Practical approximation method for firing-rate models of coupled neural networks with correlated inputs

Andrea K. Barreiro

Department of Mathematics, Southern Methodist University, P.O. Box 750235, Dallas, Texas 75275, USA

Cheng Ly*

*Department of Statistical Sciences and Operations Research, Virginia Commonwealth University, 1015 Floyd Avenue,
Richmond, Virginia 23284, USA*

Rapid experimental advances now enable simultaneous electrophysiological recording of neural activity at single-cell resolution across large regions of the nervous system. Models of this neural network activity will necessarily increase in size and complexity, thus increasing the computational cost of simulating them and the challenge of analyzing them. Here we present a method to approximate the activity and firing statistics of a general firing rate network model (of the Wilson-Cowan type) subject to noisy correlated background inputs. The method requires solving a system of transcendental equations and is fast compared to Monte Carlo simulations of coupled stochastic differential equations. We implement the method with several examples of coupled neural networks and show that the results are quantitatively accurate even with moderate coupling strengths and an appreciable amount of heterogeneity in many parameters. This work should be useful for investigating how various neural attributes qualitatively affect the spiking statistics of coupled neural networks.

I. INTRODUCTION

With advances in neural recording technologies, experimentalists can now record simultaneous activity across multiple brain regions at single cell resolution [1–4]. However, it is still a technical challenge to measure the interactions within and across brain regions that govern this multi-region activity. This challenge is heightened by the fact that cortical neurons are heterogeneous and show substantial trial-to-trial variability [5]. Numerous theoretical studies have examined how neural networks can lead to cortex-like dynamics [6–14]; however, most have been limited to a single region, leaving open the question of how interregion connection strengths contribute to network processing.

One challenge presented by analyzing multiregion neural networks is the increased number of parameters which must be specified. To survey a high-dimensional parameter space, one must have a way to efficiently simulate (as in Ref. [15]) or approximate network statistics (as in Ref. [16]). Here we present a novel approximation method for calculating the statistics of a general coupled firing rate model (based on Ref. [17]) of neural networks where we (1) assume the *activity* (not the firing statistics) are pairwise normally distributed and (2) take the entire probability distribution of the presynaptic neurons or populations (providing input) into account. Our method is fast because it requires solving nonlinear equations self-consistently rather than simulating stochastic differential equations. Several example neural networks are considered and compared with Monte Carlo simulations. A specific version of this method was presented in Ref. [18] to model the olfactory sensory pathway; here we derive formulas in a general way which is easy to evaluate and can accommodate heterogeneous networks. We also demonstrate the method's

efficacy on several example networks with much larger dimension than the specific networks examined in our previous work.

MATLAB code implementing the method is freely available at GitHub [19].

II. NEURAL NETWORK MODEL

Each cell (or homogeneous population) has a prescribed activity x_j that is modeled by the following equation [17] for $j = 1, 2, \ldots, N_c$:

$$\tau_j \frac{dx_j}{dt} = -x_j + \mu_j + \sigma_j \eta_j(t) + \sum_{k=1}^{N_c} g_{jk} F_k[x_k(t)], \quad (1)$$

where $F_k(\cdot)$ is a transfer function mapping activity to firing rate (in some units), related to the so-called F-I curve, for the kth cell or population. Thus, the instantaneous firing rate of the jth neuron is

$$F_j[x_j(t)]. \quad (2)$$

Depending on the context, the activity variable x_j may represent membrane voltage, calcium concentration, or some other quantity associated with a neuron's internal state [20]. This type of equation has historically been used to capture the average activity of a population of neurons, but hereafter we will use the term "cell" for exposition purposes. All cells receive background noise η_j, the increment of a Weiner process, uncorrelated in time but potentially correlated at each instant: $\langle \eta_j(t) \rangle = 0$, $\langle \eta_j(t) \eta_j(t') \rangle = \delta(t - t')$, and $\langle \eta_j(t) \eta_k(t') \rangle = c_{jk} \delta(t - t')$ for $j \neq k$ with $c_{jk} \in (-1,1)$. The parameters μ_j and σ_j are constants that give the background input mean and input standard deviation, respectively. The parameter g_{jk} represents coupling strength from the presynaptic kth cell and is a signed quantity; $g_{jk} < 0$ represents inhibitory coupling.

*CLy@vcu.edu

We would like to compute the following statistics:

$$\mu(j) := \langle x_j \rangle, \text{mean activity}, \tag{3}$$

$$\sigma^2(j) := \langle x_j^2 \rangle - \mu^2(j), \text{variance of activity}, \tag{4}$$

$$\text{Cov}(j,k) := \langle x_j x_k \rangle - \mu(j)\mu(k),$$
$$\text{covariance of activity}, \tag{5}$$

$$\nu_j := \langle F_j(x_j) \rangle, \text{firing rate}, \tag{6}$$

$$\text{Var}(\nu_j) := \langle F_j^2(x_j) - \nu_j^2 \rangle, \text{variance of spiking}, \tag{7}$$

$$\text{Cov}(\nu_j, \nu_k) := \langle F_j(x_j)F_k(x_k) \rangle - \nu_j \nu_k,$$
$$\text{covariance of spiking}, \tag{8}$$

$$\rho(\nu_j, \nu_k) := \frac{\text{Cov}(\nu_j, \nu_k)}{\sqrt{\text{Var}(\nu_j)\text{Var}(\nu_k)}},$$
$$\text{correlation of spiking}, \tag{9}$$

where the angular brackets $\langle \cdot \rangle$ denote averaging over time and realizations [21]. We will use the following definitions for the following normal or Gaussian probability density functions (PDF):

$$\varrho_1(y) := \frac{1}{\sqrt{2\pi}} e^{-y^2/2}, \tag{10}$$

the standard normal PDF, and

$$\varrho_{j,k}(y_1, y_2) := \frac{1}{2\pi\sqrt{1-c_{jk}^2}} \exp\left[-\frac{1}{2}\vec{y}^T \begin{pmatrix} 1 & c_{jk} \\ c_{jk} & 1 \end{pmatrix}^{-1} \vec{y} \right], \tag{11}$$

a bivariate normal distribution with $\vec{0}$ mean, unit variance, and covariance c_{jk}.

In the absence of coupling, i.e., $g_{jk} = 0$, Eq. (1) would describe a multidimensional Ornstein-Uhlenbeck process. Such a process is well understood: any pair of activity variables, (x_j, x_k), are bivariate normal random variables [22]. To see this, consider the following two equations without synaptic coupling:

$$\tau_j \frac{dx_j}{dt} = -x_j + \mu_j + \sigma_j[\sqrt{1-c_{jk}}\xi_j(t) + \sqrt{c_{jk}}\xi_c(t)], \tag{12}$$

$$\tau_k \frac{dx_k}{dt} = -x_k + \mu_k + \sigma_k[\sqrt{1-c_{jk}}\xi_k(t) + \sqrt{c_{jk}}\xi_c(t)]. \tag{13}$$

Note that we have rewritten $\eta_{j/k}(t)$ as sums of independent white noise processes $\xi(t)$. Since $x_j(t) = \frac{1}{\tau_j}\int_{-\infty}^t e^{-(t-u)/\tau_j}[\mu_j + \sigma_j \eta_j(u)]\,du$ (where we have taken the initial time to be in the far past to eliminate any impact from the initial conditions), we calculate marginal statistics using

Itô isometries:

$$\mu(j) \equiv \langle x_j \rangle = \mu_j,$$

$$\sigma^2(j) \equiv \langle [x_j - \mu(j)]^2 \rangle$$
$$= \left\langle \frac{\sigma_j^2}{\tau_j^2} \int_{-\infty}^t \int_{-\infty}^t e^{-(t-u)/\tau_j}\eta_j(u)e^{-(t-v)/\tau_j}\eta_j(v)\,du\,dv \right\rangle$$
$$= \frac{\sigma_j^2}{\tau_j^2}\int_{-\infty}^t e^{-2(t-u)/\tau_j}\,du = \frac{\sigma_j^2}{2\tau_j}. \tag{14}$$

A similar calculation shows that in general we have

$$\text{Cov}(j,k) = \frac{c_{jk}}{\tau_j + \tau_k}\sigma_j\sigma_k. \tag{15}$$

Thus, $\quad (x_j, x_k) \sim \mathbb{N}\left(\begin{pmatrix} \mu_j \\ \mu_k \end{pmatrix}, \begin{pmatrix} \frac{\sigma_j^2}{2\tau_j} & \sigma_j\sigma_k\frac{c_{jk}}{\tau_j+\tau_k} \\ \sigma_j\sigma_k\frac{c_{jk}}{\tau_j+\tau_k} & \frac{\sigma_k^2}{2\tau_k} \end{pmatrix} \right).$

Statistics for the firing rates, $F(x_j)$, are inherited from this normal distribution, since the firing rate $F(x_j)$ is simply a nonlinear function of the activity x_j.

When coupling is included, i.e., $g_{jk} \neq 0$ for some indices j and k, it may no longer be true that the activity variables x_j remain normally distributed. However, it is reasonable to suppose that, for sufficiently weak coupling, the deviations from a normal distribution will be small. Furthermore, if the firing rate function F has thresholding and saturating behavior (as does a sigmoidal function), then higher moments of x_j have limited impact on statistics of $F(x_j)$. Thus, our first assumption will be that each pair of activity variables (x_j, x_k), can be approximated by a bivariate normal, even when coupling is present. We can think of this as a weak coupling assumption, as it holds exactly only with *no* coupling.

III. REDUCTION METHOD

In our method, we assume that time is dimensionless so that the subsequent assumptions have the proper units. Note that our method can in principle be applied to systems where time has a dimension, as long they are of the form in Eq. (1) (with appropriate units for the parameters).

To compute statistics, we start by writing Eq. (1) as a low-pass filter of the right-hand side:

$$x_j(t) = x_j(t_0)e^{-(t-t_0)/\tau_j} + \frac{1}{\tau_j}\int_{t_0}^t e^{-(t-u)/\tau_j}$$
$$\times \left\{ \mu_j + \sigma_j \eta_j(u) + \sum_k g_{jk}F_k[x_k(u)] \right\} du, \tag{16}$$

used as the basis for calculating the desired moments of x_j. For example, when $\langle x_j x_k \rangle$ is desired, we use the previous equation for j and k, multiply, then take the expected value $\langle \cdot \rangle$. By letting the initial time $t_0 \to -\infty$, we eliminate transients; the resulting statistics will be stationary. The resulting exact formulas are complicated by the network coupling, so we simplify the calculation(s) as follows.

We only account for direct connections in the formulas for the first and second-order statistics, assuming the terms from the indirect connections are either small or already accounted for in the direct connections. For example: although $F_k(x_k(u))$

TABLE I. For readability, we define the following quantities. Whenever $j = k$ in the double integrals (e.g., in $\mathcal{N}_\mathcal{F}, \mathcal{S}$), the bivariate normal distribution $\varrho_{j,k}$ is replaced with the standard normal distribution ϱ_1. Note that order of the arguments matters in $\mathcal{N}_\mathcal{F}$: $\mathcal{N}_\mathcal{F}(j,k) \neq \mathcal{N}_\mathcal{F}(k,j)$ in general; all of these quantities depend on the statistics of the activity $\mu(\cdot), \sigma(\cdot)$.

Abbreviation	Definition
$\mathcal{E}_1(k)$	$\int F_k[\sigma(k)y + \mu(k)]\varrho_1(y)\,dy$
$\mathcal{E}_2(k)$	$\int F_k^2[\sigma(k)y + \mu(k)]\varrho_1(y)\,dy$
$\mathcal{V}(k)$	$\int F_k^2[\sigma(k)y + \mu(k)]\varrho_1(y)\,dy - \{\int F_k[\sigma(k)y + \mu(k)]\varrho_1(y)\,dy\}^2 = \mathcal{E}_2(k) - [\mathcal{E}_1(k)]^2$
$\mathcal{N}_\mathcal{F}(j,k)$	$\iint F_k[\sigma(k)y_1 + \mu(k)]\frac{y_2}{\sqrt{2}}\varrho_{j,k}(y_1,y_2)\,dy_1\,dy_2$, if $j \neq k$
	$\int F_j[\sigma(j)y + \mu(j)]\frac{y}{\sqrt{2}}\varrho_1(y)\,dy$, if $j = k$
$\mathcal{S}(j,k)$	$\iint F_j[\sigma(j)y_1 + \mu(j)]F_k[\sigma(k)y_2 + \mu(k)]\varrho_{j,k}(y_1,y_2)\,dy_1\,dy_2$
$\mathcal{C}_\mathcal{V}(j,k)$	$\mathcal{S}(j,k) - \mathcal{E}_1(j)\mathcal{E}_1(k)$

on the right-hand side of Eq. (16) *itself* depends on coupling terms of the form $g_{kl}F_l(x_l)$, etc., we will neglect such terms. We further make the following assumptions:

$$\left\langle \int_{-\infty}^t F_k[x_k(u)]e^{-(t-u)/\tau_l}\,du \int_{-\infty}^t F_k[x_k(v)]e^{-(t-v)/\tau_m}\,dv \right\rangle \approx \frac{\tau_l\tau_m}{\tau_l + \tau_m}\mathcal{V}_k + \tau_l\tau_m[\mathcal{E}_1(k)]^2, \tag{17}$$

$$\left\langle \int_{-\infty}^t \sigma_j\eta_j(u)e^{-(t-u)/\tau_l}\,du \int_{-\infty}^t F_k[x_k(v)]e^{-(t-v)/\tau_m}\,dv \right\rangle \approx \frac{\tau_l\tau_m}{\tau_l + \tau_m}\sigma_j\mathcal{N}_\mathcal{F}(j,k), \tag{18}$$

$$\left\langle \int_{-\infty}^t F_j[x_j(u)]e^{-(t-u)/\tau_l}\,du \int_{-\infty}^t F_k[x_k(v)]e^{-(t-v)/\tau_m}\,dv \right\rangle \approx \frac{\tau_l\tau_m}{\tau_l + \tau_m}\mathcal{C}_\mathcal{V}(j,k) + \tau_l\tau_m\mathcal{E}_1(j)\mathcal{E}_1(k). \tag{19}$$

See Table I for the definition of the symbols: $\mathcal{E}_1$, $\mathcal{N}_\mathcal{F}$, $\mathcal{V}_k$, and $\mathcal{C}_\mathcal{V}$.

Each assumption is equivalent to the assumption that two of the random variables of interest are δ-correlated in time; thus avoiding the need to compute autocorrelation functions explicitly. The first assumption [Eq. (17)] states that $F_k(x_k(t))$ is δ-correlated with itself; the second [Eq. (18)] addresses $\eta_j(t)$ and $F_k[x_k(t)]$. The final assumption [Eq. (19)] states that $F_j[x_j(t)]$ and $F_k[x_k(t)]$ are δ-correlated. We provide a detailed derivation of Eqs. (17)–(19) in Appendix A. After testing our method on several examples in Sec. IV, we will revisit the accuracy of these assumptions in Sec. V.

We arrive at the following (approximation) formulas for the statistics of the activity:

$$\mu(j) = \mu_j + \sum_k g_{jk}\mathcal{E}_1(k), \tag{20}$$

$$\sigma^2(j)\tau_j = \frac{\sigma_j^2}{2} + \sigma_j\sum_k g_{jk}\mathcal{N}_\mathcal{F}(j,k) + \frac{1}{2}\sum_k g_{jk}^2\mathcal{V}(k) + \sum_{k\neq l}g_{jk}g_{jl}\mathcal{C}_\mathcal{V}(k,l), \tag{21}$$

$$\mathrm{Cov}(j,k)\frac{\tau_j + \tau_k}{2} = \frac{1}{2}c_{jk}\sigma_j\sigma_k + \frac{1}{2}\sigma_j\sum_l g_{kl}\mathcal{N}_\mathcal{F}(j,l), + \frac{1}{2}\sigma_k\sum_l g_{jl}\mathcal{N}_\mathcal{F}(k,l) + \frac{1}{2}\sum_{l_1,l_2}g_{j,l_1}g_{k,l_2}\mathcal{C}_\mathcal{V}(l_1,l_2). \tag{22}$$

See Table I for the definition of the symbols $\mathcal{E}_1$, $\mathcal{N}_\mathcal{F}$, $\mathcal{V}$, and $\mathcal{C}_\mathcal{V}$, which all depend on the statistical quantities $\mu(\cdot)$ and $\sigma(\cdot)$ of the activity x_j. Our approximation formulas form a system of $\frac{1}{2}(N_c^2 + 3N_c)$ equations in $\mu(j)$, $\sigma(j)$, $\mathrm{Cov}(j,k)$ [i.e., for the activity only, as defined by Eqs. (3)–(5), not the firing] when considering all possible $(j,k) \in \{1, 2, \ldots, N_c\}$. This large system of equations, although nonlinear, is simple to solve because it requires a sequence of function evaluations and matrix multiplications, rather than random sampling.

Note that the normal distribution assumptions allow us to conveniently write the average quantities as integrals with respect to standard normal distributions but with shifted integrands, which leads to faster calculations because one does not have to calculate new probability density functions at each step of the iteration when solving the system self-consistently.

The resulting formulas can be written compactly with matrices; Eq. (20) for the mean activity $\mu(j)$ can easily be written as a matrix-vector equation and is thus omitted. Let **Cov** denote the $N_c \times N_c$ covariance matrix of the activity with $\mathbf{Cov}(j,k) = \mathrm{Cov}(j,k)$, **G** represent the coupling strengths $\mathbf{G}(j,k) = g_{jk}$, and **Cr** denote the correlation matrix of the background noise [i.e., $\mathbf{Cr}(j,k) = \delta_{jk} + c_{jk}(1 - \delta_{jk})$]. Then we have

$$\mathbf{Cov} = \mathbf{IT} \circ (\mathbf{Cov_0} + \mathbf{GM_{NF}} + \mathbf{M_{NF}}^T\mathbf{G}^T + \mathbf{GM_{FSq}}\mathbf{G}^T), \tag{23}$$
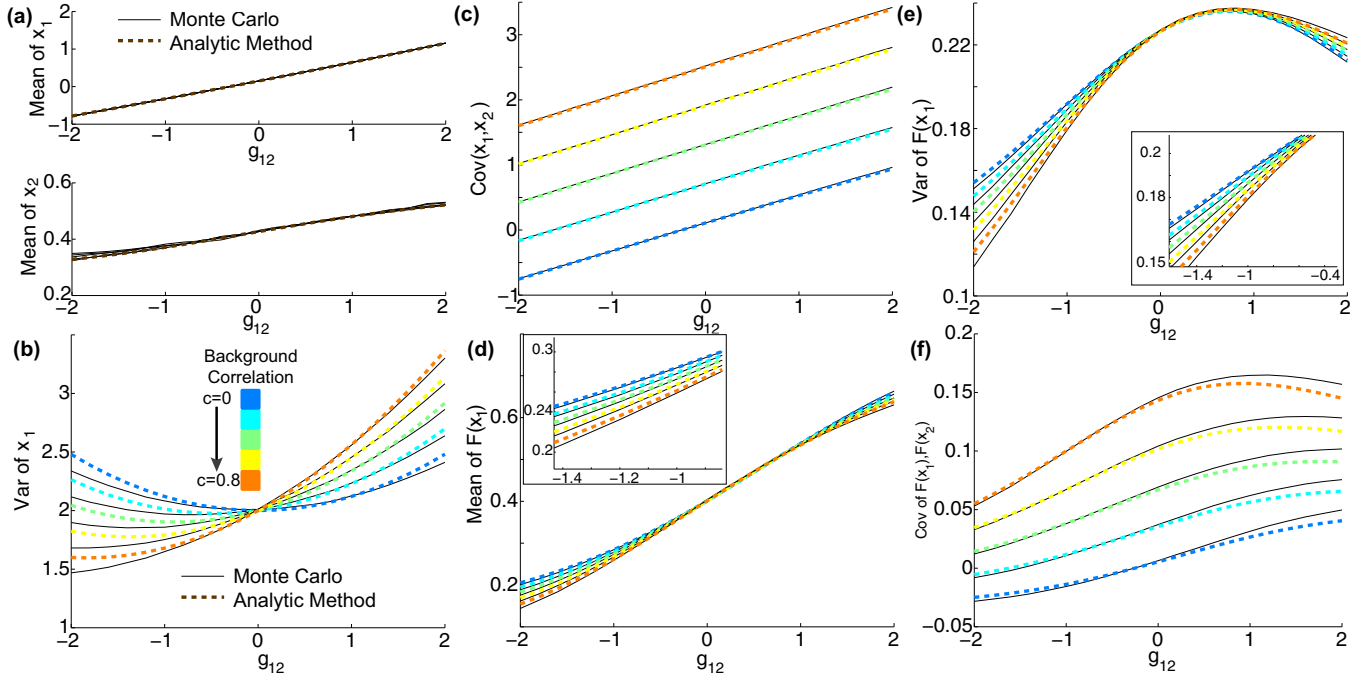
FIG. 1. Illustration of the method on a network with two neurons. In all panels, the Monte Carlo simulation results are the thin black solid lines, and the result of the analytic method [Eqs. (20)–(22) solved self-consistently, and Eqs. (28)–(30)] are the dashed colored lines representing different background correlation levels. All parameters are fixed except g_{12} and $c_{12} = c_{21} =: c$; see the main text (Network I) for values. (a) The average activity x_1 (top), x_2 (bottom) as a function of g_{12} match very well; here the analytic method is in one color (brown) because the result is independent of background correlation. (b) The variance of x_1, $\sigma^2(1)$, varies with both background correlation and input strength. The match is very good around $g_{12} = 0$ and starts to deviate as $|g_{12}| \to 2$ because with stronger coupling the normal distribution assumption is severely violated. (c) The covariance of the activity Cov(1,2). (d) Mean firing rate: $F(\nu_1)$ slightly depends on c; inset is a zoomed-in picture to show that the method captures the relationship of the curves. (e) Variance of $F(\nu_1)$. (f) The covariance of the firing rate Cov($F(\nu_1), F(\nu_2)$). The corresponding plots for x_2 [i.e., panels (b), (d), (e)] are not shown because they do not vary as much; however, the analytic method accurately captures the results from Monte Carlo simulations.

where $\circ$ represents element-wise multiplication, $(\cdot)^T$ denotes matrix transposition, and

$$\mathbf{IT}(j,k) = \frac{1}{\tau_j + \tau_k}, \qquad (24)$$

$$\mathbf{Cov_0}(j,k) = \sigma_j \sigma_k [\delta_{jk} + (1 - \delta_{jk}) c_{jk}], \qquad (25)$$

$$\mathbf{M_{NF}}(j,k) = \sigma_k \mathcal{N_F}(k,j), \qquad (26)$$

$$\mathbf{M_{FSq}}(j,k) = \mathcal{C_V}(j,k). \qquad (27)$$

Note that the matrices $\mathbf{M_{NF}}$ and $\mathbf{M_{FSq}}$ have the same nonzero entries as $\mathbf{Cr}$. Denoting $\mathbf{\Lambda_{\vec{\sigma}}}$ as the diagonal matrix with diagonal $\vec{\sigma}$, the unperturbed covariance [Eq. (25)] can also be expressed in matrix form as

$$\mathbf{Cov_0} = (\mathbf{\Lambda_{\vec{\sigma}}}) \mathbf{Cr} (\mathbf{\Lambda_{\vec{\sigma}}}).$$

Once the statistics of the activity [$\mu(j)$, $\sigma^2(j)$, and Cov(j,k)] are solved for self-consistently, the firing statistics are solved as follows:

$$\nu_j = \int F_j[\sigma(j)y + \mu(j)] \varrho_1(y)\, dy, \qquad (28)$$

$$\text{Var}(\nu_j) = \int F_j^2[\sigma(j)y + \mu(j)] \varrho_1(y)\, dy - \nu_j^2, \qquad (29)$$

$$\text{Cov}(\nu_j, \nu_k) = \iint F_j[\sigma(j)y_1 + \mu(j)] F_k[\sigma(k)y_2 + \mu(k)]$$

$$\times \mathbb{P}_{j,k}(y_1, y_2)\, dy_1 dy_2 - \nu_j \nu_k \qquad (30)$$

where $\mathbb{P}_{j,k}$ is a bivariate normal PDF with zero mean and covariance: $\begin{pmatrix} 1 & \frac{\text{Cov}(j,k)}{\sigma(j)\sigma(k)} \\ \frac{\text{Cov}(j,k)}{\sigma(j)\sigma(k)} & 1 \end{pmatrix}$. The off-diagonal terms are obtained from the second-order statistics of the activity [Eq. (21)–(22)].

IV. EXAMPLE NETWORKS AND RESULTS

A. Network I

We first consider a network that allows us to systematically explore algorithm performance as two key parameters vary. Specifically, we consider two cells ($N_c = 2$) that are reciprocally coupled without autaptic (i.e., self) coupling. For simplicity, we set the intrinsic parameters for the two cells to be identical, with $\tau_j = 1$, $F_j(x) = 0.5[1 + \tanh((x - 0.5)/0.1)] \in [0,1]$ (arbitrary units), but the mean and variance of the background input differ: $\mu_1 = 0.15$, $\mu_2 = 4/15 \approx 0.2667$, $\sigma_1 = 2$, $\sigma_2 = 3$. We vary two parameters: $g_{12} \in [-2,2]$ (input strength from x_2 to x_1), and $c_{12} = c_{21} \in [0,0.8]$, with $g_{21} = 0.4$ fixed.

In Fig. 1 we see that all of the activity *and* firing statistics are accurate compared to Monte Carlo simulations. Figure 1(a) shows the mean of x_1 as the input strength g_{12} varies from negative (inhibitory) to positive (excitatory); this statistic is independent of background correlation. Figure 1(b) shows the variance of x_1; deviations are apparent when the magnitude
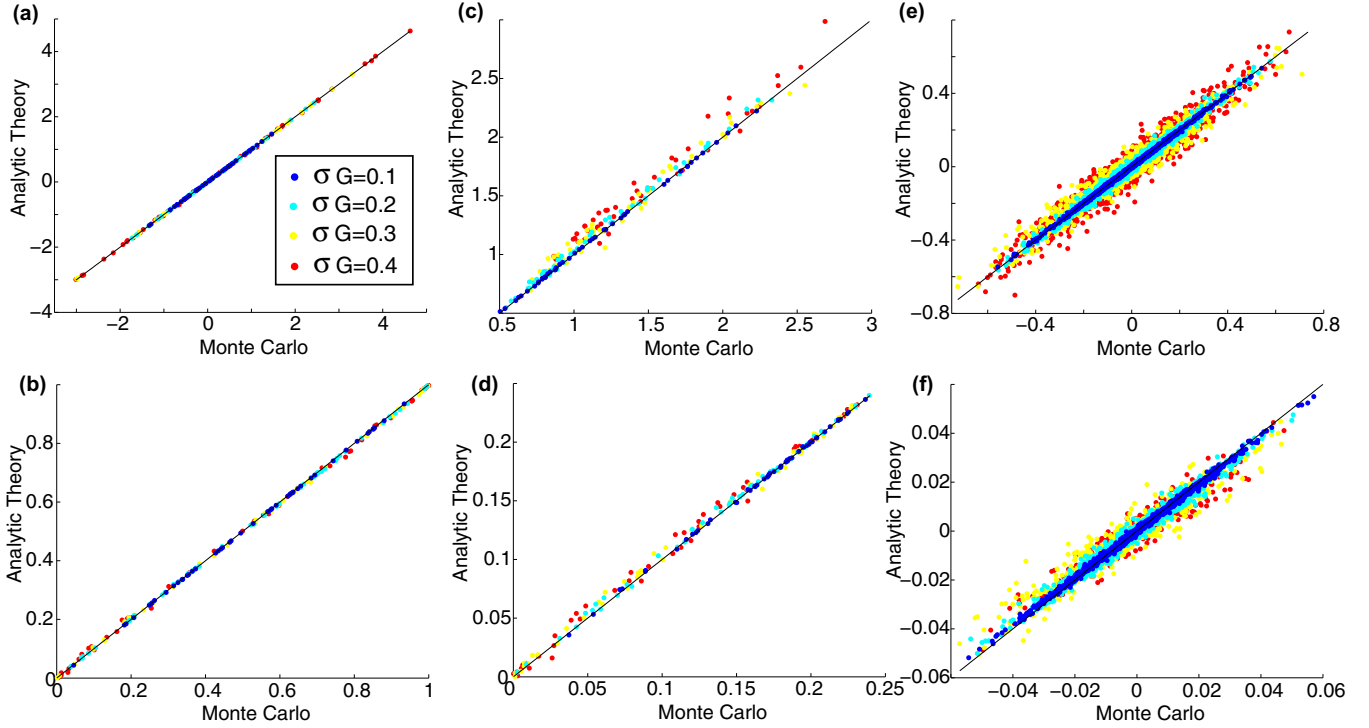
FIG. 2. A network of $N_c = 50$ neurons with heterogeneity in all parameters and all-to-all coupling (Network II). See Eqs. (31)–(37) for the distributions of the randomly selected parameters. In each panel, four different values of the variance of the distribution of the coupling matrix entries are shown, while the other parameters are held fixed. (a) Comparison of the mean activity $\mu(j)$ calculated via Monte Carlo simulations (horizontal axis) and our reduction method (vertical axis), showing all 50 values for each color (coupling matrix distribution). (b) Similar to (a) but for mean firing rate ν_j. (c) Variance of activity $\sigma^2(j)$. (d) Variance of firing rate Var(ν_j). (e) Covariance of activity Cov(j,k), showing all $50 \times 49/2 = 1225$ values for each coupling matrix. (f) Covariance of the firing rate Cov(ν_j, ν_k). The method is accurate but starts to deviate as the overall coupling strength $\|\mathbf{G}\|$ increases (from blue to red, more deviations from diagonal line).

of the coupling g_{12} is large. The covariance of the activity [Fig. 1(c)] is also accurate. Even the statistics of the firing rate are relatively accurate; the mean firing rate $F(x_1)$ [Fig. 1(d)] is only weakly dependent on background correlation, whereas the variance of $F(x_1)$ [Fig. 1(e)] appears to vary more with background correlation. In Fig. 1(f), the strong dependence of the covariance of the firing rate on background correlation is captured by our method. For brevity, we omit the corresponding statistics for x_2; the method performs equally well there.

B. Network II

We next consider an all-to-all coupled network of $N_c = 50$ neurons with heterogeneity in all parameters. The parameter values were selected from specific distributions and gave rise to quenched variability. The transfer function was set to $F_j(\bullet) = 0.5\{1 + \tanh[(\bullet - x_{\text{rev},j})/x_{sp,j}]\} \in [0,1]$, where $x_{\text{rev},j}$ and $x_{sp,j}$ are fixed parameters that depend on the the jth neuron. The distributions of the parameters for this network are

$$\tau_j \sim \mathbb{N}(1, 0.05^2), \tag{31}$$

$$\mu_j \sim 2\mathbb{U} - 1, \tag{32}$$

$$\sigma_j \sim \mathbb{U} + 1, \tag{33}$$

$$x_{\text{rev},j} \sim \mathbb{N}(0, 0.1^2), \tag{34}$$

$$x_{sp,j} \sim 0.35\mathbb{U} + 0.05, \tag{35}$$

where $\mathbb{U} \in [0,1]$ is a uniform random variable, and $\mathbb{N}$ is normally distributed with the mean and variance as the arguments. The covariance matrix $\mathbf{Cr}$ of the background noise was randomly selected as follows:

$$\mathbf{Cr} = (\mathbf{\Lambda}_{\vec{d}_s})\mathbf{A}^T\mathbf{A}(\mathbf{\Lambda}_{\vec{d}_s}), \tag{36}$$

where the entries of the $N_c \times N_c$ matrix $\mathbf{A}$ are independently chosen from a normal distribution: $a_{jk} \sim \mathbb{N}(0, 0.8^2)$ and $\vec{d}_s$ is the inverse square root of the diagonal of $\mathbf{A}^T\mathbf{A}$; i.e., if we set $\mathbf{B} := \mathbf{A}^T\mathbf{A}$ with entries b_{jk}, then $d_s(j) = 1/\sqrt{b_{jj}}$. By construction, $\mathbf{Cr}$ is symmetric positive semidefinite with 1's on the diagonal.

Finally, the entries of the coupling matrix $\mathbf{G}$ are randomly chosen, but the parameters of the distribution were varied:

$$\mathbf{G}(j,k) \sim \mathbb{N}(0, v_l), \tag{37}$$

where $v_l = (l/10)^2$ for $l = 1, 2, 3, 4$. There are no zero entries in $\mathbf{G}$ (i.e., coupling is all-to-all), with inhibition, excitation, and autaptic (self) coupling.

For each of the four values for the variance of the normal distribution, we chose a single realization of a coupling matrix $\mathbf{G}$ and computed first- and second-order statistics of x_k and $F(x_k)$. In Fig. 2 we compared our analytic versus Monte Carlo results for each cell or cell pair. Each realization is identified by a different color; in Fig. 2(a) for example, there are N_c red data points, corresponding to each $\mu(j)$ for $j = 1, \ldots, N_c$.

Points that are on the black diagonal line represent a perfect match between Monte Carlo simulations and our method.

First-order statistics $\mu(j)$ and ν_j are well captured by the analytic method, even for the largest coupling strength [Figs. 2(a) and 2(b)]. This excellent agreement is present despite the substantial amount of heterogeneity in these networks: note that $x_j = O(1)$ and that $F_j \in [0,1]$, and thus that single-cell firing rates in Fig. 2(b) have a relatively large range. Second-order statistics [variances and covariances; Figs. 2(c)–2(f)] are captured well for smaller coupling values (blue and cyan) but become less accurate for the largest coupling value (red). In particular, the analytic method appears to overestimate variance for the largest coupling strength [Fig. 2(c)].

C. Network III

Finally we consider a moderately sized network of $N_c = 100$ neurons with quenched heterogeneity in all of the intrinsic parameters, but with more physiological connectivity structure than Network II. The first 50 neurons are excitatory (**E**) ($g_{jk} \geqslant 0$ for $k = 1,2,\ldots,50$), and the last 50 are inhibitory (**I**) ($g_{jk} \leqslant 0$ for $k = 51,52,\ldots,100$). We choose a sparse (random) background correlation matrix via

$$c_{jk} = \begin{cases} 1, & \text{if } j = k \\ \mathbb{N}(0.1,0.1^2), & \text{if } k = j + 1 \text{ and } j = 1,\ldots,49 \\ \mathbb{N}(0.12,0.1^2), & \text{if } k = j + 1 \text{ and } j = 51,\ldots,99 \\ \mathbb{N}(0.3,0.1^2), & \text{if } k = 101 - j \text{ and } j = 1,\ldots,100 \\ 0 & \text{otherwise} \end{cases},$$

$$(38)$$

where as before $\mathbb{N}$ is a Gaussian random variable. That is, each cell shares correlated input with its nearest neighbors of the same type (excitatory versus inhibitory), and a single cell of the opposite type, where cell location varies along a one-dimensional line. This results in a correlation matrix which is tridiagonal, with an antidiagonal band for the E and I correlation; this sparsity structure is shown in Fig. 3(a).

In a variety of cortical areas, there is evidence that the correlation of neural activity within a population is on average positive with a wide distribution [23–25]; thus we set the distributions of excitatory and inhibitory correlation coefficients to $\mathbb{N}(0.1,0.1^2)$ and $\mathbb{N}(0.1,0.12^2)$, respectively [second and third lines of Eq. (38)]. Also, there is evidence that E and I neurons are positively correlated (i.e., the synaptic currents are negatively correlated) [26–28], so we set the average background E-I correlation [$\mathbb{N}(0.3,0.1^2)$, fourth line of Eq. (38)] to a higher value than correlations within E or I (second and third lines respectively).

In order to capture some realistic features of cortical neural networks, we impose sparse but clustered connectivity. Specifically, we have five clusters of E cells of size 10 with all-to-all connectivity and no autaptic (self-coupling) connections, and sparse random coupling within the I population (no autaptic connections) and between E and I cells (35% connection probability). See Fig. 3(b) for the sparsity structure of **G**. This is motivated by experimental evidence that E cells show clustered connectivity [29–31], and that cells tuned for specific stimulus features can be more connected, while inhibitory connections have less structure [32].
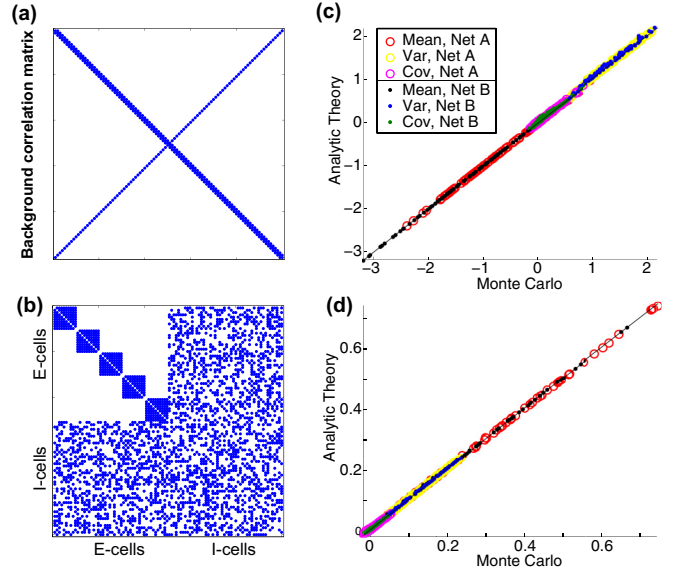

FIG. 3. A network of $N_c = 100$ neurons with heterogeneity in all parameters, but sparse background correlation and clustered and random connectivity (Network III). See Eqs. (38)–(44) for the distributions of the randomly selected parameters. Sparsity structure of the background correlation matrix **Cr** (a) and coupling matrix **G** (b). (c) Comparing all of the statistics of the activity for two realizations of the network: coupling parameters for network A are: $(g_{EE}, g_{EI}, g_{IE}, g_{II}) = (0.079, -0.24, 0.17, -0.31)$ and coupling parameters for network B are $(g_{EE}, g_{EI}, g_{IE}, g_{II}) = (0.049, -0.38, 0.16, -0.17)$. As in Fig. 2, all 100 mean and variance values are plotted, as well as all 4950 covariance values. (d) Similar to (c) but for the firing rates.

Synaptic connection strengths were chosen randomly for each realization with the following distributions:

$$g_{EE} = \mathbb{U}/10,$$
$$g_{EI} = -\frac{12}{35}\mathbb{U} - \frac{4}{35},$$
$$g_{IE} = \frac{12}{35}\mathbb{U} + \frac{4}{35}, \quad (39)$$
$$g_{II} = -\frac{12}{35}\mathbb{U} - \frac{4}{35},$$

where again $\mathbb{U} \in [0,1]$ is a uniform random variable. The value g_{EE} is used for all nonzero E to E connections: i.e., g_{jk} with $j,k \in \{1,\ldots,50\}$; g_{EI} is used for all nonzero I to E connections: i.e., g_{jk} with $j \in \{1,\ldots,50\}$ and $k \in \{51,\ldots,100\}$; g_{IE} for all nonzero E to I: i.e., g_{jk} with $j \in \{51,\ldots,100\}$ and $k \in \{1,\ldots,50\}$; similarly for g_{II}.

The distributions for the rest of the parameters were similar to Network II, with only inconsequential differences:

$$\tau_j \sim \mathbb{N}(1,0.075^2), \quad (40)$$
$$\mu_j \sim 2\mathbb{U} - 1, \quad (41)$$
$$\sigma_j \sim \mathbb{U} + 1, \quad (42)$$
$$x_{\text{rev},j} \sim \mathbb{N}(0,0.1^2), \quad (43)$$
$$x_{sp,j} \sim 0.4\mathbb{U} + 0.05. \quad (44)$$

The choices for g_{XY} and intrinsic parameters are not physiologically motivated, but rather chosen so that we can examine

how the algorithm performs on cells with a wide range of intrinsic and network parameters.

In Figs. 3(c) and 3(d) we show the results of the analytic approximation compared to Monte Carlo simulations for the activity and firing rates, respectively. In each panel, we have combined the mean, variance and covariance, and, as in Fig. 2, a data point is plotted for each cell (for means and variances) or cell pair (for covariances). Also, we show data from two instances of the network, labeled A and B; for each instance a new realization of the coupling matrix **G** and the coupling parameters [Eq. (39)] are generated (see Fig. 3 caption for values), but each of the other randomly selected parameters was kept fixed. Points that are on the black diagonal line represent a perfect match between Monte Carlo simulations and our method. As with Network II, the analytic method accurately captures the statistics cell-by-cell, despite an appreciable degree of heterogeneity.

Finally, we test how well our method approximates firing rate *correlation*, which is an important normalized measure of trial-to-trial variability (or noise correlations). The Pearson's correlation coefficient is the predominant measure in neuroscience: $\rho(\nu_j, \nu_k) = \mathrm{Cov}(\nu_j, \nu_k)/\sqrt{\mathrm{Var}(\nu_j, \nu_k)}$, i.e., the ratio of two quantities which we must estimate using the analytic method. Since this is the ratio of estimated quantities, we might expect larger errors. In Fig. 4 we show comparisons between the analytic method and Monte Carlo simulations for Network II and Network III. The method is accurate for a wide range of correlations: Fig. 4(a) shows correlations as low as -0.3 and as high as 0.3. Thus, the viability of our approximation is not limited to small correlation values, but can robustly capture the full range of correlation values observed in cortical neurons [5,33].

V. THE δ-CORRELATION ASSUMPTION

The assumptions made in deriving Eqs. (17)–(19)—each equivalent to an assumption that two random variables are δ-correlated in time—might suggest that the error of our method compared to Monte Carlo simulations would increase as τ_j increases, or perhaps that the method breaks down when the distribution of $\vec{\tau}$ has larger variance. Thus far, we have only considered relatively narrow distributions of $\vec{\tau}$. We now test this possibility in a setting where we can examine how the method performs as τ_j is increased without other confounding effects on the error.

Specifically, we simulate a network of $N_c = 50$ cells with the majority of the network parameters set to be homogenous values:

$$\mu_j = 0.7; \ \sigma_j = 1.3; \ x_{\mathrm{rev},j} = 0.1; \ x_{sp,j} = 0.35, \forall j.$$

The correlation matrix for background noise is a tridiagonal matrix with 0.3 in the upper and lower diagonal bands, and the coupling matrix is the same as in Network II: $\mathbf{G}(j,k) \sim \mathbb{N}(0, v_l)$ with $v_l = (0.1)^2$ and $v_l = (0.25)^2$. Finally, we set

$$\tau_j = (j-1)\frac{4.5}{N_c - 1} + 0.5 \qquad (45)$$

so that τ_j varies uniformly over an order of magnitude: $\tau_1 = 0.5$ and $\tau_{N_c} = 5$.

Figure 5(a) shows that the method is accurate for all possible first and second order statistics despite this large variation in
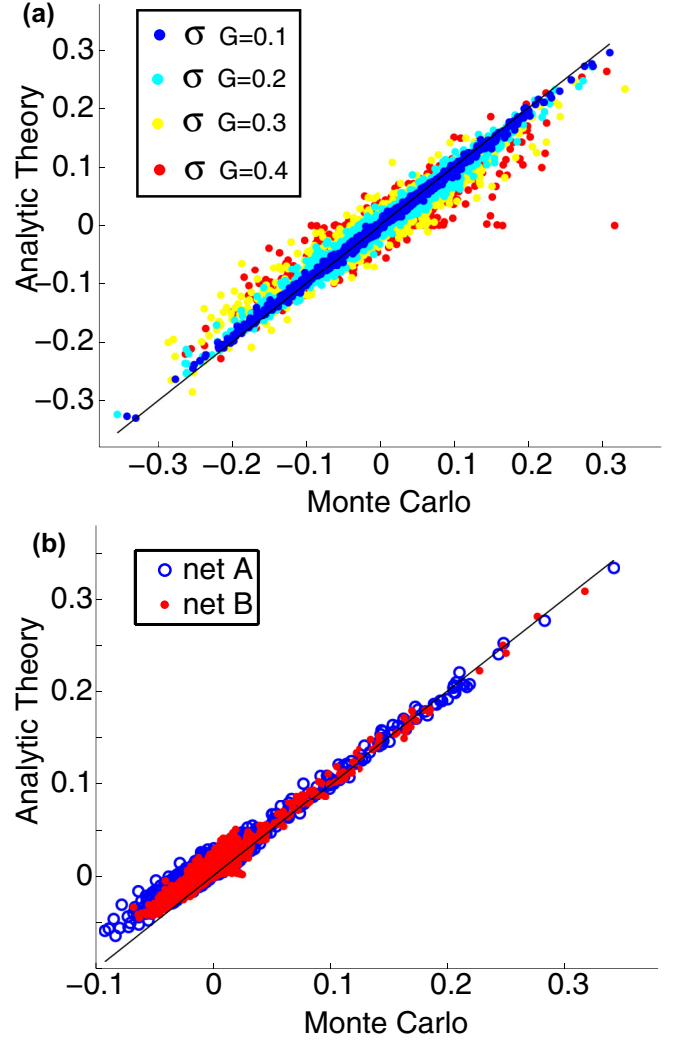

FIG. 4. Comparisons of the spike count correlation computed by our method and Monte Carlo simulations. (a) Comparing the four regimes in Network II. The results are accurate because the points predominately lie on the diagonal line. As we saw in Fig. 2, as the relative coupling strength increases, the estimation of the spike count correlation is not as accurate. (b) Comparing the two networks in Network III. In both cases, the method performs well even though *both* the numerator and denominator are estimated via the method. All $N_c(N_c - 1)/2$ firing rate correlation values are plotted for each network.

$\vec{\tau}$. Figures 5(b) and 5(c) show that the L_1-error between the Monte Carlo simulations and our method does not depend in any apparent way on the value of the time constant τ_j. Even when analyzed by a particular statistic, there is no trend with time constant. We conclude that our method is robust to large and disparate values of τ_j.

VI. ACCURACY OF THE REDUCTION METHOD

As with most calculations that assume small or large values in the parameters, an exact analytic determination of when the approximation fails is difficult, if not impossible. To capture how our method deviates as the coupling strength is increased, we performed further computations varying the
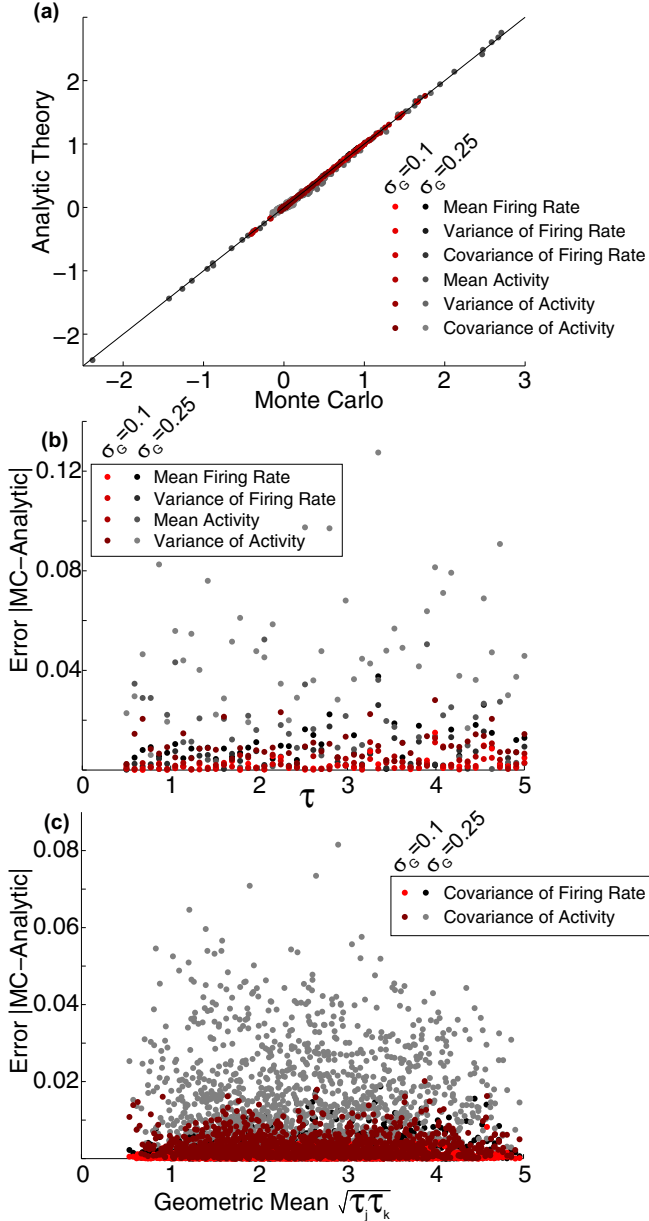
FIG. 5. A network of $N_c = 50$ neurons with heterogeneity predominately in τ_j (see main text for details). (a) Comparing our method to Monte Carlo simulations, plotting all first- and second-order statistics of the coupled network for two coupling matrices [red shade is with connection strengths $\mathbf{G} \sim \mathbb{N}(0, 0.1^2)$; black shade is with $\mathbf{G} \sim \mathbb{N}(0, 0.25^2)$ and is not as accurate]. Despite the large variation in τ, ranging in order of magnitude from 0.5 to 5, the method is accurate. (b) Plotting the L_1-error of firing rates and variances computed with our method as a function of the jth cell's τ_j value (see legend). Note that there is *no trend in error* as τ_j increases. (c) As in (b), but for all possible covariances of firing rate and activity; the horizontal axis is the geometric average of the two associated time constants: $\sqrt{\tau_j \tau_k}$. The conclusion is the same as in (b), that the τ_j values are not indicative of the error.

coupling strengths g_{jk}, system size N_c, and setting the other parameters to a variety of values.

The coupling matrix $\mathbf{G}$ was randomly chosen with half of the entries set to zero, 25% of the entries set to $g\sqrt{10}/\sqrt{N_c}$
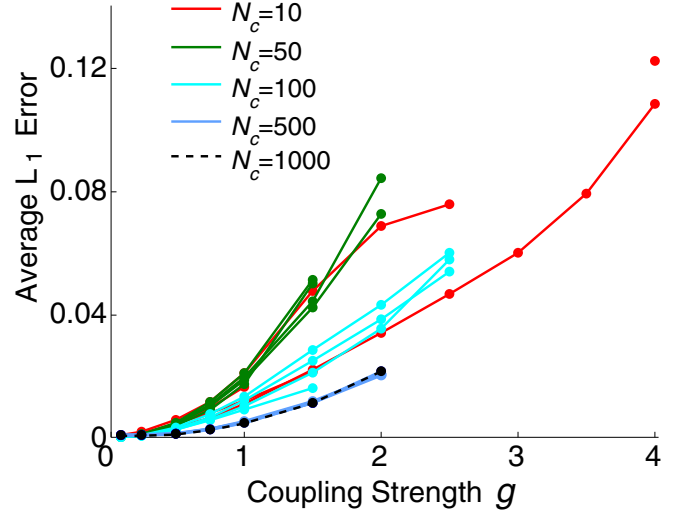

FIG. 6. Accuracy of the method as coupling strength increases. The L_1-error (averaged over the entire set of six spiking and activity statistics) of our method compared to Monte Carlo simulations increases nonlinearly as coupling strength increases (see main text for definition of g and description of networks), for a variety of network sizes. As network size N_c increases, the coupling strengths are scaled by $1/\sqrt{N_c}$.

and 25% set to $-g\sqrt{10}/\sqrt{N_c}$, where g is a scale parameter representing the magnitude of the coupling (note that $\mathbf{G}$ was chosen only once for each given network size N_c, i.e., it was held fixed while other random parameters were varied). The correlation matrix of background noise is a banded matrix with between one and four bands above and below the diagonal set to $c = 0.3$; that is, each cell shared noise with its k-nearest neighbors, for $k = 1, 2, 3$ or 4. The rest of the parameters were chosen randomly as for Network III [Eq. (41)–(44)].

Figure 6 shows these results; for each network size N_c and shared common noise footprint k, the error of our method compared to Monte Carlo simulations is plotted as a function of the magnitude of the coupling strength g. Each point represents the L_1 error between a Monte Carlo and our method, averaged over the entire set of six spiking and activity statistics (mean, variances, and covariances of both activity x_j and firing rate ν_j). For each system size N_c (except $N_c = 1000$; see explanation below), four curves show results for $k = 1, 2, 3$ and 4, respectively. We see that as the magnitude of coupling values increase, any given error curve tends to increase.

In performing computations for Fig. 6, we made the following modifications for computational tractability: for larger N_c values ($N_c = 500, 1000$), we augmented our method to use a subset of correlation and covariance values; we chose the main diagonal and the super- or subdiagonal $\mathrm{Cov}(j, j+1)$; we further computed only one instance for $N_c = 1000$ (i.e., there is only one curve). We further note that some entries are not plotted because our method did not converge to a solution and/or the resulting covariance matrices are not positive definite, which is not unexpected with randomly chosen parameters.

In summary, Fig. 6 provides a representative snapshot of the range of possible error values. While the average

error increases nonlinearly as coupling strength increases, overall the error appears to be relatively insensitive to system size. We note that the scaling of coupling strengths by the square root of system size $1/\sqrt{N}$ is considered to be "strong scaling" popularized by the theory of balanced networks [8,34], compared to the relatively weak scaling $1/N$.

VII. DISCUSSION

There has been a long history of analytic reduction methods for neural network models, both to enhance efficiency in simulation and to aid mathematical analyses. Here we summarize some of this literature and its relationship to the work presented here.

The simplest approach is a *mean-field* analysis, which would self-consistently estimate the mean values, assuming the variances $\sigma^2(j)$ to be 0 [35]. However, this neglects the fluctuations which we know to be important in neural systems; therefore many authors have augmented these theories with corrections to capture second-order statistics, higher-order statistics, or time-dependent correlations. Several authors have proposed to derive these corrections by starting with the microscopic dynamics of single neurons in the network. The microscopic dynamics in question may be given by a master equation [12,14,36–38], a generalized linear model [39,40], or the theta model [41,42]. The result is a principled theory for the second-order statistics of the network; however, the resulting calculations are often complicated and hard to execute.

Here we aimed to take a middle road between simple (but inaccurate) mean-field calculations, and principled (but complicated) theories to compute network fluctuations from microscopic dynamics. Specifically, we start with a system of coupled stochastic differential equations, each of which may represent either a single neuron or a homogeneous population, and sought to quickly and accurately estimate statistics of the coupled system. Importantly, the unperturbed state in our system is not one in which all neurons are independent; instead we perturb from a state with background noise correlations. Thus, we anticipate this approximation can be used to probe a range of neural networks, in which correlations can be significant and activity-modulated.

While the coupled firing rate models we study here were not derived directly from the microscopic dynamics of a spiking network, our results can still yield insight into spiking networks [43–45]. Our previous work [18] used the qualitative principles and intuitions gained from a simple firing rate model to characterize relationships between the analogous parameters in a full spiking model of a multi-region olfactory network. In that paper, a small system with simple coupling and background correlations was studied, whereas this paper treats networks of arbitrary size, and arbitrary coupling and input correlation structures. The work here is thus a generalization of the calculations in Ref. [18].

In other models, $F(\cdot)$ represents the function that maps firing rate to synaptic input. Here we assume that the effective synaptic input $g_{jk}F_k$ is a fixed scaling of the firing rate F_k. In other biophysical models the effective synaptic input may be a more complex transformation of the firing rate (e.g., an alpha function convolved with firing rate): the methods presented here can easily be altered to account for this. To do this, the only change would be to use $S_k(F_k)$ in Eq. (1) instead of F_k, where $S_k(\cdot)$ is some synaptic activation function.

Our method relies on the assumption that statistics are stationary in time; this assumption allows a set of statistics to be solved self-consistently. Thus we have not addressed complex network dynamics, such as oscillations or time-varying statistics. However, this limitation is not specific to our method, but also applies to related work. Previously developed approximation methods may fail when the system undergoes a bifurcation [36,37], and truncation methods (or moment closure methods) are known to fail in certain parameter regimes [46]. When the set of self-consistent equations cannot be solved, there may be other methods available to characterize the oscillatory dynamics (see Ref. [47] where this is done for the adaptive quadratic integrate-and-fire model). Likewise, we did not consider time-lagged network statistics (i.e., the entire cross-correlation functions) but rather only the instantaneous statistics. This perhaps enables the delta-correlation assumption in our method to give accurate approximations even with disparate time scales (see Sec. V). Such considerations are a fruitful path of future work.

ACKNOWLEDGMENT

APPENDIX A: DERIVATION OF EQUATIONS (17)–(19)

Here, we provide a formal derivation of the assumptions that we make to derive our main result [Eqs. (20)–(22)]. Our challenge is that while we are principally interested in zero-time-lag statistics, computing second-order statistics (such as $\langle F_j[x_j(t)]F_k[x_k(t)]\rangle$) using Eq. (16) requires us to know the autocorrelation function, i.e., $\langle F_j[x_j(t)]F_k[x_k(t+\tau)]\rangle$. Therefore we need to close the equation by making some kind of assumption about these temporal correlations.

The key assumption to justify Eq. (17) is that $F_k[x_k(t)]$ is δ-correlated in time:

$$\langle F_k[x_k(u)]F_k[x_k(v)]\rangle = \delta(u-v)\big[\langle F_k(x_k)^2\rangle - \nu_k^2\big] + \nu_k^2, \tag{A1}$$

where $\nu_k = \langle F_k(x_k)\rangle$ is the mean. Using this in the integral on the left-hand side of Eq. (17), we find that

$$\int_{-\infty}^{t} du \int_{-\infty}^{t} dv \, e^{-(t-u)/\tau_l} e^{-(t-v)/\tau_m} \langle F[x_k(u)]F[x_k(v)]\rangle$$

$$\approx \int_{-\infty}^{t} du \int_{-\infty}^{t} dv \, e^{-(t-u)/\tau_l} e^{-(t-v)/\tau_m} \delta(u-v)\big\{\langle F[x_k(u)]^2\rangle - \nu_k^2\big\} + \int_{-\infty}^{t} du \int_{-\infty}^{t} dv \, e^{-(t-u)/\tau_l} e^{-(t-v)/\tau_m} \nu_k^2$$

$$= \int_{-\infty}^{t} du\, e^{-(t-u)/\tau_l} e^{-(t-u)/\tau_m} \left(\langle F[x_k(u)]^2 \rangle - \nu_k^2 \right) + \tau_l \tau_m \nu_k^2$$

$$= \frac{\tau_l \tau_m}{\tau_l + \tau_m} \left\{ \langle F[x_k(u)]^2 \rangle - \nu_k^2 \right\} + \tau_l \tau_m \nu_k^2$$

$$= \frac{\tau_l \tau_m}{\tau_l + \tau_m} \{\mathcal{E}_2(k) - [\mathcal{E}_1(k)]^2\} + \tau_l \tau_m [\mathcal{E}_1(k)]^2 = \frac{\tau_l \tau_m}{\tau_l + \tau_m} \mathcal{V}_k + \tau_l \tau_m [\mathcal{E}_1(k)]^2.$$

The other two approximations [Eqs. (18) and (19)] are arrived at by essentially the same calculation: for completeness, we provide them here as well.

To derive Eq. (18), assume that

$$\langle \sigma_j \eta_j(u) F_k[x_k(v)] \rangle = \delta(u-v)\{\langle \sigma_j \eta_j(u) F_k[x_k(u)] \rangle - \langle \sigma_j \eta_j \rangle \langle F_k(x_k) \rangle\} + \langle \sigma_j \eta_j \rangle \langle F_k(x_k) \rangle \tag{A2}$$

$$= \delta(u-v) \langle \sigma_j \eta_j(u) F_k[x_k(u)] \rangle, \tag{A3}$$

where the last line is because $\langle \eta_j \rangle = 0$ and σ_j is a constant. Then

$$\left\langle \int_{-\infty}^{t} \sigma_j \eta_j(u) e^{-(t-u)/\tau_l}\, du \int_{-\infty}^{t} F_k[x_k(v)] e^{-(t-v)/\tau_m}\, dv \right\rangle = \int_{-\infty}^{t} du \int_{-\infty}^{t} dv\, e^{-(t-u)/\tau_l} e^{-(t-v)/\tau_m} \langle \sigma_j \eta_j(u) F_k(x_k(v)) \rangle$$

$$\approx \int_{-\infty}^{t} du \int_{-\infty}^{t} dv\, e^{-(t-u)/\tau_l} e^{-(t-v)/\tau_m} \delta(u-v) \langle \sigma_j \eta_j(u) F_k(x_k(v)) \rangle$$

$$= \int_{-\infty}^{t} du\, e^{-(t-u)/\tau_l} e^{-(t-u)/\tau_m} \langle \sigma_j \eta_j(u) F_k[x_k(u)] \rangle$$

$$= \frac{\tau_l \tau_m}{\tau_l + \tau_m} \sigma_j \mathcal{N}_F(j,k).$$

To derive Eq. (19), assume that

$$\langle F_j[x_j(u)] F_k[x_k(v)] \rangle = \delta(u-v)\{\langle F_j[x_j(u)] F_k[x_k(u)] \rangle - \nu_j \nu_k\} + \nu_j \nu_k, \tag{A4}$$

where (as before) $\nu_k = \langle F_k(x_k) \rangle$. Using this in the integral on the left-hand side of Eq. (19), we find that

$$\left\langle \int_{-\infty}^{t} F_j[x_j(u)] e^{-(t-u)/\tau_l}\, du \int_{-\infty}^{t} F_k[x_k(v)] e^{-(t-v)/\tau_m}\, dv \right\rangle$$

$$= \int_{-\infty}^{t} du \int_{-\infty}^{t} dv\, e^{-(t-u)/\tau_l} e^{-(t-v)/\tau_m} \langle F_j[x_j(u)] F_k[x_k(v)] \rangle$$

$$\approx \int_{-\infty}^{t} du \int_{-\infty}^{t} dv\, e^{-(t-u)/\tau_l} e^{-(t-v)/\tau_m} \delta(u-v)\{\langle F_j[x_j(u)] F_k[x_k(u)] \rangle - \nu_j \nu_k\} + \int_{-\infty}^{t} du \int_{-\infty}^{t} dv\, e^{-(t-u)/\tau_l} e^{-(t-v)/\tau_m} \nu_j \nu_k$$

$$= \int_{-\infty}^{t} du\, e^{-(t-u)/\tau_l} e^{-(t-u)/\tau_m} \{\langle F_j[x_j(u)] F_k[x_k(u)] \rangle - \nu_j \nu_k\} + \nu_j \nu_k \tau_l \tau_m$$

$$= \frac{\tau_l \tau_m}{\tau_l + \tau_m} \{\langle F_j[x_j(u)] F_k[x_k(u)] \rangle - \nu_j \nu_k\} + \tau_l \tau_m \nu_j \nu_k$$

$$= \frac{\tau_l \tau_m}{\tau_l + \tau_m} [\mathcal{S}(j,k) - \mathcal{E}_1(j)\mathcal{E}_1(k)] + \tau_l \tau_m \mathcal{E}_1(j)\mathcal{E}_1(k)$$

$$= \frac{\tau_l \tau_m}{\tau_l + \tau_m} \mathcal{C}_{\mathcal{V}}(j,k) + \tau_l \tau_m \mathcal{E}_1(j)\mathcal{E}_1(k).$$

APPENDIX B: ALTERNATIVE REDUCTION APPROACHES

A common method to approximate high-dimensional systems is "moment closure" methods where state variables are integrated or averaged out, and assumptions are made on various moments of the random or heterogeneous entities. Such approaches have a long history in the physical sciences [48,49] and recently in the life sciences [37,46,50]. Here we provide an alternative approach based on the probability density (or Fokker-Planck) equation of the stochastic neural network, rather than the stochastic integrals we considered

in the main text. These methods are partially related to the stochastic integral method presented earlier, but we will show that they are *different*. In a similar vein, we previously showed that the analysis of the stochastic integral is more insightful than the Fokker-Planck equation for a system of coupled noisy oscillators (compare main results and the Appendix of Ref. [51]).

The corresponding probability density function $p(\vec{x},t)$, defined by $p(\vec{x},t)\,d\vec{x} = P(X(t) \in (\vec{x}, \vec{x} + d\vec{x}))$, of the network models considered in Eq. (1) satisfies the following

Fokker-Planck equation [22,52]:

$$
\frac{\partial p(\vec{x},t)}{\partial t} = -\sum_{l=1}^{N_c} \frac{\partial}{\partial x_l} \left\{ \frac{1}{\tau_l} \left[-x_l + \mu_l + \sum_{k=1}^{N_c} g_{lk} F_k(x_k) \right] p(\vec{x},t) \right\}
$$
$$
+ \frac{1}{2} \sum_{j,k} D_{j,k} \frac{\partial^2 p(\vec{x},t)}{\partial x_j \partial x_k}. \quad \text{(B1)}
$$

where $D_{j,k} = c_{jk} \frac{\sigma_j \sigma_k}{\tau_j \tau_k}$ and the second sum is taken over all $N_c \times N_c$ pairs of (j,k). This high-dimensional partial differential equation contains all of the statistics about $\vec{X}$ and any desired transformations. We are interested in the steady-state equation $\frac{\partial p(\vec{x},t)}{\partial t} = 0$, assuming the statistics are in equilibrium. It is convenient to write the function in the curly brackets as a probability flux or current, as follows:

$$
J_l(\vec{x},t) := \frac{1}{\tau_l} \left[-x_l + \mu_l + \sum_{k=1}^{N_c} g_{lk} F_k(x_k) \right] p(\vec{x},t). \quad \text{(B2)}
$$

The steady-state equation is

$$
0 = -\sum_{l=1}^{N_c} \frac{\partial}{\partial x_l} \left\{ \frac{1}{\tau_l} \left[-x_l + \mu_l + \sum_{k=1}^{N_c} g_{lk} F_k(x_k) \right] p(\vec{x}) \right\}
$$
$$
+ \frac{1}{2} \sum_{j,k} D_{j,k} \frac{\partial^2 p(\vec{x})}{\partial x_j \partial x_k}, \quad \text{(B3)}
$$

$$
0 = -\sum_{l=1}^{N_c} \frac{\partial}{\partial x_l} J_l(\vec{x}) + \frac{1}{2} \sum_{j,k} D_{j,k} \frac{\partial^2 p(\vec{x})}{\partial x_j \partial x_k}. \quad \text{(B4)}
$$

1. Moment closure methods

We want to reduce this high-dimensional system into one that is solvable. Without coupling $g_{jk} = 0$, the solution is simply a multivariate Gaussian distribution with mean $\vec{\mu}$ and covariance matrix $\text{Cov}(j,k) = \frac{c_{jk}}{\tau_j + \tau_k} \sigma_j \sigma_k$. This motivates a closure of the system where we assume $\vec{X}$ is determined by its first two moments and is approximated by a Gaussian: $X_j = \sigma(j) + Y\mu(j)$, where Y is a standard normal random variable. We also assume the joint marginal distributions are bivariate Gaussians:

$$
\mathcal{P}(x_j,x_k) := \int p(\vec{x}) \, d\vec{x}_{\setminus j,k} \sim \mathbb{N}_2, \quad \text{(B5)}
$$

where $\mathbb{N}_2$ is the following bivariate Gaussian distribution:
$\mathbb{N}\left(\binom{\mu(j)}{\mu(k)}, \binom{\frac{\sigma(j)^2}{2\tau_j} \quad \sigma(j)\sigma(k)\frac{c_{jk}}{\tau_j+\tau_k}}{\sigma(j)\sigma(k)\frac{c_{jk}}{\tau_j+\tau_k} \quad \frac{\sigma(k)^2}{2\tau_k}} \right)$, and $d\vec{x}_{\setminus j,k}$ denotes integrating over all N_c variables except x_j and x_k. Note that these assumptions are also made in the main text.

We multiply Eq. (B4) by x_j and integrate the equation over all N_c variables, $d\vec{x} = dx_j \, d\tilde{x}$ (where $d\tilde{x} = dx_1 \ldots dx_{j-1} dx_{j+1} \ldots dx_{N_c} := d\vec{x}_{\setminus j}$):

$$
0 = -\int \sum_{l=1}^{N_c} \frac{\partial}{\partial x_l} J_l(\vec{x}) x_j \, dx_j \, d\tilde{x}
$$
$$
+ \frac{1}{2} \int \sum_{l_1,l_2} D_{l_1,l_2} \frac{\partial^2 p(\vec{x})}{\partial x_{l_1} \partial x_{l_2}} x_j \, dx_j \, d\tilde{x}. \quad \text{(B6)}
$$

Consider the first term: when $l \neq j$, we have

$$
\int \frac{\partial}{\partial x_l} J_l(\vec{x}) x_j \, dx_j \, d\tilde{x} = \int \frac{\partial}{\partial x_l} J_l(\vec{x}) dx_l \, x_j \, dx_j \, d\vec{x}_{\setminus l,j}
$$
$$
= \int J_l|_{x_l=-\infty}^{x_l=\infty} x_j \, dx_j \, d\vec{x}_{\setminus j}
$$
$$
= \int 0 \, x_j \, dx_j \, d\vec{x}_{\setminus j} = 0. \quad \text{(B7)}
$$

The last equality comes from no flux at $\pm\infty$: $J_l|_{x_l=-\infty}^{x_l=\infty} = 0$. A similar calculation applies to the second term, for all $N_c \times N_c$ values of (l_1,l_2), it is 0. When $l_1 \neq j$ and $l_2 \neq j$, integrate in x_{l_1} and x_{l_2} first and use the fact that there is no density at $\pm\infty$: $p(\vec{x})|_{x_{l_{1/2}}=-\infty}^{x_{l_{1/2}}=\infty} = 0$; when $l_{1/2} = j$, integrate in x_j first then integrate by parts, using $\partial_j p(\vec{x}) x_j|_{x_{l_{1/2}}=-\infty}^{x_{l_{1/2}}=\infty} = 0$ and $\partial_j p(\vec{x})|_{x_{l_{1/2}}=-\infty}^{x_{l_{1/2}}=\infty} = 0$. Therefore, Eq. (B6) is

$$
0 = -\int \frac{\partial}{\partial x_j} J_j(\vec{x}) x_j \, dx_j \, d\tilde{x},
$$
$$
0 = -\int J_j(\vec{x}) x_j|_{x_j=-\infty}^{x_j=\infty} \, d\tilde{x} + \int J_j(\vec{x}) \, d\vec{x},
$$
$$
0 = -0 + \frac{1}{\tau_j} \left[-\mu(j) + \mu_j + \sum_{k=1}^{N_c} g_{jk} \mathcal{E}_1(k) \right], \quad \text{(B8)}
$$

where $\mu(j) := \int x_j p(\vec{x}) \, d\vec{x}$, and we have used the approximation $\int F_k(x_k) p(\vec{x}) \, d\vec{x} \approx \mathcal{E}_1(k)$ [see Table I for definition of $\mathcal{E}_1(k)$] by assuming the marginal x_k PDF is a normal distribution with mean $\mu(k)$ and variance $\sigma^2(k)$. Rearranging Eq. (B8) gives the exact same nonlinear equation for the mean $\mu(j)$, but coupled with the variance via $\mathcal{E}_1(k)$:

$$
\mu(j) = \mu_j + \sum_{k=1}^{N_c} g_{jk} \mathcal{E}_1(k). \quad \text{(B9)}
$$

To derive a similar equation for the variance $\sigma^2(j)$, we multiply Eq. (B4) by x_j^2 and again integrate over all variables:

$$
0 = -\int \sum_{l=1}^{N_c} \frac{\partial}{\partial x_l} J_l(\vec{x}) x_j^2 \, dx_j \, d\tilde{x}
$$
$$
+ \frac{1}{2} \int \sum_{l_1,l_2} D_{l_1,l_2} \frac{\partial^2 p(\vec{x})}{\partial x_{l_1} \partial x_{l_2}} x_j^2 \, dx_j \, d\tilde{x}. \quad \text{(B10)}
$$

First, consider the diffusion (second) term: similar to before, if either $l_1 \neq j$ or $l_2 \neq j$, the term will vanish [integrate in x_{l_1} and x_{l_2} first and use the fact that there is no density at $\pm\infty$: $p(\vec{x})|_{x_{l_{1/2}}=-\infty}^{x_{l_{1/2}}=\infty} = 0$]. However, when $l_1 = l_2 = j$, integrate x_j first and use integration by parts twice to get

$$
\frac{D_{j,j}}{2} \int \frac{\partial^2 p(\vec{x})}{\partial x_j^2} x_j^2 \, dx_j = D_{j,j} \int p(\vec{x}) \, dx_j;
$$

taking into account the other $N_c - 1$ integration variables and that $\int p(\vec{x}) \, d\vec{x} = 1$ simply gives $D_{j,j}$ for the second term.

Now for the first term in Eq. (B10): if $l \neq j$, then we can see that that term in the sum vanishes by integrating x_l first and

using $J_l|_{x_l=-\infty}^{x_l=\infty} = 0$. Using integration by parts for the $l = j$ term, we get

$$-\int \frac{\partial}{\partial x_j} J_j(\vec{x}) x_j^2 \, dx_j \, d\tilde{x} = 0 + 2\int x_j J_j(\vec{x}) \, dx_j \, d\tilde{x}$$

$$= \frac{2}{\tau_j} \int \left[-x_j^2 + \mu_j x_j + \sum_{k=1}^{N_c} g_{jk} x_j F_k(x_k) \right] p(\vec{x}) \, dx_j \, d\tilde{x}.$$

Using the fact that $\int x_j^2 p(\vec{x}) \, d\vec{x} = \sigma^2(j) + \mu(j)^2$, the entire Eq. (B10) is

$$D_{j,j} = \frac{2}{\tau_j} \left[\sigma^2(j) + \mu(j)^2 - \mu_j \mu(j) \right.$$
$$\left. - \sum_{k=1}^{N_c} g_{jk} \int x_j F_k(x_k) p(\vec{x}) \, d\vec{x} \right]. \quad \text{(B11)}$$

This equation is exact thus far. We now employ our approximation: $X_j = \mu(j) + Y_1\sigma(j)$ where Y_1 is a standard normal random variable (similarly for X_k). The last term in the previous equation is

$$\mu(j) \sum_{k=1}^{N_c} g_{jk} \int F_k(x_k) p(\vec{x}) \, d\vec{x}$$

$$+\sigma(j) \sum_{k=1}^{N_c} g_{jk} \int y_1 F_k[\mu(k) + y_2\sigma(k)] p(\vec{x}) \, d\vec{x}.$$

We can approximate the first term above with Eq. (B9) to get [excluding $\mu(j)$]:

$$\sum_{k=1}^{N_c} g_{jk} \int F_k(x_k) p(\vec{x}) \, d\vec{x} \approx \sum_{k=1}^{N_c} g_{jk} \mathcal{E}_1(k) \approx \mu(j) - \mu_j.$$

Thus, this leads to a cancellation of the terms $\mu(j)^2$ and $\mu_j \mu(j)$ in Eq. (B11). We approximate the term $\int y_1 F_k[\mu(k) + y_2\sigma(k)] p(\vec{x}) \, d\vec{x}$ by assuming the joint marginal distribution of (X_j, X_k) are bivariate normal, and use the definition of $\mathcal{N}_\mathcal{F}$ in Table I to get

$$\int y_1 F_k[\mu(k) + y_2\sigma(k)] p(\vec{x}) \, d\vec{x} \approx \sqrt{2}\mathcal{N}_\mathcal{F}(j,k).$$

Therefore, the equation for the variance is

$$\sigma^2(j)\tau_j = \frac{\sigma_j^2}{2} + \sigma(j)\sqrt{2}\tau_j \sum_{k=1}^{N_c} g_{jk}\mathcal{N}_\mathcal{F}(j,k). \quad \text{(B12)}$$

This equation is similar to Eq. (21) but lacking higher order terms in F_k, as well as other differences.

To derive the analogous equation for the Cov(j,k), the procedure is almost exactly the same except Eq. (B4) is multiplied by $x_j x_k$, and there are two terms from the sum (over probability fluxes J_l) that contribute, when $l = j$ and

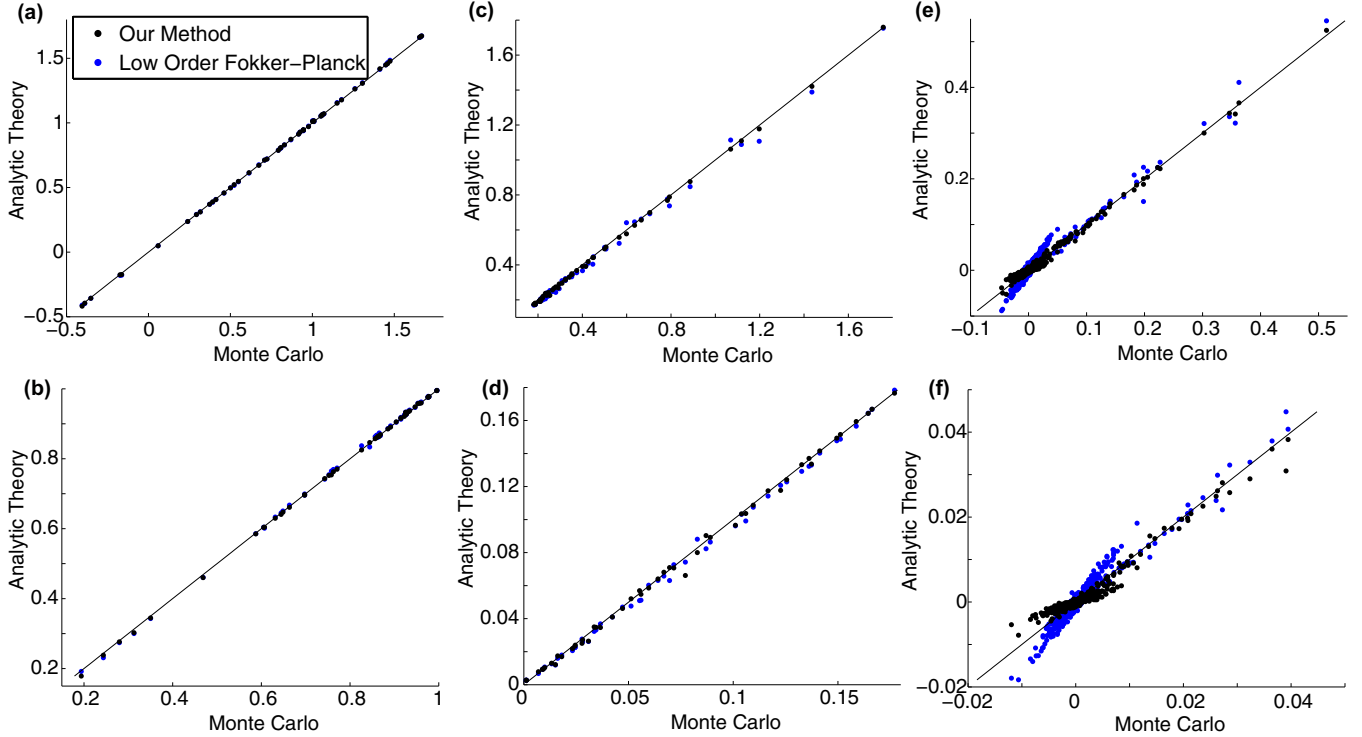**(a)** Our Method (black), Low Order Fokker–Planck (blue)

FIG. 7. Comparing our method with the lowest order approximation of the Fokker-Planck equation [Eqs. (B9), (B12), and (B13)]. The network configuration is described in Sec. V and is the same realization in Fig. 5. There are $N_c = 50$ neurons with most parameters fixed except the time scale $\vec{\tau}$ and coupling matrix **G**. (a) Comparison of the mean activity $\mu(j)$ calculated via Monte Carlo simulations (horizontal axis) and our reduction method (black) and the Fokker-Planck approximation (blue), showing all 50 values for each. (b) Similar to (a) but for mean firing rate ν_j. (c) Variance of activity $\sigma^2(j)$. (d) Variance of firing rate Var(ν_j). (e) Covariance of activity Cov(j,k), showing all $50 \times 49/2 = 1225$ values for each. (f) Covariance of the firing rate Cov(ν_j, ν_k). Our result is more accurate, especially for the second order statistics (c)–(f).

$l = k$. The result is

$$\text{Cov}(j,k)\frac{\tau_j + \tau_k}{2} = c_{jk}\frac{\sigma_j\sigma_k}{2} + \frac{\sigma(j)}{2}\sqrt{2}\tau_j \sum_{l=1}^{N_c} g_{kl}\mathcal{N}_{\mathcal{F}}(j,l)$$

$$+ \frac{\sigma(k)}{2}\sqrt{2}\tau_k \sum_{l=1}^{N_c} g_{jl}\mathcal{N}_{\mathcal{F}}(k,l). \quad \text{(B13)}$$

Again, this equation is similar to parts of Eq. (22). When $j = k$ in Eq. (B13), we recover Eq. (B12).

Together, Eqs. (B9), (B12), and (B13) form a system of transcendental equations for the complete set of first- and second-order statistics. This can be thought of as a *lowest order approximation* to the exact statistics of the coupled system. We implemented this method on the same network described in Sec. V and found that it is not as accurate as our method (Fig. 7, black dots are closer to the diagonal line than blue dots). The mean firing rates and activity perform equally well with both methods [Figs. 7(a) and 7(b)], but our method outperforms this method in calculating the variances [Figs. 7(c) and 7(d)] and even more so with the covariances [Figs. 7(e) and 7(f)].

2. Higher order moment closure methods

To derive a higher order moment closure method, one can continue the procedure described in the previous section, by multiplying $x_i x_j x_k$ with Eq. (B4) and devise a method to close the lower order equations. Equation (B9) remains the same because of the underlying normal distribution assumption:

$$\mu(j) = \mu_j + \sum_{k=1}^{N_c} g_{jk}\mathcal{E}_1(k).$$

The second set of equations (obtained by multiplying by $x_j x_k$ and integrating in $\vec{x}$) can possibly be used to better approximate higher order equations, rather than close it as was done in the *lowest order approximation*.

If one were to follow the outline of this method, the Gaussian assumptions on $\vec{x}$ reduce the higher moments $\int d\vec{x} x_i x_j x_k p(\vec{x})$ in terms of the mean and (co-)variances of $\vec{x}$, resulting in an overconstrained or redundant system. A possible way to proceed is to devise an approximation to $\int x_j F_k(x_k) p(\vec{x}) d\vec{x}$, possibly relating to $\int x_j^2 F_k(x_k) p(\vec{x}) d\vec{x}$, where an assumption beyond the Gaussian approximation of $\vec{x}$ could perhaps be used. Whether or not there exists a higher order moment closure method on the Fokker-Planck equation (B4) that is more accurate than our new method is beyond the scope of this study and an interesting area for future research.

What is very clear through all these calculations is that our method described in the main text is different than any common moment closure methods on the Fokker-Planck equation, despite some similarities in the equations. From the moment closure methods we have outlined here, we see that the resulting equations will never have second order terms in the network coupling [i.e., $F_k(x_k)^2$ or $F_j F_k$], and is thus a different approach than our method.

[1] M. B. Ahrens, M. B. Orger, D. N. Robson, J. M. Li, and P. J. Keller, Nat. Methods **10**, 413 (2013).

[2] R. Prevedel, Y.-G. Yoon, M. Hoffmann, N. Pak, G. Wetzstein, S. Kato, T. Schrödel, R. Raskar, M. Zimmer, E. S. Boyden *et al.*, Nat. Methods **11**, 727 (2014).

[3] E. R. Kandel, H. Markram, P. M. Matthews, R. Yuste, and C. Koch, Nat. Rev. Neurosci. **14**, 659 (2013).

[4] W. C. Lemon, S. R. Pulver, B. Höckendorf, K. McDole, K. Branson, J. Freeman, and P. J. Keller, Nat. Commun. **6**, 7924 (2015).

[5] M. Cohen and A. Kohn, Nat. Neurosci. **14**, 811 (2011).

[6] L. F. Abbott and C. van Vreeswijk, Phys. Rev. E **48**, 1483 (1993).

[7] C. Van Vreeswijk, L. Abbott, and G. B. Ermentrout, J. Comput. Neurosci. **1**, 313 (1994).

[8] C. van Vreeswijk and H. Sompolinsky, Science **274**, 1724 (1996).

[9] N. Brunel, J. Comput. Neurosci. **8**, 183 (2000).

[10] N. Brunel and V. Hakim, Neural Comput. **11**, 1621 (1999).

[11] M. A. Buice and C. C. Chow, Phys. Rev. E **76**, 031118 (2007).

[12] P. C. Bressloff, SIAM J. Appl. Math. **70**, 1488 (2009).

[13] A. Renart, J. de la Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. Harris, Science **327**, 587 (2010).

[14] J. Touboul and B. Ermentrout, J. Comput. Neurosci. **31**, 453 (2011).

[15] C. Stringer, M. Pachitariu, N. A. Steinmetz, M. Okun, P. Bartho, K. D. Harris, M. Sahani, and N. A. Lesica, eLife **5**, e19695 (2016).

[16] W. Gerstner and W. Kistler, *Spiking Neuron Models* (Cambridge University Press, Cambridge, 2002), pp. 147–163.

[17] H. R. Wilson and J. D. Cowan, Biophys. J. **12**, 1 (1972).

[18] A. K. Barreiro, S. H. Gautam, W. L. Shew, and C. Ly, arXiv:1702.03543 [q-bio.NC].

[19] http://github.com/chengly70/FiringRateModReduction.

[20] P. Dayan and L. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of neural Systems* (Taylor & Francis, Philadelphia, 2001).

[21] We assume the networks of interest satisfy the conditions of the ergodic theorems so that averaging over time and state are the same.

[22] C. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, New York, 1985).

[23] J. Poulet and C. Petersen, Nature (London) **454**, 881 (2008).

[24] J. Yu and D. Ferster, Neuron **68**, 1187 (2010).

[25] L. J. Gentet, M. Avermann, F. Matyas, J. F. Staiger, and C. C. Petersen, Neuron **65**, 422 (2010).

[26] L. J. Borg-Graham, C. Monier, and Y. Fregnac, Nature (London) **393**, 369 (1998).

[27] L. Borg-Graham, C. Monier, and Y. Fregnac, J. Physiol. (Paris) **90**, 185 (1996).

[28] M. Okun and I. Lampl, Nat. Neurosci. **11**, 535 (2008).

[29] S. Song, P. Sjöström, M. Reigl, S. Nelson, and D. Chklovskii, PLoS Biol. **3**, e68 (2005).

[30] R. Perin, T. K. Berger, and H. Markram, Proc. Natl. Acad. Sci. USA **108**, 5419 (2011).

[31] H. Ko, S. B. Hofer, B. Pichler, K. A. Buchanan, P. J. Sjöström, and T. D. Mrsic-Flogel, Nature (London) **473**, 87 (2011).

[32] E. Fino and R. Yuste, Neuron **69**, 1188 (2011).

[33] B. Doiron, A. Litwin-Kumar, R. Rosenbaum, G. Ocker, and K. Josić, Nat. Neurosci. **19**, 383 (2016).

[34] C. van Vreeswijk and H. Sompolinsky, Neural Comput. **10**, 1321 (1998).

[35] H. C. Tuckwell, *Introduction to Theoretic Neurobiology: Volume 2, Nonlinear and Stochastic Theories* (Cambridge University Press, New York, 2005), Vol. 8.

[36] M. A. Buice and J. D. Cowan, Phys. Rev. E **75**, 051919 (2007).

[37] M. Buice, J. Cowan, and C. Chow, Neural Comput. **22**, 377 (2010).

[38] P. C. Bressloff, J. Math. Neurosci. **5** (2015).

[39] T. Toyoizumi, K. R. Rad, and L. Paninski, Neural Comput. **21**, 1203 (2009).

[40] G. K. Ocker, K. Josić, E. Shea-Brown, and M. A. Buice, PLoS Comput. Biol. **13**, e1005583 (2017).

[41] M. Buice and C. Chow, J. Stat. Mech.: Theory Exp. (2013) P03003.

[42] M. Buice and C. Chow, PLoS Comput. Biol. **9**, e1002872 (2013).

[43] R. Rosenbaum, M. Smith, A. Kohn, J. Rubin, and B. Doiron, Nature Neurosci. **20**, 107 (2017).

[44] T. Kanashiro, G. K. Ocker, M. R. Cohen, and B. Doiron, eLife **6**, e23978 (2017).

[45] E. Ledoux and N. Brunel, Front. Comput. Neurosci. **5**, 25 (2011).

[46] C. Ly and D. Tranchina, Neural Comput. **19**, 2032 (2007).

[47] W. Nicola, C. Ly, and S. A. Campbell, SIAM J. Appl. Math. **75**, 2333 (2015).

[48] S. I. Chapman and T. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, New York, 1970).

[49] W. Dreyer, M. Junk, and M. Kunik, Nonlinearity **14**, 881 (2001).

[50] G. S. Williams, M. A. Huertas, E. A. Sobie, M. S. Jafri, and G. D. Smith, Biophys. J. **95**, 1689 (2008).

[51] C. Ly and G. B. Ermentrout, Phys. Rev. E **81**, 011911 (2010).

[52] H. Risken, *The Fokker-Planck Equation: Methods of Solutions and Applications* (Springer-Verlag, New York, 1989).