_p \sigma_{nm}^p.
\tag{71}
$$

Altogether, m_i and n_i determined by perturbation theory therefore follow Gaussian-mixture statistics, where the mean and variance depend on whether the neuron i belongs to the excitatory or inhibitory population.

Comparison with simulations. The theoretical predictions for eigenvalues obtained from Eqs. (59), (60) can be verified by comparing them with the eigenvalue outliers computed by direct eigen-decomposition of the full matrix $\mathbf{J}$. We compute the average and standard deviation of eigenvalue outliers over 30 realizations of $\mathbf{J}$.

The predictions of perturbation theory for eigenvectors given by Eq. (64) can also be verified by direct eigen-decomposition, but to compare individual entries, an appropriate normalization is required [47]. Indeed, perturbation theory assumes that $\bar{\mathbf{R}}$ is normalized and $\bar{\mathbf{L}}$ satisfies $\bar{\mathbf{L}}^\intercal \bar{\mathbf{R}} = 1$ (Eq. (50)), but the perturbed eigenvectors in Eq. (62) do not obey the same normalization. We therefore first use numerical eigen-decomposition to get the right- and left-eigenvector $\hat{\mathbf{R}}$ and $\hat{\mathbf{L}}$ of $\mathbf{J}$. We then normalize $\hat{\mathbf{R}}$ to 1, and $\hat{\mathbf{L}}$ so that $\hat{\mathbf{L}}^\intercal \hat{\mathbf{R}} = 1$. To compare $\hat{\mathbf{L}}$, $\hat{\mathbf{R}}$ with perturbation theory, we then normalize $\hat{\mathbf{L}}$, $\hat{\mathbf{R}}$ as

$$\mathbf{R} = \left(\frac{\hat{\mathbf{L}}^\intercal \bar{\mathbf{R}}}{\bar{\mathbf{L}}^\intercal \hat{\mathbf{R}}} \right)^{1/2} \hat{\mathbf{R}},$$

$$\mathbf{L} = \left(\frac{\hat{\mathbf{R}}^\intercal \bar{\mathbf{L}}}{\bar{\mathbf{R}}^\intercal \hat{\mathbf{L}}} \right)^{1/2} \hat{\mathbf{L}}, \tag{72}$$

the eigenvectors $\mathbf{L}$, $\mathbf{R}$ after normalization have the same statistics as $(\bar{\mathbf{L}} + \Delta\mathbf{L})$, $(\bar{\mathbf{R}} + \Delta\mathbf{R})$ (Eqs. (50), (64)).

2.4 Eigenvalues

Here, we apply Eqs. (59), (60) to determine the mean and variance of outlying eigenvalues for different forms of local connectivity statistics.

2.4.1 Independent random connectivity

In the case of independent random connectivity, the elements of $\mathbf{Z}$ are zero-mean, independently distributed and uncorrelated with $\bar{\mathbf{m}}$ and $\bar{\mathbf{n}}$. Averaging Eq. (60) over $\mathbf{Z}$ then leads to [32]:

$$[\theta_k] = \left[\frac{1}{N} \bar{\mathbf{n}}^\intercal \mathbf{Z}^k \bar{\mathbf{m}} \right] \tag{73}$$

$$= 0, \tag{74}$$

and therefore the mean eigenvalue $[\lambda]$ of $\mathbf{J}$ is given by the eigenvalue λ_0 of $\bar{\mathbf{J}}$. For Gaussian random connectivity, we have

$$[\lambda] = J_E - J_I, \tag{75}$$

and for sparse connectivity

$$[\lambda] = c(N_E A_E + N_I A_I). \tag{76}$$

The variance σ_λ^2 of λ can be computed by keeping only the linear term in Eq. (59), which leads to Eq. (61) under the assumption that $\lambda \approx \lambda_0$. Applying the central limit theorem then yields

$$\sigma_\lambda^2 = \frac{1}{\lambda_0^2} \left(J_E^2 \left(\sigma_{z_{EE}}^2 + \frac{N_I}{N_E} \sigma_{z_{EI}}^2 \right) + J_I^2 \left(\frac{N_E}{N_I} \sigma_{z_{IE}}^2 + \sigma_{z_{II}}^2 \right) \right). \tag{77}$$

For independent Gaussian random connectivity, we substitute $\sigma_{z_{pq}}^2$ using Gaussian variance parameters (Eq. (23), (25))

$$\sigma_\lambda^2 = \frac{1}{N\lambda_0^2}\left(J_E^2\left(g_{EE}^2 + \frac{N_I}{N_E}g_{EI}^2\right) + J_I^2\left(\frac{N_E}{N_I}g_{IE}^2 + g_{II}^2\right)\right). \tag{78}$$

For independent sparse random connectivity, we replace $\sigma_{z_{pq}}^2$ and J_p with the variances and means of the sparse model given in Eqs. (27), (28) and get

$$\sigma_\lambda^2 = \frac{1}{\lambda_0^2}\left(A_E^2 N_E + A_I^2 N_I\right)^2 (1-c)\,c^3. \tag{79}$$

2.4.2 Reciprocal motifs

In the case of connectivity with reciprocal correlations, z_{ij} and z_{ji} are correlated, so that the average $[\theta_k]$ over $\mathbf{Z}$ in Eq. (60) is non-zero for even k. Here we compute $[\theta_k]$ for $k=2$ and truncate Eq. (59) at second order to get a third-order polynomial equation for the mean eigenvalue:

$$f(\lambda) = \lambda^3 - \left(\lambda_0\lambda^2 + [\theta_1]\lambda + [\theta_2]\right) = 0. \tag{80}$$

For Gaussian connectivity we have $\theta_0 = \lambda_0 = J_E - J_I$, and θ_1 is given by

$$\theta_1 = \bar{\mathbf{n}}^\intercal \mathbf{Z}\bar{\mathbf{m}}/N = \sum_{p,q=E,I}\bar{n}^p\sum_{i\in N_p, j\in N_q} z_{ij}\bar{m}^q/N, \tag{81}$$

so that $[\theta_1] = 0$.
The next term θ_2 is

$$\theta_2 = \bar{\mathbf{n}}^\intercal \mathbf{Z}^2\bar{\mathbf{m}}/N = \sum_{p,q=E,I}\bar{n}^p\left(\sum_{i\in N_p}\sum_{j\in N_q}\sum_{k=1}^N z_{ik}z_{kj}\right)\bar{m}^q/N. \tag{82}$$

Given the reciprocal correlations defined in Eq. (29), only items with $i=j$ in θ_2 are non-zero after averaging over Gaussian realizations, so that

$$[\theta_2] = [\bar{\mathbf{n}}^\intercal \mathbf{Z}^2\bar{\mathbf{m}}]/N = \sum_{p=E,I}\bar{n}^p\left[\sum_{i\in N_p}\sum_{k=1}^N z_{ik}z_{ki}\right]\bar{m}^p/N. \tag{83}$$

We then write

$$\left[\sum_{k=1}^N z_{ik}z_{ki}\right] = \alpha_E g_{EE}^2\eta_{EE} + \alpha_I g_{EI}g_{IE}\eta_{EI},\ i\in N_E$$
$$\left[\sum_{k=1}^N z_{ik}z_{ki}\right] = \alpha_E g_{IE}g_{EI}\eta_{EI} + \alpha_I g_{II}^2\eta_{II},\ i\in N_I. \tag{84}$$

and substitute Eq. (84) and $\bar{n}^p$, $\bar{m}^p$ into Eq. (83) to obtain

$$[\theta_2] = J_E\left(\alpha_E g_{EE}^2\eta_{EE} + \alpha_I g_{EI}g_{IE}\eta_{EI}\right)$$
$$- J_I\left(\alpha_E g_{IE}g_{EI}\eta_{EI} + \alpha_I g_{II}^2\eta_{II}\right). \tag{85}$$

For sparse connectivity with reciprocal motifs, the correlations can be written as

$$\left[\sum_{k=1}^N z_{ik}z_{ki}\right] = cA_E^2 N_E(\rho_{EE} - c) - cA_E A_I N_I(c - \rho_{EI}),\ i\in N_E$$
$$\left[\sum_{k=1}^N z_{ik}z_{ki}\right] = -cA_E A_I N_E(c - \rho_{EI}) + cA_I^2 N_I(\rho_{II} - c),\ i\in N_I. \tag{86}$$

Then, combining Eqs. (36) and Eqs. (52)-(54), the second-order coefficient $[\theta_2]$ for the sparse network is

$$
\begin{aligned}
[\theta_2] = {} & A_E^3 c^2 N_E^2 \left(\rho_{EE} - c\right) + A_E^2 A_I c^2 N_E N_I \left(\rho_{EI} - c\right) \\
& + A_E A_I^2 c^2 N_I N_E \left(\rho_{EI} - c\right) + A_I^3 c^2 N_I^2 \left(\rho_{II} - c\right),
\end{aligned}
\tag{87}
$$

Using Eqs. (36), (37), it can be seen that Eq. (87) is equivalent to Eq. (85).

2.5 Eigenvectors

Here we apply Eqs. (69) and (70) to determine the variances and covariances of eigenvector entries obtained from perturbation theory for different forms of local connectivity statistics.

2.5.1 Independent random connectivity

In the case of independent random connectivity, because z_{ik} and z_{kj} are not correlated in Eq. (70), the covariances σ_{nm}^p between the eigenvector entries are zero. For independent Gaussian connectivity, introducing Eq. (24) into Eqs. (69) the variances of eigenvector entries can be written as

$$
\begin{aligned}
\sigma_{m^E}^2 &= \frac{1}{\lambda_0^2} \left(\alpha_E g_{EE}^2 + \alpha_I g_{EI}^2\right), & \sigma_{m^I}^2 &= \frac{1}{\lambda_0^2} \left(\alpha_E g_{IE}^2 + \alpha_I g_{II}^2\right), \\
\sigma_{n^E}^2 &= \frac{1}{\lambda_0^2} \left(\frac{1}{\alpha_E} J_E^2 g_{EE}^2 + \frac{1}{\alpha_I} J_I^2 g_{IE}^2\right), & \sigma_{n^I}^2 &= \frac{1}{\lambda_0^2} \left(\frac{1}{\alpha_E} J_E^2 g_{EI}^2 + \frac{1}{\alpha_I} J_I^2 g_{II}^2\right).
\end{aligned}
\tag{88}
$$

For independent sparse connectivity, substituting Eq. (28) into Eqs. (69), leads to

$$
\begin{aligned}
\sigma_{m^E}^2 = \sigma_{m^I}^2 &= \frac{1}{\lambda_0^2} \left(N_E A_E^2 c(1-c) + N_I A_I^2 c(1-c)\right), \\
\sigma_{n^E}^2 &= \frac{1}{\lambda_0^2} \left(A_E^2 N_E + A_I^2 N_I\right) A_E^2 N^2 c^3 (1-c), \\
\sigma_{n^I}^2 &= \frac{1}{\lambda_0^2} \left(A_E^2 N_E + A_I^2 N_I\right) A_I^2 N^2 c^3 (1-c).
\end{aligned}
\tag{89}
$$

2.5.2 Reciprocal motifs

In the case of connectivity with reciprocal correlations, the variances of eigenvector entries are identical to the independent case.

As we have shown in Eq. (70), noise correlation between the rank-one vectors arises from the correlation between pairwise random connectivity weights in the situation with reciprocal motifs, only items with $i = j$ (for $z_{ik}, \ z_{kj}$) in the same population q are non-zero, so that we have

$$
\sigma_{nm} = \sum_{p=E,I} \alpha_p \sigma_{nm}^p
\tag{90}
$$

with

$$
\sigma_{nm}^p = \frac{1}{N_p \lambda_0^2} \sum_{q=E,I} \left(\bar{n}^q \sum_{i \in N_q} \sum_{k \in N_p} z_{ik} z_{ki} \bar{m}^q\right).
\tag{91}
$$

For Gaussian connectivity with reciprocal correlations, the covariances between entries on $\mathbf{Z}$ matched by population can be written as

$$
\begin{aligned}
\Big[\sum_{i \in N_E} \sum_{k \in N_E} z_{ik} z_{ki} \Big] &= N_E \alpha_E g_{EE}^2 \eta_{EE} \\
\Big[\sum_{i \in N_E} \sum_{k \in N_I} z_{ik} z_{ki} \Big] &= N_E \alpha_I g_{EI} g_{IE} \eta_{EI} \\
\Big[\sum_{i \in N_I} \sum_{k \in N_E} z_{ik} z_{ki} \Big] &= N_I \alpha_E g_{IE} g_{EI} \eta_{EI} \\
\Big[\sum_{i \in N_I} \sum_{k \in N_I} z_{ik} z_{ki} \Big] &= N_I \alpha_I g_{II}^2 \eta_{II}.
\end{aligned}
\tag{92}
$$

Combining Eq. (92) and the mean rank-one connectivity loadings Eqs. (52)-(54), we obtain the population covariances as

$$
\begin{aligned}
\sigma_{nm}^E &= \frac{1}{\lambda_0^2} \left(J_E g_{EE}^2 \eta_{EE} - J_I g_{EI} g_{IE} \eta_{EI} \right), \\
\sigma_{nm}^I &= \frac{1}{\lambda_0^2} \left(J_E g_{EI} g_{IE} \eta_{EI} - J_I g_{II}^2 \eta_{II} \right).
\end{aligned}
\tag{93}
$$

We note that the large deviation of the dominant eigenvalue λ in the network with reciprocal motifs also increases the nonlinearity of the vector perturbations. To account for this nonlinearity, we start from Eq. (40) for λ and get $\sigma_{nm} = \lambda - \sum_{p=E,I} \alpha_p \bar{m}^p \bar{n}^p = \lambda - \lambda_0$, then we compare with Eq. (80) and get the approximation relationship

$$
\sigma_{nm} \approx \frac{\theta_2}{\lambda^2} = \frac{1}{N} \frac{(\bar{\mathbf{n}}^\mathsf{T} \mathbf{Z})(\mathbf{Z} \bar{\mathbf{m}})}{\lambda^2}.
\tag{94}
$$

Similarly, for the covariance of each population we have

$$
\begin{aligned}
\sigma_{nm}^E &= \frac{1}{\lambda^2} \left(J_E g_{EE}^2 \eta_{EE} - J_I g_{EI} g_{IE} \eta_{EI} \right), \\
\sigma_{nm}^I &= \frac{1}{\lambda^2} \left(J_E g_{EI} g_{IE} \eta_{EI} - J_I g_{II}^2 \eta_{II} \right).
\end{aligned}
\tag{95}
$$

For sparse connectivity with reciprocal correlations, the calculations are similar, with entries of $\mathbf{Z}$ being Bernoulli-distributed

$$
\begin{aligned}
\Big[\sum_{i \in N_E} \sum_{k \in N_E} z_{ik} z_{ki} \Big] &= N_E^2 A_E^2 c (\rho_{EE} - c) \\
\Big[\sum_{i \in N_E} \sum_{k \in N_I} z_{ik} z_{ki} \Big] &= -N_E N_I A_E A_I c (c - \rho_{EI}) \\
\Big[\sum_{i \in N_I} \sum_{k \in N_E} z_{ik} z_{ki} \Big] &= -N_I N_E A_E A_I c (c - \rho_{EI}) \\
\Big[\sum_{i \in N_I} \sum_{k \in N_I} z_{ik} z_{ki} \Big] &= N_I^2 A_I^2 c (\rho_{II} - c).
\end{aligned}
\tag{96}
$$

and we have the population covariance

$$
\begin{aligned}
\sigma_{nm}^E &= \frac{1}{\lambda^2} N A_E c^2 \left(A_E^2 N_E (\rho_{EE} - c) + A_I^2 N_I (\rho_{EI} - c) \right), \\
\sigma_{nm}^I &= \frac{1}{\lambda^2} N A_I c^2 \left(A_E^2 N_E (\rho_{EI} - c) + A_I^2 N_I (\rho_{II} - c) \right).
\end{aligned}
\tag{97}
$$

Using Eqs. (36), (37), it can be seen that Eq. (97) is equivalent to Eq. (95).

2.6 Dynamics

In this section, we show how approximating locally-defined connectivity by a global low-rank structure allows us to analyse the emerging low-dimensional dynamics. We first summarize the mean-field theory (MFT) for Gaussian-mixture low-rank networks [31, 33]. We then apply it to unit-rank connectivity obtained as an approximation of locally-defined connectivity. We finally compare the resulting description of the dynamics with an alternate mean-field approach for random connectivity consisting of a superposition of low-rank and full-rank random parts as in Eq. (22) [30, 32].

Throughout this study, we consider recurrent networks of rate units with recurrent interactions defined by a connectivity matrix $\mathbf{J}$. The dynamical activity of unit i is represented by a variable $x_i(t)$, which we interpret as the total synaptic input current. The firing rate of unit i is given by $r_i(t) = \phi(x_i(t))$ where $\phi(x) = 1 + \tanh(x - \theta)$ is a positive transfer function. We focus on networks without external inputs, so that the dynamics of synaptic input to neuron i is given by

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1}^{N} J_{ij}\phi(x_i(t)). \tag{98}$$

In Figs 5-7, we compare the dynamics determined by direct simulations for a locally-defined connectivity matrix with a mean-field description obtained for a unit-rank approximation.

2.6.1 Mean-field theory for Gaussian-mixture low-rank connectivity

Here we review the mean-field theory for networks in which the connectivity matrix is exactly low-rank, with components of connectivity vectors moreover drawn from Gaussian-mixture distribution. Previous works have shown that in this case, the dynamics of the collective activity $\mathbf{x}(t) = \{x_i\}_{i=1\dots N}$ are embedded in a linear subspace of dimension R spanned by the connectivity vectors $\mathbf{m}^{(r)}$ [30–33]. Thus, $\mathbf{x}(t)$ can be expressed as

$$\mathbf{x}(t) = \sum_{r=1}^{R} \kappa_r(t)\mathbf{m}^{(r)}, \tag{99}$$

where $\kappa_r(t)$ for $r = 1 \dots R$ are collective latent variables that quantify the components of $\mathbf{x}(t)$ along the connectivity vectors $\mathbf{m}^{(r)}$. We assume that $\mathbf{m}^{(r)}$ are orthogonal to each other, so that $\kappa_r(t)$ can be expressed as

$$\kappa_r(t) = \frac{\mathbf{x}(t)^\mathsf{T}\mathbf{m}^{(r)}}{||\mathbf{m}^{(r)}||^2}. \tag{100}$$

For simplicity, here we moreover assume that the initial value of $\mathbf{x}(t)$ lies in the subspace spanned by the vectors $\mathbf{m}^{(r)}$. More generally, the initial state can be included as an additional input to the dynamics [31, 33].

For a unit rank connectivity $\mathbf{J} = \mathbf{m}\mathbf{n}^\mathsf{T}/N$, there is a single latent variable κ corresponding to the connectivity vector $\mathbf{m}$, and the dynamics of $\mathbf{x}(t)$ is expressed as

$$\mathbf{x}(t) = \kappa(t)\mathbf{m}, \tag{101}$$

with $\kappa(t)$ given by

$$\kappa(t) = \frac{\mathbf{x}(t)^\mathsf{T}\mathbf{m}}{||\mathbf{m}||^2}. \tag{102}$$

Substituting Eq. (101) into Eq. (98) and inserting the unit-rank connectivity, the dynamics of the latent variable κ can be expressed as

$$\dot{\kappa}(t) = -\kappa(t) + \kappa_{rec}(t) \tag{103}$$

where

$$\kappa_{rec}(t) = \frac{1}{N}\sum_{i=1}^{N} n_i \phi(\kappa(t)m_i). \tag{104}$$

The quantity $\kappa_{rec}(t)$ represents the total recurrent input to κ. The sum in the r. h. s. of Eq. (104) can moreover be interpreted as the empirical average of $n_i\phi(\kappa(t)m_i)$ over the neurons in the network. In the limit of large network size N, this average converges to the integral of $n\phi(\kappa(t)m)$ over the distribution $P(m,n)$ of the components of connectivity vectors. For low-rank networks, the mean-field limit corresponds to replacing $\kappa_{rec}(t)$ with this integral [31, 33]:

$$\kappa_{rec} = \int \mathrm{d}m\mathrm{d}n P(m,n)n\phi(\kappa m). \tag{105}$$

In the Gaussian-mixture low-rank model, each neuron i is assigned to a population p for $p = 1\ldots P$. Within each population, the components (m_i, n_i) are generated from a multivariate Gaussian distribution $P^p(m,n)$, that is

$$P^p(m,n) = \mathcal{N}\left(\begin{pmatrix} \bar{m}^p \\ \bar{n}^p \end{pmatrix}, \begin{pmatrix} \sigma_{m^p}^2 & \sigma_{nm}^p \\ \sigma_{nm}^p & \sigma_{n^p}^2 \end{pmatrix} \right). \tag{106}$$

In the mean-field limit, κ_{rec} is therefore given by

$$\kappa_{rec} = \sum_{p=1}^{P} \alpha_p \int \mathrm{d}m\mathrm{d}n P^p(m,n)n\phi(\kappa m), \tag{107}$$

where α_p is the fraction of neurons in population p.

Integrating by parts, κ_{rec} can be re-expressed as (S1 Appendix)

$$\kappa_{rec} = \sum_{p=1}^{P} \alpha_p(\bar{n}^p\langle\phi(\mu_x^p, \Delta_x^p)\rangle + \langle\phi'(\mu_x^p, \Delta_x^p)\rangle\sigma_{nm}^p\kappa). \tag{108}$$

Here μ_x^p, Δ_x^p are the mean and variance of the inputs to population p, given by

$$\begin{aligned} \mu_x^p &= \kappa\bar{m}^p, \\ \Delta_x^p &= \kappa^2\sigma_{m^p}^2, \end{aligned} \tag{109}$$

and the symbol $\langle f(\mu, \Delta)\rangle$ stands for the expected value of a function $f(x)$ with respect to a Gaussian variable x with mean and variance μ, Δ, that is

$$\langle f(\mu, \Delta)\rangle = \int dx(2\pi)^{-1/2}\exp\left(-x^2/2\right)f\left(\mu + \sqrt{\Delta}x\right). \tag{110}$$

Altogether, using MFT for Gaussian-mixture low-rank networks gives the closed dynamics of the latent variable κ:

$$\dot{\kappa} = -\kappa + \sum_{p=1}^{P} \alpha_p(\bar{n}^p\langle\phi(\kappa\bar{m}^p, \kappa^2\sigma_{m^p}^2)\rangle + \langle\phi'(\kappa\bar{m}^p, \kappa^2\sigma_{m^p}^2)\rangle\sigma_{nm}^p\kappa). \tag{111}$$

In particular, the corresponding steady state is given by

$$\kappa = \sum_{p=1}^{P} \alpha_p (\bar{n}^p \langle \phi(\kappa \bar{m}^p, \kappa^2 \sigma^2_{m^p}) \rangle + \langle \phi'(\kappa \bar{m}^p, \kappa^2 \sigma^2_{m^p}) \rangle \sigma^p_{nm} \kappa). \tag{112}$$

Note that the first and second terms on the r. h. s. respectively correspond to the mean and covariance of the entries of the unit-rank connectivity vectors $\mathbf{m}$ and $\mathbf{n}$.

2.6.2 Approximate dynamics for locally-defined connectivity

We next apply the MFT to unit-rank connectivity obtained as an approximation of locally-defined connectivity for the different considered cases.

Independent connectivity. We start from the network with independent connectivity, in which case the unit-rank connectivity vectors obtained by approximating locally-defined connectivity have no covariance, i. e. $\sigma^p_{nm} = 0$ (Methods Sec. 2.5).

The dynamical system for the latent variable κ therefore contains only the mean term

$$\dot{\kappa} = -\kappa + \sum_{p=1}^{P} \alpha_p \bar{n}^p \langle \phi(\kappa \bar{m}^p, \kappa^2 \sigma^2_{m^p}) \rangle. \tag{113}$$

For the Gaussian random model, inserting the expressions for $\bar{m}^p$ and $\bar{n}^p$ (Eqs. (52)-(54)), the fixed point obeys

$$\kappa = J_E \langle \phi(\kappa, \kappa^2 \sigma^2_{m^E}) \rangle - J_I \langle \phi(\kappa, \kappa^2 \sigma^2_{m^I}) \rangle, \tag{114}$$

where the variance $\sigma^2_{m^p}$ of connectivity components m_i is given by Eq. (88).

For the sparse random model, we further consider $\bar{n}^p$ given by Eqs. (36), (52)-(54) and the fixed point is

$$\kappa = c N_E A_E \langle \phi(\kappa, \kappa^2 \sigma^2_{m^E}) \rangle + c N_I A_I \langle \phi(\kappa, \kappa^2 \sigma^2_{m^I}) \rangle, \tag{115}$$

where $\sigma^2_{m^p}$ is obtained from Eq. (89).

Reciprocal motifs. Correlations between reciprocal connections lead to non-zero covariance σ^p_{nm} between the unit-rank connectivity vectors obtained by approximating locally-defined connectivity (Methods Sec. 2.3, Eq. (70)). The dynamical system for the latent variable κ therefore contains both the mean and covariance terms (Eq. (112)).

For the Gaussian random model, combining Eqs. (52)-(54), (88), (95) the fixed point obeys

$$\kappa = \sum_{p=E,I} \alpha_p \Big(\bar{n}^p \langle \phi(\kappa \bar{m}^p, \kappa^2 \sigma^2_{m^p}) \rangle$$
$$+ \frac{1}{\lambda^2} \langle \phi'(\kappa \bar{m}^p, \kappa^2 \sigma^2_{m^p}) \rangle (J_E g_{pE} g_{Ep} \eta_{Ep} - J_I g_{Ip} g_{pI} \eta_{Ip}) \kappa \Big) \tag{116}$$

with the variance $\sigma^2_{m^p}$ of connectivity components m_i given by Eq. (88). For the sparse model, combining Eqs. (36), (52)-(54), (89), (97), the fixed point obeys

$$\kappa = \sum_{p=E,I} \alpha_p \Big(\bar{n}^p \langle \phi(\kappa \bar{m}^p, \kappa^2 \sigma^2_{m^p}) \rangle$$
$$+ \frac{1}{\lambda^2} \langle \phi'(\kappa \bar{m}^p, \kappa^2 \sigma^2_{m^p}) \rangle N A_p c^2 \big(A_E^2 N_E (\rho_{Ep} - c) + A_I^2 N_I (\rho_{Ip} - c) \big) \kappa \Big). \tag{117}$$

with the variance $\sigma^2_{m^p}$ of connectivity components m_i given by Eq. (89).

2.6.3 Mean-field theory for superpositions of low-rank and full rank random connectivity

Here we review an alternate form of mean-field theory for random connectivity consisting of a superposition of a low-rank structure and full-rank random part [30,32]. This form of MFT can be directly applied to independently generated connections, where the connectivity matrix consists precisely of a superposition of a low-rank part corresponding to the mean, and a full-rank random part corresponding to fluctuations (Eqs. (22), (119)). Extending this type of MFT to the situation where reciprocal connections are present is however challenging [18]. Moreover, in contrast to the case where connectivity is exactly low-rank, when the additional full-rank random part is present the mean-field theory describes only the steady-state activity (and linearized dynamics around it), but not the full dynamics as in Eq. (98).

The key assumption of MFT for randomly connected networks is that the total input x_i to each unit can be approximated as a stochastic Gaussian process [52]. The first two cumulants (mean and variance) of that Gaussian process are then computed self-consistently to characterize the steady-state activity.

At a fixed point, the total input x_i obeys

$$x_i = \sum_{j=1}^{N} J_{ij}\phi(x_j). \tag{118}$$

Replacing J_{ij}, where i, j belong to populations p, q respectively, by the superposition of rank-one mean and full-rank random connectivity components $\bar{m}^p\bar{n}^q/N + z_{ij}$ we get

$$x_i = \frac{\bar{m}^p}{N} \sum_{q=1...P} \bar{n}^q \sum_{j \in N_q} \phi(x_j(t)) + \sum_{j=1}^{N} z_{ij}\phi(x_j). \tag{119}$$

Denoting by $[\cdot]$ the average over the distribution of x_i, the mean of x_i can then be expressed as

$$[x_i] = \bar{m}^p\bar{\kappa} \tag{120}$$

where we introduced

$$\bar{\kappa} = \sum_{p=1}^{P} \bar{n}^p \sum_{i \in N_p} [\phi(x_i)]/N, \tag{121}$$

$$\tag{122}$$

and we assumed that the zero-mean random connectivity z_{ij} is uncorrelated with the firing rate $\phi(x_j)$, so that

$$\sum_{j=1}^{N} [z_{ij}\phi(x_j)] = 0. \tag{123}$$

Similarly, the correlation between x_i and x_j, where $i \in N_p$ and $j \in N_q$, is given by $\quad$ 925

$$
\begin{aligned}
[x_i x_j] &= \frac{\bar{m}^p}{N} \sum_{s=1}^{P} \bar{n}^s \sum_{k \in N_s} [\phi(x_k)] \frac{\bar{m}^q}{N} \sum_{t=1}^{P} \bar{n}^t \sum_{l \in N_t} [\phi(x_l)] \\
&+ \frac{\bar{m}^p}{N} \sum_{s=1}^{P} \bar{n}^s \sum_{k \in N_s} [\phi(x_k)] \sum_{l=1}^{N} z_{jl} [\phi(x_l)] \\
&+ \frac{\bar{m}^q}{N} \sum_{t=1}^{P} \bar{n}^t \sum_{l \in N_t} [\phi(x_l)] \sum_{k=1}^{N} z_{ik} [\phi(x_k)] \\
&+ \Big[\sum_{l=1}^{N} z_{jl} \phi(x_l) \sum_{k=1}^{N} z_{ik} \phi(x_k) \Big] \\
&= \bar{m}^p \bar{m}^q \bar{\kappa}^2 + \delta_{ij} \sum_{k=1}^{N} [z_{ik}^2][\phi^2(x_k)]
\end{aligned}
\tag{124}
$$

where we assume the neuronal activities are decorrelated $[\phi(x_i)\phi(x_j)] = [\phi(x_i)][\phi(x_j)]$ $\quad$ 926
when $i \neq j$. This assumption holds for independently-generated connections, but not in $\quad$ 927
presence of reciprocal correlations [18]. The covariance between x_i and x_j therefore $\quad$ 928
becomes $\quad$ 929

$$
[x_i x_j] - [x_i][x_j] = \delta_{ij} \sum_{k=1}^{N} [z_{ik}^2][\phi^2(x_k)].
\tag{125}
$$

Within the mean-field approximation, neuronal activation x_i are therefore $\quad$ 930
uncorrelated Gaussian variables with mean and variance given by Eqs. 120 and 125 $\quad$ 931

$$
\begin{aligned}
\mu_{x_i} &:= [x_i] = \bar{m}^p \bar{\kappa}, \\
\Delta_{x_i} &:= [x_i^2] - [x_i]^2 = \delta_{ij} \sum_{k=1}^{N} [z_{ik}^2][\phi^2(x_k)].
\end{aligned}
\tag{126}
$$

To determine $\bar{\kappa}$ and $[\phi(x_k)^2]$, we finally express Eqs. (121) and (126) as Gaussian $\quad$ 932
integrals over x_i in population p: $\quad$ 933

$$
\begin{aligned}
\bar{\kappa} &= \sum_{q=1}^{P} \alpha_q \bar{n}^q \langle \phi(\mu_x^q, \Delta_x^q) \rangle, \\
\sum_{k=1}^{N} [z_{ik}^2][\phi(x_k)^2] &= \sum_{q=1}^{P} N_q \sigma_{z_{pq}}^2 \langle \phi(\mu_x^q, \Delta_x^q)^2 \rangle.
\end{aligned}
\tag{127}
$$

Here we replaced $\bar{m}^p = 1$ and $\sum_{k=1}^{N} [z_{ik}^2] f(\cdot) = \sum_{q=1}^{P} N_q \sigma_{z_{pq}}^2 f(\cdot)$ given the eigenvector $\quad$ 934
normalization in Eq. (52), and the assumption that variances $[z_{ik}^2]$ depend on the $\quad$ 935
populations the units i and k belong to (Eqs. (24), (28), (52)). Therefore, the $\quad$ 936
stationary mean and variance of the dynamics of synaptic inputs in population p are $\quad$ 937

$$
\begin{aligned}
\mu_x^p &:= \bar{m}^p \bar{\kappa}, \\
\Delta_x^p &:= \sum_{q=1}^{P} N_q \sigma_{z_{pq}}^2 \langle \phi(\mu_x^q, \Delta_x^q)^2 \rangle.
\end{aligned}
\tag{128}
$$

Eqs. (128), (127) give the self-consistent equations for the stationary solutions of the $\quad$ 938
dynamics. $\quad$ 939

More specifically, in the Gaussian random model, we combine connectivity statistics given by Eqs. (24), (52)-(54), so that we have

$$
\begin{aligned}
\mu_x^p &:= J_E \langle \phi(\mu_x^E, \Delta_x^E) \rangle - J_I \langle \phi(\mu_x^I, \Delta_x^I) \rangle, \\
\Delta_x^p &:= \sum_{q=E,I} \alpha_q g_{pq}^2 \langle \phi^2(\mu_x^q, \Delta_x^q) \rangle,
\end{aligned}
\tag{129}
$$

while for the sparse random model, we combine connectivity statistics given by Eqs. (27), (28), (36), (52)-(54), so that we have

$$
\begin{aligned}
\mu_x^p &:= cN_E A_E \langle \phi(\mu_x^E, \Delta_x^E) \rangle + cN_I A_I \langle \phi(\mu_x^I, \Delta_x^I) \rangle, \\
\Delta_x^p &:= \sum_{q=E,I} N_q c(1-c) A_q^2 \langle \phi^2(\mu_x^q, \Delta_x^q) \rangle.
\end{aligned}
\tag{130}
$$

Code Availability

Code will be made available upon publication.

Supporting information

S1 Appendix. Dynamics in Gaussian-mixture low-rank networks. Here, we provide the derivation for the dynamics of the latent variable κ (Eq. (111)) in the Gaussian-mixture low-rank network model. We consider a rank-one connectivity consisting of P populations, with the neurons in each population accounting for a α_p percentage of all neurons. The entries on the left and right eigenvectors $\mathbf{n}^p$, $\mathbf{m}^p$ assigned to population p are sampled from a multivariate Gaussian distribution with mean $\bar{n}^p$, $\bar{m}^p$, variance $\sigma_{n^p}^2$, $\sigma_{m^p}^2$ and covariance σ_{nm}^p (Methods Sec. 2.3) So, κ_{rec} in Eq. (107) is further decomposed into two integrals involving the contributions from the mean and random connected components, respectively,

$$
\begin{aligned}
\kappa_{rec} = &\sum_{p=1}^{P} \alpha_p \int \mathrm{d}m P^p(m) \bar{n}^p \phi(\kappa m) \\
&+ \sum_{p=1}^{P} \alpha_p \int \mathrm{d}m \mathrm{d}n P^p(m,n)(n - \bar{n}^p) \phi(\kappa m),
\end{aligned}
\tag{131}
$$

here, $P^p(m)$ represents the marginal Gaussian distribution for connectivity loadings m_i of neurons in population p

$$
P^p(m) = \mathcal{N}(\bar{m}^p, \sigma_{m^p}^2),
\tag{132}
$$

and $P^p(m,n)$ represents the multivariate Gaussian distribution for connectivity loadings m_i, n_i of neurons in population p (Eq. (106)).

Note that the synaptic input to neurons in population p is a scaled Gaussian variable $x_i = \kappa m_i$ corresponding to the Gaussian loadings $m_i \sim \mathcal{N}(\bar{m}^p, \sigma_{m^p}^2)$, so it conforms to a Gaussian distribution $\mathcal{N}(\mu_x^p, \Delta_x^p)$ with the population-averaged mean and variance

$$
\begin{aligned}
\mu_x^p &= \kappa \bar{m}^p, \\
\Delta_x^p &= \kappa^2 \sigma_{m^p}^2.
\end{aligned}
\tag{133}
$$

Thus, the first term in Eq. (131) is

$$
\int \mathrm{d}m \mathcal{N}(\bar{m}^p, \sigma_{m^p}^2) \phi(\kappa m) = \langle \phi(\kappa \bar{m}^p, \kappa^2 \sigma_{m^p}^2) \rangle = \langle \phi(\mu_x^p, \Delta_x^p) \rangle,
\tag{134}
$$

a Gaussian integral term.

We then use Stein's lemma

$$\int \mathcal{D}z\, z f(z) = \int \mathcal{D}z \frac{\mathrm{d}}{\mathrm{d}z} f(z) \tag{135}$$

and further replace the multivariate Gaussian distribution by Eq. (106), to compute the second term in Eq. (131) attributed to the random connectivity component as

$$\int \mathrm{d}m \mathrm{d}n \mathcal{N}\left(\begin{pmatrix} \bar{m}^p \\ \bar{n}^p \end{pmatrix}, \begin{pmatrix} \sigma_{m^p}^2 & \sigma_{nm}^p \\ \sigma_{nm}^p & \sigma_{n^p}^2 \end{pmatrix} \right)(n - \bar{n}^p)\phi(\kappa m)$$
$$= \langle \phi(\kappa\bar{m}^p, \kappa^2\sigma_{m^p}^2)\rangle \sigma_{nm}^p \kappa \tag{136}$$
$$= \langle \phi(\mu_x^p, \Delta_x^p)\rangle \sigma_{nm}^p \kappa.$$

Finally, combining the contributions from both the mean and random components, we re-express κ_{rec} in Eq. (107) and retrieve Eq. (111).

S2 Appendix. Linear stability at fixed points in rank-one networks. To determine the stability of fixed points of the network with rank-one connectivity structure, we consider the rank-one connectivity $\mathbf{J}_{R1}$ and study the stability of the one-dimensional latent dynamical variable κ in the neighbourhood area of its fixed point κ^0. We set the fixed point of synaptic input x_i^0, define the perturbation of the latent variable κ^1 and the perturbation of the synaptic input $x_i^1 = \kappa^1 m_i$, the temporal evolution of κ^1 is expressed as

$$\dot{\kappa}^1 = -\kappa^1 + \langle n_i[\phi'(x_i^0)x_i^1]\rangle$$
$$= -\kappa^1 + \sum_{p=E,I} \alpha_p \langle n_i^p \phi'(x_i^{p,0})x_i^{p,1}\rangle, \tag{137}$$

for the rank-one approximation network $\mathbf{J}_{R1} = \mathbf{m}\mathbf{n}^\intercal/N$, we remove symbol $[\cdot]$. Next, in the Gaussian-mixture low-rank framework, the entries n_i^p and m_i^p are jointly sampled from a bivariate Gaussian distribution characterized by means, variances and covariances, $\bar{m}^p$, $\bar{n}^p$ and $\sigma_{m^p}^2$, $\sigma_{n^p}^2$, σ_{nm}^p. Using a similar approach to previous studies [30, 31], we set

$$n_i^p = \bar{n}^p + \left(\sqrt{1-\gamma}\zeta_n^p + \sqrt{\gamma}\zeta^p\right)\sigma_{n^p}$$
$$m_i^p = \bar{m}^p + \left(\sqrt{1-\gamma}\zeta_m^p + \sqrt{\gamma}\zeta^p\right)\sigma_{m^p} \tag{138}$$
$$\sigma_{nm}^p = \gamma\sigma_{n^p}\sigma_{m^p}$$

where ζ_m^p, ζ_n^p and ζ^p are three independent normal random variables $\mathcal{N}(0,1)$. By substituting variables in Eq. (138), after some linear algebra, we get

$$\langle n_i^p \phi'(x_i^{p,0})x_i^{p,1}\rangle = \langle n_i^p \phi'(x_i^{p,0})\kappa^1 m_i^{p,1}\rangle$$
$$= \left((\bar{n}^p\bar{m}^p + \sigma_{nm}^p)\langle\phi'\rangle + \bar{n}^p\sigma_{m^p}^2\langle\phi''\rangle\kappa^0 + \bar{m}^p\sigma_{nm}^p\langle\phi''\rangle\kappa^0 \right. \tag{139}$$
$$\left. + \sigma_{m^p}^2\sigma_{nm}^p\langle\phi'''\rangle(\kappa^0)^2\right)\kappa^1.$$

We finally obtain the time evolution of κ

$$\dot{\kappa}^1 = -\kappa^1 + \sum_{p=E,I} \alpha_p\big((\bar{n}^p\bar{m}^p + \sigma_{nm}^p)\langle\phi'\rangle + \bar{n}^p\sigma_{m^p}^2\langle\phi''\rangle\kappa^0$$
$$+ \bar{m}^p\sigma_{nm}^p\langle\phi''\rangle\kappa^0 + \sigma_{m^p}^2\sigma_{nm}^p\langle\phi'''\rangle(\kappa^0)^2\big)\kappa^1. \tag{140}$$

The Jacobian for the latent dynamical variable's fixed point κ^0 is

$$
S_\kappa = -1 + \sum_{p=E,I} \alpha_p \big((\bar{n}^p \bar{m}^p + \sigma_{nm}^p) \langle \phi' \rangle + \bar{n}^p \sigma_{m^p}^2 \langle \phi'' \rangle \kappa^0 \\
+ \bar{m}^p \sigma_{nm}^p \langle \phi'' \rangle \kappa^0 + \sigma_{m^p}^2 \sigma_{nm}^p \langle \phi''' \rangle (\kappa^0)^2 \big),
\tag{141}
$$

which is a scalar in our rank-one network. The fixed point is stable when $S_\kappa < 0$ and unstable when $S_\kappa > 0$. All the above parameters are expressed and calculated in the text, so that we can check the stability of each fixation point using the formulas above.

S3 Appendix.   Comparison between dynamics in full-rank connectivity with rank-one approximation. Here, we compare the dynamics solved in the full-rank network with dynamics solved by the rank-one approximation, considering the limitations of applying the classical MFT, we only discuss networks with independent connections.

Using the rank-one approximation, global dynamics is characterized by the latent dynamical variable κ, which satisfies

$$
\kappa = J_E \langle \phi(\kappa, \kappa^2 \sigma_{m^E}^2) \rangle - J_I \langle \phi(\kappa, \kappa^2 \sigma_{m^I}^2) \rangle.
\tag{142}
$$

Combining the connectivity statistics given in Eqs. (52)-(54), (88), we thus express the self-consistent equations for the population mean and variance of the synaptic inputs x_i using κ (Eq. (109))

$$
\mu_x^E = \mu_x^I = J_E \langle \phi(\mu_x^E, \Delta_x^E) \rangle - J_I \langle \phi(\mu_x^I, \Delta_x^I) \rangle
\tag{143}
$$

and

$$
\Delta_x^E = \frac{1}{\lambda_0^2} \big(J_E \langle \phi(\mu_x^E, \Delta_x^E) \rangle - J_I \langle \phi(\mu_x^I, \Delta_x^I) \rangle \big)^2 \big(\alpha_E g_{EE}^2 + \alpha_I g_{EI}^2 \big) \\
\Delta_x^I = \frac{1}{\lambda_0^2} \big(J_E \langle \phi(\mu_x^E, \Delta_x^E) \rangle - J_I \langle \phi(\mu_x^I, \Delta_x^I) \rangle \big)^2 \big(\alpha_E g_{IE}^2 + \alpha_I g_{II}^2 \big).
\tag{144}
$$

Comparing with the self-consistent equations for population mean and variance in the full-rank network Eq. (129), we find in particular that the expression for μ_x^p is the same in both representations, but the expressions for the variance Δ_x^p are different. Specifically, the low-rank approximation shows that the different heterogeneity between excitatory and inhibitory populations depends only on the block-structured variances, i. e. , $\sum_{q=E,I} \alpha_p g_{pq}^2 / \lambda_0^2$, and independent of the historical population activity. The full-rank dynamics, on the other hand, shows that the different heterogeneity depends on the local relationship between block structured variance and the structure of historical population activity [68].

For a simplified network example, where the locally defined connectivity has homogeneous random parameters $g_{pq} = g, \ p = E, I$, the variance of the rank-one perturbation eigenvector $\sigma_{m^p}^2$ is thereby the same for both excitatory and inhibitory populations. Because $\lambda_0 = J_E - J_I$, means and variances in Eqs. (143), (144) are

$$
\mu_x = \mu_x^E = \mu_x^I = \lambda_0 \langle \phi(\mu_x, \Delta_x) \rangle, \\
\Delta_x = \Delta_x^E = \Delta_x^I = g^2 \langle \phi(\mu_x, \Delta_x) \rangle^2.
\tag{145}
$$

So that, for this simple example, the heterogeneity of dynamics in the full-rank approximation is $g^2 \langle (\phi(\mu_x, \Delta_x))^2 \rangle$ (Eq. (129)), while the heterogeneity in the rank-one approximation is $g^2 \langle \phi(\mu_x, \Delta_x) \rangle^2$, the difference does not substantially change the bistable transition and the performance of the low-rank approximation.

S4 Appendix. Chaotic dynamical transition point for networks with independent connectivity. Conventionally, the collective dynamics of random neural network becomes chaotic when the effect of random component is too strong. Based on previous study [11], when the effective random gain, also known as the radius of the random eigenvalue bulk r_g, exceed the critical threshold 1, the network enters the chaotic dynamical state.

Considering the block-like E-I random network with i. i. d. random connectivity, we define a $\mathbb{R}^{2 \times 2}$ matrix with elements $\alpha_q g_{pq}^2$

$$\mathbf{M}_g = \begin{bmatrix} \alpha_E g_{EE}^2 & \alpha_I g_{EI}^2 \\ \alpha_E g_{IE}^2 & \alpha_I g_{II}^2 \end{bmatrix}, \tag{146}$$

the first eigenvalue of $\mathbf{M}_g$ determines the radius of the continuous eigenvalues bulk of $\mathbf{J}$, that is

$$r_g = \sqrt{\lambda_{M_g}}. \tag{147}$$

For the sparse E-I network, considering the relationships given by Eq. (36), we calculate the radius for the sparse network as

$$r_g = \sqrt{(N_E A_E^2 + N_I A_I^2)c(1-c)}. \tag{148}$$

S5 Appendix. Local connectivity statistics in rank-R Gaussian-mixture models. We show that the statistical properties of the entries on the rank-R connectivity vectors $\bar{m}^{(r),p}$, $\bar{n}^{(r),p}$, $\sigma'_{m_r^p}$ $\sigma_{n_r^p}^2$ and $\sigma_{n_r m_{r'}}^p$ directly determine the means $\bar{J}_{pq}$, variances $\sigma_{z_{pq}}^2$ and reciprocal correlations η_{pq} of the resulting local synaptic weights J_{ij}, where $i \in N_p^r$, $j \in N_q$. Considering Eq. (2), the cell-type-dependent mean

$$\bar{J}_{pq} = \frac{1}{N} \sum_{r=1}^{R} \bar{m}_i^{(r)} \bar{n}_j^{(r)} = \frac{1}{N} \sum_{r=1}^{R} \bar{m}^{(r),p} \bar{n}^{(r),q}, \tag{149}$$

and the cell-type-dependent variance of locally defined connections is

$$\sigma_{z_{pq}}^2 = \frac{1}{N^2} \Big[\Big(\sum_{r=1}^{R} m_i^{(r)} n_j^{(r)} - \sum_{r=1}^{R} \bar{m}^{(r),p} \bar{n}^{(r),q} \Big)^2 \Big]$$
$$= \frac{1}{N^2} \Big[\Big(\sum_{r=1}^{R} \Delta m_i^{(r)} \bar{n}^{(r),q} + \bar{m}^{(r),p} \Delta n_j^{(r)} + \Delta m_i^{(r)} \Delta n_j^{(r)} \Big)^2 \Big] \tag{150}$$
$$= \frac{1}{N^2} \sum_{r=1}^{R} \big(\sigma_{m_r^p}^2 (\bar{n}^{(r),q})^2 + (\bar{m}^{(r),p})^2 \sigma_{n_r^q}^2 + \sigma_{m_r^p}^2 \sigma_{n_r^q}^2 \big).$$

Finally the correlation between the pairwise weights J_{ij}, J_{ji} is computed as

$$\big[(J_{ij} - [J_{ij}])(J_{ji} - [J_{ji}]) \big]_J$$
$$= \frac{1}{N^2} \Big[\sum_{r,r'=1}^{R} m_i^{(r)} n_j^{(r)} m_j^{(r')} n_i^{(r')} - \sum_{r,r'=1}^{R} m_i^{(r)} n_j^{(r)} \bar{m}^{(r'),q} \bar{n}^{(r'),p} $$
$$- \sum_{r,r'=1}^{R} \bar{m}^{(r),p} \bar{n}^{(r),q} m_j^{(r')} n_i^{(r')} + \sum_{r,r'=1}^{R} \bar{m}^{(r),p} \bar{n}^{(r),q} \bar{m}^{(r'),q} \bar{n}^{(r'),p} \Big] \tag{151}$$
$$= \frac{1}{N^2} \sum_{r,r'=1}^{R} \big(\sigma_{n_{r'} m_r}^p \bar{m}^{(r'),q} \bar{n}^{(r),q} + \sigma_{n_r m_{r'}}^q \bar{m}^{(r),p} \bar{n}^{(r'),p} + \sigma_{n_r m_{r'}}^q \sigma_{n_{r'} m_r}^p \big).$$

Substituting Eqs. (150), (151) into Eq. (29) leads to the resulting reciprocal correlation η_{pq}.

Acknowledgments

References

1. Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB. Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS biology. 2005;3(3):e68.

2. Perin R, Berger TK, Markram H. A synaptic organizing principle for cortical neuronal groups. Proceedings of the National Academy of Sciences. 2011;108(13):5419–5424.

3. Barbour B, Brunel N, Hakim V, Nadal JP. What can we learn from synaptic weight distributions? TRENDS in Neurosciences. 2007;30(12):622–629.

4. Ko H, Hofer SB, Pichler B, Buchanan KA, Sjöström PJ, Mrsic-Flogel TD. Functional specificity of local synaptic connections in neocortical networks. Nature. 2011;473(7345):87–91.

5. Harris KD, Mrsic-Flogel TD. Cortical connectivity and sensory coding. Nature. 2013;503(7474):51–58.

6. Seeman SC, Campagnola L, Davoudian PA, Hoggarth A, Hage TA, Bosma-Moody A, et al. Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex. elife. 2018;7:e37349.

7. Campagnola L, Seeman SC, Chartrand T, Kim L, Hoggarth A, Gamlin C, et al. Local Connectivity and Synaptic Dynamics in Mouse and Human Neocortex. bioRxiv. 2021;.

8. Fino E, Yuste R. Dense inhibitory connectivity in neocortex. Neuron. 2011;69(6):1188–1203.

9. Vegué M, Perin R, Roxin A. On the structure of cortical microcircuits inferred from small sample sizes. Journal of Neuroscience. 2017;37(35):8498–8510.

10. Trousdale J, Hu Y, Shea-Brown E, Josić K. Impact of network structure and cellular response on spike time correlations. PLoS computational biology. 2012;8(3):e1002408.

11. Aljadeff J, Stern M, Sharpee T. Transition to chaos in random networks with cell-type-specific connectivity. Physical review letters. 2015;114(8):088101.

12. Aljadeff J, Renfrew D, Stern M. Eigenvalues of block structured asymmetric random matrices. Journal of Mathematical Physics. 2015;56(10):103502.

13. Aljadeff J, Renfrew D, Vegué M, Sharpee TO. Low-dimensional dynamics of structured random networks. Physical Review E. 2016;93(2):022302.

14. Recanatesi S, Ocker GK, Buice MA, Shea-Brown E. Dimensionality in recurrent spiking networks: Global trends in activity and local origins in connectivity. PLoS computational biology. 2019;15(7):e1006446.

15. Dahmen D, Recanatesi S, Jia X, Ocker GK, Campagnola L, Jarsky T, et al. Strong and localized coupling controls dimensionality of neural activity across brain areas. bioRxiv. 2021; p. 2020–11.

16. Potjans TC, Diesmann M. The cell-type specific cortical microcircuit: relating structure and activity in a full-scale spiking network model. Cerebral cortex. 2014;24(3):785–806.

17. Brunel N. Is cortical connectivity optimized for storing information? Nature neuroscience. 2016;19(5):749–755.

18. Martí D, Brunel N, Ostojic S. Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks. Physical Review E. 2018;97(6):062314.

19. Guzman SJ, Schlögl A, Frotscher M, Jonas P. Synaptic mechanisms of pattern completion in the hippocampal CA3 network. Science. 2016;353(6304):1117–1123.

20. Lizier JT, Atay FM, Jost J. Information storage, loop motifs, and clustered structure in complex networks. Physical Review E. 2012;86(2):026110.

21. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature reviews neuroscience. 2009;10(3):186–198.

22. Hu Y, Trousdale J, Josić K, Shea-Brown E. Motif statistics and spike correlations in neuronal networks. Journal of Statistical Mechanics: Theory and Experiment. 2013;2013(03):P03012.

23. Hu Y, Brunton SL, Cain N, Mihalas S, Kutz JN, Shea-Brown E. Feedback through graph motifs relates structure and function in complex networks. Physical Review E. 2018;98(6):062312.

24. Ocker GK, Hu Y, Buice MA, Doiron B, Josić K, Rosenbaum R, et al. From the statistics of connectivity to the statistics of spike times in neuronal networks. Current opinion in neurobiology. 2017;46:109–119.

25. Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the national academy of sciences. 1982;79(8):2554–2558.

26. Amit DJ, Gutfreund H, Sompolinsky H. Storing infinite numbers of patterns in a spin-glass model of neural networks. Physical Review Letters. 1985;55(14):1530.

27. Boerlin M, Machens CK, Denève S. Predictive coding of dynamical variables in balanced spiking networks. PLoS computational biology. 2013;9(11):e1003258.

28. Eliasmith C, Anderson CH. Neural engineering: Computation, representation, and dynamics in neurobiological systems. MIT press; 2003.

29. Sussillo D, Abbott LF. Generating coherent patterns of activity from chaotic neural networks. Neuron. 2009;63(4):544–557.

30. Mastrogiuseppe F, Ostojic S. Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. Neuron. 2018;99(3):609–623.

31. Beiran M, Dubreuil A, Valente A, Mastrogiuseppe F, Ostojic S. Shaping dynamics with multiple populations in low-rank recurrent networks. Neural Computation. 2021;33(6):1572–1615.

32. Schuessler F, Dubreuil A, Mastrogiuseppe F, Ostojic S, Barak O. Dynamics of random recurrent networks with correlated low-rank structure. Physical Review Research. 2020;2(1):013111.

33. Dubreuil A, Valente A, Beiran M, Mastrogiuseppe F, Ostojic S. The role of population structure in computations through neural dynamics. Nature Neuroscience. 2022; p. 1–12.

34. Pereira U, Brunel N. Attractor dynamics in networks with learning rules inferred from in vivo data. Neuron. 2018;99(1):227–238.

35. Landau ID, Sompolinsky H. Coherent chaos in a recurrent neural network with structured connectivity. PLoS computational biology. 2018;14(12):e1006309.

36. Landau ID, Sompolinsky H. Macroscopic fluctuations emerge in balanced networks with incomplete recurrent alignment. Physical Review Research. 2021;3(2):023171.

37. Beiran M, Meirhaeghe N, Sohn H, Jazayeri M, Ostojic S. Parametric control of flexible timing through low-dimensional neural manifolds. Available at SSRN 3967676. 2021;.

38. Kadmon J, Timcheck J, Ganguli S. Predictive coding in balanced neural networks with noise, chaos and delays. Advances in neural information processing systems. 2020;33:16677–16688.

39. Logiaco L, Abbott L, Escola S. Thalamic control of cortical dynamics in a model of flexible motor sequencing. Cell reports. 2021;35(9):109090.

40. Ahmadian Y, Fumarola F, Miller KD. Properties of networks with partially structured and partially random connectivity. Physical Review E. 2015;91(1):012820.

41. Murphy BK, Miller KD. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. Neuron. 2009;61(4):635–648.

42. Pettine WW, Louie K, Murray JD, Wang XJ. Excitatory-inhibitory tone shapes decision strategies in a hierarchical neural network model of multi-attribute choice. PLoS computational biology. 2021;17(3):e1008791.

43. Brunel N. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. Journal of computational neuroscience. 2000;8(3):183–208.

44. Rajan K, Abbott LF. Eigenvalue spectra of random matrices for neural networks. Physical review letters. 2006;97(18):188104.

45. Tao T. Outliers in the spectrum of iid matrices with bounded rank perturbations. Probability Theory and Related Fields. 2013;155(1):231–263.

46. Rivkind A, Barak O. Local dynamics in trained recurrent neural networks. Physical review letters. 2017;118(25):258101.

47. Greenbaum A, Li Rc, Overton ML. First-order perturbation theory for eigenvalues and eigenvectors. SIAM review. 2020;62(2):463–482.

48. Ginibre J. Statistical ensembles of complex, quaternion, and real matrices. Journal of Mathematical Physics. 1965;6(3):440–449.

August 25, 2022

49. Tao T, Vu V, Krishnapur M. Random matrices: Universality of ESDs and the circular law. The Annals of Probability. 2010;38(5):2023–2065.

50. Girko VL. Circular law. Theory of Probability & Its Applications. 1985;29(4):694–706.

51. Sommers HJ, Crisanti A, Sompolinsky H, Stein Y. Spectrum of Large Random Asymmetric Matrices. Phys Rev Lett. 1988;60:1895–1898. doi:10.1103/PhysRevLett.60.1895.

52. Sompolinsky H, Crisanti A, Sommers HJ. Chaos in random neural networks. Physical review letters. 1988;61(3):259.

53. Ostojic S. Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. Nature neuroscience. 2014;17(4):594–600.

54. Herbert E, Ostojic S. The impact of sparsity in low-rank recurrent neural networks. bioRxiv. 2022;.

55. Markovsky I. Low rank approximation: algorithms, implementation, applications. vol. 906. Springer; 2012.

56. Eckart C, Young G. The approximation of one matrix by another of lower rank. Psychometrika. 1936;1(3):211–218.

57. Farrell BF, Ioannou PJ. Accurate low-dimensional approximation of the linear dynamics of fluid flow. Journal of the atmospheric sciences. 2001;58(18):2771–2789.

58. Durstewitz D. A state space approach for piecewise-linear recurrent neural networks for identifying computational dynamics from neural measurements. PLoS computational biology. 2017;13(6):e1005542.

59. Schaeffer R, Khona M, Meshulam L, Fiete IR, et al. Reverse-engineering recurrent neural network solutions to a hierarchical inference task for mice. bioRxiv. 2020;.

60. Bondanelli G, Deneux T, Bathellier B, Ostojic S. Network dynamics underlying OFF responses in the auditory cortex. Elife. 2021;10:e53151.

61. Langdon C, Engel TA. Latent circuit inference from heterogeneous neural responses during cognitive tasks. bioRxiv. 2022;.

62. Hu Y, Sompolinsky H. The spectrum of covariance matrices of randomly connected recurrent neuronal networks with linear dynamics. PLoS computational biology. 2022;18(7):e1010327.

63. Clark DG, Abbott L, Litwin-Kumar A. Dimension of Activity in Random Neural Networks. arXiv preprint arXiv:220712373. 2022;.

64. Stewart GW. Matrix Algorithms: Volume II: Eigensystems. SIAM; 2001.

65. Buzsáki G, Mizuseki K. The log-dynamic brain: how skewed distributions affect network operations. Nature Reviews Neuroscience. 2014;15(4):264–278.

66. Perin R, Berger TK, Markram H. A synaptic organizing principle for cortical neuronal groups. Proceedings of the National Academy of Sciences. 2011;108(13):5419–5424.

67. Kuczala A, Sharpee TO. Eigenvalue spectra of large correlated random matrices. Physical Review E. 2016;94(5):050101.

68. Valente A, Ostojic S, Pillow J. Probing the relationship between linear dynamical systems and low-rank recurrent neural network models. Neural Computation. 2022;34(9):1871–1892.