



Investigating the Correlation–Firing Rate Relationship in Heterogeneous Recurrent Networks

Andrea K. Barreiro¹ · Cheng Ly²

Received: 5 January 2018 / Accepted: 21 May 2018 / Published online: 06 June 2018

© The Author(s) 2018. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Abstract The structure of spiking activity in cortical networks has important implications for how the brain ultimately codes sensory signals. However, our understanding of how network and intrinsic cellular mechanisms affect spiking is still incomplete. In particular, whether cell pairs in a neural network show a positive (or no) relationship between pairwise spike count correlation and average firing rate is generally unknown. This relationship is important because it has been observed experimentally in some sensory systems, and it can enhance information in a common population code. Here we extend our prior work in developing mathematical tools to succinctly characterize the correlation and firing rate relationship in heterogeneous coupled networks. We find that very modest changes in how heterogeneous networks occupy parameter space can dramatically alter the correlation–firing rate relationship.

Keywords Recurrent network · Spike count correlation · Linear response · Heterogeneity

Abbreviations

E excitatory

I inhibitory

Asyn asynchronous

Strong Asyn strong asynchronous

✉ A.K. Barreiro
abarreiro@smu.edu

C. Ly
CLy@vcu.edu

¹ Department of Mathematics, Southern Methodist University, Dallas, USA

² Department of Statistical Science and Operations Research, Virginia Commonwealth University, Richmond, USA

1 Introduction

One prominent goal of theoretical neuroscience is to understand how spiking statistics of cortical networks are modulated by network attributes [9, 28, 42]. This understanding is essential to the larger question of how sensory information is encoded and transmitted, because the statistics of neural activity impact population coding [7, 15–17, 37]. One family of statistics that is implicated in nearly all population coding studies is trial-to-trial variability (and co-variability) in spike counts; there is now a rich history of studying how these statistics arise, and how they effect coding of stimuli [1, 10, 18, 25, 33]. Recent work has demonstrated that the information content of spiking neural activity depends on spike count correlations and its relationship (if any) with stimulus tuning [1, 6, 18, 25, 44].

An important relationship observed in many experimental studies is that pairwise correlations on average increase with firing rates. This has been observed *in vitro* [8] and in several visual areas: area MT [2], V4 [5] (when measured between cells in the same attentional state), V1 [19, 36], and notably, in ON–OFF directionally sensitive retinal ganglion cells [11, 44]. The retinal studies involved cells with a clearly identified function, and therefore allowed investigation of the coding consequences of the observed correlation–firing rate relationship. These studies found that the *stimulus-dependent* correlation structure observed compared favorably to a structure in which *stimulus-independent* correlations were matched to their (stimulus-)averaged levels. This finding reflects a general principle articulated in other studies [18, 25], that stimulus-dependent correlations are beneficial when they serve to spread the neural response in a direction *orthogonal* to the signal space.

These findings thus provide strong motivation for understanding what network mechanisms can produce this positive (and perhaps beneficial) correlation–firing rate relationship. This correlation–firing rate trend has been explained theoretically in feedforward networks driven by common input [8, 26, 38]; however, many cortical networks are known to be dominated by strong recurrent activity [24, 34, 39]. On the other hand, theoretical studies of the mechanisms for correlations in recurrent networks have largely analyzed homogeneous networks (i.e., identical cells, aside from excitatory/inhibitory cell identity) [9, 13, 27, 28, 40, 41], and have not considered how correlations vary with firing rates with realistic intrinsic heterogeneity. Thus, how spike count correlations can vary across a population of heterogeneously-tuned, recurrently connected neurons is not yet well understood despite its possible implications for coding.

In a previous paper, we addressed this gap by developing a mathematical method to compactly describe how second-order spike count statistics vary with both intrinsic and network attributes [4]. Specifically, we adapted network linear response theory [14, 27, 43] to account for heterogeneous and recurrent networks, which in turn allows us to identify important network connections that contribute to correlations via a single-cell susceptibility function [8]. Here, we will use this method to survey a broad family of recurrent networks to understand how three factors influence the relationship between correlations and firing rates; how the neurons occupy parameter space, the strength of recurrent excitation, and the strength of background noise. This work thus provides a more complete picture of how even modest changes in important circuit parameters can alter the correlation–firing rate relationship.

2 Results

First, we summarize how a network linear response theory allows us to use the single-neuron firing rate function to approximate correlations. We then apply this theory to examine three factors that can modulate the correlation–firing rate relationship: direction in effective parameter space, strength of recurrent excitation, and strength of background noise.

2.1 Using the Single-Neuron Firing Rate Function to Characterize Correlation Susceptibility

The response of a neuron is generally a nonlinear function of model parameters. However, background noise linearizes this response, so that we can use a linear theory to describe the change in response to small changes in parameters. We assume we have some way to approximate the change in firing rate which occurs as a result of a change in parameter X :

$$v_i(t) = v_{i,0} + (A_{X,i} * X_i)(t); \tag{1}$$

$v_{i,0}$ is the baseline rate (when $X = 0$) and $A_{X,i}(t)$ is a *susceptibility function* that characterizes the firing rate response [8, 20, 41]. The parameter which is modulated has often been chosen to be a current bias μ [8, 41]; however, it could also be the mean or variance of a conductance, a time constant, or a spike generation parameter [29, 30].

In a coupled network, the parameter change X_i arises from inter-neuron coupling; substituting $X_i(t) \rightarrow \sum_j (\mathbf{J}_{X,ij} * v_j)$ and moving to the spectral domain, we find

$$\langle \tilde{y}(\omega) \tilde{y}^*(\omega) \rangle = (\mathbf{I} - \tilde{\mathbf{K}}(\omega))^{-1} \langle \tilde{y}^0(\omega) \tilde{y}^{0*}(\omega) \rangle (\mathbf{I} - \tilde{\mathbf{K}}^*(\omega))^{-1}, \tag{2}$$

where $\tilde{y}_i = \mathcal{F}[y_i - v_i]$ is the Fourier transform of the mean-shifted process (v_i is the average firing rate of cell i) and $\tilde{f} = \mathcal{F}[f]$ for all other quantities; $\tilde{\mathbf{K}}_{ij}(\omega) = \tilde{A}_{X,i}(\omega) \tilde{\mathbf{J}}_{X,ij}(\omega)$ is the interaction matrix in the frequency domain (which may depend on the parameter being varied, i.e. X); $\langle \tilde{y}^0(\omega) \tilde{y}^{0*}(\omega) \rangle$ is the power spectrum of the uncoupled neural response. Using the usual series expansion for $(\mathbf{I} - \tilde{\mathbf{K}}(\omega))^{-1}$, we see that the covariance $\tilde{\mathbf{C}}(\omega) \equiv \langle \tilde{y}(\omega) \tilde{y}^*(\omega) \rangle$ naturally decomposes into contributions from different graph motifs:

$$\begin{aligned} \tilde{\mathbf{C}}(\omega) &= (\mathbf{I} - \tilde{\mathbf{K}}(\omega))^{-1} \tilde{\mathbf{C}}^0(\omega) (\mathbf{I} - \tilde{\mathbf{K}}^*(\omega))^{-1} \\ &= \tilde{\mathbf{C}}^0(\omega) + \tilde{\mathbf{K}}(\omega) \tilde{\mathbf{C}}^0(\omega) + \tilde{\mathbf{C}}^0(\omega) \tilde{\mathbf{K}}^*(\omega) + \tilde{\mathbf{K}}(\omega) \tilde{\mathbf{C}}^0(\omega) \tilde{\mathbf{K}}^*(\omega) + \dots \end{aligned} \tag{3}$$

Each instance of $\tilde{\mathbf{K}}$ includes the susceptibility function in the spectral domain, $A_X(\omega)$, which modulates the effect of each directed connection by the responsiveness of the target neuron to the parameter of interest. As noted by previous authors [27], the validity of the expansion in Eq. (3) relies on the spectral radius of $\tilde{\mathbf{K}}$, $\rho(\tilde{\mathbf{K}})$, remaining less than one. We checked that this remains true for all networks we examined in this paper, with a maximum of $\rho(\tilde{\mathbf{K}}) = 0.9564$.

We next justify why—at least for long-time correlations—we can estimate each of these terms using the single-neuron firing rate function. First, in the small frequency limit, $A_{X,i}(\omega)$ will coincide with the derivative of the firing rate with respect to the parameter:

$$\lim_{\omega \rightarrow 0} \tilde{A}_{X,i}(\omega) = \frac{dv_i}{dX}.$$

For the common input motif, $\tilde{\mathbf{K}}(\omega)\tilde{\mathbf{C}}^0(\omega)\tilde{\mathbf{K}}^*(\omega)$, we can write

$$(\tilde{\mathbf{K}}\tilde{\mathbf{C}}^0\tilde{\mathbf{K}}^*)_{ij} = \sum_{k \rightarrow i, k \rightarrow j} \tilde{\mathbf{K}}_{ik}\tilde{\mathbf{C}}_{kk}^0\tilde{\mathbf{K}}_{jk} \tag{4}$$

$$= \sum_{k \rightarrow (i,j), k \in E} (\tilde{A}_{g_E,i}\tilde{\mathbf{J}}_{ik})\tilde{\mathbf{C}}_{kk}^0(\tilde{A}_{g_E,j}\tilde{\mathbf{J}}_{jk}) + \sum_{k \rightarrow (i,j), k \in I} (\tilde{A}_{g_I,i}\tilde{\mathbf{J}}_{ik})\tilde{\mathbf{C}}_{kk}^0(\tilde{A}_{g_I,j}\tilde{\mathbf{J}}_{jk}) \tag{5}$$

$$= |\tilde{\mathbf{J}}_E|^2 \sum_{k \rightarrow (i,j), k \in E} \tilde{A}_{g_E,i}\tilde{A}_{g_E,j}\tilde{\mathbf{C}}_{kk}^0 + |\tilde{\mathbf{J}}_I|^2 \sum_{k \rightarrow (i,j), k \in I} \tilde{A}_{g_I,i}\tilde{A}_{g_I,j}\tilde{\mathbf{C}}_{kk}^0 \tag{6}$$

$$= |\tilde{\mathbf{J}}_E|^2 \tilde{A}_{g_E,i}\tilde{A}_{g_E,j} \sum_{k \rightarrow (i,j), k \in E} \tilde{\mathbf{C}}_{kk}^0 + |\tilde{\mathbf{J}}_I|^2 \tilde{A}_{g_I,i}\tilde{A}_{g_I,j} \sum_{k \rightarrow (i,j), k \in I} \tilde{\mathbf{C}}_{kk}^0, \tag{7}$$

where we have separated excitatory (E) and inhibitory (I) contributions, using g_E and g_I to denote the mean conductance of each type, and assumed that the synaptic coupling kernels, $\tilde{\mathbf{J}}_{jk}$, depend only on E/I cell identity (and have thus dropped the first subscript, which adds no additional information).

This provides a formula for the long-time *covariance*, but we are typically interested in the *correlation*; fortunately, the small frequency limit also allows us to obtain a normalized correlation measure from the cross-spectrum, because

$$\lim_{T \rightarrow \infty} \rho_{T,ij} = \lim_{T \rightarrow \infty} \frac{\text{Cov}_T(n_i, n_j)}{\sqrt{\text{Var}_T(n_i)\text{Var}_T(n_j)}} = \frac{\tilde{\mathbf{C}}_{ij}(0)}{\sqrt{\tilde{\mathbf{C}}_{ii}(0)\tilde{\mathbf{C}}_{jj}(0)}}, \tag{8}$$

where $\text{Cov}_T(n_i, n_j)$ and $\text{Var}_T(n_i)$ denote covariance and variance of spike counts in a time window T ; that is, $\rho_{T,ij}$ is the Pearson correlation coefficient (which varies between -1 and 1).

Finally, letting $\omega \rightarrow 0$ and normalizing with the assumption that spiking is close to a Poisson process, as expected for an asynchronously firing network, so that

$\tilde{C}_{kk} \approx v_k$:

$$\frac{(\tilde{\mathbf{K}}\tilde{\mathbf{C}}^0\tilde{\mathbf{K}}^*)_{ij}}{\sqrt{\tilde{C}_{ii}\tilde{C}_{jj}}} = \underbrace{\frac{1}{\sqrt{v_i v_j}} \frac{dv_i}{dg_E} \frac{dv_j}{dg_E}}_{\text{modulation by firing rate function}} |\tilde{\mathbf{J}}_E|^2 \underbrace{\left[\sum_{\substack{k \rightarrow (i,j), \\ k \in E}} \tilde{C}_{kk}^0 \right]}_{\text{total E common input}} + \underbrace{\frac{1}{\sqrt{v_i v_j}} \frac{dv_i}{dg_I} \frac{dv_j}{dg_I}}_{\text{modulation by firing rate function}} |\tilde{\mathbf{J}}_I|^2 \underbrace{\left[\sum_{\substack{k \rightarrow (i,j), \\ k \in I}} \tilde{C}_{kk}^0 \right]}_{\text{total I common input}}. \tag{9}$$

The above expression summarizes how the impact of excitatory and inhibitory common input are modulated by the single-neuron firing rate function, v , and its derivatives.

Thus far, we have presented results previously obtained by others [27, 40, 41]. We now depart from these authors by focusing specifically on the behavior of the term in Eq. (9); and for simplicity, the behavior of this modulating factor for two identical neurons; e.g.

$$\frac{1}{\sqrt{v_i v_j}} \frac{dv_i}{dg_I} \frac{dv_j}{dg_I} \rightarrow \frac{1}{v} \left(\frac{dv}{dg_I} \right)^2. \tag{10}$$

In principle, analogous expressions govern larger motifs, such as chains, that involve additional evaluations of v and its derivatives; for example, excitatory length-3 chains arising from $\tilde{\mathbf{K}}^3\tilde{\mathbf{C}}^0$ would be:

$$\frac{(\tilde{\mathbf{K}}^3\tilde{\mathbf{C}}^0)_{ij}}{\sqrt{\tilde{C}_{ii}\tilde{C}_{jj}}} = \underbrace{\left[\frac{1}{\sqrt{v_i v_j}} \frac{dv_i}{dg_E} \frac{dv_l}{dg_E} \frac{dv_k}{dg_E} \right]}_{\text{modulation by firing rate function}} \times |\tilde{\mathbf{J}}_E|^3 \times \underbrace{\left[\sum_{l \rightarrow i} \sum_{k \rightarrow l} \sum_{j \rightarrow k} \tilde{C}_{jj}^0 \right]}_{\text{all } E \rightarrow E \rightarrow E \rightarrow E \text{ paths}}. \tag{11}$$

However, we found that, for a wide range of networks, direct common input—and *inhibitory* common input in particular—was the dominant contributor to pairwise correlations (Fig. 6(A)).

We now examine how different network mechanisms modulate the correlation–firing rate relationship, focusing on three factors: direction in effective parameter space, strength of recurrent excitation, and strength of background noise.

2.2 Direction in Parameter Space Modulates the Correlation–Firing Rate Relationship

Suppose we have a firing rate function of one or more intrinsic parameters (for exposition purposes, assume a function of two parameters (x_1, x_2)), i.e.

$$v = F(x_1, x_2).$$

The parameters x_j might include input bias, membrane time constant, ionic/synaptic reversal potentials, or a spiking threshold. Based on our arguments from the previous section, we will approximate *correlation susceptibility* by the quantity

$$\hat{S} = \frac{1}{F} \left(\frac{\partial F}{\partial x_1} \right)^2,$$

where x_1 is an appropriately chosen parameter. Specifically, we will find, empirically, that the inhibitory common input is the dominant term, and therefore will use $x_1 = g_I$, the mean inhibitory conductance. Thus, the units of \hat{S} in all figures are $\text{Hz}/[g_I^2]$, where g_I is dimensionless (see Eq. (17)).

Heterogeneous firing rates can arise when each neuron occupies a different location in parameter space (i.e. a different (x_1, x_2) point), thus causing both firing rate F and susceptibility \hat{S} to vary from neuron to neuron. We now ask: how does \hat{S} change with firing rate?

Note that this question, as stated, is ill-posed because F and \hat{S} are both functions of two parameters (x_1 and x_2): there is not necessarily a one-to-one or even a functional relationship between these quantities. Suppose that, locally, a population of neurons can be described as lying along a parameterized path in the (x_1, x_2) plane: i.e., $(x_1(s), x_2(s))$, for $s_{\min} < s < s_{\max}$. Then we can compute the directional derivative:

$$\frac{d\hat{S}}{dF} = \frac{d\hat{S}/ds}{dF/ds} = \frac{\nabla \hat{S} \cdot d\mathbf{x}}{\nabla F \cdot d\mathbf{x}}, \tag{12}$$

where we have expressed the directional derivatives in terms of the local direction of the path: i.e. $d\mathbf{x} = (\frac{dx_1}{ds}, \frac{dx_2}{ds})$ and the gradients of F and \hat{S} . However, this depends not only on the functions F and \hat{S} , but also on the direction $d\mathbf{x}$.

To gain some intuition for why (and when) direction in (x_1, x_2) space matters, we consider the networks studied in [4]. Previously, we simplified the firing rate function by setting all but two parameters (inhibitory conductance, $\langle g_{I,i} \rangle$, and threshold, θ_i) to their population average; i.e.

$$F(\langle g_{I,i} \rangle, \theta_i) \equiv f(\langle g_{I,i} \rangle, \langle \sigma_{g_{I,i}} \rangle_p, \langle g_{E,i} \rangle_p, \langle \sigma_{g_{E,i}} \rangle_p, \langle \sigma_i \rangle_p, \theta_i), \tag{13}$$

and $\langle \cdot \rangle_p$ denotes the population average. In Figs. 1(A) and (B), we show F and $\hat{S} \equiv (\frac{\partial F}{\partial x_1})^2/F$ thus computed, for the asynchronous network studied in that paper. We then surveyed a sequence of diagonal paths through the center (i.e., midpoint of the ranges of threshold and inhibitory conductance), with each path plotted in a different color. In Fig. 1(C) we plot firing rate (solid lines) and susceptibility (dashed-dotted lines) along each curve. Finally, in Fig. 1(D) we plot the susceptibility versus the firing rate, along each path.

This last panel shows that there is striking variability in the susceptibility-firing rate relationship, depending on the direction the path takes through the center of the $(\theta, \langle g_I \rangle)$ plane. Any given firing rate (below ~ 15 Hz) is consistent with either increase or decrease of susceptibility. All curves go through a single point in the $(\theta, \langle g_I \rangle)$ plane, and therefore a single point in the (F, \hat{S}) plane; here the direction

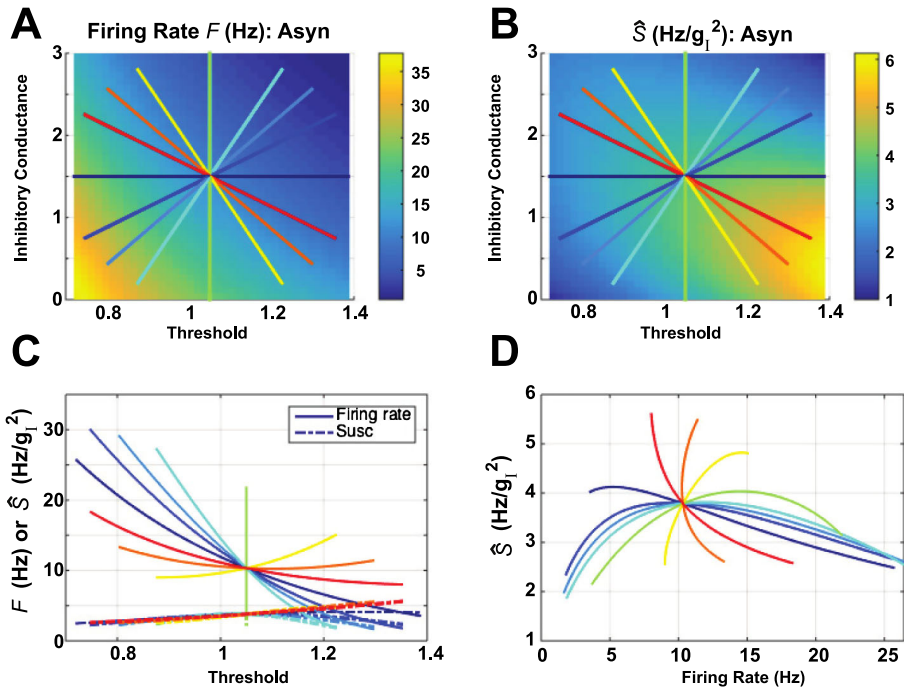


Fig. 1 Firing rate and susceptibility (\hat{S}), computed for the asynchronous (Asyn) network studied in [4], with all other parameters besides threshold θ and mean inhibitory conductance (g_I) set to their average values (thereby leaving a two-dimensional parameter space). Here, we traverse the plane on straight-line paths defined by their angle through the origin. Although the units of g_I are dimensionless, they are shown as the units for \hat{S} for completeness. The units of θ (i.e., voltage) are also scaled to be dimensionless

of the \hat{S} – F relationship (i.e., whether \hat{S} increases or decreases with F) can change rapidly with angle, as for the red, orange, and yellow curves.

We then extended these observations by traversing the phase space with two additional families of straight-line paths (Fig. 2); the radial paths are repeated in Figs. 2(A) and (B). For horizontal (Figs. 2(C) and (D)) and vertical (Figs. 2(E) and (F)) families, paths no longer intersect at a single point; nevertheless, a given firing rate is consistent with both increases and decreases in susceptibility. This is pronounced in Fig. 2(F), where at 15 Hz susceptibility decreases with firing rate in the orange, yellow and light green paths, but increases for the light blue, medium blue, royal blue, and indigo paths.

We performed the same computations on the *strong asynchronous* network studied in [4] that has larger excitatory coupling strength: results are shown in Fig. 3. A given firing rate could be consistent with either increase or decrease of susceptibility; we see this in the radial paths (Figs. 3(A) and (B)) and horizontal paths (Figs. 3(C) and (D)) for rates between 5–10 Hz. However, vertical paths (Figs. 3(E) and (F)) always have susceptibility increasing with firing rate (except perhaps at the highest firing rates). As in the asynchronous network, direction of susceptibility (increase vs. decrease) can change rapidly with angle, for paths that intersect a single point; see Figs. 3(A)–(B), red to orange to yellow.

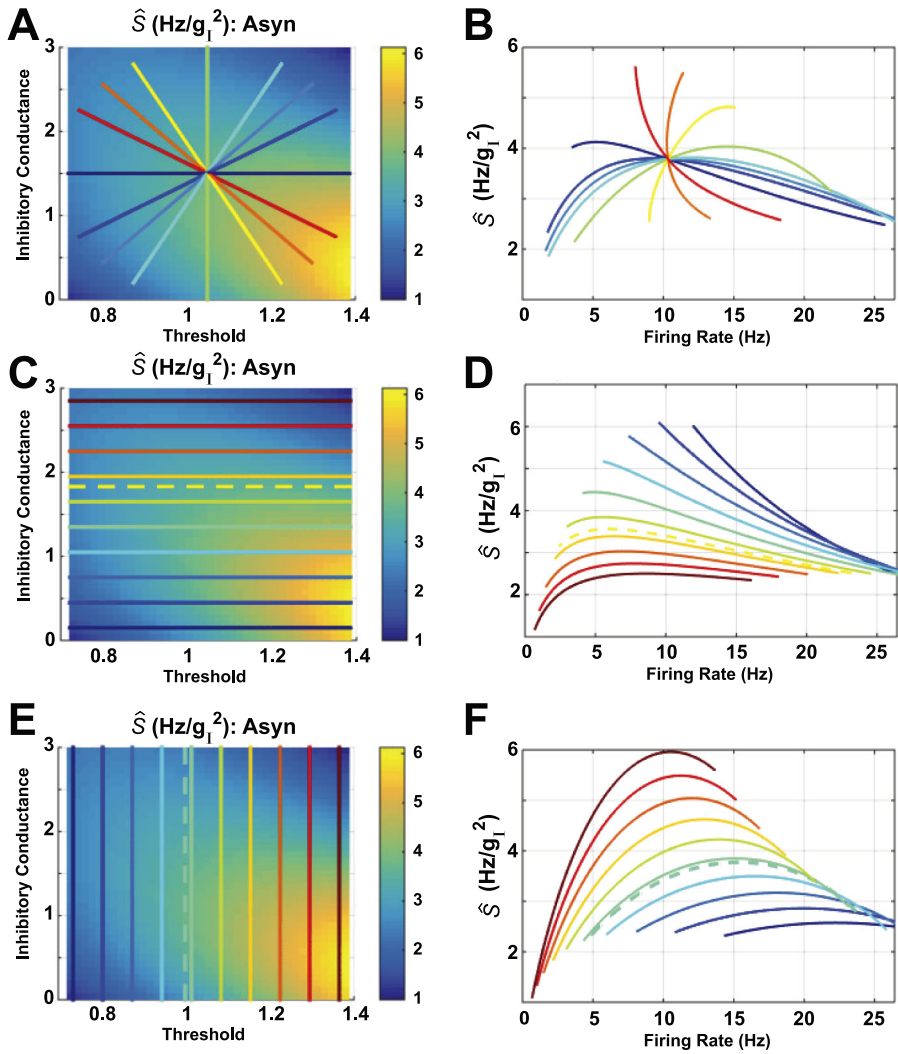


Fig. 2 Susceptibility (\hat{S}) vs. firing rate, computed for the asynchronous network studied in [4], with all other parameters besides threshold θ and mean inhibitory conductance $\langle g_I \rangle$ set to their average values (thereby leaving a two-dimensional parameter space: the other (averaged) parameters are $\langle \sigma_{g_{I,i}} \rangle_p = 0.6602$ (see Eq. (30)), $\langle \sigma_{g_{E,i}} \rangle_p = 0.0026$ (see Eq. (29)), $\langle \sigma_i \rangle_p = 6.3246$, $\langle \langle g_{E,i} \rangle \rangle_p = 0.0053$ (see Eq. (17)). Here, we traverse the plane on three different families of straight-line paths. The dashed lines show paths visualized in [4]: $\theta = 1$ (vertical, aqua blue) and $\langle g_I \rangle = 1.83$ (horizontal, orange/yellow)

2.3 Quantifying the Likelihood of a Positive Correlation–Firing Rate Relationship

In the previous section, we saw that the path a network occupies in effective parameter space can have a dramatic effect on the correlation–firing rate relationship: we next seek to formalize these observations. Let $d\mathbf{x}$ be the local direction in which we want to

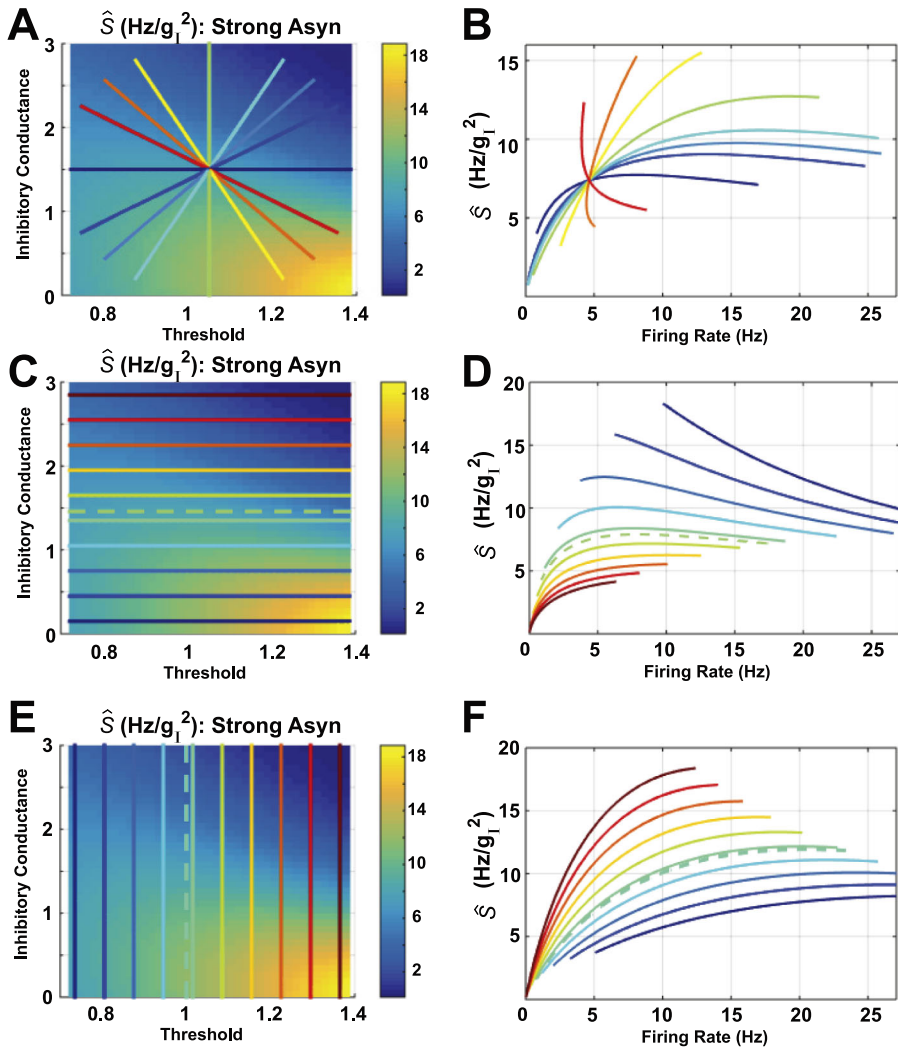


Fig. 3 Susceptibility (\hat{S}) vs. firing rate, computed for the strong asynchronous (Strong Asyn) network studied in [4], with all other parameters besides threshold θ and mean inhibitory conductance ($\langle g_I \rangle$) set to their average values (thereby leaving a two-dimensional parameter space: the other (averaged) parameters are $\langle \sigma_{g_I, i} \rangle_p = 0.5884$ (see Eq. (30)), $\langle \sigma_{g_E, i} \rangle_p = 0.0378$ (see Eq. (29)), $\langle \sigma_i \rangle_p = 4.7434$, $\langle \langle g_E, i \rangle \rangle_p = 0.0611$, see Eq. (17)). Here, we traverse the plane on three different families of straight-line paths. Dashed lines show paths visualized in [4]: $\theta = 1$ (vertical, aqua blue) and $\langle g_I \rangle = 1.46$ (horizontal, yellow/green)

consider the behavior of F and \hat{S} . If $\nabla \hat{S} \cdot d\mathbf{x}$ and $\nabla F \cdot d\mathbf{x}$ are of the same sign, then \hat{S} increases with F ; if $\nabla \hat{S} \cdot d\mathbf{x}$ and $\nabla F \cdot d\mathbf{x}$ have opposite signs, then \hat{S} decreases with F . The more aligned $\nabla \hat{S}$ and ∇F , the more paths lead to $\frac{d\hat{S}}{dF} > 0$; the more anti-aligned $\nabla \hat{S}$ and ∇F , the more paths lead to $\frac{d\hat{S}}{dF} < 0$. For example, consider the limiting cases where: (i) if $\nabla \hat{S}$ and ∇F point exactly in the same direction, then $\nabla \hat{S} \cdot d\mathbf{x}$ and

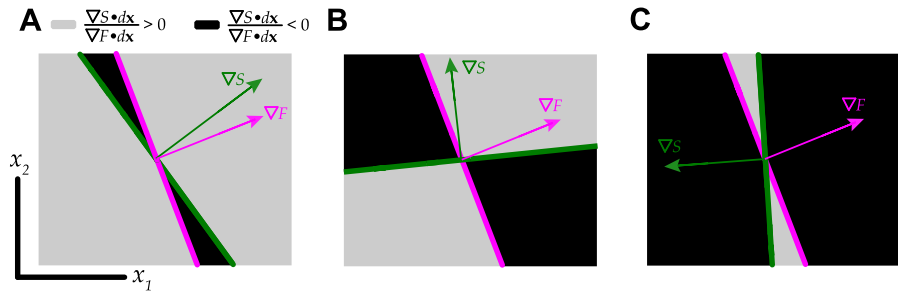


Fig. 4 Where $\nabla \hat{S}$ and ∇F are similarly aligned, \hat{S} and F will both increase along most paths through that point. In each panel, gray shows the part of the \mathbf{x} -plane where $\frac{d\hat{S}}{dF} = \frac{\nabla \hat{S} \cdot d\mathbf{x}}{\nabla F \cdot d\mathbf{x}} > 0$, black where $\frac{d\hat{S}}{dF} < 0$. From left to right: $\nabla \hat{S}$ and ∇F positively aligned; $\nabla \hat{S}$ and ∇F orthogonal; $\nabla \hat{S}$ and ∇F negatively aligned

$\nabla F \cdot d\mathbf{x}$ are always same-signed for any $d\mathbf{x}$; (ii) if $\nabla \hat{S}$ and ∇F point in opposite directions, then $\nabla \hat{S} \cdot d\mathbf{x}$ and $\nabla F \cdot d\mathbf{x}$ are never same-signed. Figure 4 illustrates how the alignment of these two quantities alters the region where correlation increases with firing rate.

To examine the utility of this idea, we reconsider the radial paths along which we previously computed susceptibility. The paths all go through a single point, so we can check the sign of $(\nabla \hat{S} \cdot \mathbf{x})(\nabla F \cdot \mathbf{x})$ for all directions through this point. In Figs. 5(A) and (C), white indicates positive and black negative. Figures 5(B) and (D) repeat the susceptibility-firing rate curves from Fig. 2(B) and Fig. 3(B). For the asynchronous network (Fig. 5(A)), the red, indigo, and royal blue paths are predicted to have negative $d\hat{S}/dF$, as we can confirm in Fig. 5(B). Yellow, green, and light blue curves are predicted to have positive $d\hat{S}/dF$. The orange curve is close to $dF = 0$; the true blue curve is close to $dS = 0$. For the strong asynchronous network, only the red curve is in the negative region of Fig. 5(C); this is also the only path with $d\hat{S}/dF < 0$ in Fig. 5(D).

Of course, this prediction only applies to portions of the path near the point at which we computed the gradients; away from this point, gradients will change along with the direction of the \hat{S} vs. F curve. For example, the royal blue curve in Fig. 5(B) increases with firing rate for small firing rates, and the light blue, true blue, and royal blue curves in Fig. 5(D) decrease with firing rate, (for large firing rates).

This motivates a direction-independent measure to quantify the fraction of paths that lead to an increase of correlation with firing rate. This is given exactly in terms of the angle between $\nabla \hat{S}$ and ∇F :

$$\cos \theta = \frac{\nabla \hat{S} \cdot \nabla F}{\|\nabla \hat{S}\| \|\nabla F\|} \tag{14}$$

and in particular the fraction of \mathbf{x} directions in which \hat{S} increases with F is given by

$$\frac{1}{\pi} \left(\pi - \cos^{-1} \left(\frac{\nabla \hat{S} \cdot \nabla F}{\|\nabla \hat{S}\| \|\nabla F\|} \right) \right). \tag{15}$$

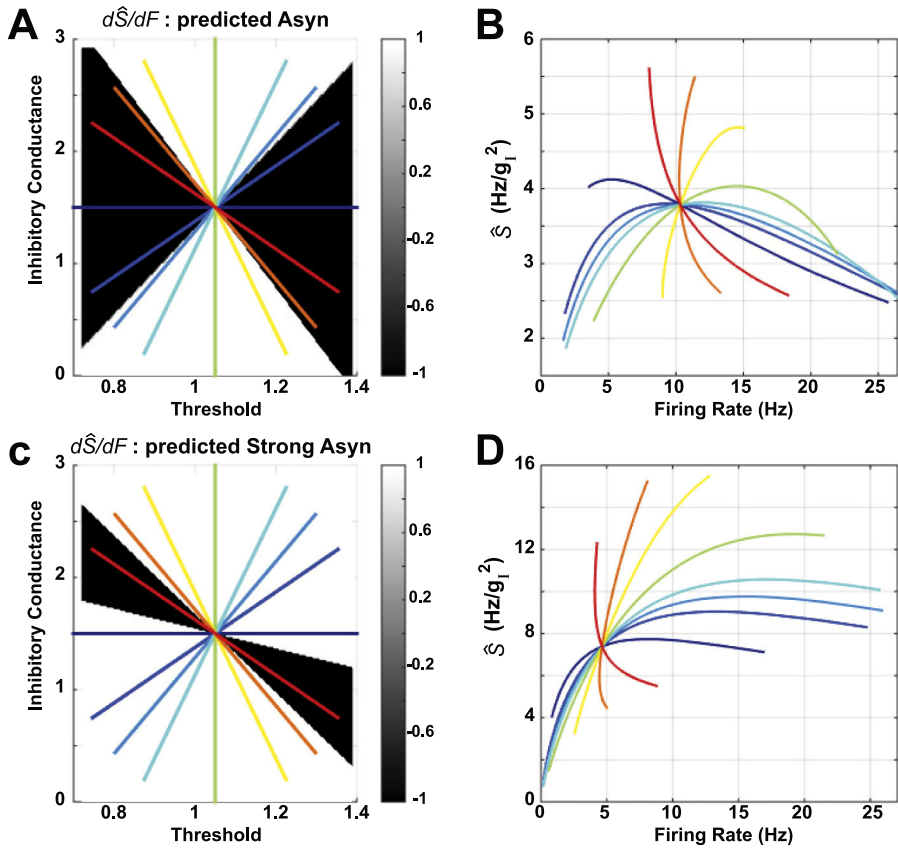


Fig. 5 Using a single number to predict $d\hat{S}/dF$. (A) Paths through parameter space for the asynchronous network: white shows the part of the x-plane where $\frac{d\hat{S}}{dF} = \frac{\nabla\hat{S} \cdot d\mathbf{x}}{\nabla F \cdot d\mathbf{x}} > 0$, black where $\frac{d\hat{S}}{dF} < 0$, where $\nabla\hat{S}$ and ∇F are computed at the center of the diagram. (B) Correlation susceptibility vs. firing rate, for each path illustrated in (A). (C)–(D) As in (A)–(B), but for the strong asynchronous network

Because \cos^{-1} has range $[0, \pi]$, this varies between 0 and 1. The more aligned $\nabla\hat{S}$ and ∇F , the more paths lead to $\frac{d\hat{S}}{dF} > 0$; the more anti-aligned $\nabla\hat{S}$ and ∇F , the more paths lead to $\frac{d\hat{S}}{dF} < 0$.

2.4 Strength of Recurrent Excitation Modulates the Correlation–Firing Rate Relationship

Our use of inhibitory susceptibility (i.e. Eq. (10)) to characterize the relationship between correlations and firing rates relied on intermediate assumptions, specifically:

- Second-order motifs dominate pairwise correlations.
- Inhibitory common input is the dominant second-order motif.
- Asynchronous spiking assumption: $\text{Var}_T(n_i) = T v_i \Rightarrow \tilde{C}_{ii} = v_i$.

Here, we check that these conditions are still satisfied for a range of neural network models.

In [4], we examined two spiking regimes achieved by varying the strength of excitation: both recurrent excitation W_{EE} and excitatory input into the inhibitory population W_{IE} . We next applied our theory to a dense grid of parameters (different networks), each identified by its location on the (W_{EE}, W_{IE}) plane. Both excitatory strengths were varied from a minimum of their values for the asynchronous network ($W_{EE} = 0.5$ and $W_{IE} = 5$) to a maximum of 1.4 times their value in the strong asynchronous network (to $W_{EE} = 12.6$ and $W_{IE} = 11.2$). The parameter W_{XY} is a dimensionless scale factor (see Eqs. (17)–(20)); for example in an all-to-all homogeneous network, $W_{XY} = 1$ is when the presynaptic input is exactly the average population firing rate (filtered by the synapse).

To quantify the importance of paths of different length through the network, we can define the contributions at any specific order k by using the interaction network \mathbf{K} :

$$\tilde{\mathbf{R}}_{ij}^k = \frac{(\sum_{l=0}^k \tilde{\mathbf{K}}^l \tilde{\mathbf{C}}^0 (\tilde{\mathbf{K}}^*)^{k-l})_{ij}}{\sqrt{\tilde{\mathbf{C}}_{ii} \tilde{\mathbf{C}}_{jj}}}. \tag{16}$$

Then we regressed the total correlation $(\tilde{\mathbf{C}}_{ij} / \sqrt{\tilde{\mathbf{C}}_{ii} \tilde{\mathbf{C}}_{jj}})$ against the contributions at each specific order $(\tilde{\mathbf{R}}_{ij}^k)$; the corresponding fraction of variance explained (R^2 value) gives a quantitative measure of how well the total correlation can be predicted from each term.

Similarly, we quantify the importance of specific second-order motif types, by regressing the total contribution from second-order motifs $(\tilde{\mathbf{R}}_{ij}^2)$ against the contribution from specific motifs. We performed this computation for each network (a total of 225 networks), and summarize the results in Fig. 6; to present the data compactly, we collapse R^2 values (all values are $\in [0, 1]$) for a fixed W_{EE} and varying W_{IE} by showing mean and standard deviation only (standard deviation as error bars). Second-order contributions dominate for small to moderate W_{EE} (Fig. 6(A)); other orders only compete when W_{EE} has already exceeded the level of the strong asynchronous network (where the network is close to the edge of instability, and at the limit of validity for mean-field, asynchronous assumptions).

To illustrate the meaning of this statistic, in Fig. 6(B) we show contributions up to fourth order $(\tilde{\mathbf{R}}_{ij}^k, \text{ for } k = 1, \dots, 4)$ vs. total correlation $(\tilde{\mathbf{C}}_{ij} / \sqrt{\tilde{\mathbf{C}}_{ii} \tilde{\mathbf{C}}_{jj}})$ for all E–E cell pairs in a network, for two individual networks included in the summary panel. Note that the second-order contributions cluster near the unity line in both cases, indicating that second-order contributions are the best predictor of total correlations, consistent with the R^2 values stated.

Of the second-order motifs, inhibitory common input is the dominant contribution at any value of W_{EE} , except perhaps the last, at which point the excitatory population has unrealistically high firing rates (Fig. 6(C)). To summarize, we have confirmed that far from being limited to a few examples, the conditions we identified in [4] as

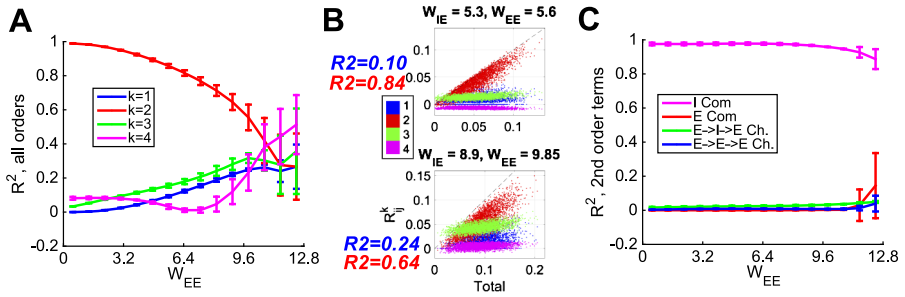


Fig. 6 Second-order motifs dominate pairwise correlations in a wide range of networks; inhibitory common input is the dominant second-order motif. (A) Fraction of variance explained (R^2) from linear regressions of total correlation ($\tilde{C}_{ij}/\sqrt{\tilde{C}_{ii}\tilde{C}_{jj}}$) against contributions from first order (blue), second order (red), third-order (green), and fourth-order (magenta) motifs. (B) Contributions up to fourth order (\tilde{R}_{ij}^k , for $k = 1, \dots, 4$) vs. total correlation ($\tilde{C}_{ij}/\sqrt{\tilde{C}_{ii}\tilde{C}_{jj}}$) for all E-E cell pairs in a network, for two individual networks included in panel A. (C) Fraction of variance explained (R^2) from linear regressions of contributions to pairwise correlations from second-order motifs (\tilde{R}_{ij}^2) against contributions from the distinct types of second-order motifs: inhibitory common input (magenta), excitatory common input (red), decorrelating chains (green), and correlating chains (blue). Both (A,C): Each data point represents the mean value from 15 networks with W_{IE} between 5 and 11.2; error bars show standard deviation across these values. $W_{XY} = 1$ is when the presynaptic input is exactly the average population firing rate (filtered by the synapse) in an all-to-all homogeneous network

allowing us to focus on susceptibility to inhibition to explain pairwise correlations, appear to hold up for a variety of networks.

We note that our findings about the relative magnitudes of terms of different orders are purely empirical; that is, they are based on concrete numerical observations, rather than a priori estimates. Thus, it should be reassessed if anything about the parameters or network connectivity is changed. In particular, a likely reason for the relative weakness of first-order terms is that in these networks excitation is almost always weaker than inhibition; while first-order terms are always excitatory, second-order terms can involve excitation and/or (comparatively strong) inhibition.

Having confirmed the validity of our approach, we computed the susceptibility with respect to inhibition, for each of the networks examined in the previous section (because of instability, we restricted these computations to excitatory strengths $\times 1.2$ the values used in [4]). Because background noise values varied slightly, we actually examined two network families; one in which we chose σ values as in the asynchronous network, another in which we chose σ values as in the strong asynchronous network. Also as in [4], we estimated susceptibility using network-averaged values of g_E, g_I, σ_{g_E} , and σ_{g_I} .

Surprisingly, the difference in background noise dwarfed the recurrent excitation strengths, at least in accessing the relationship between \hat{S} and firing rate. In Fig. 7, we show \hat{S} vs. F curves, for a set of representative networks, on a single plot. Color indicates W_{EE} (shade of blue) and W_{IE} (shade of red); values of W_{EE} are 0.50 (as in the asynchronous network from [4]), 6.45, 9 (as in the strong asynchronous network from [4]), and 10.7, values of W_{IE} are 5 (as in the asynchronous network from [4]), 7.1, 8 (as in the strong asynchronous network from [4]), and 8.6. For reference,

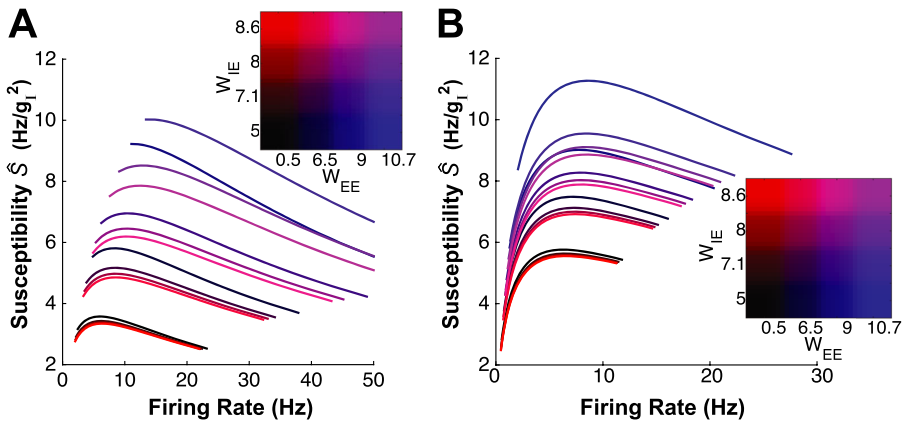


Fig. 7 Firing rate vs. susceptibility (\hat{S}), computed for a family of networks generated by modulating the strength of excitation (W_{EE} and W_{IE}). (A) Background noise values σ_E, σ_I set as in the asynchronous network from [4]. (B) Background noise values σ_E, σ_I set as in the strong asynchronous network from [4]. Sixteen curves are chosen, for a survey of the range of networks achievable by varying strength of recurrent excitation. Values of W_{EE} are 0.50 (as in the asynchronous network from [4]), 6.45, 9 (as in the strong asynchronous network from [4]), and 10.7. Values of W_{IE} are 5 (as in the asynchronous network from [4]), 7.1, 8 (as in the strong asynchronous network from [4]), and 8.6. Again, $W_{XY} = 1$ is when the presynaptic input is exactly the average population firing rate (filtered by the synapse) in an all-to-all homogeneous network, so the coupling strengths vary significantly

$W_{XY} = 1$ is when the presynaptic input is exactly the average population firing rate (filtered by the synapse) in an all-to-tall homogeneous network, so these coupling strengths vary significantly. We see that, for the full range of recurrent excitation values, \hat{S} vs. F curves in Fig. 7(A) are mostly decreasing; \hat{S} vs. F curves in Fig. 7(B) are mostly increasing for low F , and saturating for high F .

2.5 Background Noise Modulates the Correlation–Firing Rate Relationship

To further explore the role of background noise, we repeated the susceptibility calculation on additional families of networks, now allowing background noise strengths σ_E and σ_I (i.e. to the excitatory and inhibitory populations) to vary separately. Input to excitatory cells was varied between $\sigma_E = 1.5$ and 2.5; input to inhibitory cells was varied between $\sigma_I = 1.5$ and 3. These noise values are relatively large; see Eq. (17) and note that voltage is of order 1. In Fig. 8(A) we display susceptibility vs. firing rate curves for 12 (σ_E, σ_I) pairs; asterisks indicate σ_E and σ_I by color (green intensity for σ_E and blue intensity for σ_I). Within each panel curves are colored as in Fig. 7: red intensity for W_{IE} and blue intensity for W_{EE} .

Surprisingly, most network families (i.e. (σ_E, σ_I)) were associated with a *decrease* in correlation with firing rate. The exceptions are (0.15, 0.25) (as in the strong asynchronous network in [4]) and (0.15, 0.3). This latter was most robustly associated with a positive correlation–firing rate relationship. Furthermore, the shape of susceptibility–firing curves did not appear to vary much with the strength of recurrent excitation (i.e., curves within each panel are similar).

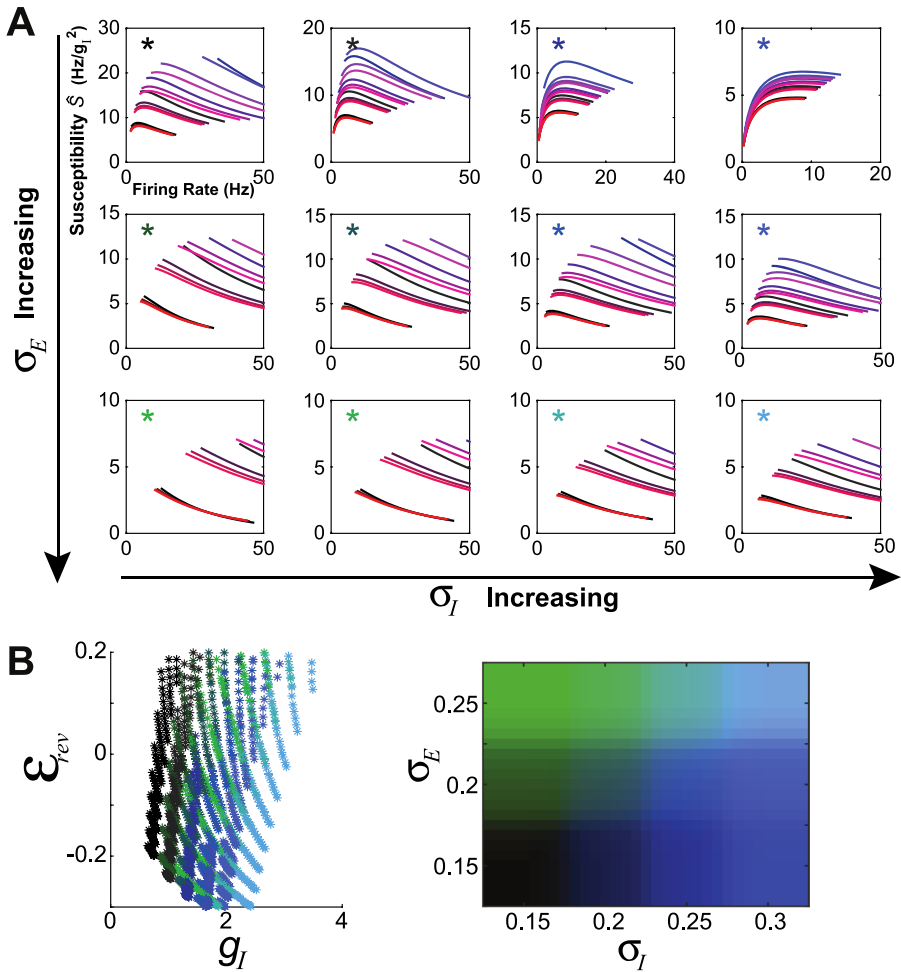


Fig. 8 The strength of background noise modulates the correlation–firing rate relationship. (A) Each panel shows firing rate vs. susceptibility (\hat{S}), computed for a family of networks generated by modulating the strength of excitation (W_{EE} and W_{IE}) with various background noise levels (see Eq. (17) for σ_E and σ_I definitions). (B) Population-averaged effective parameters (g_I) and \mathcal{E}_{rev} , for each network displayed in (A); see Eq. (28) for \mathcal{E}_{rev}

We next sought to investigate possible mechanisms for a positive correlation–firing rate relationship, by examining the effective parameters that govern the neural response: in essence, the network’s “operating point” (see Eq. (26)). Possible choices include g_I , g_E , σ_{g_E} , σ_{g_I} , and the effective reversal potential \mathcal{E}_{rev} ; we found σ_{g_E} and σ_{g_I} to be largely functions of g_E and g_I , while \mathcal{E}_{rev} has a (nonlinear) functional relationship with g_E and g_I . Thus two parameters suffice, and here we chose to use g_I and \mathcal{E}_{rev} . In Fig. 8(B), we plot average parameter values for each curve, color-coded by (σ_E, σ_I) . Any given color is consistent with a relatively tight range of g_I and (comparatively) broad range of \mathcal{E}_{rev} . As σ_I increases (increasing blue intensity), inhibitory conductance g_I increases and reversal potential \mathcal{E}_{rev} decreases. However,

it was not apparent that any particular region in $(g_I, \mathcal{E}_{\text{rev}})$ parameter space was associated with a positive correlation–firing rate relationship, in that the values of g_I and \mathcal{E}_{rev} that supported a positive relationship supported negative relationships as well.

3 Discussion

In this paper, we showed that using a single-cell firing rate function to examine the relationship between correlations and firing rates is feasible for a wide range of heterogeneous, recurrent networks. We focused on three factors that can modulate the correlation–firing rate relationship: how the network occupies effective parameter space, strength of recurrent excitation, and strength of background noise. Although there are many sets of parameters known to vary within a heterogeneous network of neurons, we have already demonstrated vastly different correlation–firing rate relationships with our methods, with a theory that can be readily applied to other model networks.

One possible application of this work is in designing neural networks for computational experimentation; just as modelers now modify cortical networks to obey experimental constraints on firing rates [3, 35], we could also include a constraint on the desired correlation–firing rate relationship. Here we showed that we can quickly assess a wide range of possible network configurations for a positive correlation–firing rate relationship: in Sect. 2.5, for example, we performed calculations on $15 \times 15 \times 12 = 2700$ heterogeneous networks, with a nominal amount of computing time.

One surprising finding in our computations was the relative insensitivity of the slope of the correlation–firing rate relationship to recurrent excitation (W_{EE} , W_{IE}), as demonstrated in Figs. 7 and 8. This is striking in contrast to the strong sensitivity on display in Figs. 2(B) and 3(B). This difference is explained as follows: in every case where we computed the susceptibility for a self-consistent network (i.e. a solution of Eqs. (26)–(30) and (32)–(33)), the source of heterogeneous firing rates was neural excitability, set via a spiking threshold θ . The resulting effective parameters, such as inhibitory conductance $\langle g_I \rangle$, did not deviate strongly from their population mean values. In essence, all of these networks took a horizontal path through $(\theta, \langle g_I \rangle)$ parameter space, as in Figs. 2(C), (D) and Figs. 3(C), (D). If we were to generate networks where heterogeneity arises from another source—such as the strength or frequency of inhibitory connections [23]—we might see different results. We look forward to exploring this in future work.

A priori, there is no reason to expect that the correlation–firing rate relationship in these recurrent networks can be simplified to a feedforward motif based on shared inhibitory input; this was purely an empirical observation (see Fig. 6(B)). We remark that others have shown that the effective input correlation can be canceled to have near zero input (and thus output) correlation on average in balanced networks [28, 40], in contrast to some of the models considered here (i.e., *strong asynchronous* regime with more net excitation). The conditions for correlation cancellation in these model networks is beyond the scope of this study, but note that others have shown

correlation cancellation does not always hold ([21, 22] via altering connection probabilities). The studies that demonstrate correlation cancellation often have faster (or equal) inhibitory synaptic time scales than excitatory: $\tau_I \leq \tau_E$ [21, 28, 32] ([40] used current-based instantaneous synapses or $\tau_I = \tau_E = 10$ ms) while in our networks the inhibitory synapses have longer time scales (Table 1). Note that Fig. S4 of [28] shows that having effectively zero input correlation does not hold as the inhibitory time scales increase beyond the excitatory time scales. Finally, system size is another key parameter that could certainly affect the magnitude of the recurrent feedback. In contrast to [28, 40], we did not account for how system size would affect correlation cancellation in these heterogeneous networks.

Although affirmative answers to whether correlations increase with firing rate in experiments were cited in the Introduction [2, 5, 8, 11, 19, 36, 44] we also note that many experiments have shown that the average correlation across a population can decrease with firing rate when a global state changes or a stimulus is presented. A recent review [9] shows that stimulus-induced decorrelation (with increased firing rate) occurs in a variety of brain regions and animals. This is slightly different from the situation we examine here, where we consider the relationship between correlations and firing rates within a stimulus condition. Regardless, the fact that the relationship between correlation and firing rate is not obvious points to the continued need for theoretical studies into the mechanisms of spike statistics modulation.

Finally, our finding that correlations often *decrease*, rather than increase, with firing rate stands in apparent contradiction to earlier work on feedforward networks [8, 38]. On closer inspection, we can identify several reasons why our results differ; with conductance inputs (rather than currents) we have a quantitatively different relationship between input parameters and firing rates; furthermore with more adjustable single-neuron parameters, the same sets of firing rates may be observed with single-neuron parameters set in different ways. While the current clamp experiments described in [8] found a consistent increase of correlations with firing rates, we can hypothesize that the parallel dynamic clamp experiments in which pairwise correlations arise from common inhibitory input, would in fact show a decrease with firing rate. However, we also predict that whether an increase or decrease with firing rate is observed would depend on whether firing rates are modulated by varying the level of inhibitory input, or by otherwise varying the excitability of the cells (perhaps pharmacologically).

4 Methods

4.1 Neuron Model and Network Setup

We considered randomly connected networks of excitatory and inhibitory neurons. Each cell is a linear integrate-and-fire model with second-order alpha-conductances, i.e. membrane voltage v_i was modeled with a stochastic differential equation, as long as the voltage is below threshold θ_i :

$$\tau_m \frac{dv_i}{dt} = -v_i - g_{E,i}(t)(v_i - \mathcal{E}_E) - g_{I,i}(t)(v_i - \mathcal{E}_I) + \sigma_i \sqrt{\tau_m} \xi_i(t). \quad (17)$$

Table 1 Intrinsic parameters that are fixed throughout

	Parameter									
	τ_m	\mathcal{E}_E	\mathcal{E}_I	τ_{ref}	α_E	α_I	$\tau_{r,E}$	$\tau_{d,E}$	$\tau_{r,I}$	$\tau_{d,I}$
Value	20 ms	6.5	-0.5	2 ms	1	2	1 ms	5 ms	2 ms	10 ms

Fixed parameters for model networks; see Eqs. (17)–(20). All are dimensionless except the time scales.

When v_i reaches θ_i , a spike is recorded and voltage is reset to 0 following a refractory period:

$$v_i(t) \geq \theta_i \implies v_i(t + \tau_{\text{ref}}) = 0, \tag{18}$$

Each neuron receives Gaussian white background noise with magnitude σ_i depending only on the cell type; that is, $\langle \xi_i(t) \rangle = 0$ and $\langle \xi_i(t)\xi_i(t+s) \rangle = \delta(s)$. The membrane time constant, τ_m , and excitatory and inhibitory synaptic reversal potentials, \mathcal{E}_E and \mathcal{E}_I , are the same for every cell in the network (see Table 1). The thresholds θ_i are a significant source of heterogeneity, and they are selected from a log-normal distribution with mean 1 and variance $e^{(0.2)^2} - 1$; since the system size is moderate, the θ_i 's were set to have C.D.F. (cumulative distribution function) values equally spaced from 0.05 to 0.95 for both E and I cells.

Each cell responds to synaptic input through conductance terms, $g_{E,i}$ and $g_{I,i}$, which are each governed by a pair of differential equations:

$$\tau_{d,X} \frac{dg_{X,i}}{dt} = -g_{X,i} + g_{X,i}^{(1)}, \tag{19}$$

$$\tau_{r,X} \frac{dg_{X,i}^{(1)}}{dt} = -g_{X,i}^{(1)} + \tau_{r,X} \alpha_X \left(\frac{W_{YX}}{N_{YX}} \right) \sum_{j \in X, j \rightarrow i} \sum_k \delta(t - t_{j,k}), \tag{20}$$

where $Y = \{E, I\}$ denotes the type of cell i and $X = \{E, I\}$ denotes the type of the source neuron j . Each spike is modeled as a delta-function that impacts the auxiliary variable $g_{X,i}^{(1)}$; here $t_{j,k}$ is the k th spike of cell j . The rise and decay time constants $\tau_{r,X}$ and $\tau_{d,X}$ and pulse amplitude α_X depend only on the type of the source neuron, that is they are otherwise the same across the population. The parameter W_{YX} denotes the strength of $X \rightarrow Y$ synaptic connections, which are (once given the type of source and target neurons) identical across the population. The ‘‘raw’’ synaptic weight (listed in Table 2) is divided by N_{YX} , the total number of $X \rightarrow Y$ connections received by each Y -type cell.

Table 2 show connectivity parameters for the two example networks we discuss in Sect. 2.2. For Figs. 1–3, five parameters are set as stated in this table. In Sect. 2.4 and Figs. 6–7, W_{EE} was varied between 0.5 and 12.6 and W_{IE} between 5 and 11.2. In Sect. 2.5 and Fig. 8, W_{EE} was varied between 0.5 and 10.8 and W_{IE} between 5 and 9.6; σ_E was varied between 1.5 and 2.5 and σ_I between 1.5 and 3.

Table 2 Excitatory connection strengths mediate different firing regimes

	Parameter						Figures
	$W_{EI} (I \rightarrow E)$	$W_{IE} (E \rightarrow I)$	W_{EE}	W_{II}	σ_E	σ_I	
Asynchronous	10	5	0.5	5	$2/\sqrt{2}$	$3/\sqrt{2}$	Figs. 1, 2, 5(A) and (B)
Str. Asynch.	10	8	9	5	$1.5/\sqrt{2}$	$2.5/\sqrt{2}$	Figs. 3, 5(C) and (D)
% connectivity	35 %	20 %	40 %	40 %			

Connectivity details for networks examined in Sect. 2.2. Here W_{YX} denotes $X \rightarrow Y$ connections; e.g. W_{IE} denotes the strength of excitatory connections onto inhibitory neurons. The parameter σ_i denotes the strength of background noise in units of (scaled) voltage, and depends only on cell type (E or I). The parameters W_{EE} , W_{IE} , σ_E and σ_I will vary in Sects. 2.4 and 2.5; see text for details. For reference, $W_{XY} = 1$ is when the presynaptic input is exactly the average population firing rate (filtered by the synapse) in an all-to-tall homogeneous network.

4.2 Linear Response Theory

In general, computing the response of even a single neuron to an input requires solving a complicated, nonlinear stochastic process. However, it often happens that the presence of background noise linearizes the response of the neuron, so that we can describe this response as a perturbation from a background state. This response is furthermore linear in the perturbing input and thus referred to as *linear response* theory [31]. The approach can be generalized to yield the dominant terms in the coupled network response as well. We will use the theory to predict the covariance matrix of spiking activity. The derivation is presented in full in [20, 29, 30]; here, we present only the main points.

We assume we have some way to approximate the change in firing rate which occurs as a result of a change in parameter:

$$v_i(t) = v_{i,0} + (A_{X,i} * \epsilon X_i)(t); \tag{21}$$

$v_{i,0}$ is the baseline rate (when $X = 0$) and $A_{X,i}(t)$ is a *susceptibility function* that characterizes this firing rate response up to order ϵ [8, 20, 41].

In order to consider joint statistics, we need the trial-by-trial response of the cell. First, we propose to approximate the response of each neuron by

$$y_i(t) \approx y_i^0(t) + \left(A_{X,i} * \sum_j (\mathbf{J}_{X,ij} * y_j) \right)(t); \tag{22}$$

that is, each input X_i has been replaced by a filtered version of the presynaptic firing rates y_j .

In the frequency domain this becomes

$$\tilde{y}_i(\omega) = \tilde{y}_i^0 + \tilde{A}_{X,i}(\omega) \left(\sum_j \tilde{\mathbf{J}}_{X,ij}(\omega) \tilde{y}_j(\omega) \right), \tag{23}$$

where $\tilde{y}_i = \mathcal{F}[y_i - v_i]$ is the Fourier transform of the mean-shifted process (v_i is the average firing rate of cell i) and $\tilde{f} = \mathcal{F}[f]$ for all other quantities. In matrix form, this yields a self-consistent equation for \tilde{y} in terms of \tilde{y}^0 :

$$(\mathbf{I} - \tilde{\mathbf{K}}(\omega))\tilde{y} = \tilde{y}^0 \Rightarrow \tilde{y} = (\mathbf{I} - \tilde{\mathbf{K}}(\omega))^{-1}\tilde{y}^0, \tag{24}$$

where $\tilde{\mathbf{K}}_{ij}(\omega) = \tilde{A}_{X,i}(\omega)\tilde{\mathbf{J}}_{X,ij}(\omega)$ is the interaction matrix in the frequency domain. The cross-spectrum is then computed via

$$\langle \tilde{y}(\omega)\tilde{y}^*(\omega) \rangle = (\mathbf{I} - \tilde{\mathbf{K}}(\omega))^{-1}\langle \tilde{y}^0(\omega)\tilde{y}^{0*}(\omega) \rangle(\mathbf{I} - \tilde{\mathbf{K}}^*(\omega))^{-1}. \tag{25}$$

To compute the interaction matrix for a network of conductance-based neurons, we use the effective time constant approximation (as in the supplemental for [41]). We first separate each conductance into mean and fluctuating parts, e.g., $g_{E,i} \rightarrow \langle g_{E,i} \rangle + (g_{E,i} - \langle g_{E,i} \rangle)$ (see the discussion in [12]). Next we identify an effective conductance $g_{0,i}$ and potential $\mathcal{E}_{\text{rev},i}$, and treat the fluctuating part of the conductances as noise, i.e. $g_{E,i} - \langle g_{E,i} \rangle \rightarrow \sigma_{g_{E,i}}\xi_{E,i}(t)$, so that Eq. (17) becomes

$$\begin{aligned} \tau_m \frac{dv_i}{dt} = & -g_{0,i}(v_i - \mathcal{E}_{\text{rev},i}) + \sigma_{g_{E,i}}\xi_{E,i}(t)(v_i - \mathcal{E}_E) \\ & + \sigma_{g_{I,i}}\xi_{I,i}(t)(v_i - \mathcal{E}_I) + \sqrt{\sigma_i^2\tau_m}\xi_i(t), \end{aligned} \tag{26}$$

where

$$g_{0,i} = 1 + \langle g_{E,i} \rangle + \langle g_{I,i} \rangle, \tag{27}$$

$$\mathcal{E}_{\text{rev},i} = \frac{\langle g_{E,i} \rangle \mathcal{E}_E + \langle g_{I,i} \rangle \mathcal{E}_I}{g_{0,i}}, \tag{28}$$

$$\sigma_{g_{E,i}}^2 = \text{Var}[g_{E,i}(t)] = \text{E}[(g_{E,i}(t) - \langle g_{E,i} \rangle)^2], \tag{29}$$

$$\sigma_{g_{I,i}}^2 = \text{Var}[g_{I,i}(t)] = \text{E}[(g_{I,i}(t) - \langle g_{I,i} \rangle)^2]. \tag{30}$$

The parameters which govern the firing rate response will now be the conductance mean and variance, e.g. $\langle g_{E,i} \rangle$ and $\sigma_{g_{E,i}}^2$; we next compute how these depend on incoming firing rates for second-order α -function synapses (Eqs. (19) and (20)). We first simplify the equation for the auxiliary variable (Eq. (20)):

$$\tau_{r,X} \frac{dg_{X,i}^{(1)}}{dt} = -g_{X,i}^{(1)} + \tau_{r,X}\hat{\alpha}_{X,i} \sum_k \delta(t - t_k) \tag{31}$$

so that $\hat{\alpha}_{X,i}$ includes all factors that contribute to the pulse size in Eq. (20), including synapse strength and pulse amplitude. The time constants $\tau_{r,X}$, $\tau_{d,X}$ and synapse jump sizes $\hat{\alpha}_{X,i}$ generally depend on cell type. Then assuming that each spike train is a Poisson process with a constant mean firing rate: i.e., each spike train is modeled as a stochastic process $S(t)$ with

$$\langle S(t) \rangle = \nu; \quad \langle S(t)S(t + \tau) \rangle - \nu^2 = \nu\delta(\tau),$$

a straightforward but lengthy calculation shows that

$$\begin{aligned} \langle g_{X,i}(t) \rangle &= \hat{\alpha}_{X,i} \nu_{X,i} \tau_{r,X}, \\ \text{Var}[g_{X,i}(t)] &= \left(\frac{1}{2} \hat{\alpha}_{X,i}^2 \nu_{X,i} \tau_{r,X} \right) \left(\frac{\tau_{r,X}}{\tau_{r,X} + \tau_{d,X}} \right) \\ &= \langle g_{X,i}(t) \rangle \times \frac{\hat{\alpha}}{2} \times \left(\frac{\tau_{r,X}}{\tau_{r,X} + \tau_{d,X}} \right), \end{aligned} \tag{32}$$

where $\nu_{X,i}$ is the total rate of type- X spikes incoming to cell i . Notice that modulating the rate of an incoming spike train will impact *both* the mean and variance of the input to the effective equation, Eq. (26) (via $\mathcal{E}_{\text{rev},i}$ and $\sigma_{g_{X,i}}$). Furthermore, this impact may differ for excitatory and inhibitory neurons, giving us a total of *four* parameters that can be varied in the effective equation.

Therefore, we have four susceptibility functions to compute, $\tilde{A}_{\langle g_E \rangle,i}(\omega)$, $\tilde{A}_{\langle g_I \rangle,i}(\omega)$, $\tilde{A}_{\sigma_{g_E}^2,i}(\omega)$, and $\tilde{A}_{\sigma_{g_I}^2,i}(\omega)$. The first two capture the change in firing rate as a result of a change in mean conductance— $\langle g_{E,i} \rangle \rightarrow \langle g_{E,i} \rangle_0 + \langle g_{E,i} \rangle_1 \exp(i\omega t)$ or $\langle g_{I,i} \rangle \rightarrow \langle g_{I,i} \rangle_0 + \langle g_{I,i} \rangle_1 \exp(i\omega t)$ —while the final two address a change in variance— $\sigma_{g_{E,i}}^2 \rightarrow (\sigma_{g_{E,i}}^2)_0 + (\sigma_{g_{E,i}}^2)_1 \exp(i\omega t)$ or $\sigma_{g_{I,i}}^2 \rightarrow (\sigma_{g_{I,i}}^2)_0 + (\sigma_{g_{I,i}}^2)_1 \exp(i\omega t)$. Since the corresponding Fokker–Planck equation required to obtain these entities is linear, we can compute both susceptibilities separately and combine them to get the net effect. With these pieces, we now have the interaction matrix:

$$\tilde{\mathbf{K}}_{ij}(\omega) = \begin{cases} \tilde{A}_{\langle g_E \rangle,i}(\omega) \tilde{\mathbf{J}}_{ij}(\omega) + \tilde{A}_{\sigma_{g_E}^2,i}(\omega) \tilde{\mathbf{L}}_{ij}(\omega), & j \text{ excitatory,} \\ \tilde{A}_{\langle g_I \rangle,i}(\omega) \tilde{\mathbf{J}}_{ij}(\omega) + \tilde{A}_{\sigma_{g_I}^2,i}(\omega) \tilde{\mathbf{L}}_{ij}(\omega), & j \text{ inhibitory,} \end{cases} \tag{34}$$

where $\tilde{\mathbf{L}}(\omega)$ plays a similar role as $\tilde{\mathbf{J}}$, but for the effect of incoming spikes on the *variance* of conductance. Its relationship to $\tilde{\mathbf{J}}$ (either in the frequency or time domain) is given by the same simple scaling shown in Eq. (33): i.e., for j excitatory,

$$\tilde{\mathbf{L}}_{ij}(\omega) = \tilde{\mathbf{J}}_{ij}(\omega) \times \left(\frac{\hat{\alpha}_{E,i}}{2} \right) \times \left(\frac{\tau_{r,E}}{\tau_{r,E} + \tau_{d,E}} \right), \tag{35}$$

where the first factor comes from the effective spike amplitude $\hat{\alpha}_{E,i}$ (and is the scale factor proposed in [29], Eq. (64)), and the second arises from using second-order (vs. first-order) alpha-functions.

To implement this calculation, we first solve for a self-consistent set of firing rates: that is, ν_i is the average firing rate of Eq. (26), along with Eqs. (27)–(30) and (32)–(33). We then apply Richardson’s threshold integration method [29, 30] directly to Eq. (26) to compute the unperturbed power spectrum ($\tilde{y}^0(\omega) \tilde{y}^{0*}(\omega)$) and susceptibility functions. The software we used to implement this calculation is described more fully in [4] and can be found at https://github.com/andreakbarreiro/LR_CondBased.

4.3 Computing Statistics from Linear Response Theory

Linear response theory yields the cross-spectrum of the spike train, $\langle \tilde{y}_i(\omega)\tilde{y}_j^*(\omega) \rangle$, for each distinct pair of neurons i and j (see Eq. (25)). The cross-correlation function, $\mathbf{C}_{ij}(\tau)$, measures the similarity between two processes at time lag τ , while the cross-spectrum measures the similarity between two processes at frequency ω :

$$\mathbf{C}_{ij}(\tau) \equiv \langle (y_i(t) - v_i)(y_j(t + \tau) - v_j) \rangle, \tag{36}$$

$$\tilde{\mathbf{C}}_{ij}(\omega) \equiv \langle \tilde{y}_i(\omega)\tilde{y}_j^*(\omega) \rangle. \tag{37}$$

The Weiner–Khinchin theorem [31] implies that $\{\mathbf{C}_{ij}, \tilde{\mathbf{C}}_{ij}\}$ are a Fourier transform pair: that is,

$$\tilde{\mathbf{C}}_{ij}(\omega) = \int_{-\infty}^{\infty} \mathbf{C}_{ij}(t)e^{-2\pi i\omega t} dt. \tag{38}$$

In principle, the cross-correlation $\mathbf{C}(t)$ and cross-spectrum $\tilde{\mathbf{C}}(\omega)$ matrices are functions on the real line, reflecting the fact that correlation can be measured at different time scales. In particular, for a stationary point process the covariance of spike counts over a window of length T , n_i and n_j , can be related to the cross-correlation function \mathbf{C}_{ij} by the following formula [17]:

$$\text{Cov}_T(n_i, n_j) = \int_{-T}^T \mathbf{C}_{ij}(\tau)(T - |\tau|) d\tau. \tag{39}$$

The variance of spike counts over a time window of length T , n_i , is likewise given by integrating the autocorrelation function \mathbf{C}_{ii} :

$$\text{Var}_T(n_i) = \int_{-T}^T \mathbf{C}_{ii}(\tau)(T - |\tau|) d\tau. \tag{40}$$

By normalizing by the time window and taking the limit as $T \rightarrow \infty$,

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{\text{Cov}_T(n_i, n_j)}{T} &= \lim_{T \rightarrow \infty} \int_{-T}^T \mathbf{C}_{ij}(\tau) \left(1 - \frac{|\tau|}{T}\right) d\tau \\ &= \int_{-\infty}^{\infty} \mathbf{C}_{ij}(\tau) d\tau = \tilde{\mathbf{C}}_{ij}(0), \end{aligned} \tag{41}$$

we can see that, for an integrable cross-correlation function, we can use $\tilde{\mathbf{C}}_{ij}(0)$ as a measure of long-time covariance.

Similarly, the long-time limit of the Pearson correlation coefficient of the spike counts,

$$\lim_{T \rightarrow \infty} \rho_{T,ij} = \lim_{T \rightarrow \infty} \frac{\text{Cov}_T(n_i, n_j)}{\sqrt{\text{Var}_T(n_i)\text{Var}_T(n_j)}} = \frac{\tilde{\mathbf{C}}_{ij}(0)}{\sqrt{\tilde{\mathbf{C}}_{ii}(0)\tilde{\mathbf{C}}_{jj}(0)}}, \tag{42}$$

gives us a normalized measure of long-time correlation.

Acknowledgements This work was motivated in part by useful feedback we received at the Third Annual International Conference on Mathematical Neurosciences, held in Boulder CO in May–June 2017. We would like to thank the organizers for an enjoyable and stimulating conference.

Funding Not applicable.

Availability of data and materials Software used to generate the computational results shown here can be found at: https://github.com/andreakarreiro/LR_CondBased.

Ethics approval and consent to participate Not applicable.

Competing interests The authors declare that they have no competing interests.

Consent for publication Not applicable.

Authors' contributions AKB and CL designed the project. AKB and CL wrote software. AKB performed simulations. AKB and CL designed figures. AKB and CL wrote the paper. All authors read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nat Rev Neurosci*. 2006;7:358–66.
2. Bair W, Zohary E, Newsome WT. Correlated firing in macaque visual area mt: time scales and relationship to behavior. *J Neurosci*. 2001;21(5):1676–97.
3. Barreiro AK, Gautam SH, Shew W, Ly C. A theoretical framework for analyzing coupled neuronal networks: application to the olfactory system. *PLoS Comput Biol*. 2017;13(10):e1005780.
4. Barreiro AK, Ly C. When do correlations increase with firing rate in recurrent networks? *PLoS Comput Biol*. 2017;13(4):e1005506.
5. Cohen MR, Maunsell JHR. Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci*. 2009;12:1594–600.
6. da Silveira RA, Berry MJ. High-fidelity coding with correlated neurons. *PLoS Comput Biol*. 2014;10(11):e1003970.
7. Dayan P, Abbott LF. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. London: Taylor & Francis; 2001.
8. de la Rocha J, Doiron B, Shea-Brown E, Josić K, Reyes A. Correlation between neural spike trains increases with firing rate. *Nature*. 2007;448:802–6.
9. Doiron B, Litwin-Kumar A, Rosenbaum R, Ocker GK, Josić K. The mechanics of state-dependent neural correlations. *Nat Neurosci*. 2016;19(3):383–93.
10. Ecker AS, Berens P, Tolias AS, Bethge M. The effect of noise correlations in populations of diversely tuned neurons. *J Neurosci*. 2011;31(40):14272–83.
11. Franke F, Fiscella M, Sevelev M, Roska B, Hierlemann A, Azeredo da Silveira R. Structures of neural correlation and how they favor coding. *Neuron*. 2016;89:409–22.
12. Gerstner W, Kistler WM, Naud R, Paninski L. *Neuronal dynamics: from single neurons to networks and models of cognition*. Cambridge: Cambridge University Press; 2014.
13. Helias M, Tetzlaff T, Diesmann M. The correlation structure of local neuronal networks intrinsically results from recurrent dynamics. *PLoS Comput Biol*. 2014;10(1):e1003428.

14. Hu Y, Trousdale J, Josic K, Shea-Brown E. Motif statistics and spike correlations in neuronal networks. *J Stat Mech Theory Exp.* 2013;2013(3):03012.
15. Hu Y, Zylberberg J, Shea-Brown E. The sign rule and beyond: boundary effects, flexibility, and noise correlations in neural population codes. *PLoS Comput Biol.* 2014;10(2):e1003469.
16. Josić K, Shea-Brown E, Doiron B, de la Rocha J. Stimulus-dependent correlations and population codes. *Neural Comput.* 2009;21:2774–804.
17. Kay SM. Fundamentals of statistical signal processing, volume 1: estimation theory. New York: Prentice Hall; 1993.
18. Kohn A, Coen-Cagli R, Kanitscheider I, Pouget A. Correlations and neuronal population information. *Annu Rev Neurosci.* 2016;39:237–56.
19. Lin IC, Okun M, Carandini M, Harris KD. The nature of shared cortical variability. *Neuron.* 2015;87:644–56.
20. Lindner B, Doiron B, Longtin A. Theory of oscillatory firing induced by spatially correlated noise and delayed inhibitory feedback. *Phys Rev E.* 2005;72(6):061919.
21. Litwin-Kumar A, Doiron B. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat Neurosci.* 2012;15(11):1498–505.
22. Litwin-Kumar A, Doiron B. Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nat Commun.* 2014;5:5319.
23. Ly C. Firing rate dynamics in recurrent spiking neural networks with intrinsic and network heterogeneity. *J Comput Neurosci.* 2015;39:311–27.
24. Martin KAC. Microcircuits in visual cortex. *Curr Opin Neurobiol.* 2002;12(4):418–25.
25. Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A. Information-limiting correlations. *Nat Neurosci.* 2014;17:1410–7.
26. Ostojic S, Brunel N, Hakim V. How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *J Neurosci.* 2009;29:10234–53.
27. Pernice V, Staude B, Cardanobile S, Rotter S. How structure determines correlations in neuronal networks. *PLoS Comput Biol.* 2011;7(5):e1002059.
28. Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A, Harris KD. The asynchronous state in cortical circuits. *Science.* 2010;327:587–90.
29. Richardson MJE. Firing-rate response of linear and nonlinear integrate-and-fire neurons to modulated current-based and conductance-based synaptic drive. *Phys Rev E.* 2007;76:021919.
30. Richardson MJE. Spike-train spectra and network response functions for non-linear integrate-and-fire neurons. *Biol Cybern.* 2008;99:381–92.
31. Risken H. The Fokker–Planck equation: methods of solutions and applications. New York: Springer; 1989.
32. Rosenbaum R, Smith MA, Kohn A, Rubin JE, Doiron B. The spatial structure of correlated neuronal variability. *Nat Neurosci.* 2017;20:107–14.
33. Ruff DA, Cohen MR. Attention can either increase or decrease spike count correlations in visual cortex. *Nat Neurosci.* 2014;17(11):1591–7.
34. Sanchez-Vives MV, McCormick DA. Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nat Neurosci.* 2000;3(10):1027–34.
35. Schuecker J, Schmidt M, van Albada SJ, Diesmann M, Helias M. Fundamental activity constraints lead to specific interpretations of the connectome. *PLoS Comput Biol.* 2017;13(2):e1005179.
36. Schulz DPA, Sahani M, Carandini M. Five key factors determining pairwise correlations in visual cortex. *J Neurophysiol.* 2015;114:1022–33.
37. Shamir M, Sompolinsky H. Implications of neuronal diversity on population coding. *Neural Comput.* 2006;18:1951–86.
38. Shea-Brown E, Josic K, de la Rocha J, Doiron B. Correlation and synchrony transfer in integrate-and-fire neurons: basic properties and consequences for coding. *Phys Rev Lett.* 2008;100:108102.
39. Shu Y, Hasenstaub A, McCormick DA. Turning on and off recurrent balanced cortical activity. *Nature.* 2003;423(6937):288–93.
40. Tetzlaff T, Helias M, Einevoll GT, Diesmann M. Decorrelation of neural-network activity by inhibitory feedback. *PLoS Comput Biol.* 2012;8:e1002596.
41. Trousdale J, Hu Y, Shea-Brown E, Josic K. Impact of network structure and cellular response on spike time correlations. *PLoS Comput Biol.* 2012;8(3):e1002408.
42. van Vreeswijk C, Sompolinsky H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science.* 1996;274:1724–6.

43. Zhao L, Beverlin B II, Netoff T, Nykamp DQ. Synchronization from second order network connectivity statistics. *Front Comput Neurosci*. 2011;5:1–16.
44. Zylberberg J, Cafaro J, Turner MH, Shea-Brown E, Rieke F. Direction-selective circuits shape noise to ensure a precise population code. *Neuron*. 2016;89(2):369–83.