How strong are correlations in strongly recurrent neuronal networks?

Ran Darshan[1,2,3], Carl van Vreeswijk[3] and David Hansel[3*]

1. ELSC, The Hebrew University of Jerusalem, Israel;
2. Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, USA;
3. Center for Neurophysics, Physiology and Pathology; Cerebral Dynamics,
Plasticity and Learning Research Team; CNRS-UMR8119, Paris, France

ABSTRACT

Cross-correlations in the activity in neural networks are commonly used to characterize their dynamical states and their anatomical and functional organizations. Yet, how these latter network features affect the spatiotemporal structure of the correlations in recurrent networks is not fully understood. Here, we develop a general theory for the emergence of correlated neuronal activity from the dynamics in strongly recurrent networks consisting of several populations of binary neurons. We apply this theory to the case in which the connectivity depends on the anatomical or functional distance between the neurons. We establish the architectural conditions under which the system settles into a dynamical state where correlations are strong, highly robust and spatially modulated. We show that such strong correlations arise if the network exhibits an *effective* feedforward structure. We establish how this feedforward structure determines the way correlations scale with the network size and the degree of the connectivity. In networks lacking an effective feedforward structure correlations are extremely small and only weakly depend on the number of connections per neuron. Our work shows how strong correlations can be consistent with highly irregular activity in recurrent networks, two key features of neuronal dynamics in the central nervous system.

I. INTRODUCTION

Two-point correlations are commonly used to characterize collective dynamics in extended systems [1–7]. Recent technical advances [8–10] make it possible to simultaneously record the activity of many neurons in networks in the brain. This allows for the measurement of two-point correlations for large numbers of neuronal pairs in spontaneous activity as well as upon sensory stimulation or in controlled behavioral conditions.

Correlations in neuronal activity impact the ability of networks to encode information [11–14]. Correlations are also functionally important in performing sensory, motor, or cognitive tasks [15]. For instance, correlated oscillatory activity has been hypothesized to be involved in visual perception [16]. In a recent study combining modeling, electrophysiology and analysis of behavior, we

argued that non-oscillatory correlated neuronal activity in the central nervous system is essential in the generation of exploratory behavior [17]. Correlations are also important for the self-organization of neuronal networks through activity dependent plasticity [18]. Indeed, changes in synaptic strength are thought to depend on the temporal correlation of the activity of the pre and postsynaptic neurons [19, 20].

Correlation strengths depend on the brain area, [21, 22], the layer in cortex [23], stimulus conditions, behavioral states [24] and experience [25–27]. A wide range of values for correlation coefficients, from negligible [21, 28] to substantial [11, 29–33] have been reported in the last two decades. Correlation coefficients are usually higher for close-by neurons than for neurons that are far apart [31, 34–37]. In cortex, they drop significantly over distances of $200 - 400~\mu m$ [35, 37]. Recent works have reported correlations varying non-monotonically with distance [38] or correlations which are positive for close-by neurons but negative for neurons farther apart [39]. Correlation coefficients also depend on functional properties of the neurons. Neurons which code for similar features of sensory stimuli are more correlated [21, 31, 34, 37, 40].

Neurons in cortex receive recurrent inputs from several hundreds [41] to a few thousands [42] of other neurons. Individual connections can induce post-synaptic potentials in a range of 0.1 mV to several mV [35]. Thus, 50 simultaneous inputs are sufficient to trigger or suppress a spike in a neuron. These facts have been incorporated in model networks with strongly recurrent connectivity in which connection strengths are $\mathcal{O}(1/\sqrt{K})$, where K is the average number of inputs per neuron. This scaling is in contrast to the one used in standard mean-field models (e.g. [43]) where connections are $\mathcal{O}(1/K)$ and thus are much weaker. Recent experiments in cortical cultures are consistent with a $1/\sqrt{K}$ scaling of connection strength [44].

van Vreeswijk and Sompolinsky [45, 46] showed that strongly recurrent networks consisting of two populations of neurons, one excitatory (E) and one inhibitory (I), randomly connected on a directed Erdös-Rényi graph, operate in a state in which the strong excitation is balanced by the strong inhibition. In this *balanced regime*, neurons receive strong excitatory and inhibitory inputs, each $\mathcal{O}(\sqrt{K})$, but due to the recurrent dynamics these inputs cancel each other at the leading order. This cancellation, which does not require fine-tuning, results in $\mathcal{O}(1)$ net inputs into the neurons whereas spatial heterogeneities and temporal fluctuations in the inputs are also $\mathcal{O}(1)$. As a result, the activity of the neurons remains

———————
* Correspondance author:david.hansel@univ-paris5.fr

finite and exhibits strong temporal irregularity and heterogeneity [47, 48].

The latter results were established in two-population sparsely connected networks. Renart et al. [28] showed that they also hold for densely connected networks. This is because in unstructured and strongly recurrent networks the dynamics suppress correlations [49]. Despite the fact that in these networks neurons share a finite fraction of their inputs, they operate in an asynchronous state with $\mathcal{O}(1/N)$ correlations, where N is the number of neurons in the network [50].

These previous studies focused on strongly recurrent unstructured networks with two neuronal populations. In the brain, however, neural networks comprise a diversity of excitatory and inhibitory cell types, which differ in their morphology, molecular signature and importantly for our purpose, in their connectivity. These networks are also structured at many levels. In particular, the probability of connection falls off with distance and depends on functional properties of pre- and postsynaptic neurons. For example, in mouse primary auditory cortex the probability of excitatory neurons to be connected decays to zero within $\sim 300\mu m$ [35]. In cat primary visual cortex neurons interact locally on range of $\sim 500\mu m$, whereas long range patchy connections are observed up to several mms [51, 52].

In the present paper we investigate how *structure* in network connectivity can give rise to strongly correlated activity. Our goal is to explore the general architectural features that control the strength of pair-wise correlations in strongly recurrent neural circuits consisting of several excitatory and inhibitory neuronal populations. In Section II, we define the network architectures and the neuronal model we use. In Section III, we establish a set of constraints, *the balanced correlation equations*, that need to be satisfied in any strongly recurrent network if firing rates do not saturate. We derive in Section IV explicit expressions for the Fourier components of the correlations in two-population networks with spatially modulated connectivity. We establish the conditions under which correlations are strong and show that these do not violate the balanced correlation equations. Section V is devoted to networks with an arbitrary number of populations. We prove two theorems, which state for which network architectures correlations are $\mathcal{O}(1/N)$ and for which they increase with K, when K is large. In Section VI, we apply these theorems to specific examples. In Section VII, we assume that $K = \mathcal{O}(N^\gamma)$ and derive a bound on γ for which the scaling theorems still apply. The paper closes with a discussion of our results.

II.   THE NETWORK MODEL

II.1.   Architecture

We consider a neuronal network with a ring architecture, comprising D neuronal populations, some excitatory and others inhibitory (Fig. 1; [53, 54]). For simplicity, we assume that all populations have the same number of neurons, N. Neuron i, in population α (neuron (i,α)) is located at angle $\theta_i^\alpha = \frac{2\pi i}{N}$ with, $i = 1, \ldots, N$ and $\alpha = 1, \ldots, D$. The probability, $P_{ij}^{\alpha\beta}$, that a neuron (j,β) projects to neuron (i,α) depends on their distance on the ring (Fig. 1b). We write: $P_{ij}^{\alpha\beta} = \frac{K}{N} f_{\alpha\beta}(|\theta_i^\alpha - \theta_j^\beta|)$, where $\sum_j f_{\alpha\beta}(|\theta_i^\alpha - \theta_j^\beta|) = N$. In this paper we assume a finite number of non-zero Fourier modes in $f_{\alpha\beta}$.

Thus, a neuron in population α receives, on average, K inputs from neurons in each of the populations β. We denote by $\boldsymbol{\Lambda}$ the adjacency matrix of the network connectivity

$$\Lambda_{ij}^{\alpha\beta} = \begin{cases} 1 & \text{if } (j,\beta) \text{ is presynaptic to } (i,\alpha) \\ 0 & \text{otherwise} \end{cases}$$

For simplicity we assume that all connections from population β to population α have the same strength, $j_{\alpha\beta}$. We thus define the connectivity matrix, $\boldsymbol{J}$, as

$$J_{ij}^{\alpha\beta} = j_{\alpha\beta} \Lambda_{ij}^{\alpha\beta}.$$

Note that $j_{\alpha\beta}$ is positive (negative) if population β is excitatory (inhibitory).

In all this paper we focus on strongly recurrent networks, characterized by interactions which are $\mathcal{O}(\frac{1}{\sqrt{K}})$ [45, 46]. Thus, we scale the synaptic strengths with the mean connectivity as

$$j_{\alpha\beta} = \frac{J_{\alpha\beta}}{\sqrt{K}}. \tag{1}$$

where $J_{\alpha\beta}$ is $\mathcal{O}(1)$.

II.2.   Single neuron dynamics

The state of neuron (i,α) is characterized by a binary variable, S_i^α. When the neuron is quiescent, $S_i^\alpha = 0$, while if it is active, $S_i^\alpha = 1$. The total synaptic current into neuron (i,α) at time t, $h_i^\alpha(t)$, is the result of all its recurrent interactions with the other neurons in the network, as well as the feedforward inputs coming from outside the network. It is given by

$$h_i^\alpha(t) = \sum_{\beta=1}^D \sum_{j=1}^N J_{ij}^{\alpha\beta} S_j^\beta(t) + I_i^\alpha \tag{2}$$

where the external input $I_i^\alpha = \sqrt{K} I_\alpha(\theta)$ [46] is assumed to be constant in time. Here, I_α is $\mathcal{O}(1)$ and thus I_i^α is $\mathcal{O}(\sqrt{K})$.

The neurons are updated with Glauber dynamics at zero temperature [1, 28, 43]. Specifically, update times for neuron (i,α) are Poisson distributed with rate $1/\tau$: if neuron (i,α) is updated at time t, S_i^α is set to 1 if $h_i^\alpha(t) \geq T_\alpha$ and to 0 otherwise (for simplicity, we assume the same threshold, T_α, for all neurons in population α,
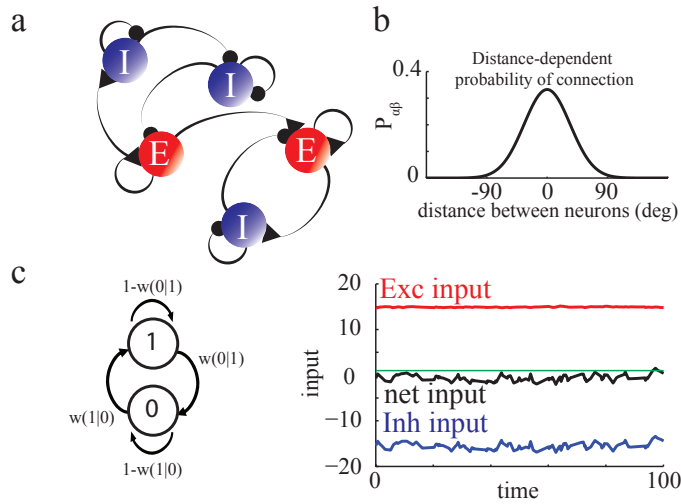
a

b

c

FIG. 1. **Network architectures and dynamics. a.** Networks consist of D populations of neurons, excitatory (E,red) and inhibitory (I, blue), recurrently coupled and receiving external drives. Triangles: Excitatory connections. Circles: Inhibitory connections. **b.** The neurons in each population are located on a ring. The probability that neuron (j, β) is connected to neuron (i, α) depends on their distance. **c.** Left: Neurons are binary units with zero temperature Glauber dynamics. Right: The network operates in the balanced regime. Red: Excitatory input into a neuron. Blue: Inhibitory input into the same neuron. Black: The net input is on the order of the threshold (green, $T = 1$). Time is given in units of $\tau = 1$.

and the same update rate for all neurons). Accordingly, the transition probability, w, (Fig. 1c) can be written as

$$w(S_i^\alpha(t) \to 1 - S_i^\alpha(t)) = \frac{1}{\tau}[S_i^\alpha(t) - \Theta(h_i^\alpha(t) - T_\alpha)]^2.$$

where Θ is the Heaviside function. We normalize time so that $\tau = 1$.

III. SPATIO-TEMPORAL PROFILE AND CORRELATIONS OF THE ACTIVITY FOR $N, K \to \infty$

We write the state of neuron (i, α) as

$$S_i^\alpha(t) = m_\alpha(\theta_i^\alpha) + \Delta S_i^\alpha + \delta S_i^\alpha(t), \qquad (3)$$

where $m_\alpha(\theta) = [\langle S_i^\alpha(t)\rangle_t]_J$ is the average of $S_i^\alpha(t)$ over realizations of the network connectivity matrix and over time. In Eq. (3), the term ΔS_i^α is the quenched disorder, $\Delta S_i^\alpha = \langle S_i^\alpha(t)\rangle_t - m_\alpha(\theta_i^\alpha)$, whereas $\delta S_i^\alpha(t)$ represents the temporal fluctuations in the activity of neuron (i, α), $\delta S_i^\alpha(t) = S_i^\alpha(t) - \langle S_i^\alpha(t)\rangle_t$.

Similarly, we can write:

$$h_i^\alpha(t) = h_\alpha(\theta_i^\alpha) + \Delta h_i^\alpha + \delta h_i^\alpha(t), \qquad (4)$$

where $h_\alpha(\theta_i^\alpha) = [\langle h_i^\alpha \rangle_t]_J$ is a smooth function of its argument. For large N it is given by:

$$h_\alpha(\theta) = \sqrt{K}\left[\sum_{\beta=1}^{D} J_{\alpha\beta}\int \frac{d\theta'}{2\pi}f_{\alpha\beta}(\theta - \theta')m_\beta(\theta') + I_\alpha(\theta)\right].$$
$$(5)$$

The second term in Eq. (4) is the quenched disorder in the input. It satisfies

$$\Delta h_i^\alpha = \sum_{\beta=1}^{D} \frac{J_{\alpha\beta}}{\sqrt{K}}\sum_j \left[\Delta\Lambda_{ij}^{\alpha\beta}m_\beta(\theta_j^\beta) + \Lambda_{ij}^{\alpha\beta}\Delta S_j^\beta\right], \qquad (6)$$

where $\Delta\Lambda_{ij}^{\alpha\beta} = \Lambda_{ij}^{\alpha\beta} - [\Lambda_{ij}^{\alpha\beta}]_J$.

Finally, the temporal fluctuations in the inputs are

$$\delta h_i^\alpha(t) = \sum_{\beta=1}^{D} \frac{J_{\alpha\beta}}{\sqrt{K}}\sum_j \Lambda_{ij}^{\alpha\beta}\delta S_j^\beta(t). \qquad (7)$$

Because we scale the synaptic strength as $1/\sqrt{K}$, these temporal fluctuations are $\mathcal{O}(1)$ when correlations are weak. At first sight, $h_\alpha(\theta)$ is $\mathcal{O}(\sqrt{K})$. This would imply that, depending on whether $h_\alpha(\theta)$ is positive or negative, the neurons fire either at very high or at very low rate. This happens unless the network settles into a state in which excitation and inhibition are, on average, balanced. In this case, the net average inputs to the neurons are in fact $\mathcal{O}(1)$. On the other hand, large spatial and temporal correlations could in principle lead to temporal fluctuations and heterogeneities in the inputs of $\mathcal{O}(\sqrt{K})$. Thus in presence of strong correlations it is not sufficient to require that the mean input, $h_\alpha(\theta)$, is $\mathcal{O}(1)$, for the network to operate in a biologically relevant regime. For that, we need both the mean net inputs and the input fluctuations to be $\mathcal{O}(1)$ at any time and for all neurons. When this happens we say that the network operates in the balanced regime.

III.1. Balance equations for the quenched averaged population activities

As in [55, 56], the requirement that the mean input is $\mathcal{O}(1)$ yields

$$\sum_{\beta=1}^{D} J_{\alpha\beta}\int \frac{d\theta'}{2\pi}f_{\alpha\beta}(\theta - \theta')m_\beta(\theta') + I_\alpha(\theta) = \mathcal{O}(1/\sqrt{K}),$$
$$(8)$$

for all $\alpha \in \{1, \dots, D\}$ and all $\theta \in [0, 2\pi)$. In the large K limit this yields a set of linear equations which determines the functions $m_\alpha(\theta)$ to leading order. These equations can be written in Fourier space

$$\sum_{\beta} J_{\alpha\beta}^{(n)}m_\beta^{(n)} + I_\alpha^{(n)} = \mathcal{O}\left(\frac{1}{\sqrt{K}}\right). \qquad (9)$$

where the superscript n denotes the nth Fourier mode and we have used the short hand notation

$$J_{\alpha\beta}^{(n)} = J_{\alpha\beta}f_{\alpha\beta}^{(n)}. \qquad (10)$$

which is the nth Fourier mode of the connectivity matrix. In what follows, we consider $J_{\alpha\beta}^{(n)}$ as the elements of a $D \times D$ matrix, $\boldsymbol{J}^{(n)}$.

The spatial average of the activities, $m_\alpha^{(0)}$, must be non-negative for the balanced state to exist. This implies that the parameters $J_{\alpha\beta}$ and the external inputs, $I_\alpha^{(0)}$, must satisfy a set of inequalities. For example, for networks with two populations, one excitatory ($\alpha = E$) and one inhibitory ($\alpha = I$), these inequalities are [46]

$$\frac{I_E^{(0)}}{I_I^{(0)}} > \frac{J_{EI}}{J_{II}} > \frac{J_{EE}}{J_{IE}}. \tag{11}$$

In general, Eq. (9), also implies additional inequalities that must be satisfied by $f_{\alpha\beta}^{(n)}$ to guarantee that $m_\alpha(\theta) \geq 0$ for all α and θ [55, 56]. In the present work we focus on the case where the external inputs are spatially homogeneous. Thus, $m_\alpha^{(n)} = 0$ for $n \geq 1$, and the only condition which is required for the balanced state to exist is: $m_\alpha^{(0)} \geq 0$.

To study the stability of the homogeneous balanced state it is useful to introduce the *interaction* matrix

$$\bar{\mathcal{J}}_{ij}^{\alpha\beta} = g_\alpha J_{\alpha\beta} f_{\alpha\beta}(\theta_i^\alpha - \theta_j^\beta) \tag{12}$$

where g_α is the population averaged gain (see Appendixes A,D; [46, 49]).

Small perturbations, $\delta m_\alpha(\theta, t)$, of the activity profile around this homogeneous state evolve according to [46]:

$$\frac{d\delta m_\alpha^{(n)}}{dt} = -\delta m_\alpha^{(n)} + \sqrt{K} \sum_\beta \bar{\mathcal{J}}_{\alpha\beta}^{(n)} \delta m_\beta^{(n)}, \tag{13}$$

where $\delta m_\alpha^{(n)}(t)$ and $\bar{\mathcal{J}}_{\alpha\beta}^{(n)}$ are the nth Fourier modes of $\delta m_\alpha(\theta, t)$ and $\bar{\mathcal{J}}_{ij}^{\alpha\beta}$. Since each row of $\bar{\mathcal{J}}_{ij}^{\alpha\beta}$ has $\mathcal{O}(K)$ non-zero elements which are $\mathcal{O}(1/K)$, $\bar{\mathcal{J}}_{\alpha\beta}^{(n)}$ is $\mathcal{O}(1)$.

The stability of the balanced state with respect to perturbations in $m_\alpha(\theta)$ requires that, for all n, all the eigenvalues of the matrices $\boldsymbol{\mathcal{J}}^{(n)}$ have real parts smaller than 1. For instance, for a two population (E,I) network one must have in the large K limit

$$J_{EE}|J_{II}|f_{EE}^{(n)} f_{II}^{(n)} - J_{IE}|J_{EI}|f_{EI}^{(n)} f_{IE}^{(n)} \leq 0. \tag{14}$$

for all n. Note that the gain of the neurons, g_α, dropped from this equation. The balanced state undergoes a Turing bifurcation when for some $n \geq 1$, $J_{EE}|J_{II}|f_{EE}^{(n)} f_{II}^{(n)} - J_{IE}|J_{EI}|f_{EI}^{(n)} f_{IE}^{(n)}$ crosses 0 and becomes positive.

III.2. Balance equation for pair-wise correlations

The time-lagged auto- and cross-correlation functions of the activities of a pair of neurons, $(i, \alpha), (j, \beta)$, for $(j, \beta) \neq (i, \alpha)$, are

$$a_i^\alpha(\tau) = \langle \delta S_i^\alpha(t) \delta S_i^\alpha(t+\tau) \rangle_t \tag{15}$$

$$c_{ij}^{\alpha\beta}(\tau) = \langle \delta S_i^\alpha(t) \delta S_j^\beta(t+\tau) \rangle_t \tag{16}$$

In what follows, it is convenient to define $c_{ii}^{\alpha\alpha} = 0$.

Due to the randomness of the connectivity, the number of excitatory and inhibitory inputs varies from neuron to neuron, resulting in heterogeneous firing rates between neurons even when they belong to the same population [46, 48]. This structural randomness also results in heterogeneity of the auto-correlations of single neuron activities. It also contributes to the heterogeneity in pair cross-correlations. The latter is further enhanced by the spatial variability in the number of inputs shared by pair of neurons. A full characterization of the distributions of the auto- and cross-correlations is beyond the scope of this paper. Instead, here we will focus on their population averages.

The population average auto-correlation, A_α, is given by $A_\alpha(\tau) = \frac{1}{N} \sum_i a_i^\alpha(\tau)$, which in the thermodynamic limit is also $A_\alpha(\tau) = [a_i^\alpha(\tau)]_J$. With the architecture we use, the probability that neurons (i, α) and (j, β) share common inputs from a third neuron, (k, γ) is

$$\Pr(\Lambda_{ik}^{\alpha\gamma} = 1 \wedge \Lambda_{jk}^{\beta\gamma} = 1) =$$
$$\frac{K^2}{N^2} f_{\alpha\gamma}(\theta_i^\alpha - \theta_k^\gamma) f_{\beta\gamma}(\theta_j^\beta - \theta_k^\gamma) \tag{17}$$

As is evident from this equation, the number of shared inputs averaged over all pairs of neurons separated by the same distance on the ring, Δ, depends only on Δ. Thus we define the average cross-correlations as $C_{\alpha\beta}(\Delta, \tau) = \langle c_{ij}^{\alpha\beta}(\tau) \rangle$, where $\langle . \rangle$ denotes the average over pairs with $|\theta_i^\alpha - \theta_j^\beta| = \Delta$. In the thermodynamic limit, this quantity does not depend on the specific realization of the network connectivity matrix. The Fourier expansion of this function is

$$C_{\alpha\beta}(\Delta, \tau) = \sum_{n=0}^{N-1} C_{\alpha\beta}^{(n)}(\tau) e^{-in\Delta}, \tag{18}$$

where $C_{\alpha\beta}^{(n)}(\tau) = \frac{1}{N^2} \sum_{kj} c_{kj}^{\alpha\beta}(\tau) e^{in(\theta_k^\alpha - \theta_j^\beta)}$.

Equation (9), which determines the population averaged firing rates, stems from the constraint that the net input in every neuron must be $\mathcal{O}(1)$ when K is large. The condition that for any pair of neurons the correlation in their inputs is finite in that limit leads to another constraint, which we now derive.

Let us consider the $ND \times ND$ matrix $\boldsymbol{Q}$ defined by

$$Q_{ij}^{\alpha\beta} = [\langle \delta h_i^\alpha(t) \delta h_j^\beta(t) \rangle_t]_J, \tag{19}$$

for $(i, \alpha) \neq (j, \beta)$ and $Q_{ii}^{\alpha\alpha} = 0$. Using Eq.(7), one finds

$$Q_{ij}^{\alpha\beta} = \sum_{\gamma\gamma'} \frac{J_{\alpha\gamma} J_{\beta\gamma'}}{K} \sum_{kk'} [\Lambda_{ik}^{\alpha\gamma} \Lambda_{jk'}^{\beta\gamma'}]_J [\langle \delta S_k^\gamma(t) \delta S_{k'}^{\gamma'}(t) \rangle_t]_J \tag{20}$$

for $(i, \alpha) \neq (j, \beta)$. Similarly to $C_{\alpha\beta}(\Delta)$, $Q_{ij}^{\alpha\beta}$ is a function of $\Delta = |\theta_i^\alpha - \theta_j^\beta|$ only.

Expanding Eq.(20) in Fourier ($Q_{\alpha\beta}^{(n)} \equiv$

$\frac{1}{N^2} \sum_{kj} Q_{kj}^{\alpha\beta} e^{in(\theta_k^\alpha - \theta_j^\beta)})$, yields

$$Q_{\alpha\beta}^{(n)} = K \sum_{\gamma\gamma'} J_{\alpha\gamma}^{(n)} J_{\beta\gamma'}^{(-n)} C_{\gamma\gamma'}^{(n)} + \sum_\gamma \frac{K}{N} J_{\alpha\gamma}^{(n)} J_{\beta\gamma}^{(-n)} A_\gamma$$

which can be written in matrix form

$$\boldsymbol{Q}^{(n)} = K \boldsymbol{J}^{(n)} \boldsymbol{C}^{(n)} [\boldsymbol{J}^\dagger]^{(n)} + \frac{K}{N} \boldsymbol{J}^{(n)} \boldsymbol{A} [\boldsymbol{J}^\dagger]^{(n)} \qquad (21)$$

Here $\boldsymbol{X}$ denotes the $D \times D$ matrix $X_{\alpha\beta}$, $\boldsymbol{A}_{\alpha\beta} = A_\alpha \delta_{\alpha\beta}$ and the superscript $\dagger$ denotes Hermitian conjugation ($[\boldsymbol{J}_{\alpha\beta}^\dagger]^{(n)} = [\boldsymbol{J}_{\beta\alpha}^{(-n)}]$).

Note that the diagonal of the matrix $\boldsymbol{Q}$ is not the variance of the inputs. The latter is (see Appendix E4)

$$\sigma_\alpha^2 = K \sum_{n\gamma\gamma'} J_{\alpha\gamma}^{(n)} J_{\alpha\gamma'}^{(n)} C_{\gamma\gamma'}^{(n)} + \sum_\gamma J_{\alpha\gamma}^2 A_\gamma \qquad (22)$$

The requirement that crosscorrelations of the inputs into pair of neurons are at most $\mathcal{O}(1)$ implies that all the quantities $Q_{\alpha\beta}^{(n)}$ are also at most $\mathcal{O}(1)$. This yields

$$\boldsymbol{J}^{(n)} \boldsymbol{C}^{(n)} [\boldsymbol{J}^\dagger]^{(n)} = \mathcal{O}(1/K) \qquad (23)$$

We call Eq. (23), the *balanced correlation equation*. It implies that for all n for which the matrix $\boldsymbol{J}^{(n)}$ is invertible and has entries $\mathcal{O}(1)$, $\boldsymbol{C}^{(n)}$ is smaller or equal to $\mathcal{O}(1/K)$ and is thus weak. For a broad class of network architectures, we show below that correlations are in fact $\mathcal{O}(1/N)$ and barely depend on K, for large K. When, however, $\boldsymbol{J}^{(n)}$ is singular for some $n \geq 0$, correlations

can be larger than $\mathcal{O}(1/K)$. In fact we will see that in those cases correlations can even increase with K.

Finally, we note that one can also write a balanced correlation equation for the quenched disorder of the neural activity, using the fact that $[\Delta h_i^\alpha \Delta h_j^\beta]_J^{(n)}$ also must be at most $\mathcal{O}(1)$. This requirement leads to the condition

$$\boldsymbol{J}^{(n)} \boldsymbol{\Gamma}^{(n)} [\boldsymbol{J}^\dagger]^{(n)} = \mathcal{O}(1/K) \qquad (24)$$

with

$$[\boldsymbol{\Gamma}^{(n)}]_{\alpha\beta} \equiv [\Delta S_i^\alpha \Delta S_j^\beta]_J^{(n)} \qquad (25)$$

IV. CORRELATIONS IN TWO-POPULATION NETWORKS

In Appendix A we derive an equation for the spatial Fourier modes of the equal-time quenched average correlation functions, $C_{\alpha\beta}(\Delta, 0)$. It yields (omitting the second argument)

$$2\boldsymbol{C}^{(n)} = \sqrt{K} \left(\bar{\boldsymbol{\mathcal{J}}}^{(n)} \boldsymbol{C}^{(n)} + \boldsymbol{C}^{(n)} [\bar{\boldsymbol{\mathcal{J}}}^\dagger]^{(n)} \right) \qquad (26)$$
$$+ \frac{\sqrt{K}}{N} \left(\bar{\boldsymbol{\mathcal{J}}}^{(n)} \boldsymbol{A} + \boldsymbol{A} [\bar{\boldsymbol{\mathcal{J}}}^\dagger]^{(n)} \right),$$

There we also show that the solution of this equation is a fixed point of the dynamics of the correlations which is always stable when the balanced state is stable with respect to perturbation in the population rates.

This equation holds for a network with an arbitrary number of neuronal populations. In the case of two populations, it can be solved explicitly, yielding after a straightforward calculation (see Appendix E)

$$C_{EE}^{(n)} = -\frac{A_E}{N} + \frac{1}{N} \frac{-2A_E + A_E\left(T^{(n)} + 2\bar{\mathcal{J}}_{II}^{(n)}\right)\sqrt{K} - \left(A_E(\Delta^{(n)} + (\bar{\mathcal{J}}_{II}^{(n)})^2) + A_I(\bar{\mathcal{J}}_{EI}^{(n)})^2\right)K}{-2 + 3T^{(n)}\sqrt{K} - \left((T^{(n)})^2 + 2\Delta^{(n)}\right)K + T^{(n)}\Delta^{(n)}K^{3/2}}$$

$$C_{EI}^{(n)} = -\frac{1}{N} \frac{\left(A_E\bar{\mathcal{J}}_{IE}^{(n)} + A_I\bar{\mathcal{J}}_{EI}^{(n)}\right)\sqrt{K} - \left(A_I\bar{\mathcal{J}}_{EE}^{(n)}\bar{\mathcal{J}}_{EI}^{(n)} + A_E\bar{\mathcal{J}}_{II}^{(n)}\bar{\mathcal{J}}_{IE}^{(n)}\right)K}{-2 + 3T^{(n)}\sqrt{K} - \left((T^{(n)})^2 + 2\Delta^{(n)}\right)K + T^{(n)}\Delta^{(n)}K^{3/2}}$$

$$C_{II}^{(n)} = -\frac{A_I}{N} + \frac{1}{N} \frac{-2A_I + \left(A_I(T^{(n)} + 2\bar{\mathcal{J}}_{EE}^{(n)})\right)\sqrt{K} - \left(A_I(\Delta^{(n)} + (\bar{\mathcal{J}}_{EE}^{(n)})^2)) + A_E(\bar{\mathcal{J}}_{IE}^{(n)})^2\right)K}{-2 + 3T^{(n)}\sqrt{K} - \left((T^{(n)})^2 + 2\Delta^{(n)}\right)K + T^{(n)}\Delta^{(n)}K^{3/2}}$$

$$(27)$$

where $T^{(n)}$ and $\Delta^{(n)}$ are the trace and the determinant of the matrix $\bar{\boldsymbol{\mathcal{J}}}^{(n)}$.

The expansion of these expressions for large K gives

$$C_{EE}^{(n)} = -\frac{A_E}{N} - \frac{1}{N\sqrt{K}} \frac{A_I(\bar{\mathcal{J}}_{EI}^{(n)})^2 + A_E(\Delta^{(n)} + (\bar{\mathcal{J}}_{II}^{(n)})^2)}{T^{(n)}\Delta^{(n)}} + \mathcal{O}(\frac{1}{NK})$$

$$C_{EI}^{(n)} = \frac{1}{N\sqrt{K}} \frac{A_E\bar{\mathcal{J}}_{II}^{(n)}\bar{\mathcal{J}}_{IE}^{(n)} + A_I\bar{\mathcal{J}}_{EE}^{(n)}\bar{\mathcal{J}}_{EI}^{(n)}}{T^{(n)}\Delta^{(n)}} + \mathcal{O}(\frac{1}{NK})$$

$$C_{II}^{(n)} = -\frac{A_I}{N} - \frac{1}{N\sqrt{K}} \frac{A_E(\bar{\mathcal{J}}_{IE}^{(n)})^2 + A_I(\Delta^{(n)} + (\bar{\mathcal{J}}_{EE}^{(n)})^2)}{T^{(n)}\Delta^{(n)}} + \mathcal{O}(\frac{1}{NK})$$

$$(28)$$

Thus in general when K is large $C_{EE}^{(n)}$ and $C_{II}^{(n)}$ are very small, namely $\mathcal{O}(1/N)$, with negative prefactors which do not depend on K. As for $C_{EI}^{(n)}$, it is smaller than $C_{EE}^{(n)}$ and $C_{II}^{(n)}$ by a factor $1/\sqrt{K}$.

It should be noted that to derive Eq. (28) we assumed that $T^{(n)} \neq 0$ and $\Delta^{(n)} \neq 0$. Equation (27) indicates, however, that when $T^{(n)} = 0$ and $\Delta^{(n)} \neq 0$, or $T^{(n)} \neq 0$ and $\Delta^{(n)} = 0$, $C_{EE}^{(n)}$, $C_{II}^{(n)}$ and $C_{EI}^{(n)}$ are $\mathcal{O}(1/N)$.

The situation is different if $T^{(n)} = 0$ *and* $\Delta^{(n)} = 0$. Equation (27) shows that in this case it is possible to get correlations which are $\mathcal{O}(K/N)$.

In the rest of this section we consider in detail two-population networks in which for the probabilities of connection only the first two Fourier modes are non-zero

$$P_{ij}^{\alpha\beta} = \frac{K}{N}\left[1 + 2f_{\alpha\beta}^{(1)}\cos\left(\theta_i^\alpha - \theta_j^\beta\right)\right], \qquad (29)$$

with $\alpha = E, I, \beta = E, I$.

IV.1. $T^{(1)} \neq 0$ and $\Delta^{(1)} \neq 0$

For $f_{\alpha\beta}^{(1)}$ such that $T^{(1)} \neq 0$ and $\Delta^{(1)} \neq 0$ we have in the large N, K limit

$$C_{EE}(\Delta) = -\frac{A_E}{N}(1 + 2\cos\Delta) + \mathcal{O}(\frac{1}{N\sqrt{K}}) \qquad (30)$$

$$C_{EI}(\Delta) = \frac{1}{N\sqrt{K}}(\bar{C}_{EI}^{(0)} + 2\bar{C}_{EI}^{(1)}\cos(\Delta)) + \mathcal{O}(\frac{1}{NK})$$

$$C_{II}(\Delta) = -\frac{A_I}{N}(1 + 2\cos\Delta) + \mathcal{O}(\frac{1}{N\sqrt{K}})$$

where $\bar{C}_{EI}^{(n)} \equiv N\sqrt{K}C_{EI}^{(n)}$ are $\mathcal{O}(1)$. Thus, in that limit, the spatial average and modulation of the correlations within the E and I populations do not depend on K, at the leading order. Moreover, $C_{EE}(\Delta)$ and $C_{II}(\Delta)$ depend on the synaptic strengths only because the auto-correlations A_E and A_I depend on these parameters.

Figure 2 depicts simulation results for $N = 40000$ and $K = 2000$. Figure 2a plots $C_{EE}(\Delta)$ for $f_{\alpha\beta}^{(1)} = 0.25$ ($\alpha, \beta \in E, I$) and two sets of values for the interaction strengths (solid and dashed lines). For comparison we also plots the results of a simulation when the connectivity is unstructured ($f_{\alpha\beta}^{(1)} = 0; \alpha, \beta \in E, I$; gray line). In all these cases $C_{EE}(\Delta)$ is very small (note the scale on thr y-axis). When the connectivity is spatially modulated $C_{EE}(\Delta)$ varies with distance. However, the spatial averages are comparable in the three cases considered ($\frac{1}{2\pi}\int C_{EE}(\Delta)d\Delta \sim -0.2 \times 10^{-5}$). This is in agreement with Eq. (30) since, to the leading order, the auto-correlations, A_E and A_I, are not expected to depend on whether the connectivity is spatially modulated or not.

Note that according to Eq. (30), $C_{EE}(\Delta)$ and $C_{II}(\Delta)$ are negative for close-by neurons (Δ small) and positive for neurons that are far apart. This is the case for the set of parameters corresponding to the solid line in Fig. 2a (see also Fig. 2b-d). However, when finite K corrections

are not negligible, which happens when $|T^{(1)}\Delta^{(1)}|$ is sufficiently small, short range correlations can be positive and longer range correlations negative (Eq. (30)). This is the case in Fig. 2a, dashed line.

Figure 2b compares simulations (circles) and analytical results (solid lines) for the dependence on N of $NC_{\alpha\beta}^{(1)}$ (parameters as in panel a, red solid line). It shows that the spatial modulation of the correlations in the simulation is close to the large K analytical results. Figure 2c-d depict the dependence of $NC_{\alpha\beta}$ on K. In the whole range of K considered here simulations and analytical results are close. For $|C_{\alpha\alpha}|$ the nearby correlations and modulation amplitudes increase with K and are much larger than that of $|C_{EI}|$.
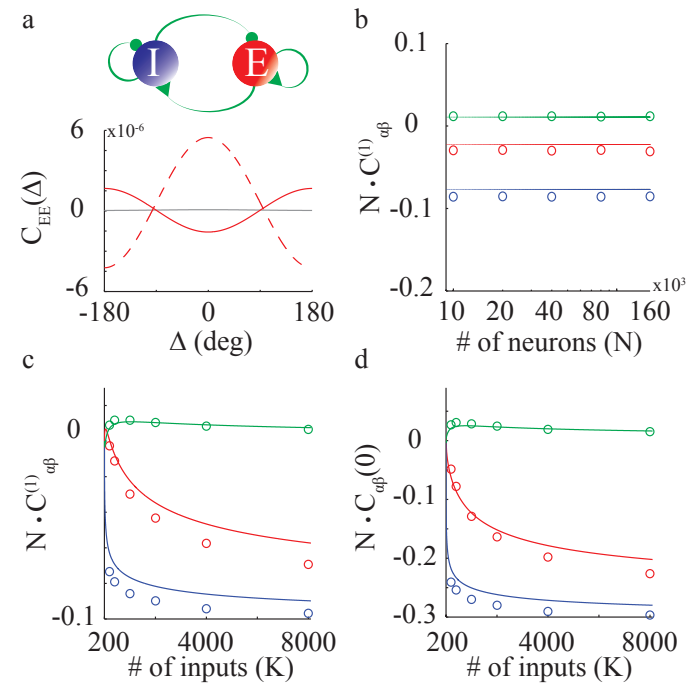

FIG. 2. **Correlations in two-population E-I networks.** Connection probabilities are as in Eq. (29). Panels a, c, d: $N = 40000$. **a.** Top panel: The network architecture. All connections are spatially modulated ($f_{\alpha\beta}^{(1)} \neq 0$). Bottom panel: Simulation results for $C_{EE}(\Delta)$ with $K = 2000$. Red: $f_{EE}^{(1)} = f_{EI}^{(1)} = f_{IE}^{(1)} = f_{II}^{(1)} = 0.25$. For comparison the correlation is also plotted for $f_{EE}^{(1)} = f_{EI}^{(1)} = f_{IE}^{(1)} = f_{II}^{(1)} = 0$ (Gray). For red-solid and gray lines other parameters are $J_{EE} = 0.3, J_{IE} = 3, J_{EI} = 2.5, J_{II} = 5; I_E = 0.3, I_I = 0.3, T_E = 1, T_I = 0.7$. With these parameters $m_E^{(0)} \simeq 0.12, m_I^{(0)} \simeq 0.13$, $g_E \simeq 0.22, g_I \simeq 0.1$ and $A_E \simeq 0.1, A_I \simeq 0.1$. For red-dashed line other vparameters are: $J_{EE} = 1; J_{IE} = 2; J_{EI} = 1; J_{II} = 1.5; I_E = 0.2; I_I = 0.08$. **b.** $NC_{\alpha\beta}^{(1)}$ vs. N for $K = 1000$. **c.** $NC_{\alpha\beta}^{(1)}$ vs. K. **d.** $NC_{\alpha\beta}(0)$ vs. K ($N = 40000$). In panels b, c, d: Solid line: Analytics, Eq. (27). Circles: Simulations. Red: $\alpha = \beta = E$. Blue: $\alpha = \beta = I$. Green: $\alpha = E, \beta = I$. Parameters are as for red-solid line in panel a.

IV.2. $T^{(1)} = \Delta^{(1)} = 0$

We now consider a network in which $f_{EI}^{(1)} \neq 0$ and $f_{EE}^{(1)} = f_{II}^{(1)} = f_{IE}^{(1)} = 0$. The spatially modulated component of the interaction has therefore an explicit feedforward structure (Fig. 3a, top panel) and $T^{(1)} = \Delta^{(1)} = 0$.

Solving Eq. (27) shows that the correlations are on average $\mathcal{O}(1/N)$ and that their modulations are

$$C_{EE}^{(1)} = \frac{K}{N} \frac{A_I}{2} |\bar{\mathcal{J}}_{EI}^{(1)}|^2 \qquad (31)$$

$$C_{EI}^{(1)} = -\frac{\sqrt{K}}{N} \frac{A_I}{2} |\bar{\mathcal{J}}_{EI}^{(1)}|$$

$$C_{II}^{(1)} = 0$$

As a result, correlations in the E population are spatially modulated and $\mathcal{O}(K/N)$. They are positive for short range and negative for long range. This is in contrast to the correlations in the inhibitory population which are not spatially modulated and $\mathcal{O}(1/N)$ while the EI correlations are spatially modulated and $\mathcal{O}(\sqrt{K}/N)$.

Figure 3 compares analytical results with simulations. Panel a plots simulation results for $C_{\alpha\beta}(\Delta)$ for $f_{EI}^{(1)} = 0.25$. Other parameters are as in Fig. 2a (grey solid line). Thus, the locally averaged firing rates are the same as in simulations in the latter figure. Correlations in the inhibitory neurons (blue) are extremely weak and are not spatially modulated. In contrast, the correlations of excitatory pairs (red) are larger by two orders of magnitude compared to those in Fig. 2. Correlations between E and I neurons (green) are weaker than those between E neurons. They are negative for nearby neurons while for nearby excitatory pairs they are positive. All these features are in agreement with our analytical results,

The spatial modulation of the EE correlations increases with K (Fig. 3b, circles). There is quantitative agreement between simulations and theory (Eq. (31)) up to $K \approx 4000$ for $N = 40000$. In this range, $C_{EE}^{(1)}$ in the simulations varies linearly with K. For larger values of K, $C_{EE}^{(1)}$ is larger than predicted by Eq. (31). This is because this equation was derived by linearizing the dynamics, which is only valid when correlations are not too large. In fact, simulation results for fixed K deviate less from Eq. (31) when N is increased (Fig.3b-c). When K increases, $C_{EI}^{(1)}$, becomes more negative (Fig. 3c; green). Here too, for $N = 40000$ simulations agree well with Eq. (31) up to $K \approx 4000$ and deviations are smaller when N is larger. The spatial modulation of the II correlations in the simulations are extremely small (Fig. 3c; blue) as the theory predicts.

Finally, according to Eq. (31), the spatial modulation, $C_{EE}^{(1)}$, increases quadratically with $f_{EI}^{(1)}$ whereas $C_{EI}^{(1)}$ varies linearly with this parameter. Our simulations are in very good agreement with these analytical results (Fig. 3d).
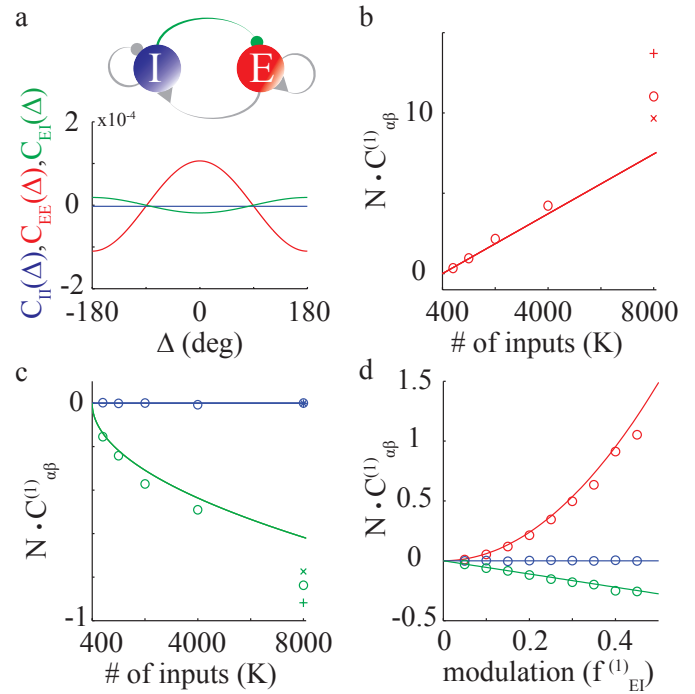

FIG. 3. **Correlations in a two population network with an explicit feedfworward structure.** Connection probabilities as in Eq. (29). Parameters: $f_{EE}^{(1)} = f_{IE}^{(1)} = f_{II}^{(1)} = 0$. Panels a, b, c: $f_{EI}^{(1)} = 0.25$. Interaction strengths as in Fig. 2. **a.** Top panel: The network architecture. Green: the connections which are spatially modulated ($f_{EI}^{(1)} \neq 0$). Connections in gray are unstructured ($f_{\alpha\beta}^{(1)} = 0$). Bottom panel: Simulation results for $C_{\alpha\beta}(\Delta)$; $N = 40000$, $K = 2000$.**b, c.** $NC_{\alpha\beta}^{(1)}$ vs. K. Solid lines: Analytical results, Eq. (31). Simulation results are plotted for $N = 20000$ (plus), $N = 40000$ (circles) and $N = 80000$ (crosses). **b.** $NC_{EE}^{(1)}$. **c.** Blue; $NC_{II}^{(1)}$. Green: $NC_{EI}^{(1)}$. **d.** $NC_{\alpha\beta}^{(1)}$ vs. $f_{EI}^{(1)}$. Solid lines: Analytical results, Eq. (31); $N = 40000$, $K = 400$.

IV.3. $\Delta^{(1)} = 0$, $T^{(1)} \neq 0$

The network investigated in IV.2 has an explicit feedforward structure. Adding any spatial modulation to the II connectivity ($f_{II}^{(1)} \neq 0$, top panel of Fig. 4a) destroys this structure and now $T^{(1)} \neq 0$ (while $\Delta^{(1)} = 0$). Solving Eq. (26) for that case, one finds:

$$C_{EE}^{(1)} = \frac{K}{N} A_I \frac{|\bar{\mathcal{J}}_{EI}^{(1)}|^2}{(2 + \sqrt{K}|\bar{\mathcal{J}}_{II}^{(1)}|)(1 + \sqrt{K}|\bar{\mathcal{J}}_{II}^{(1)}|)} \qquad (32)$$

$$C_{EI}^{(1)} = -\frac{K}{N} A_I \frac{|\bar{\mathcal{J}}_{EI}^{(1)}|}{\sqrt{K}(2 + \sqrt{K}|\mathcal{J}_{II}^{(1)}|)(1 + \sqrt{K}|\mathcal{J}_{II}^{(1)}|)|}$$

$$C_{II}^{(1)} = -\frac{K}{N} A_I \frac{|\bar{\mathcal{J}}_{II}^{(1)}|}{\sqrt{K}(1 + \sqrt{K}|\bar{\mathcal{J}}_{II}^{(1)}|)}$$

In the large N, K limit, $C_{EE}(\Delta)$ and $C_{II}(\Delta)$ are both $\mathcal{O}(1/N)$ and do not depend on K, whereas $C_{EI}(\Delta)$ is

$\mathcal{O}(1/(N\sqrt{K}))$. Therefore the addition of a spatial modulation in the II interactions suppresses the correlations that inhibitory projections induce in the E population.
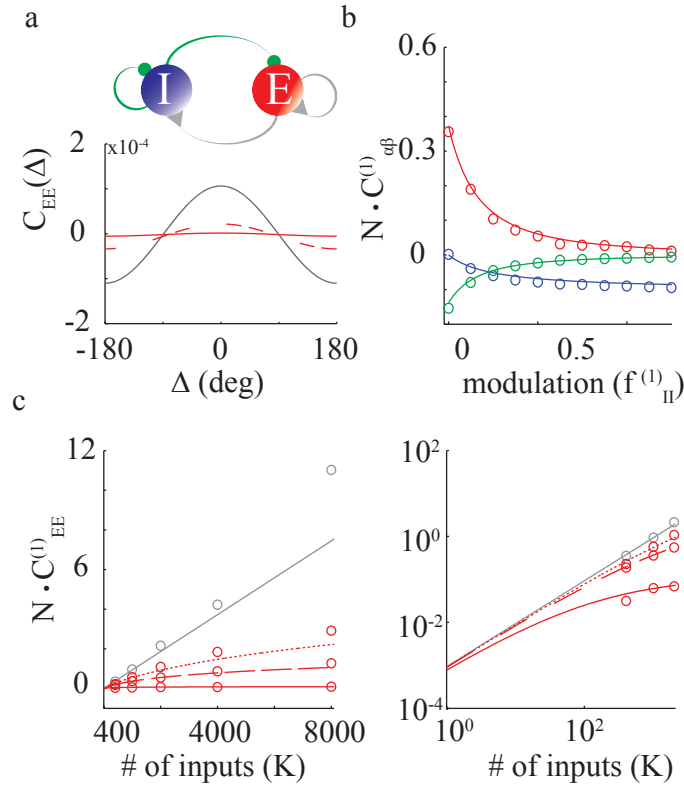

FIG. 4. **Spatial modulation in the II interactions suppresses the correlations in the E populations.** Same network as in Fig. 3 but $f_{II}^{(1)} = 0.25$. $N = 40000$. **a.** Top panel: The network architecture. Interactions plotted in green are spatially modulated. Bottom panel: Simulation results for $C_{EE}(\Delta)$. $N = 40000$, $K = 2000$. Gray line: $f_{II}^{(1)} = 0$; Dashed- red: $f_{II}^{(1)} = 0.05$; Solid-red : $f_{II}^{(1)} = 0.25$. **b.** $NC_{\alpha\beta}^{(1)}$ vs. $f_{II}^{(1)}$. Circles: Simulations. Solid lines: Solution of Eq. (32). $N = 40000$, $K = 400$. Colors are as in Fig. 3. **c.** Left panel: $NC_{EE}^{(1)}$ vs. K. Solid lines: Eq. (32). Circles: Simulations. Gray line: $f_{II}^{(1)} = 0$; Dotted red: $f_{II}^{(1)} = 0.025$; Dashed red: $f_{II}^{(1)} = 0.05$; Solid red : $f_{II}^{(1)} = 0.25$. Right panel: Same as left panel but in a log-log scale.

To understand further the origin of this, let us consider the quenched average correlations of the inputs, $Q_{\alpha\beta}(\Delta)$ (Eq. (19)). Using Eq. (21), one finds

$$Q_{EE}^{(1)} = K(\bar{\mathcal{J}}_{EI}^{(1)})^2(C_{II}^{(1)} + \frac{A_I}{N})$$

$$Q_{EI}^{(1)} = K\bar{\mathcal{J}}_{EI}^{(1)}\bar{\mathcal{J}}_{II}^{(1)}(C_{II}^{(1)} + \frac{A_I}{N})$$

$$Q_{II}^{(1)} = K(\bar{\mathcal{J}}_{II}^{(1)})^2(C_{II}^{(1)} + \frac{A_I}{N})$$

When $\bar{\mathcal{J}}_{II}^{(1)} = 0$, $C_{II}^{(1)} = 0$, therefore $Q_{EE}^{(1)} = \frac{K}{N}(\bar{\mathcal{J}}_{EI}^{(1)})^2 A_I$. On the other hand, when $\bar{\mathcal{J}}_{II}^{(1)} \neq 0$, $C_{II}^{(1)}$ is given by

Eq.(32). This yields

$$Q_{EE}^{(1)} = \frac{A_I}{N} \frac{K(\bar{\mathcal{J}}_{EI}^{(1)})^2}{1 + \sqrt{K}|\bar{\mathcal{J}}_{II}^{(1)}|} \tag{33}$$

Thus, when II interactions are spatially modulated, a cancellation between terms which are $\mathcal{O}(K/N)$, reduces $Q_{EE}^{(1)}$ by a factor of $\mathcal{O}(1/\sqrt{K})$. As a result, $C_{EE}^{(1)}$ is much smaller than when II interactions are not modulated. A similar argument explains the suppression in $C_{EI}^{(1)}$.

Fig. 4a depicts simulation results for $f_{EI}^{(1)} = 0.25$ and three values of $f_{II}^{(1)}$. It demonstrates the suppression of correlations in the excitatory population when II interactions are also modulated. The dependence of this effect on $f_{II}^{(1)}$ is depicted in Fig. 4b. Increasing $f_{II}^{(1)}$, decreases the modulation of all the correlations in very good agreement with the analytical results (compare circles and solid lines).

According to Eq. (32), the correlation in the E population always increases linearly with K, for small K. The cross-over between this regime, where $C_{EE}(\Delta)$ is $\mathcal{O}(K/N)$, and the large K regime, where $C_{EE}(\Delta)$ is $\mathcal{O}(1/N)$, occurs for $\sqrt{K} \simeq 1/\bar{\mathcal{J}}_{II}^{(1)}$. Figure 4c depicts this crossover in numerical simulations. Thus, although in the large N, K limit a transition from $\mathcal{O}(K/N)$ to $\mathcal{O}(1/N)$ correlations occurs as soon as $f_{II}^{(1)} \neq 0$, this qualitative difference is significant only when K is sufficiently large. In other words, for moderately large number of inputs per neuron, correlations can exhibit a close to linear increase even if the structure of spatial modulation of the interaction matrix is not completely feedforward.

V. SCALING CORRELATION THEOREMS

In the previous section we studied networks with two neuronal populations. In this case, it is straightforward to analytically derive explicit expressions for the correlations. These expressions are simple enough to fully classify how the network structure affects the scaling (with K and N) of the correlations. For networks consisting of more than two populations, analytical expressions for the correlations can in principle be derived. However, dealing with these expressions becomes rapidly impractical as the number of populations increases. In the following we adopt an alternative approach. We prove two general theorems, which, given the network architecture, allow us to determine how correlations vary with K, without computing these correlations explicitly.

To prove these theorems, we rewrite Eq. (26) in the Jordan basis of $\bar{\mathcal{J}}^{(n)}$. We write

$$\bar{\mathcal{J}}^{(n)} = U^{(n)} \mathcal{J}_{\text{jor}}^{(n)} [U^{(n)}]^{-1} \tag{34}$$

$$[\bar{\mathcal{J}}^{(n)}]^\dagger = V^{(n)} [\mathcal{J}_{\text{jor}}^{(n)}]^* [V^{(n)}]^{-1}$$

Here x^* denotes the complex conjugate of x, $U^{(n)}$ ($V^{(n)}$) are matrices whose rows are the generalized eigenvectors of $\bar{\mathcal{J}}^{(n)}$, and $\mathcal{J}_{\text{jor}}^{(n)}$ is the Jordan normal form of

$\bar{\mathcal{J}}^{(n)}$. This implies that $[\mathcal{J}_{\text{jor}}^{(n)}]^*$ is the Jordan form of

$[\bar{\mathcal{J}}^{(n)}]^\dagger$ is the $[\bar{\mathcal{J}}^{(n)}]^\dagger$. We can write

$$[\mathcal{J}_{\text{jor}}^{(n)}]_{\mu\nu} = \lambda_\mu^{(n)}\delta_{\mu\nu} + \epsilon_\mu^{(n)}\delta_{\mu,\nu-1}$$

with $\mu,\nu \in 1,...,D$ and $\epsilon_\mu^{(n)} = 1$ within a Jordan block and zero otherwise.

Equation (26) then yields

$$\left[2 - \sqrt{K}(\lambda_\mu^{(n)} + [\lambda_\nu^{(n)}]^*)\right]\hat{C}_{\mu\nu}^{(n)} = \sqrt{K}(\epsilon_\mu^{(n)}\hat{C}_{\mu+1,\nu}^{(n)} + \epsilon_{\nu-1}^{(n)}\hat{C}_{\mu,\nu-1}^{(n)}) \qquad (35)$$
$$+ \frac{\sqrt{K}}{N}(\epsilon_\mu^{(n)}\hat{A}_{\mu+1,\nu}^{(n)} + \epsilon_{\nu-1}^{(n)}\hat{A}_{\mu,\nu-1}^{(n)}) + \frac{\sqrt{K}}{N}(\lambda_\mu^{(n)} + [\lambda_\nu^{(n)}]^*)\hat{A}_{\mu\nu}^{(n)}$$

where we have defined

$$\hat{C}^{(n)} = [U^{(n)}]^{-1}C^{(n)}V^{(n)} \qquad (36)$$
$$\hat{A}^{(n)} = [U^{(n)}]^{-1}A^{(n)}V^{(n)}$$

Note that while the matrix $C^{(n)}$ is symmetric and the matrix $A^{(n)}$ is diagonal, this is not in general the case for $\hat{C}^{(n)}$ and $\hat{A}^{(n)}$.

Let us assume that the network is in a stable balanced state in which the matrix $\hat{A}^{(n)}$ has no zero elements. In Appendix B we prove

Correlation Theorem 1: The nth Fourier mode of the correlation matrix scales as $\mathcal{O}(1/N)$ if and only if $\mathcal{J}_{\text{jor}}^{(n)}$ does not have a Jordan block whose real part is a shift matrix (A shift matrix, S, of dimension P is a $P \times P$ matrix of the form $S_{\mu\nu} = \delta_{\mu+1,\nu}$).

Correlation Theorem 2: If $\bar{\mathcal{J}}^{(n)}$ has at least one Jordan block whose real part is a shift matrix, the nth Fourier mode of the correlation matrix is $\mathcal{O}(K^{P(n)-1}/N)$, where $P(n)$ is the dimension of the largest block in $\mathcal{J}_{\text{jor}}^{(n)}$, whose real part is a shift matrix.

Corollary: The nth Fourier mode of the correlation matrix is $\mathcal{O}(K^{D-1}/N)$ if and only if $\mathcal{J}^{(n)}$ is nilpotent of degree D.

In Appendix B we also show how to extend these resukts to case where $\hat{A}^{(n)}$ has zero elements.

The matrix $U^{(n)}$ can be viewed as a transformation of the original network of D populations into a network of D effective populations. The condition that $\bar{\mathcal{J}}^{(n)}$ has a Jordan block which is a shift matrix of size P, can be interpreted as the existence of P effective populations whose effective interaction matrix is feedforward in its nth Fourier mode. In other words, the original network exhibits a hidden feedforward structure, which is embedded in the nth Fourier mode of its connectivity. Theorems 1 and 2 therefore implies that only when such a structure exists, the nth mode of the correlation matrix increases with K. To know which elements in this matrix increase

with K one has to compute the matrices $U^{(n)}$ and $V^{(n)}$ (see Eq. (36)).

VI. APPLICATIONS OF THE CORRELATION THEOREMS

In this section we consider networks comprising D populations with connection probabilities

$$P_{ij}^{\alpha\beta} = \frac{K}{N}\left(1 + 2f_{\alpha\beta}^{(1)}\cos(\theta_i^\alpha - \theta_j^\beta)\right) \qquad (37)$$

with $\alpha = 1,...,D$ and $\beta = 1,...,D$.

VI.1. Two population networks

For a network of two populations the Jordan form of the matrix $\bar{\mathcal{J}}^{(1)}$ has the form: $\mathcal{J}_{\text{jor}}^{(1)} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ if $\bar{\mathcal{J}}^{(1)}$ is diagonalizable. Otherwise, it has the form: $\mathcal{J}_{\text{jor}}^{(1)} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ with λ real. Theorem 1 and 2 imply that only in the second case with $\lambda = 0$, some of the correlations $C_{\alpha\beta}(\Delta)$ are $\mathcal{O}(K/N)$. Otherwise, all correlations are $\mathcal{O}(1/N)$. It is equivalent to say that some correlations are $\mathcal{O}(K/N)$ if and only if $\bar{\mathcal{J}}^{(1)} \neq \mathbf{0}$ and $T^{(1)} = \Delta^{(1)} = 0$. We therefore recover the result we derived in Section IV without explicit computation of the correlations.

As noted above, the correlation theorems do not tell us which of the elements of the correlation matrix are $\mathcal{O}(K/N)$ when $\mathcal{J}_{\text{jor}}^{(1)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. However, since a 2×2 matrix has such a Jordan form if and only if it is nilpotent $((\bar{\mathcal{J}}^{(1)})^2 = 0)$ we have to consider two types of connectivity:

Type 1: The network has an explicit feedforward structure, *i.e.*, the interaction matrix is either $\bar{\mathcal{J}}^{(1)} = $

$$\begin{pmatrix} 0 & \bar{\mathcal{J}}_{EI}^{(1)} \\ 0 & 0 \end{pmatrix} \text{ or } \bar{\mathbf{\mathcal{J}}}^{(1)} = \begin{pmatrix} 0 & 0 \\ \bar{\mathcal{J}}_{IE}^{(1)} & 0 \end{pmatrix}.$$

In the former case, the matrix $\mathbf{U}^{(1)}$ is $\mathbf{U}^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1/\bar{\mathcal{J}}_{EI}^{(1)} \end{pmatrix}$, whereas $\mathbf{V}^{(1)} = \begin{pmatrix} 0 & 1 \\ 1/\bar{\mathcal{J}}_{EI}^{(1)} & 0 \end{pmatrix}$. Then Eq. (36) gives (see Appendix B, Eq. (B8))

$$\hat{\mathbf{C}}^{(1)} = \begin{pmatrix} \mathcal{O}(\sqrt{K}/N) & \mathcal{O}(K/N) \\ 0 & \mathcal{O}(\sqrt{K}/N) \end{pmatrix} \quad (38)$$

Using Eq. (36), one finds that $C_{EE}^{(1)} = \mathcal{O}(K/N)$, $C_{II}^{(1)} = 0$, whereas $C_{EI}^{(1)} = \mathcal{O}(\sqrt{K}/N)$, in agreement with Eq. (31). A similar calculation in the latter case (when the modulation is in the IE interactions) gives $C_{EE}^{(1)} = 0$, $C_{II}^{(1)} = \mathcal{O}(K/N)$, and $C_{EI}^{(1)}$ is $\mathcal{O}(\sqrt{K}/N)$, all in agreement with Eq. (27).

Type 2: $\bar{\mathbf{\mathcal{J}}}^{(1)} = c \begin{pmatrix} 1 & -a \\ 1/a & -1 \end{pmatrix}$ where $a, c > 0$.

In this case the network has no *explicit* feedforward structure since all four interactions (EE, EI, IE, II) are spatially modulated. It has, however, a *hidden* feedforward structure, as revealed by the Jordan form of the interaction matrix.

For this network, the transformation matrices are $\mathbf{U}^{(1)} = \begin{pmatrix} a & a/c \\ 1 & 0 \end{pmatrix}$ and $\mathbf{V}^{(1)} = \begin{pmatrix} -1 & -1/c \\ a & 0 \end{pmatrix}$. Using Eq. (36) and Eq. (38), it is clear that the transformation from $\hat{\mathbf{C}}^{(1)}$ to $\mathbf{C}^{(1)}$ mixes elements which are 0 and $\mathcal{O}(\sqrt{K}/N)$, with those which are $\mathcal{O}(K/N)$. Thus, while in $\hat{\mathbf{C}}^{(1)}$ only the element $\hat{C}_{12}^{(1)}$ is $\mathcal{O}(K/N)$, all the elements of the correlation matrix $\mathbf{C}^{(1)}$ are $\mathcal{O}(K/N)$. This is also in line with Eq. (27).

We consider an example of such a network in Fig. 5. The parameters $J_{\alpha\beta}$ and the external inputs are as in Fig. 2-4. Therefore, to leading order, the population averaged activities, m_α, the autocorrelations, A_α, and the population gains, g_α are the same as in Figs. 2-4. The modulation of the connection probability, $f_{EE}^{(1)}, f_{IE}^{(1)}, f_{EI}^{(1)}, f_{II}^{(1)}$, are all non-zero (Fig. 5a) and tuned so that:

$$\bar{\mathbf{\mathcal{J}}}^{(1)} = \frac{1}{20} \begin{pmatrix} 1 & -1/2 \\ 2 & -1 \end{pmatrix}$$

The Jordan form of this matrix is graphically represented in Fig. 5b. The top panel in Fig. 5c depicts simulation results for the correlations in this network. They are all positive for close-by neurons and their spatial modulations increase approximately linearly with K (Fig. 5c, bottom). The top panel of Fig. 5d shows that most of the power in these correlations results from the element $\hat{C}_{12}$ (red). The latter increases linearly with K (Fig. 5d, bottom), while $\hat{C}_{11}, \hat{C}_{22} = \mathcal{O}(\sqrt{K}/N)$ and $\hat{C}_{21}$ is two order of magnitude smaller and does not exhibit significant change with K (note the log-log scale in this

figure). These simulations are in quantitative agreement with Eq. (27) up to $K \approx 4000$ and the deviations for larger K decrease with N (as in Fig 4).

The interaction matrix, $\bar{\mathbf{\mathcal{J}}}^{(1)}$, depends on the matrix $\mathbf{J}^{(1)}$, and on the population gains, g_E and g_I. If the connectivity matrix has an explicit feedforward structure, the interaction matrix has also such a structure, independantly of the population gains. Thus, although changing the external inputs, I_E, I_I, modifies this gains, this does not destroy the feedforward structure and thus does not change the scaling of the correlations with K and N. In contrast, in networks with a hidden feedforward structure, this scaling is sensitive to perturbations in the external inputs since the hidden feedforward structure is destroyed unless the ratio g_E/g_I remains constant.
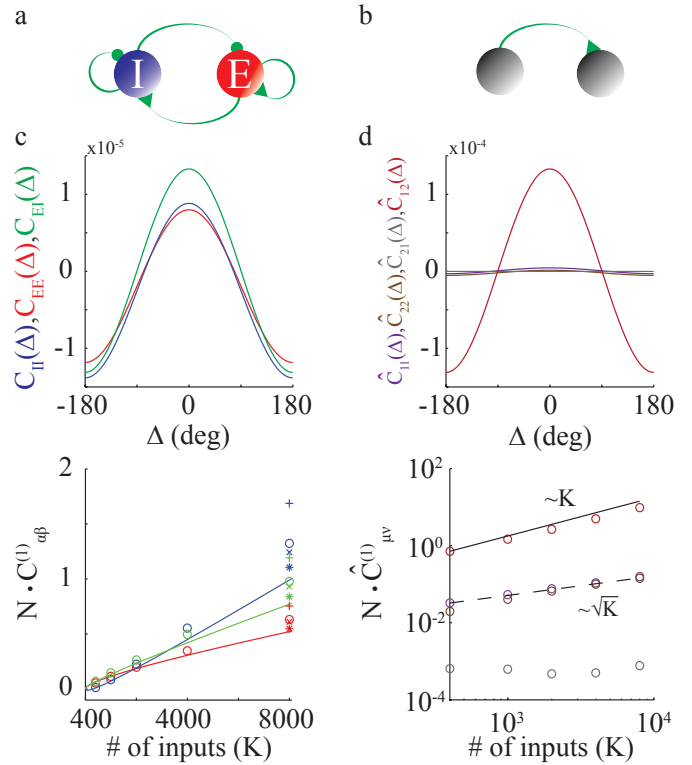

FIG. 5. **Correlations in a two population network with a hidden feedforward structure. a.** Network architecture. All connections are modulated. Connection probabilities are as in Eq. (29) with $f_{EE}^{(1)} = 0.373, f_{EI}^{(1)} = 0.0224, f_{II}^{(1)} = 0.0487, f_{IE}^{(1)} = 0.1623$. Connection strengths are as in Figs. 2-3. **b.** Hidden feedforward structure of the mode $n = 1$ as revealed in the Jordan basis. **c.** Top panel: Simulation results for $C_{\alpha\beta}(\Delta)$. $N = 40000, K = 2000$. Bottom panel: $NC_{\alpha\beta}^{(1)}$ vs. K. Solid lines: Analytical results (Eq. (27)). Simulations are plotted for $N = 20000$ (plus), $N = 40000$ (circles), $N = 80000$ (cross) and $N = 160000$ (asterisks). **d.** Top panel: Correlation matrix in the Jordan basis of $\bar{\mathbf{\mathcal{J}}}^{(1)}$. Purple: $\hat{C}_{11}(\Delta)$; Brown: $\hat{C}_{22}(\Delta)$; Gray: $\hat{C}_{21}(\Delta)$; Red: $\hat{C}_{12}(\Delta)$. Bottom panel: $N\hat{C}_{\mu\nu}^{(1)}$ vs. K in log-log scale. Solid line: Linear fit. Dashed line: Fit to a square root function.

VI.2. Examples with three populations or more

The two networks depicted in Fig. 6 consist of one excitatory and two inhibitory populations. The positivity of the spatial averaged activity of these three populations, $m_\alpha^{(0)}$ ($\alpha = 1, 2, 3$), constrains the parameters (see Section III), $J_{\alpha\beta}$, through a set of inequalities. We leave the calculation of these conditions to the reader.

In both cases, the first Fourier mode of the population average connectivity matrix has the form

$$\boldsymbol{J}^{(1)} = \begin{pmatrix} 0 & J_{12}^{(1)} & J_{13}^{(1)} \\ 0 & J_{22}^{(1)} & J_{23}^{(1)} \\ 0 & 0 & 0 \end{pmatrix} \quad (39)$$

with $J_{12}^{(1)}, J_{13}^{(1)}, J_{23}^{(1)} < 0$ (left panels of Fig 6a-c).

In the network of Fig. 6a, $J_{22}^{(1)} = 0$. Therefore $\boldsymbol{J}^{(1)}$ and $\bar{\boldsymbol{\mathcal{J}}}^{(1)}$ are nilpotent of degree 3. According to the Corollary of Section V, the first mode of the correlation matrix is $\mathcal{O}(K^2/N)$.

The Jordan form of $\bar{\boldsymbol{\mathcal{J}}}^{(1)}$ is

$$\boldsymbol{\mathcal{J}}_{\text{jor}}^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

This is graphically represented in Fig. 6a, middle panel. The matrix $\hat{\boldsymbol{C}}^{(1)}$ satisfies (see Appendix B, Eq.(B8))

$$\hat{\boldsymbol{C}}^{(1)} = \begin{pmatrix} \mathcal{O}(K/N) & \mathcal{O}(K^{3/2}/N) & \mathcal{O}(K^2/N) \\ \mathcal{O}(K^{1/2}/N) & \mathcal{O}(K/N) & \mathcal{O}(K^{3/2}/N) \\ 0 & \mathcal{O}(K^{1/2}/N) & \mathcal{O}(K/N) \end{pmatrix}$$

Using the transformation matrices (Eq.(34)), one can show that correlations are $\mathcal{O}(K^2/N)$ only within the excitatory population.

In the network in Fig. 6b, $J_{22}^{(1)} < 0$. Therefore, the interaction matrix, $\bar{\boldsymbol{\mathcal{J}}}^{(1)}$, is not nilpotent. Its Jordan form is

$$\boldsymbol{\mathcal{J}}_{\text{jor}}^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & J_{22}^{(1)} \end{pmatrix}.$$

The upper Jordan block is a shift matrix of degree 2. The corresponding feedforward structure is graphically represented in Fig. 6b (middle panel). According to Theorem 2, $\hat{\boldsymbol{C}}^{(1)}$ is $\mathcal{O}(K/N)$. It satisfies (see Appendix B, Eq. (B8))

$$\hat{\boldsymbol{C}}^{(1)} = \begin{pmatrix} \mathcal{O}(\sqrt{K}/N) & \mathcal{O}(K/N) & \mathcal{O}(1/N) \\ 0 & \mathcal{O}(\sqrt{K}/N) & \mathcal{O}(1/N) \\ \mathcal{O}(1/N) & \mathcal{O}(1/N) & \mathcal{O}(1/N) \end{pmatrix}$$

and using the transformation matrices one can show that only $\boldsymbol{C}_{11}^{(1)}$ is $\mathcal{O}(K/N)$. Other correlations are either $\mathcal{O}(\sqrt{K}/N)$ or $\mathcal{O}(1/N)$.
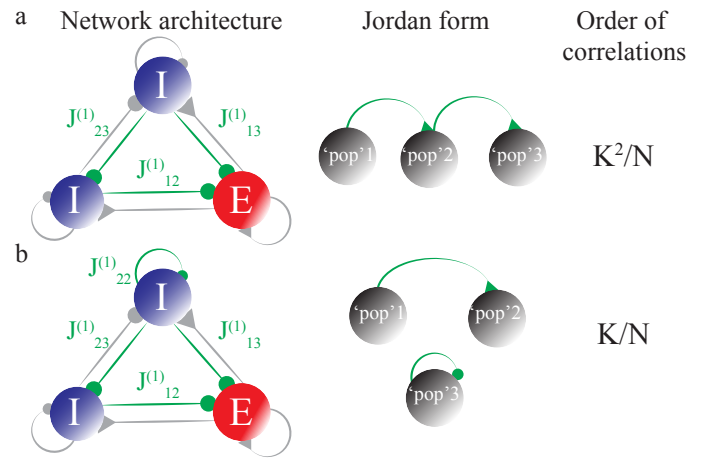

FIG. 6. **Examples of networks with three populations and their Jordan representations.** Probability of connections are as in Eq. (37). **a-b.** Left: A network of three populations, two inhibitory and one excitatory. Gray: Unstructured connections. Green: Spatially modulated connections. Middle: The Jordan representation of the $n = 1$ Fourier mode of the network connectivity (left panel). Right: The scaling of the strongest correlations. **b.** $J_{12}^{(1)}, J_{13}^{(1)}, J_{23}^{(1)} \neq 0$. Other entries of the matrix $\boldsymbol{J}^{(1)}$ are zero (see main text). **b.** Same as (a), but with $J_{22}^{(1)} \neq 0$.

This approach can be generalized to arbitrary number of populations to classify the scaling of the correlation matrix for different architectures. Examples of networks with four populations are depicted in Fig 7, together with the graphic representations of their Jordan forms and the maximum order of the correlations.

VII. CONSTRAINT ON SCALING OF THE NUMBER OF INPUTS WITH THE NETWORK SIZE

In this section we assume that K and N scale together

$$K \propto N^\gamma \quad (40)$$

with $0 < \gamma \leq 1$.

Equation (26), which determines the correlations of the neuronal activities, is obtained under the Ansatz that the correlations between the inputs $h_i^\alpha(t)$ and $h_j^\beta(t)$ are sufficiently small, namely $o(1)$ (see Appendix A). This condition is more stringent than the balance correlation equation, Eq. (23). It can be written as

$$\boldsymbol{J}^{(n)} \boldsymbol{C}^{(n)} [\boldsymbol{J}^{(n)}]^\dagger = o(1/N^\gamma) \quad (41)$$

This condition constrains γ as we now show.

According to Theorem 1, if the Jordan form, $\boldsymbol{\mathcal{J}}_{\text{jor}}^{(n)}$, has no Jordan block whose real part is a shift matrix for any n, correlations will be $\mathcal{O}(1/N)$. This will also be
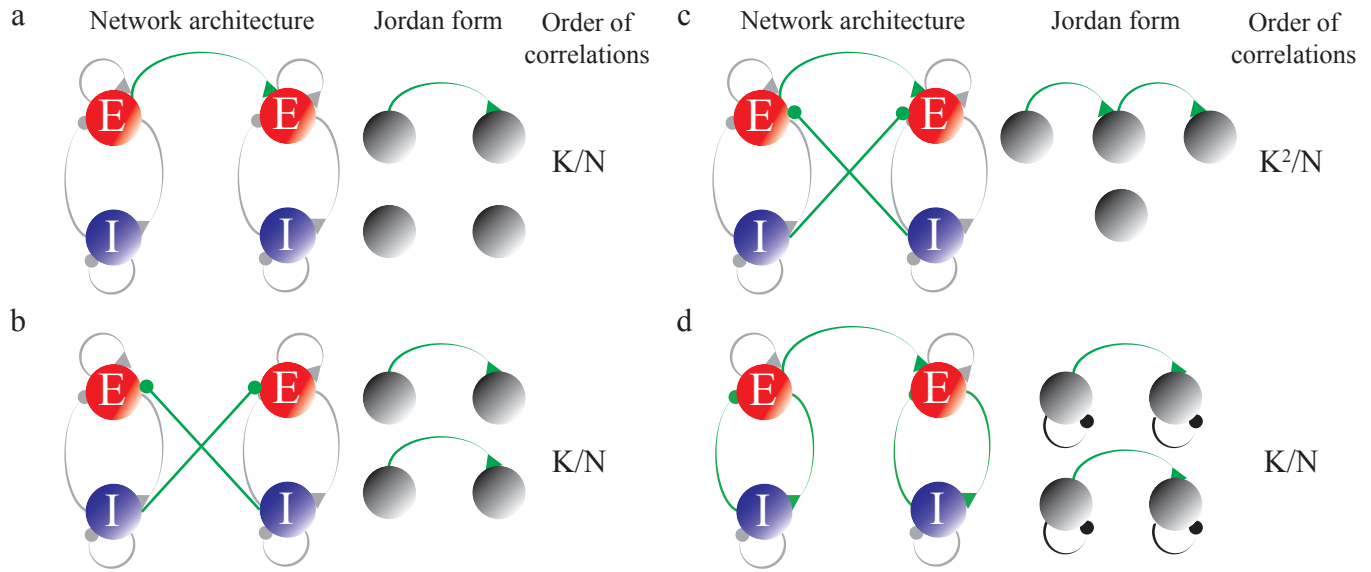
FIG. 7. **Examples of networks with four populations and their Jordan representations.** Probability of connections are as in Eq. (37). **a-c.** Left: The network consists of two coupled E-I networks. Gray: Unstructured connections. Green: Spatially modulated connections. Middle: The Jordan representation of the $n = 1$ Fourier mode of the connectivity (network of the left panel. Right: The scaling of the strongest correlation. **d.** Same as (a-c), with a population averaged connectivity matrix as in Eq. (46). Middle: The Jordan form is complex (Eq. (47)). Black lines corresponds to complex eigenvalues.

the order of $\boldsymbol{J}^{(n)}\boldsymbol{C}^{(n)}[\boldsymbol{J}^{(n)}]^{\dagger}$. Therefore, Eq. (41) only requires $\gamma < \gamma_{\max} = 1$.

If the Jordan form, $\boldsymbol{\mathcal{J}}^{(n)}_{\text{jor}}$, contains a Jordan whose real part in the shift matrix, we have to apply Theorem 2. In this theorem, $P(n)$ is the dimension of the largest block in $\boldsymbol{\mathcal{J}}^{(n)}_{\text{jor}}$, with a real part which is a shift matrix (see Section V). Let us denote by $P_{\max}$ the largest $P(n)$ over all Fourier modes, i.e.,

$$P_{\max} = \max_n P(n) \tag{42}$$

Equation (41) implies

$$\bar{\boldsymbol{\mathcal{J}}}^{(n)}\boldsymbol{C}^{(n)}[\bar{\boldsymbol{\mathcal{J}}}^{(n)}]^{\dagger} = o(N^{-\gamma})$$

for all n. This yields in the Jordan basis

$$\sum_{m,k}[\mathcal{J}^{(n)}_{\text{jor}}]_{\mu m}\hat{C}^{(n)}_{mk}[[\mathcal{J}^{(n)}_{\text{jor}}]^{*}]_{k\nu} = o(N^{-\gamma}) \tag{43}$$

for all μ, ν.

By definition of $P_{\max}$, for at least one Fourier mode, n, the matrix $\boldsymbol{\mathcal{J}}^{(n)}_{\text{jor}}$ has at least one block whose real part is a shift matrix of degree $P_{\max}$. In general, there can be several such Jordan blocks. For example, in the network depicted in Fig. 7b, for which $P_{\max} = 2$, there are two Jordan blocks with $P = 2$.

We first assume that all blocks which are a shift matrix of size P_{max} are real. Since for such blocks in $\boldsymbol{\mathcal{J}}_{\text{jor}}\hat{C}\boldsymbol{\mathcal{J}}^{*}_{\text{jor}}$ scale the same with K, it is sufficient to consider the case where there is only one such block. We denote it by $\boldsymbol{S}$

and by $\hat{\boldsymbol{C}}_{\max}$ the corresponding block in $\hat{\boldsymbol{C}}$. Equation (41) then yields

$$\boldsymbol{S}\hat{\boldsymbol{C}}_{\max}\boldsymbol{S} = o(N^{-\gamma}) \tag{44}$$

For instance, for $P_{\max} = 2$, we have (see Eq. (B8))

$$\hat{\boldsymbol{C}}_{\max} = \begin{pmatrix} \mathcal{O}(N^{\gamma/2-1}) & \mathcal{O}(N^{\gamma-1}) \\ 0 & \mathcal{O}(N^{\gamma/2-1}) \end{pmatrix}$$

and thus

$$\boldsymbol{S}\hat{\boldsymbol{C}}_{\max}\boldsymbol{S} = 0$$

Therefore, for this block Eq. (41) is always satisfied. The latter equation, however, also applies to other Jordan blocks and Fourier modes. This implies that $\gamma < \gamma_{\max} = 1$

For $P_{\max} = 3$, we have

$$\hat{\boldsymbol{C}}_{\max} = \begin{pmatrix} \mathcal{O}(N^{\gamma-1}) & \mathcal{O}(N^{3\gamma/2-1}) & \mathcal{O}(N^{2\gamma-1}) \\ \mathcal{O}(N^{\gamma/2-1}) & \mathcal{O}(N^{\gamma-1}) & \mathcal{O}(N^{3\gamma/2-1}) \\ 0 & \mathcal{O}(N^{\gamma/2-1}) & \mathcal{O}(N^{\gamma-1}) \end{pmatrix}$$

and

$$\boldsymbol{S}\hat{\boldsymbol{C}}_{\max}\boldsymbol{S} = \begin{pmatrix} 0 & \mathcal{O}(N^{\gamma/2-1}) & \mathcal{O}(N^{\gamma-1}) \\ 0 & 0 & \mathcal{O}(N^{\gamma/2-1}) \\ 0 & 0 & 0 \end{pmatrix}$$

Thus, Eq. (41) is satisfied only if $\gamma < \gamma_{\max} = 1/2$.

In general, for a $P_{\max} \times P_{\max}$ shift matrix $\boldsymbol{S}\hat{\boldsymbol{C}}_{\max}\boldsymbol{S}$ is $\mathcal{O}(N^{\gamma(P_{\max}-2)-1})$. This implies that $\gamma < \gamma_{\max}$ with

$$\gamma_{\max} = \frac{1}{P_{\max} - 1} \tag{45}$$

.

Let us now consider networks in which there is at least one pair of complex conjugate Jordan blocks whose real parts are a shift matrix of size $P_{\max}$. An example, of such a network is depicted in Fig. 7d. The first Fourier mode of the population averaged connectivity matrix in this example is

$$\boldsymbol{J}^{(1)} = \begin{pmatrix} 0 & -a & c & 0 \\ b & 0 & 0 & 0 \\ 0 & 0 & 0 & -a \\ 0 & 0 & b & 0 \end{pmatrix}, \qquad (46)$$

with a, b, c real and positive. For this network

$$\boldsymbol{\mathcal{J}}_{\text{jor}} = \begin{pmatrix} i\omega & 1 & 0 & 0 \\ 0 & i\omega & 0 & 0 \\ 0 & 0 & -i\omega & 1 \\ 0 & 0 & 0 & -i\omega \end{pmatrix} \qquad (47)$$

with $\omega = \sqrt{ab}$.

These complex conjugate blocks can in general be written as $\pm i\omega \boldsymbol{I} + \boldsymbol{S}$, where $\boldsymbol{I}$ is the identity matrix of size $P_{\max}$. Thus

$$\boldsymbol{J}_{\text{jor}}^{\max} \hat{\boldsymbol{C}}_{\max} [\boldsymbol{J}_{\text{jor}}^{\max}]^* = \omega^2 \hat{\boldsymbol{C}}_{\max} \pm i\omega [\hat{\boldsymbol{C}}_{\max}\boldsymbol{S} - \boldsymbol{S}\hat{\boldsymbol{C}}_{\max}]$$
$$+ \boldsymbol{S}\hat{\boldsymbol{C}}_{\max}\boldsymbol{S} \qquad (48)$$

As shown above, $\hat{\boldsymbol{C}}_{\max} = \mathcal{O}(N^{\gamma(P_{\max}-1)-1})$ and $\boldsymbol{S}\hat{\boldsymbol{C}}_{\max}\boldsymbol{S} = \mathcal{O}(N^{\gamma(P_{\max}-2)-1})$. It is straightforward to also show that $\hat{\boldsymbol{C}}_{\max}\boldsymbol{S} - \boldsymbol{S}\hat{\boldsymbol{C}}_{\max} = \mathcal{O}(N^{\gamma(P_{\max}-3/2)-1})$. Therefore $\boldsymbol{J}_{\text{jor}}^{\max}\hat{\boldsymbol{C}}_{\max}[\boldsymbol{J}_{\text{jor}}^{\max}]^* = \mathcal{O}(N^{\gamma(P_{\max}-1)-1})$. Equation (41) is then satisfied only if $\gamma < \gamma_{\max}$, with

$$\gamma_{\max} = \frac{1}{P_{\max}} \qquad (49)$$

According to Theorem 2, if $\mathcal{J}_{\text{jor}}^{(n)}$ has a block whose real part is a shift matrix for at least one mode n, $\boldsymbol{C} = \mathcal{O}(N^{-\alpha})$ where $\alpha = 1 - \gamma(P_{\max} - 1)$. If $\gamma < \gamma_{\max}$, correlations in the activity will decrease more slowly than $1/N$, when N is increased. If $\gamma > \gamma_{\max}$ correlations will increase with N and the network will not operate in the balanced regime. Finally, if $\gamma = \gamma_{\max}$, our theory will give $\mathcal{O}(1)$ correlations in the input which is inconsistent with the Ansatz in Eq. (41). In this case substantial corrections to the Eq. (26) should be taken into account. A different approach, similar to the one in [28] must be adopted to self-consistently determine these correlations.

VIII.   DISCUSSION

VIII.1.   Main results

We developed a theory for the emergence of correlations in strongly recurrent networks of binary neurons. Each neuron receives on average K inputs from each of D populations, with probabilities which are spatially modulated. The synaptic strengths scale as $1/\sqrt{K}$ and the network operates in the balanced regime [45]. For simplicity we considered networks with a one dimensional ring architecture with connection probabilities solely dependent on distance and on the nature (excitatory or inhibitory) of the pre- and postsynaptic populations.

We present a balanced correlation equation, which together with the balanced rate equation, define the balanced regime and insure that mean inputs to the neurons and their fluctuations are both $\mathcal{O}(1)$. We derive a set of equations that determine the equilibrium values of the quenched averaged correlations and we show that the solution of these equations is stable provided that the solution of the balanced rate equations is stable.

Key results of our work are two scaling correlation theorems. The first shows that generically, all the Fourier modes of the quenched average correlations are small when N and K are large. They are of $\mathcal{O}(1/N)$, and independent of K to leading order. This is true in the large N limit even if we take $K = pN$, provided that p is not too large. However, the second theorem states that there are recurrent network architectures in which some of the Fourier modes of the quenched averaged correlations increase with K. These architectures are characterized by an explicit, or a hidden, feedforward structure in those modes. This structure is revealed by the Jordan form of the interaction matrix averaged over realizations. If this Jordan form contains a block whose real part is a shift matrix of size $P > 1$, the corresponding mode in the correlation increases at least as $\mathcal{O}(\frac{K^{P-1}}{N})$. Importantly, in these cases the network still operates in the balanced regime provided that $K \lesssim N^{\frac{1}{P-1}}$ (or $K \lesssim N^{\frac{1}{P}}$, see Section VII). Corrections to the theory become important when K approaches $N^{\frac{1}{P-1}}$ (or $N^{\frac{1}{P}}$).

VIII.2.   Generality of the results

For simplicity, we assumed that synaptic weights depend only on the identities of the populations to which pre and postsynaptic neurons belong. The results, however, will not change if the weights are heterogeneous with distributions that depend on the pre and postsynaptic populations, as long as the mean and the variance of all these distributions are finite.

For notational simplicity we assumed that all populations have the same number of neurons, N, and each population receives, on average, inputs from K neurons in every population. The theory can be easily extended to networks in which population α has $N_\alpha = \nu_\alpha N$ neurons and the average number of connections from population β to population α is $K_{\alpha\beta} = \kappa_{\alpha\beta}K$. This will not affect the scaling of the correlations with N and K. Prefactors, however, will be different. For instance, assuming four times fewer inhibitory than excitatory neurons in the two-population networks of subsection IV.1, without changing the number of connections per neuron, the correlations of the inhibitory neurons will increase by a factor of 4.

We focused on networks with a one-dimensional ring

architecture and connection probabilities which are solely distance dependent. This greatly simplifies the problem, because when averaged over realizations, correlations depend solely on distance. Furthermore, because of the linearity of the self-consistent equation for the correlations, the different Fourier modes decouple, allowing us to analyze each mode separately. However, our analytical approach does not require rotation invariance. It can be extended to any network architecture for which the Jordan normal form of the interaction matrix, averaged over realizations, can be established.

VIII.3. Robustness and self-consistency of the results

The theory presented here makes the Anzatz that correlations are sufficiently small so that the dynamics of the crosscorrelations can be linearized (Appendix A). If this Anzatz is correct, the theory is self-consistent. When the theory predicts correlations which are $\mathcal{O}(1)$, non-linear terms contribute and the correlations start to deviate from the theoretical value (see for example simulation results for $N = 40000$ in Fig. 3b). Nevertheless, the order of the correlations is still correctly predicted.

In the large N, K limit, only networks with feedforward structures -hidden or explicit- can exhibit correlations that increase with the average number of inputs. However, when K and N are only moderately large, as is the case in biological systems, a strict tuning of the architecture is not necessary. This is because there is a crossover between the regimes of strong and weak correlations as K is increased (see Fig. 4c). As shown in Appendix B1, the value of K for which this crossover occurs depends on the eigenvalues of the interaction matrix.

VIII.4. Relation to previous works

Non-interacting neurons can exhibit correlations if they share feedforward inputs [57, 58]. For instance, Litvak et al. [59] investigated a chain of layers of integrate-and-fire neurons lacking any recurrent interactions and coupled only feedforwardly. In their model, each neuron in a layer receives inputs from the same number of excitatory and inhibitory neurons in the previous layer in such a way that their temporal averages *exactly* balance. They found a build up of correlations along the chain. This is because the correlations induced by shared feedforward inputs are not suppressed during the activity propagation since the network lacks any recurrent interactions and is thus purely feedforward.

Cortes and van Vreeswijk [60] studied a chain of strongly recurrent unstructured E-I subnetworks coupled through excitatory unstructured feedforward projections. They found a gradual build up of correlations along the chain. These correlations, however, decrease if the connectivity, K, and the sub-networks size, N, increase together. This result is in agreement with our theory which

predicts that the correlations are $\mathcal{O}(1/N)$ through the whole chain.

In [43] Ginzburg and Sompolinsky considered networks of binary neurons with finite temperature Glauber dynamics, unstructured, dense ($K = pN$) connectivity and weak interactions, i.e., of the order of $\mathcal{O}(1/K)$ and not $\mathcal{O}(1/\sqrt{K})$ as in our work. Mean Field theory shows that in these networks correlations are $\mathcal{O}(1/N)$ with a prefactor which diverges in the zero temperature limit. This is in contrast to what happens in the strongly recurrent unstructured networks we considered here where correlations also scale as $\mathcal{O}(1/N)$ but with a prefactor, which is finite despite the fact that we assumed zero temperature Glauber dynamics. This is because in strongly recurrent networks, intrinsic noise emerges from the deterministic dynamics of the network.

They also demonstrated that in their model the correlations amplify up to $\mathcal{O}(1)$ at Hopf bifurcation onsets [43]. At such onsets the dynamics exhibit critical slowing down and thus this amplification is accompanied by a divergence of the decorrelation times. Our work demonstrates a different amplification mechanism: it occurs at a point where the Jordan form of the interaction matrix, $\bar{\mathcal{J}}$, contains a block which is a shift matrix. Since there is no critical slowing down at such a point, the decorrelation times are finite and on the order of the update time constant (data not shown). Our theory may thus account for substantial correlations with short time scales, as frequently observed in the brain [28, 31, 61].

Renart et al. [28] and Helias et al. [49] investigated strongly recurrent unstructured networks with one excitatory and one inhibitory population of binary neurons which are densely connected, i.e with $K \propto N$. They found that in these networks mean pair correlations were $\mathcal{O}(1/N)$. As they only considered dense connectivity they could not, however, disentangle the dependence on N and K. Here, we considere different relations between K and N and show that in unstructured networks the correlations are $\mathcal{O}(1/N)$, and in practice do not depend on K. This last result is remarkable since one would expect correlations to increase with the degree of connectivity. This is not the case: the balance of excitation and inhibition prevents that to occur in unstructured networks.

On the other hand, and somewhat surprisingly, we found that even if the fraction of common inputs shared by neurons is very small, a build up of correlations can still occur for some network architectures. For example, in a network of four populations with a feedforward structure, correlations would be of $\mathcal{O}(K^3/N)$, and thus, to satisfy the balanced correlation equation, the scaling of K with the network size can be at most $K = \mathcal{O}(N^{1/3})$. However, with this architecture and scaling, the probability of two neurons to share their inputs is $\mathcal{O}(N^{-4/3})$ (Eq. (17)) and the number of inputs shared by two neurons is therefore $\mathcal{O}(N^{-1/3})$. Thus, although the number of shared inputs goes to zero in the large N limit, correlations get amplified up to $\mathcal{O}(1)$ thanks to the feedforward architecture.

Rosenbaum et al. [39] have recently investigated

how feedforward excitation can drive correlations in spatially structured E-I networks operating in the balanced regime. The specific architecture they considered is reminiscent of the particular example presented in Fig. 7a. In their study the fluctuations which drove the correlated activity were those in the feedforward inputs and the contribution to the correlations of the fluctuations generated by the recurrent dynamics was neglected. In contrast, our work focuses on the role of the *recurrent* dynamics in the emergence of correlations.

We recently studied the emergence of correlations in a network consisting of two strongly recurrent E-I subnetworks, the first projecting to the second with topographically organized feedforward connections [17]. The architecture in that work is also reminiscent of the example presented in Fig. 7a. We showed that while in the first subnetwork correlations were weak, $\mathcal{O}(1/N)$, in the second subnetwork the activity was self-organized in such a way that correlations in macroscopic sub-populations of excitatory neurons were finite and did not depend on N and K when the latter were sufficiently large. This does not contradict our theory. In the architecture considered in [17] subsets of neurons in the first subnetwork are projecting to a macroscopic fraction of neurons in the second subnetwork. This organization generates correlations between elements of the connectivity matrix, unlike in the models considered here. Generalizing our theory to such cases is possible but beyond the scope of this paper.

VIII.5.   Directions for future work

The theorems of Section V, tell us how the Fourier modes of the averaged correlations scale for large N and K. In the examples considered in Sections IV, VI and VII, we focused on connectivities whose Fourier expansion involve only two modes. In these cases, the scaling of the spatial correlations can be immediately deduced from that of the Fourier modes. This will also be the case for interactions described by a finite number of Fourier modes. However, if the interactions are described by an infinite number of modes, inferring the scaling of the correlations from those of their modes can be more complicated. For instance, if one takes the large N limit with $K \propto N^\gamma$, the convergence of the Fourier series may not be uniform in N. In that case the scaling with N of the spatial correlations may be highly non-trivial. We will address this issue in an upcoming paper.

The present paper focuses on locally averaged two-point correlation functions. Recent progress in experimental techniques will create large data sets of neuronal activities from which distributions of pairwise correlations can be extracted. The power of theoretical approaches to interpret such data will be greatly enhanced if they provide not only locally averaged correlations but also higher order statistics of their distributions. Thus

it would be interesting to extend our approach to estimate the scaling of higher order moments of correlations in binary networks.

Several previous studies investigated EI networks (*e.g.* [49, 62]) in which inhibition and excitation were unstructured and their strengths were only a function of the presynaptic neurons, i.e., $J_{EE} = J_{IE}$ and $J_{II} = J_{EI}$. Other studies assumed unstructured connectivity with $J_{\alpha E} = -J_{\alpha I}$, $\alpha \in \{E, I\}$ (*e.g.* [63]). In both cases the interaction parameters are on the edge of the region where the network evolves towards the balanced state (see Eq. (11)). The network dynamics may be qualitatively different on the edge of this region than inside it. It is thus not clear that the scaling theorems presented here apply to these tuned cases. A further investigation of the correlation structure in such networks is a subject for future research.

Do the conclusions derived here for networks of binary neurons hold for networks with more realistic single neuron dynamics? To approach this question we performed extensive numerical simulations of strongly recurrent networks consisting of one inhibitory and one excitatory population of leaky integrate-and-fire (LIF) neurons. The detailed analysis of these simulations will be presented elsewhere. In brief, we found in our simulations, that in these networks the averaged pairwise correlations scale with K and N in manner that is consistent with the theory presented here for binary networks. It would be very interesting to extend our analytical approach to these type of networks.

To conclude, van Vreeswijk and Sompolinsky [45, 46] investigated strongly recurrent networks of binary neurons with unstructured sparse connectivity. They showed how these network dynamics evolve into a balanced state. Due to the sparseness of the connectivity, correlations are negligible in these networks. Subsequent studies [28, 49] extended these results to unstructured networks with dense connectivity, and found that here too the network dynamics evolve to a balance state in which reverberations keep the correlations very small. In contrast, as shown here, in networks with structured connectivity correlations can be large. The theory presented here gives the conditions on the network architecture to evolve into a balanced state with strong correlations and show how they depend on the network size and number of connections.

[1] Roy J Glauber, "Time-dependent statistics of the ising model," Journal of mathematical physics **4**, 294–307 (1963).

[2] Mark C Cross and Pierre C Hohenberg, "Pattern formation outside of equilibrium," Reviews of modern physics **65**, 851 (1993).

[3] David Hansel and Haim Sompolinsky, "Chaos and synchrony in a model of a hypercolumn in visual cortex," Journal of computational neuroscience **3**, 7–34 (1996).

[4] Udo Seifert, "Stochastic thermodynamics, fluctuation theorems and molecular machines," Reports on Progress in Physics **75**, 126001 (2012).

[5] Fotis K Diakonos, AK Karlis, and P Schmelcher, "A universal mechanism for long-range cross-correlations," EPL (Europhysics Letters) **105**, 26004 (2014).

[6] Brent Doiron, Ashok Litwin-Kumar, Robert Rosenbaum, Gabriel K Ocker, and Krešimir Josić, "The mechanics of state-dependent neural correlations," Nature neuroscience **19**, 383–393 (2016).

[7] Tatjana Tchumatchenko, Aleksey Malyshev, Theo Geisel, Maxim Volgushev, and Fred Wolf, "Correlations and synchrony in threshold neuron models," Physical Review Letters **104**, 058102 (2010).

[8] Edwin M Maynard, Craig T Nordhausen, and Richard A Normann, "The utah intracortical electrode array: a recording structure for potential brain-computer interfaces," Electroencephalography and clinical neurophysiology **102**, 228–239 (1997).

[9] Simon Peron, Tsai-Wen Chen, and Karel Svoboda, "Comprehensive imaging of cortical networks," Current opinion in neurobiology **32**, 115–123 (2015).

[10] Cyrille Rossant, Shabnam N Kadir, Dan FM Goodman, John Schulman, Maximilian LD Hunter, Aman B Saleem, Andres Grosmark, Mariano Belluscio, George H Denfield, Alexander S Ecker, *et al.*, *Spike sorting for large, dense electrode arrays*, Tech. Rep. (Nature Publishing Group, 2016).

[11] Ehud Zohary, Michael N Shadlen, and William T Newsome, "Correlated neuronal discharge rate and its implications for psychophysical performance," (1994).

[12] LF Abbott and Peter Dayan, "The effect of correlated variability on the accuracy of a population code," Neural computation **11**, 91–101 (1999).

[13] Haim Sompolinsky, Hyoungsoo Yoon, Kukjin Kang, and Maoz Shamir, "Population coding in neuronal systems with correlated noise," Physical Review E **64**, 051904 (2001).

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

[15] Gyorgy Buzsaki, *Rhythms of the Brain* (Oxford University Press, 2006).

[16] Charles M Gray, Peter König, Andreas K Engel, and Wolf Singer, "Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties," Nature **338**, 334–337 (1989).

[17] R Darshan, WE Wood, S Peters, A Leblois, and D Hansel, "A canonical neural mechanism for behavioral variability." Nature communications **8**, 15415 (2017).

[18] Neta Ravid Tannenbaum and Yoram Burak, "Shaping neural circuits by high order synaptic interactions," PLoS computational biology **12**, e1005056 (2016).

[19] Donald Olding Hebb, *The organization of behavior: A neuropsychological theory* (Psychology Press, 2005).

[20] Guo-qiang Bi and Mu-ming Poo, "Synaptic modification by correlated activity: Hebb's postulate revisited," Annual review of neuroscience **24**, 139–166 (2001).

[21] Alexander S Ecker, Philipp Berens, Georgios A Keliris, Matthias Bethge, Nikos K Logothetis, and Andreas S Tolias, "Decorrelated neuronal firing in cortical microcircuits," science **327**, 584–587 (2010).

[22] Marlene R Cohen and Adam Kohn, "Measuring and interpreting neuronal correlations," Nature neuroscience **14**, 811–819 (2011).

[23] Bryan J Hansen, Mircea I Chelaru, and Valentin Dragoi, "Correlated variability in laminar cortical circuits," Neuron **76**, 590–602 (2012).

[24] Maria C Dadarlat and Michael P Stryker, "Locomotion enhances neural encoding of visual stimuli in mouse v1," Journal of Neuroscience **37**, 3764–3775 (2017).

[25] Gideon Rothschild, Lior Cohen, Adi Mizrahi, and Israel Nelken, "Elevated correlations in neuronal ensembles of mouse auditory cortex following parturition," Journal of Neuroscience **33**, 12851–12861 (2013).

[26] Takaki Komiyama, Takashi R Sato, Daniel H OConnor, Ying-Xin Zhang, Daniel Huber, Bryan M Hooks, Mariano Gabitto, and Karel Svoboda, "Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice," Nature **464**, 1182–1186 (2010).

[27] James M Jeanne, Tatyana O Sharpee, and Timothy Q Gentner, "Associative learning enhances population coding by inverting interneuronal correlation patterns," Neuron **78**, 352–363 (2013).

[28] Alfonso Renart, Jaime de la Rocha, Peter Bartho, Liad Hollender, Néstor Parga, Alex Reyes, and Kenneth D Harris, "The asynchronous state in cortical circuits," science **327**, 587–590 (2010).

[29] Timothy J Gawne and Barry J Richmond, "How independent are the messages carried by adjacent inferior temporal cortical neurons?" Journal of Neuroscience **13**, 2758–2771 (1993).

[30] Timothy J Gawne, Troels W Kjaer, John A Hertz, and Barry J Richmond, "Adjacent visual cortical complex cells share about 20% of their stimulus-related information," Cerebral Cortex **6**, 482–489 (1996).

[31] Matthew A Smith and Adam Kohn, "Spatial and temporal scales of neuronal correlation in primary visual cortex," The Journal of Neuroscience **28**, 12591–12603 (2008).

[32] Diego A Gutnisky and Valentin Dragoi, "Adaptive coding of visual information in neural populations," Nature **452**, 220–224 (2008).

[33] Wyeth Bair, Ehud Zohary, and William T Newsome, "Correlated firing in macaque visual area mt: time scales and relationship to behavior," Journal of Neuroscience **21**, 1676–1697 (2001).

[34] Daeyeol Lee, Nicholas L Port, Wolfgang Kruse, and Apostolos P Georgopoulos, "Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex," The Journal of neuroscience **18**, 1161–1170 (1998).

[35] Robert B Levy and Alex D Reyes, "Spatial profile of excitatory and inhibitory synaptic connectivity in mouse

primary auditory cortex," The Journal of Neuroscience **32**, 5609–5619 (2012).

[36] Elodie Fino and Rafael Yuste, "Dense inhibitory connectivity in neocortex," Neuron **69**, 1188–1203 (2011).

[37] Hiroki Tanaka, Hiroshi Tamura, and Izumi Ohzawa, "Spatial range and laminar structures of neuronal correlations in the cat primary visual cortex," Journal of neurophysiology **112**, 705–718 (2014).

[38] Shervin Safavi, Abhilash Dwarakanath, Vishal Kapoor, Joachim Werner, Nicholas G Hatsopoulos, Nikos K Logothetis, and Theofanis I Panagiotaropoulos, "Non-monotonic spatial structure of interneuronal correlations in prefrontal microcircuits," bioRxiv , 128249 (2017).

[39] Robert Rosenbaum, Matthew A Smith, Adam Kohn, Jonathan E Rubin, and Brent Doiron, "The spatial structure of correlated neuronal variability," Nature Neuroscience **20**, 107–114 (2017).

[40] Daniel J Denman and Diego Contreras, "The structure of pairwise correlation in mouse primary visual cortex reveals functional organization in the absence of an orientation map," Cerebral Cortex **24**, 2707–2720 (2014).

[41] Tom Binzegger, Rodney J Douglas, and Kevan AC Martin, "A quantitative map of the circuit of cat primary visual cortex," Journal of Neuroscience **24**, 8441–8453 (2004).

[42] Moshe Abeles, *Corticonics: Neural circuits of the cerebral cortex* (Cambridge University Press, 1991).

[43] Iris Ginzburg and Haim Sompolinsky, "Theory of correlations in stochastic neural networks," Physical review E **50**, 3171 (1994).

[44] Jérémie Barral and Alex D Reyes, "Synaptic scaling rule preserves excitatory-inhibitory balance and salient neuronal network dynamics," Nature neuroscience **19**, 1690–1696 (2016).

[45] Carl van Vreeswijk and Haim Sompolinsky, "Chaos in neuronal networks with balanced excitatory and inhibitory activity," Science **274**, 1724–1726 (1996).

[46] Carl van Vreeswijk and Haim Sompolinsky, "Chaotic balanced state in a model of cortical circuits," Neural computation **10**, 1321–1371 (1998).

[47] Michael Monteforte and Fred Wolf, "Dynamic flux tubes form reservoirs of stability in neuronal circuits," Physical Review X **2**, 041007 (2012).

[48] Alex Roxin, Nicolas Brunel, David Hansel, Gianluigi Mongillo, and Carl van Vreeswijk, "On the distribution of firing rates in networks of cortical neurons," The Journal of neuroscience **31**, 16217–16226 (2011).

[49] Moritz Helias, Tom Tetzlaff, and Markus Diesmann, "The correlation structure of local neuronal networks intrinsically results from recurrent dynamics," PLoS Comput Biol **10**, e1003428 (2014).

[50] David Hansel and Haim Sompolinsky, "Synchronization and computation in a chaotic neural network," Physical Review Letters **68**, 718 (1992).

[51] Zoltán F Kisvárday, E Toth, Martin Rausch, and Ulf T Eysel, "Orientation-specific relationship between populations of excitatory and inhibitory lateral connections in the visual cortex of the cat." Cerebral cortex (New York, NY: 1991) **7**, 605–618 (1997).

[52] Armen Stepanyants, Luis M Martinez, Alex S Ferecskó, and Zoltán F Kisvárday, "The fractions of short-and long-range connections in the visual cortex," Proceedings of the National Academy of Sciences **106**, 3555–3560 (2009).

[53] Rani Ben-Yishai, David Hansel, and Haim Sompolinsky, "Traveling waves and the processing of weakly tuned inputs in a cortical network module," Journal of computational neuroscience **4**, 57–77 (1997).

[54] David Hansel and Haim Sompolinsky, "13 modeling feature selectivity in local cortical circuits," (1998).

[55] C Van Vreeswijk and H Sompolinsky, "Irregular activity in large networks of neurons," Methods and models in neurophysics. Amsterdam: Elsevier (2005).

[56] Robert Rosenbaum and Brent Doiron, "Balanced networks of spiking neurons with spatially dependent recurrent connections," Physical Review X **4**, 021039 (2014).

[57] Eric Shea-Brown, Krešimir Josić, Jaime de La Rocha, and Brent Doiron, "Correlation and synchrony transfer in integrate-and-fire neurons: basic properties and consequences for coding," Physical review letters **100**, 108102 (2008).

[58] Jaime De La Rocha, Brent Doiron, Eric Shea-Brown, Krešimir Josić, and Alex Reyes, "Correlation between neural spike trains increases with firing rate," Nature **448**, 802–806 (2007).

[59] Vladimir Litvak, Haim Sompolinsky, Idan Segev, and Moshe Abeles, "On the transmission of rate code in long feedforward networks with excitatory–inhibitory balance," The Journal of neuroscience **23**, 3006–3015 (2003).

[60] Nelson Cortes and Carl van Vreeswijk, "Pulvinar thalamic nucleus allows for asynchronous spike propagation through the cortex," Frontiers in computational neuroscience **9** (2015).

[61] Matthew A Smith, Xiaoxuan Jia, Amin Zandvakili, and Adam Kohn, "Laminar dependence of neuronal correlations in visual cortex," Journal of neurophysiology **109**, 940–947 (2013).

[62] Nicolas Brunel, "Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons," Journal of computational neuroscience **8**, 183–208 (2000).

[63] Kanaka Rajan and LF Abbott, "Eigenvalue spectra of random matrices for neural networks," Physical review letters **97**, 188104 (2006).

Appendix A: Correlations in binary networks

Here we calculate the equilibrium value and the stability of the quenched average correlations.

We define, for $(j, \beta) \neq (i, \alpha)$, the out of equilibrium auto- and crosscorrelations, $c_{ij}^{\alpha\beta}(t, \tau)$ as

$$c_{ij}^{\alpha\beta}(t, \tau) \equiv \langle \delta S_i^{\alpha}(t) \delta S_j^{\beta}(t + \tau) \rangle_{\text{init}}$$
$$a_i^{\alpha}(t, \tau) \equiv \langle \delta S_i^{\alpha}(t) \delta S_i^{\alpha}(t + \tau) \rangle_{\text{init}}$$

where $\langle \cdot \rangle_{\text{init}}$ denotes averaging over many initial conditions choosen with a probability measure that, for simplicity, we choose such that $\langle S_i^{\alpha}(0) \rangle_{\text{init}} = \langle S_i^{\alpha}(t) \rangle_t$. It is also convenient for the notation to define $c_{ii}^{\alpha\alpha}(t, \tau) = 0$. In this paper we focus on the equal time correlations, $c_{ij}^{\alpha\beta}(t) \equiv c_{ij}^{\alpha\beta}(t, 0)$.

For networks of binary neurons the dynamics of the

equal-time crosscorrelations is given by [1, 28, 43]

$$\frac{dc_{ij}^{\alpha\beta}(t)}{dt} = -2c_{ij}^{\alpha\beta}(t) + \langle \delta\Theta[h_i^\alpha(t) - T]\delta S_j^\beta(t)\rangle_{\text{init}}$$
$$+ \langle \delta S_i^\alpha(t)\delta\Theta[h_j^\beta(t) - T]\rangle_{\text{init}},$$

where $\delta\Theta[h_i^\alpha(t) - T] \equiv \Theta[h_i^\alpha(t) - T] - \langle\Theta[h_i^\alpha(t) - T]\rangle_{\text{init}}$.

If we make the *Ansatz* that correlations are weak, we can, to leading order, take $\langle\delta\Theta[h_i^\alpha(t) - T]\delta S_j^\beta(t)\rangle_{\text{init}} = g_i^\alpha\langle\delta h_i^\alpha(t)\delta S_j^\beta(t)\rangle_{\text{init}}$, where g_i^α is the gain of neuron (i, α), $g_i^\alpha = \partial_h\langle\Theta[h_i^\alpha(t) - T]\rangle_{\text{init}}$, which is, to leading order, independent of the correlations [28, 43, 49].

Thus,

$$\frac{dc_{ij}^{\alpha\beta}(t)}{dt} = -2c_{ij}^{\alpha\beta}(t) + g_i^\alpha\langle\delta h_i^\alpha(t)\delta S_j^\beta(t)\rangle_{\text{init}}$$
$$+ g_j^\beta\langle\delta S_i^\alpha(t)\delta h_j^\beta(t)\rangle_{\text{init}}$$
$$= -2c_{ij}^{\alpha\beta}(t) + g_i^\alpha\sum_{\gamma,k}J_{ik}^{\alpha\gamma}\langle\delta S_k^\gamma(t)\delta S_j^\beta(t)\rangle_{\text{init}} +$$
$$+ g_j^\beta\sum_{\gamma,k}J_{jk}^{\beta\gamma}\langle\delta S_k^\gamma(t)\delta S_i^\alpha(t)\rangle_{\text{init}}.$$

Using $\langle\delta S_i^\alpha(t)\delta S_j^\beta(t)\rangle_{\text{init}} = c_{ij}^{\alpha\beta}(t) + \delta_{\alpha,\beta}\delta_{i,j}a_i^\alpha(t)$ yields

$$\frac{dc_{ij}^{\alpha\beta}(t)}{dt} = -2c_{ij}^{\alpha\beta}(t) + \sum_{\gamma,k}\left[\tilde{J}_{ik}^{\alpha\gamma}c_{kj}^{\gamma\beta}(t) + \tilde{J}_{jk}^{\beta\gamma}c_{ik}^{\alpha\gamma}(t)\right] +$$
$$+ \tilde{J}_{ij}^{\alpha\beta}a_j^\beta(t) + \tilde{J}_{ji}^{\beta\alpha}a_i^\alpha(t), \qquad \text{(A1)}$$

where $\tilde{J}_{ij}^{\alpha\beta} = g_i^\alpha J_{ij}^{\alpha\beta}$. Note that, for weak correlations $\langle S_i^\alpha(t)\rangle_{\text{init}}$ does not depend on the correlations so that, $\langle S_i^\alpha(t)\rangle_{\text{init}} = \langle S_i^\alpha(t)\rangle_t$, for our choice of initial conditions $(\langle S_i^\alpha(0)\rangle_{\text{init}} = \langle S_i^\alpha(t)\rangle_t)$. Hence, the equal time autocorrelation a_i^α is independent of time and given by $a_i^\alpha = \langle S_i^\alpha(t)\rangle_t - \langle S_i^\alpha(t)\rangle_t^2$.

We now average over the quenched disorder. Due to the rotational symmetry of the connection probabilities, $P_{ij}^{\alpha\beta}$, $[a_i^\alpha]_J$ is a constant

$$[a_i^\alpha]_J = A_\alpha$$

whereas $[c_{ij}^{\alpha\beta}(t)]_J$ and $[\tilde{J}_{ij}^{\alpha\beta}]_J$ are functions of the difference in the location of neurons (i, α) and (j, β)

$$[c_{ij}^{\alpha\beta}(t)]_J = C_{\alpha\beta}(\theta_i^\alpha - \theta_j^\beta, t)$$

$$[\tilde{J}_{ij}^{\alpha\beta}]_J = \frac{1}{N}\mathcal{J}_{\alpha\beta}(\theta_i^\alpha - \theta_j^\beta)$$

where $\mathcal{J}_{\alpha\beta}(\Delta) = \sqrt{K}g_\alpha J_{\alpha\beta}f_{\alpha\beta}(\Delta)$. Here we have assumed that the correlations in the quenched disorder in the inputs to the neurons are small, such that, to leading order, the expected value of the gain does not depend on the neuronal position. We comment on this Ansatz in Appendix C.

Thus, for large N, the quenched average of Eq. (A1) yields

$$\frac{d}{dt}C_{\alpha\beta}(\Delta, t) = -2C_{\alpha\beta}(\Delta, t) + \sum_\gamma\int\frac{d\Delta'}{2\pi}$$
$$\left[\mathcal{J}_{\alpha\gamma}(\Delta - \Delta')C_{\gamma\beta}(\Delta', t) +\right.$$
$$\left.+ \mathcal{J}_{\gamma\beta}(-\Delta')C_{\alpha\gamma}(\Delta - \Delta', t)\right] +$$
$$+ \mathcal{J}_{\alpha\beta}(\Delta)\frac{A_\beta}{N} + \mathcal{J}_{\beta\alpha}(-\Delta)\frac{A_\alpha}{N}. \quad \text{(A2)}$$

Here we have not taken into account that $c_{ii}^{\alpha\alpha}(t) = 0$. It is easy to see, however, that in the cross-correlations this neglects $\mathcal{O}(1/N^2)$ corrections. We show below that these corrections are indeed negligeable in the large N limit.

The nth Fourier mode of $C_{\alpha\beta}(\Delta, t)$, $C_{\alpha\beta}^{(n)}(t)$, satisfies

$$\frac{d}{dt}C_{\alpha\beta}^{(n)}(t) = -2C_{\alpha\beta}^{(n)}(t)$$
$$+ \sum_\gamma\left[\mathcal{J}_{\alpha\gamma}^{(n)}C_{\gamma\beta}^{(n)}(t) + C_{\alpha\gamma}^{(n)}(t)[\mathcal{J}_{\gamma\beta}^{(n)}]^\dagger\right]$$
$$+ \mathcal{J}_{\alpha\beta}^{(n)}\frac{A_\beta}{N} + \frac{A_\alpha}{N}[\mathcal{J}_{\alpha\beta}^{(n)}]^\dagger,$$

where $\mathcal{J}_{\alpha\beta}^{(n)}$ and $[\mathcal{J}_{\alpha\beta}^{(n)}]^\dagger$ are the nth Fourier mode of $\mathcal{J}_{\alpha\beta}(\Delta)$ and $\mathcal{J}_{\alpha\beta}^\top(\Delta)$, respectively.

This can be written more compactly as

$$\tau\frac{d}{dt}\boldsymbol{C}^{(n)}(t) = -2\boldsymbol{C}^{(n)}(t) + \boldsymbol{\mathcal{J}}^{(n)}\boldsymbol{C}^{(n)}(t) + \boldsymbol{C}^{(n)}(t)[\boldsymbol{\mathcal{J}}^{(n)}]^\dagger$$
$$+ \boldsymbol{\mathcal{J}}^{(n)}\frac{\boldsymbol{A}}{N} + \frac{\boldsymbol{A}}{N}[\boldsymbol{\mathcal{J}}^{(n)}]^\dagger. \quad \text{(A3)}$$

where $\boldsymbol{X}$ denotes the $D \times D$ matrix, $X_{\alpha\beta}$, and $\boldsymbol{A}$ is the $D \times D$ matrix $A_{\alpha\beta} \equiv \delta_{\alpha,\beta}A_\alpha$. The equilibrium values of the Fourier components of equal-time correlation functions thus satisfy Eq. (26).

Appendix B: Correlation theorems

In Appendix A we derived N^2D^2 coupled linear differential equations that determine the evolution of equal-time cross-correlations for all neuronal pairs in the network. Averaging these equations over the quenched disorder and using the rotation invariance of the connection probabilities yields a set of ND^2 coupled equations for the quenched averaged correlations. In Fourier space these equations lead to N independent sets of D^2 coupled equations for the correlations. Here we prove Theorems 1 and 2 (see Section V) which state how these correlations scale with N and K.

To leading order, $\boldsymbol{A}$ is independent of K and N, but $\boldsymbol{\mathcal{J}}^{(n)}$ is proportional to $\sqrt{K}$. Accordingly, we define $\bar{\boldsymbol{\mathcal{J}}}^{(n)} \equiv \boldsymbol{\mathcal{J}}^{(n)}/\sqrt{K}$ and $[\bar{\boldsymbol{\mathcal{J}}}^{(n)}]^\dagger \equiv [\boldsymbol{\mathcal{J}}^{(n)}]^\dagger/\sqrt{K}$. Thus, we can rewrite the evolution equation of the nth mode of

the correlations as

$$\frac{d}{dt}\boldsymbol{C}^{(n)}(t) = -2\boldsymbol{C}^{(n)}(t)$$
$$+ \sqrt{K}\left[\bar{\boldsymbol{\mathcal{J}}}^{(n)}\boldsymbol{C}^{(n)}(t) + \boldsymbol{C}^{(n)}(t)[\bar{\boldsymbol{\mathcal{J}}}^{(n)}]^\dagger\right]$$
$$+ \frac{\sqrt{K}}{N}\left[\bar{\boldsymbol{\mathcal{J}}}^{(n)}\boldsymbol{A}(t) + \boldsymbol{A}(t)[\bar{\boldsymbol{\mathcal{J}}}^{(n)}]^\dagger\right]. \quad (B1)$$

The $D \times D$ matrix, $\bar{\boldsymbol{\mathcal{J}}}^{(n)}$ can be written as

$$\bar{\boldsymbol{\mathcal{J}}}^{(n)} = \boldsymbol{U}^{(n)}\boldsymbol{\mathcal{J}}_{\text{jor}}^{(n)}[\boldsymbol{U}^{(n)}]^{-1}.$$

where $\boldsymbol{\mathcal{J}}_{\text{jor}}^{(n)}$ is the Jordan normal form of $\bar{\boldsymbol{\mathcal{J}}}^{(n)}$ and $[\boldsymbol{U}^{(n)}]^{-1}$ is the transformation matrix to the Jordan basis.

The matrix $\boldsymbol{\mathcal{J}}_{\text{jor}}^{(n)}$ can be written as

$$[\boldsymbol{\mathcal{J}}_{\text{jor}}^{(n)}]_{\mu\nu} = \lambda_\mu^{(n)}\delta_{\mu,\nu} + \epsilon_\mu^{(n)}\delta_{\mu,\nu-1}.$$

where $\lambda_\mu^{(n)}$ are the eigenvalues of $\bar{\boldsymbol{\mathcal{J}}}^{(n)}$ and $\epsilon_\mu^{(n)} = 1$ inside a Jordan block and is 0 otherwise (for clarity, in the Jordan basis we use the subscripts μ and ν, rather than α and β that we use in the original basis).

Importantly, the Jordan form of a matrix and of its Hermitian conjugate are complex conjugate. We thus can write

$$[\bar{\boldsymbol{\mathcal{J}}}^{(n)}]^\dagger = \boldsymbol{V}^{(n)}[\boldsymbol{\mathcal{J}}_{\text{jor}}^{(n)}]^*[\boldsymbol{V}^{(n)}]^{-1},$$

For notational convenience we will suppress the superscript (n) in the rest of this Appendix. Defining $\hat{\boldsymbol{C}}$ as

$$\hat{\boldsymbol{C}} = \boldsymbol{U}^{-1}\boldsymbol{C}\boldsymbol{V}$$

and inserting into Eq. (B1) yields

$$\frac{d}{dt}\hat{C}_{\mu\nu}(t) = -\Lambda_{\mu\nu}\hat{C}_{\mu\nu}(t)+$$
$$\sqrt{K}\left[\epsilon_\mu\hat{C}_{\mu+1,\nu}(t) + \epsilon_{\nu-1}\hat{C}_{\mu,\nu-1}(t)\right]$$
$$+ \frac{\sqrt{K}}{N}\left[(\lambda_\mu + \lambda_\nu^*)\hat{A}_{\mu\nu} +\right.$$
$$\left.+ \epsilon_\mu\hat{A}_{\mu+1,\nu} + \epsilon_{\nu-1}\hat{A}_{\mu,\nu-1}\right], \quad (B2)$$

where we have defined $\Lambda_{\mu\nu} = 2 - \sqrt{K}(\lambda_\mu + \lambda_\nu^*)$ and

$$\hat{\boldsymbol{A}} = \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{V}.$$

We now assume that the connectivity is such that the system is stable to perturbations of the locally averaged rates. This implies that the real part of all eigenvalues of $\bar{\boldsymbol{\mathcal{J}}}$ are less than $1/\sqrt{K}$. The real part of $\Lambda_{\mu\nu}$ is therefore positive and thus $\hat{\boldsymbol{C}}(t)$ converges to an equilibrium value, $\hat{\boldsymbol{C}}^\infty$.

To see this, first consider $\hat{C}_{D,1}(t)$ which satisfies

$$\frac{d}{dt}\hat{C}_{D,1}(t) = -\Lambda_{D,1}\hat{C}_{D,1}(t) + \frac{\sqrt{K}}{N}(\lambda_D + \lambda_1^*)\hat{A}_{D,1}.$$

Therefore $\hat{C}_{D,1}(t)$, converges to its equilibrium value.

The evolution equations of $\hat{C}_{D,2}$ can be written as

$$\frac{d}{dt}\hat{C}_{D,2}(t) = -\Lambda_{D,2}\hat{C}_{D,2}(t) + M_{D,2}\hat{C}_{D,2}(t)$$

where $M_{D,2}$ depends on $\hat{C}_{D,1}(t)$, $\hat{A}_{D,1}$ and $\hat{A}_{D,2}$. Since $\hat{C}_{D,1}(t)$ converges to its equilibrium value, $M_{D,2}$ converges to a constant. Because $\text{Re}(\Lambda_{D,2}) > 0$, $\hat{C}_{D,2}$ also converges to its equilibrium value, $\hat{\boldsymbol{C}}_{D,1}^\infty$. A similar argument shows that likewise $\hat{C}_{D-1,1}$ converges to $\hat{C}_{D-1,1}^\infty$. One then sees by recursion that the whole matrix $\hat{\boldsymbol{C}}$ converges to its equilibrium value, $\hat{\boldsymbol{C}}^\infty$. Hence, Eq. (26) determines the *stable* equilibrium values of the correlations.

From here we only consider the correlations at equilibrium and, for notational simplicity, we drop the superscript ∞.

In the Jordan basis the equilibrium values of the correlations satisfy

$$\Lambda_{\mu\nu}\hat{C}_{\mu\nu} = \sqrt{K}\left[\epsilon_\mu\hat{C}_{\mu+1,\nu} + \epsilon_{\nu-1}\hat{C}_{\mu,\nu-1}\right] +$$
$$+ \frac{\sqrt{K}}{N}\left[(\lambda_\mu + \lambda_\nu^*)\hat{A}_{\mu\nu}+\right.$$
$$\left.+\epsilon_\mu\hat{A}_{\mu+1,\nu} + \epsilon_{\nu-1}\hat{A}_{\mu,\nu-1}\right], \quad (B3)$$

Let us now consider the case where $\bar{\boldsymbol{\mathcal{J}}}$ is diagonizable, so that $\epsilon_\mu = 0$ for all μ. In this case $\hat{C}_{\mu\nu} = 0$ for $\mu \neq \nu$ and

$$\hat{C}_{\mu\mu} = \frac{\sqrt{K}(\lambda_\mu + \lambda_\mu^*)}{N[2 - \sqrt{K}(\lambda_\mu + \lambda_\mu^*)]}\hat{A}_{\mu\mu},$$

Thus

$$\lim_{K\to\infty}\left(\lim_{N\to\infty}N\hat{C}_{\mu\mu}\right) = -\hat{A}_{\mu\mu} \quad (B4)$$

unless $|\lambda_\mu + \lambda_\nu^*|$ is $o(1/\sqrt{K})$, in which case

$$\lim_{K\to\infty}\left(\lim_{N\to\infty}N\hat{C}_{\mu\nu}\right) = 0 \quad (B5)$$

In the first situation $\hat{C}_{\mu\mu} = \mathcal{O}(1/N)$. In the second situation $\hat{C}_{\mu\mu} = o(1/N)$.

Let us now assume that $\bar{\boldsymbol{\mathcal{J}}}$ is not diagonizable. Then the $D \times D$ Jordan form of $\bar{\boldsymbol{\mathcal{J}}}$ consists of B Jordan blocks $(1 \leq B < D)$ that we denote by $[\bar{\boldsymbol{\mathcal{J}}}_{\text{jor}}]^i$, $i = 1, ..., B$. The size of the ith block will be denoted by $s(i) \times s(i)$. Without loss of generality, we can assume that the blocks are ordered in increasing size.

The indices μ and ν of the elements of the Jordan block i, take values between $l(i)$ and $h(i)$, with $l(i) = 1 + \sum_{j=1}^{i-1}s(j)$ and $h(i) = \sum_{j=1}^{i}s(j)$. All the diagonal elements of this Jordan block are equal one of the eigenvalues of $\boldsymbol{\mathcal{J}}$, which we denote by $\hat{\lambda}_i$. The off-diagonal elements, ϵ_μ, are all equal to 1 except for $\mu = h(i)$ $(i = 1, .., B-1)$ for which $\epsilon_\mu = 0$.

The matrix $\hat{C}$ consists of B^2 sectors that we denote by S_{ij}. In the sector S_{ij}, $\mu \in \{l(i), \ldots, h(i)\}$ and $\nu \in \{l(j), \ldots, h(j)\}$.

Let us consider Eq. (B3) for μ, ν in sector S_{ij}. Since $\epsilon_{h(i)}$ and $\epsilon_{l(j)-1}$ are zero, the equation in this sector does not depend on elements of $\hat{C}$ outside of it. Thus, we can solve Eq. (B3) recursively to determine all the elements of $\hat{C}$ in this sector. The recursion goes as follows. First, one solves for $\hat{C}_{h(i),l(j)}$

$$\hat{C}_{h(i),l(j)} = \frac{1}{N}\frac{\sqrt{K}\left(\hat{\lambda}_i + \hat{\lambda}_j^*\right)}{2 - \sqrt{K}\left(\hat{\lambda}_i + \hat{\lambda}_j^*\right)}\hat{A}_{h(i)l(j)} \qquad (B6)$$

We can the solve Eq. (B3) to get $\hat{C}_{\mu\nu}$ for $\mu, \nu = h(i) - 1, l(j)$ and $\mu, \nu = h(i), l(j) + 1$. This process can be repeated until all the elements of $\hat{C}$ in the sector S_{ij} are determined.

1. $\hat{A}_{h(i),l(j)}$ are all non-zero

A similar recursion can be performed to estimate the order of magnitude of all the elements of $\hat{C}$. First we obtain

$$\hat{C}_{h(i),l(j)} = \mathcal{O}\left(\frac{\hat{A}_{h(i),l(j)}}{N}\right) = \mathcal{O}\left(\frac{1}{N}\right).$$

For $\hat{\lambda}_i + \hat{\lambda}_j^* = \mathcal{O}(1)$ the recursion shows that we have

$$\hat{C}_{\mu\nu} = \mathcal{O}\left(\frac{1}{N}\right) \qquad (B7)$$

for all μ, ν in sector S_{ij}.

If $\hat{\lambda}_i + \hat{\lambda}_j^* = 0$, $\hat{C}_{h(i),l(j)} = 0$ and by recursion all the other elements in sector S_{ij} are

$$\hat{C}_{\mu\nu} = \mathcal{O}\left(K^{\ell/2}/N\right), \qquad (B8)$$

where $\ell = \nu - \mu + h(i) - l(j)$. Thus, in this case $\hat{C}_{l(i),h(j)}$ is the largest entry in the sector. It satisfies

$$\hat{C}_{l(i),h(j)} = \mathcal{O}\left(K^{(s(i)+s(j)-2)/2}/N\right) \qquad (B9)$$

The scaling of the elements in sector S_{ij} thus depends $\hat{\lambda}_i + \hat{\lambda}_j^*$. The stability of the dynamics imposes that $\mathcal{R}e(\hat{\lambda}_i) < 0$ for all blocks (or positive but at most $\mathcal{O}(1/\sqrt{K})$, see next subsection) and thus the condition $\hat{\lambda}_i + \hat{\lambda}_j^* = 0$ implies that, for large K, the real part of $\hat{\lambda}_i$ and $\hat{\lambda}_j$ are zero. In other words, the real part of $[\mathcal{J}_{\mathrm{jor}}]^i$ and $[\mathcal{J}_{\mathrm{jor}}]^j$ are shift matrices of size $s(i) \times s(i)$ and $s(j) \times s(j)$, respectively. For the sector S_{ii} it means that the real part of $[\mathcal{J}_{\mathrm{jor}}]^i$ is a shift matrix of size $s(i) \times s(i)$.

Proving Correlation Theorem 1 is now straightforward. If the real part of all the B Jordan blocks of $\bar{\mathcal{J}}$ are different from a shift matrix, one finds that $\hat{\lambda}_i + \hat{\lambda}_j^* = \mathcal{O}(1) \neq 0$

for all $i, j \in \{1, \ldots, B\}$. According to Eq. (B7), correlations are at most $\mathcal{O}(1/N)$ in all of the B^2 sectors of the matrix $\hat{C}$. As a result, $\boldsymbol{C}^{(n)} = \mathcal{O}(1/N)$. On the other hand, if in each of the B^2 sectors of the matrix $\boldsymbol{C}$ correlations are at most $\mathcal{O}(1/N)$, this is also the case for the element of $\hat{C}$ in all the sectors. In this case, Eq. (B7) implies that there is no Jordan block in $\bar{\mathcal{J}}$ for which $\mathcal{R}e[\hat{\lambda}_i] = 0$.

Restoring the index n of the Fourier mode, one sees that if $\bar{\mathcal{J}}^{(n)}$ has at least one Jordan block whose real part is a shift matrix and denoting by $P(n)$ the size of the largest shift, Eq. (B9) implies that $\hat{C}_{l(i),h(i)} = \mathcal{O}(K^{P(n)-1}/N)$, and thus also $\hat{\boldsymbol{C}}^{(n)} = \mathcal{O}(K^{P(n)-1}/N)$. This proves Correlation Theorem 2.

According to Eq.(B9), the scaling of the correlation is $\hat{\boldsymbol{C}}^{(n)} = \mathcal{O}(K^{D-1}/N)$ if and only if $P(n) = D$. When $\bar{\mathcal{J}}^{(n)}$ is a real matrix (when the probability of connections are symmetric in Δ), it means that the Jordan form of $\bar{\mathcal{J}}^{(n)}$ is a shift matrix of size D. This is equivalent for saying that $\bar{\mathcal{J}}^{(n)}$ is nilpotent of degree D. On the other hand, if $\bar{\mathcal{J}}^{(n)}$ is a shift matrix of degree D, it has a Jordan form which is a shift matrix of degree D. According to Eq.(B9), the scaling of the correlation is then $\hat{\boldsymbol{C}}^{(n)} = \mathcal{O}(K^{D-1}/N)$. This proves the Corollary in section V.

When $K = \mathcal{O}(1/|\hat{\lambda}_i+\hat{\lambda}_j^*|^2)$ there is a crossover in sector S_{ij} of the matrix $\hat{C}$ between weak correlations ($\mathcal{O}(1/N)$) and strong correlations ($\mathcal{O}(K^{(s(i)+s(j)-2)/2}/N)$). To see this, we note that if $K = |\hat{\lambda}_i + \hat{\lambda}_j^*|^{-\frac{1}{\Gamma}}$, for $\Gamma \leq 1/2$

$$\hat{C}_{\mu\nu} = \mathcal{O}\left(\frac{K^{\Gamma\ell}}{N}\right)$$

whereas for $1/2 < \Gamma \leq 1$

$$\hat{C}_{\mu\nu} = \mathcal{O}\left(\frac{K^{\ell/2}}{N}\right)$$

for $\ell \neq 0$ and

$$\hat{C}_{\mu\nu} = \mathcal{O}\left(\frac{K^{1/2-\Gamma}}{N}\right).$$

for $\ell = 0$

2. Some elements of $\hat{A}$ are zero

So far, we have assumed that all the elements in $\hat{A}$ are non zero. The derivation can, however, be extended to include also situations where this is not the case as follows.

Elements $\hat{A}_{\mu\nu} = 0$ in a sector where $\hat{\lambda}_i + \hat{\lambda}_j^* = \mathcal{O}(1)$ might result in some elements of $\hat{C}$ in that sector to be equal to 0 rather than $\mathcal{O}\left(\frac{1}{N}\right)$. However, this will not affect the overall scaling of the correlations.

Elements $\hat{A}_{\mu\nu} = 0$ in a sector S_{ij} where $\hat{\lambda}_i + \hat{\lambda}_j^* = 0$, do not change the scaling in the matrix, provided that

$\hat{A}_{h(i),l(j)} \neq 0$. If $\hat{A}_{h(i),l(j)} = 0$ the effect depends on $\hat{A}_{h(i)-1,l(j)}$ and $\hat{A}_{h(i),l(j)+1}$. If at least one of them is nonzero, the order of the largest entry in the sector is decreased by $\sqrt{K}$. If both of them are zero, the order of this entry decreases at least by a factor of K. Reflecting on further element being zero, one reaches the following conclusion: Let $\{\mu_{ij}, \nu_{ij}\}$ be the indices $\{\mu, \nu\}$ in S_{ij} for which $\hat{A}_{\mu\nu} \neq 0$, which maximize $\mu - \nu$. Then, the maximal order of $\hat{C}_{\mu\nu}$ in the sector is $\mathcal{O}(K^{P_{ij}-1}/N)$, where

$$P_{ij} = 1 + (\mu_{ij} - l(i) + h(i) - \nu_{ij})/2. \qquad (B10)$$

In conclusion: The highest order in $\hat{C}$ is $\mathcal{O}(K^{P-1}/N)$, where P is the maximum value of the P_{ij} $(i,j \in \{1,\ldots,B\})$ defined by: 1) $P_{ij} = 1$ for sector S_{ij} for which $\hat{\lambda}_i + \hat{\lambda}_j^* \neq 0$. 2) $P_{ij} = P$ for sector S_{ij} in which $\hat{\lambda}_i + \hat{\lambda}_j^* = 0$ and all $\hat{A}_{\mu\nu} \neq 0$. 3) P_{ij} is given by Eq. (B10) if in sector S_{ij}, $\hat{\lambda}_i + \hat{\lambda}_j^* = 0$ but some $\hat{A}_{\mu\nu}$ are zero.

Appendix C: Correlations of the quenched disorder

Let us define the correlation of the quenched disorder in the outputs of the neurons as

$$\Gamma_{ij}^{\alpha\beta} = [\Delta S_i^\alpha \Delta S_j^\beta]_J \qquad (C1)$$

for $(i,\alpha) \neq (j,\beta)$ and ΔS_i^α defined in Eq.(3). A derivation similar to that in Appendix A yields

$$2\Gamma_{ij}^{\alpha\beta} = \sum_{\gamma,k} \left[\mathcal{J}_{ik}^{\alpha\gamma}\Gamma_{kj}^{\gamma\beta} + \mathcal{J}_{jk}^{\beta\gamma}\Gamma_{ik}^{\alpha\gamma} \right] + \\ + \mathcal{J}_{ij}^{\alpha\beta}q_\beta/N + \mathcal{J}_{ji}^{\beta\alpha}q_\alpha/N, \qquad (C2)$$

where $q_\alpha = [(\Delta S_i^\alpha)^2]_J$. This equation can be solved using the same approach as in Appendices A-B. This analysis shows that correlations of the quenched disorder and correlations of the temporal fluctuations are of the same order. Thus, when the temporal fluctuations are small, the Ansatz in Appendix A, were we neglected the spatial fluctuations in the neuronal gain, is justified. If the correlations are too strong, the Ansatz may no longer be satisfied, but neither is the linearization assumed to derive equation (A2).

Appendix D: Self-consistent equations for the autocorrelation and the gain

We follow the notations from [46]:

$$m_\alpha = \frac{1}{2}\text{erfc}\left(\frac{T_\alpha - h_\alpha}{\sqrt{2(\sigma_\alpha^2 + \sigma_{q\alpha}^2)}} \right) \qquad (D1)$$

$$q_\alpha = \frac{1}{2}\int Dx\, \text{erfc}^2\left(\frac{T_\alpha - h_\alpha - \sigma_{q\alpha}x}{\sqrt{2\sigma_\alpha^2}} \right) \qquad (D2)$$

$$A_\alpha = m_\alpha - q_\alpha \qquad (D3)$$

where the variance of the input noise, $\sigma_\alpha^2 = [\langle(h_i^\alpha - \langle h_i^\alpha\rangle_t)^2\rangle_t]_J$, is

$$\sigma_\alpha^2 = \sum_\beta J_{\alpha\beta}^2(m_\beta - q_\beta) + K\sum_{\beta\beta'} \frac{J_{\alpha\beta}J_{\alpha\beta'}}{N^2} \sum_{j\neq j'} f_{\alpha\beta}(\theta_i^\alpha - \theta_j^\beta)f_{\alpha\beta'}(\theta_i^\alpha - \theta_{j'}^{\beta'})C_{\beta\beta'}(\theta_j^\beta - \theta_{j'}^{\beta'}) \qquad (D4)$$

and the quenched disorder in the inputs, $\sigma_{q\alpha}^2 = [(\langle h_i^\alpha\rangle_t - [\langle h_i^\alpha\rangle_t]_J)^2]_J$, is

$$\sigma_{q\alpha}^2 = \sum_\beta J_{\alpha\beta}^2\left(q_\beta - p\int \frac{d\theta}{2\pi}f^2(\theta)m_\beta^2\right) + K\sum_{\beta\beta'} J_{\alpha\beta}J_{\alpha\beta'}\frac{1}{N^2} \sum_{j\neq j'} f_{\alpha\beta}(\theta_i^\alpha - \theta_j^\beta)f_{\alpha\beta'}(\theta_i^\alpha - \theta_{j'}^{\beta'})[\Delta S_j^\beta \Delta S_{j'}^{\beta'}]_J \qquad (D5)$$

Finally, the gain of the neurons is

$$g_\alpha = \frac{1}{\sqrt{2\pi}\sigma_{T\alpha}}e^{\frac{-(T_\alpha - h_\alpha)^2}{2(\sigma_\alpha^2 + \sigma_{q\alpha}^2)}} \qquad (D6)$$

Equations (D1)-(D3) need to be solved self consistently. For simplicity, when we solved them we neglect $\mathcal{O}(p)$ terms (but see [49]).

Appendix E: Cross-correlations in two-population networks

For a two-population network Eq. (26) yields

$$\boldsymbol{B}^{(n)} \begin{pmatrix} C_{EE}^{(n)} \\ C_{EI}^{(n)} \\ C_{II}^{(n)} \end{pmatrix} = \frac{1}{N} \boldsymbol{D}^{(n)} \begin{pmatrix} A_E \\ A_I \end{pmatrix} \qquad \text{(E1)}$$

with

$$\boldsymbol{B}^{(n)} = \begin{pmatrix} 2(1 - \mathcal{J}_{EE}^{(n)}) & -2\mathcal{J}_{EI}^{(n)} & 0 \\ -\mathcal{J}_{IE}^{(n)} & 2 - (\mathcal{J}_{EE}^{(n)} + \mathcal{J}_{II}^{(n)}) & -\mathcal{J}_{EI}^{(n)} \\ 0 & -2\mathcal{J}_{IE}^{(n)} & 2(1 - \mathcal{J}_{II}^{(n)}) \end{pmatrix}$$

$$C_{EE}^{(n)} = \frac{1}{N} \frac{-2A_E \bar{\mathcal{J}}_{EE}^{(n)} \sqrt{K} + (-A_I (\bar{\mathcal{J}}_{EI}^{(n)})^2 + A_E (\bar{\mathcal{J}}_{EE}^{(n)} T^{(n)} + \Delta^{(n)} + \bar{\mathcal{J}}_{EE}^{(n)} \bar{\mathcal{J}}_{II}^{(n)})) K - A_E T^{(n)} \Delta^{(n)} K^{3/2}}{-2 + 3T^{(n)} \sqrt{K} - ((T^{(n)})^2 + 2\Delta^{(n)}) K + T^{(n)} \Delta^{(n)} K^{3/2}}$$

$$C_{EI}^{(n)} = -\frac{1}{N} \frac{(A_E \bar{\mathcal{J}}_{IE}^{(n)} + A_I \bar{\mathcal{J}}_{EI}^{(n)}) \sqrt{K} - (A_I \bar{\mathcal{J}}_{EE}^{(n)} \bar{\mathcal{J}}_{EI}^{(n)} + A_E \bar{\mathcal{J}}_{II}^{(n)} \bar{\mathcal{J}}_{IE}^{(n)}) K}{-2 + 3T^{(n)} \sqrt{K} - ((T^{(n)})^2 + 2\Delta^{(n)}) K + T^{(n)} \Delta^{(n)} K^{3/2}}$$

$$C_{II}^{(n)} = \frac{1}{N} \frac{-2A_I \bar{\mathcal{J}}_{II}^{(n)} \sqrt{K} + (-A_E (\bar{\mathcal{J}}_{IE}^{(n)})^2 + A_I (\bar{\mathcal{J}}_{II}^{(n)} T^{(n)} + \Delta^{(n)} + \bar{\mathcal{J}}_{EE}^{(n)} \bar{\mathcal{J}}_{II}^{(n)})) K - A_I T^{(n)} \Delta^{(n)} K^{3/2}}{-2 + 3T^{(n)} \sqrt{K} - ((T^{(n)})^2 + 2\Delta^{(n)}) K + T^{(n)} \Delta^{(n)} K^{3/2}}$$

and

$$\boldsymbol{D}^{(n)} = \begin{pmatrix} 2\mathcal{J}_{EE}^{(n)} & 0 \\ \mathcal{J}_{IE}^{(n)} & \mathcal{J}_{EI}^{(n)} \\ 0 & 2\mathcal{J}_{II}^{(n)} \end{pmatrix}$$

Solving this equation one gets

where $T^{(n)} = Tr\bar{\boldsymbol{\mathcal{J}}}^{(n)}, \Delta^{(n)} = \det \bar{\boldsymbol{\mathcal{J}}}^{(n)}$. After some algebra, this equation can be rewritten as Eq. (27).