Diagrammatic expansion of pulse-coupled network dynamics in terms of subnetworks

Aaditya V. Rangan
Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, New York 10012, USA
(Received 29 October 2008; revised manuscript received 11 May 2009; published 1 September 2009)

We introduce a framework wherein various measurements of a pulse-coupled network's stationary dynamics can be expanded in terms of the network's connectivity. Such measurements include the occurrence rate of pulses (e.g., firing rates within a neuronal network) as well as higher-order correlations in activity between various nodes in the network. The various terms in this expansion can be interpreted as diagrams corresponding to subnetworks of the original network, which span both space (in terms of the network's graph) as well as time (in the sense of causality).

DOI: 10.1103/PhysRevE.80.036101 PACS number(s): 89.75.Fb, 11.10.−z, 87.19.L−, 87.85.dq

The study of dynamics on networks is becoming increasingly more relevant within biology. An important subclass of biological networks is "pulse-coupled" networks, which encompass many types of ecological, gene-regulatory, and neuronal networks. What do these pulse-coupled networks do, and how can we analyze them? This is a perplexing question whose answer depends on the pulse-coupled network and dynamic regime under consideration. To begin to answer this larger question, one must first understand the relationship (or map) between the architecture of a pulse-coupled network and certain statistical features of that network's dynamics—e.g., the relationship between the input to a neuronal network (i.e., a specific architectural feature) and the average firing rate of that network (i.e., a specific dynamical feature). Typically, these types of relationships cannot be understood easily, and theorists and modelers often resort to simulations in order to probe the properties of these maps. In this paper we present a framework which takes a first step towards systematically linking the dynamics of a general pulse-coupled network to the underlying architecture of that network. Our approach consists of two majors steps. First, we invoke a weak-coupling limit to formally expand the evolution operator and equilibrium distribution for the full network. Second, by using the evolution operator and equilibrium distribution for the network, we can construct a series expansion for any particular dynamical feature of the network (e.g., firing rates or correlations). The various terms in the series expansion of any dynamical feature incorporate successively larger subnetworks of the original network, thus capturing the specific architecture of the original network.

To provide a simple example which motivates our approach, we consider an idealized pulse-coupled network composed of N nodes $\{Z_1,\ldots,Z_N\}$, each of which is either "firing" or "inactive" at any point in (discrete) time. Thus, at each time the system is in one of 2^N possible system states $\{s_1,\ldots,s_{2^N}\}$, with each system state s_i corresponding to a unique set of firing nodes. This network (which we will describe below) is similar to the McCulloch-Pitts neuronal networks, which have been used to study neuronal network dynamics, learning, memory formation, and retrieval [1,2]. We will assume that the dynamics of this network is Markov, and that the activity of each node at a given time is only dependent on the system state at the previous time step, and is independent of the activity of the other nodes at the current time step. Specifically, given that the system is in state $\mathbb{S}(t)=s_i$ at time t, we assume that the activity of any node z at time $t+1$ is a Bernoulli random variable assuming values of either 1 (firing) or 0 (inactive), with probabilities

$$P(z(t+1)=1|\mathbb{S}(t)=s_i)=f\left(\eta_z+\sum_a \Delta_{za}a(t)\right)$$
$$=f\left(\eta_z+\sum_{a\in s_i}\Delta_{za}\right), \qquad (1)$$

$$P(z(t+1)=0|\mathbb{S}(t)=s_i)=1-P(z(t+1)=1|\mathbb{S}(t)=s_i)$$

where $f:\mathbb{R}\to[0,1]$ is a monotonically increasing function, η_z is the background input to the node z, and the coupling coefficient Δ_{za} is the coupling strength from node $a\to z$. The sum $\Sigma_{a\in s_i}$ is only over nodes a which are firing within the state s_i. Thus, in this simple system, the activation of any given node at time step t only affects other nodes in the system during the time step $t+1$. This system is fully described by its state-transition operator L^f, which can be written as

$$L^f_{ij}=P(\mathbb{S}(t+1)=s_i|\mathbb{S}(t)=s_j)$$
$$=\prod_z\begin{cases}P(z(t+1)=1|\mathbb{S}(t)=s_j), & \text{if } z\in s_i\\ P(z(t+1)=0|\mathbb{S}(t)=s_j), & \text{if } z\notin s_i.\end{cases}$$

Now, given a particular coupling matrix Δ, how does this network behave? What sorts of firing events and correlations will this network produce, and how will the network activity depend on the vector of inputs $\vec{\eta}=[\eta_{Z_1},\eta_{Z_2},\ldots,\eta_{Z_N}]^{\mathsf{T}}$?

For the particular simple example given in Eq. (1), if we assume that both $\vec{\eta}$ and Δ are very small, then by renormalizing $\vec{\eta}$ and Δ appropriately we can assume that $f(\cdot)$ is given by $f(\cdot)\sim\alpha+(\cdot)$ up to first order. This involves the assumption that each node fires independently with rate α in the absence of coupling or input. Using this first-order expression for $f(\cdot)$, we can expand the state-transition operator

$$L^f_{ij}\sim\gamma_i\left\{1+\sum_b\tilde{\mu}(b,s_i)\left[\eta_b+\sum_a\mu(a,s_j)\Delta_{ba}\right]\right\}, \qquad (2)$$

where $\beta=1-\alpha$, and we have introduced

$$\mu(a,s) = 1 \text{ if } a \in s, \quad 0 \text{ if } a \notin s,$$

$$\tilde{\mu}(b,s_i) = +1/\alpha \text{ if } a \in s, \quad -1/\beta \text{ if } a \notin s,$$

$$\gamma_j = \prod_a \begin{cases} \alpha, & \text{if } a \in s_j \\ \beta, & \text{if } a \notin s_j. \end{cases}$$

Throughout the rest of this introduction, we will use L^F_{ij} to describe a first-order expansion of the dynamics of this network. If the input $\vec{\eta}$ is fixed, then the behavior of the network is captured by the equilibrium distribution of system states ρ^F, which is the dominant eigenvector of L^F [3]. This equilibrium distribution is given by

$$\rho^F_j = \langle P(\mathcal{S}(t) = s_j)\rangle_t \sim \gamma_j \left[1 + \sum_b \tilde{\mu}(b,s_j)\left(\eta_b + \alpha \sum_a \Delta_{ba} \right) \right].$$

The equilibrium distribution ρ^F, in combination with the state-transition operator L^F, contains all the information about the stationary dynamics of the network. For example, the equilibrium firing rate associated with any node z in the network (denoted by $[z]$) can be obtained by appropriately projecting the equilibrium distribution,

$$[z] = \langle P(z(t) = 1)\rangle_t = \sum_j \mu(z,s_j)\rho^F_j \sim \alpha + \eta_z + \alpha \sum_a \Delta_{za}. \tag{3}$$

These firing rates are exactly what one would expect from a mean-field approach—the average input to node z is equal to $\eta_z + \alpha \Sigma_a \Delta_{za}$ (where we assume every other node in the network fires at rate α). Since we know both ρ^F and L^F, we can also compute higher-order statistics, such as correlations between firing events within this network. As an example, let us denote by $y \to z$ the event that node z fires one time step after node y, and let $[y \to z]$ denote the observation rate (in equilibrium) of these events. This can be computed as follows:

$$[y \to z] = \langle P(z(t+1) = 1, y(t) = 1)\rangle_t = \sum_{i,j} \mu(z,s_i)\mu(y,s_j)L^F_{ij}\rho^F_j \sim [y][z] + \alpha\beta\Delta_{zy}, \tag{4}$$

implying that the correlation $[y \to z] - [y][z]$ between the activity of the two nodes y and z is proportional to Δ_{zy}.

Working within this framework, we may ask how these particular dynamic observables shift as we perturb the input. Given any Markov system, we know that if the state-transition matrix L changes to $L + \varepsilon L'$ then to first order in ε the equilibrium distribution ρ shifts to $\rho + \varepsilon\rho'$, satisfying the equation

$$(L + \varepsilon L')(\rho + \varepsilon\rho') = \rho + \varepsilon\rho', \quad (I - L)\rho' = L'\rho.$$

Now we should note that the matrix $(I-L)$ is a square $2^N \times 2^N$ matrix with column sums equal to 0, and thus is not invertible. The column space of $(I-L)$ is perpendicular to the vector $e = [1, \ldots, 1]$. Fortunately, the column sum of L' is also 0 (since $L + \varepsilon L'$ must be a stochastic matrix), and the only component of ρ' we are interested in is the component perpendicular to e, since we can always assume that the eigenvector $\rho + \varepsilon\rho'$ is normalized to have sum 1. Thus, $(I-L)$ can be viewed as a linear operator from $e_\perp$ to $e_\perp$. For the simple system described in Eq. (2) such an inverse exists,

$$[(I - L^F)^{-1}]_{ij} \sim \delta_{ij} + 2^{-N} + L^F_{ij}.$$

Using this inverse we can readily calculate the shift in the equilibrium firing rates of the system with respect to changes in the input $\vec{\eta}$. This firing-rate shift is given by the derivative

$$\partial_{\eta_y}[z] = \sum_j \mu(z,s_j)\partial_{\eta_y}\rho^F_j \sim \delta_{yz} + \Delta_{yz}.$$

Note that this expression cannot be obtained by simply taking a derivative of the first-order terms shown in Eq. (3), as higher-order terms come into play when computing $\partial_{\eta_y}\rho_j$. An immediate consequence of this calculation is that, to first-order, the infinitesimal shift in $\vec{\eta}$ which gives rise to the greatest (Euclidian) shift in the equilibrium firing-rate vector for the system (i.e., the change in input to which the system is most responsive) is given by the first right singular vector of the matrix $(I+\Delta)$, which is equivalent to the dominant eigenvector of the symmetric matrix $(\Delta + \Delta^{\mathbf{T}})$. In a similar fashion, we can compute the shift in correlations with respect to changes in input,

$$\partial_{\eta_z}\{[y \to x] - [y][x]\} \sim (\beta - \alpha)\delta_{zy}\Delta_{xy} - \Delta_{yz}([x] - \alpha) - \Delta_{xz}([y] - \alpha).$$

In the simple example described above, the relationship between the architecture of the network (i.e., $\vec{\eta}$ and Δ) and the dynamics of the network (e.g., $[z]$, $[y \to z]$, $\partial_\eta[z]$, etc.) depends strongly on the state-transition matrix L^F, and the equilibrium distribution of system states ρ^F. Using first-order approximations to L^F and ρ^F limits us to first-order approximations of $[z]$ and $[y \to z]$. It is natural to ask if this approach can be generalized—can higher-order contributions to various dynamic observables be systematically computed? Moreover, can these same techniques be applied to more realistic systems, such as the types of neuronal networks typically used in large-scale computational modeling? Within the rest of this paper we will show that, indeed, this framework can be generalized to any order for a very large class of pulse-coupled networks. The key elements of this framework will be an approximation to L^F, an approximation to ρ^F, and an approximation to $(I-L^F)^{-1}$.

The basic idea underlying our framework is as follows. By describing the internal dynamics of each node appropriately, a pulse-coupled network can be reduced to a Markov system [3]. If the dynamics of each node is ergodic, and the coupling is sufficiently weak, then the dynamics of the network will be ergodic, and there will be a unique time-averaged distribution of system states ρ^F, which will depend on the system's architecture (i.e., the dynamics associated with each node, the external input to each node, and the coupling between nodes). Note that ρ^F can also be interpreted as the distribution of system states in an ensemble average. Given ρ^F, we can compute various projections of the system's dynamics, such as the system's firing rate. Higher-order statistics, such as event-chains [4] and two-point correlations require understanding ρ^F as well as the

state-transition operator L^F for the system. In practice, precisely determining ρ^F as well as L^F for a complicated network is not feasible. Note, however, that if there is no coupling between nodes then the state-transition operator can be obtained by taking the direct product of several uncoupled single node state-transition operators, and the distribution ρ^F is simply the direct product of the uncoupled single node equilibrium distributions. In what follows we derive a formal expansion (with respect to coupling strength) of the distribution ρ^F (as well as various projections of the system's dynamics) for a large class of heterogeneous and topologically distinct pulse-coupled networks. Due to the pulse-coupling between different components of the network, the terms of this expansion each can be interpreted as a particular flow of information within different subnetworks of the full network. Using this expansion we can derive several useful relationships, and analyze pulse-coupled networks in a new way.

I. DISCRETE-TIME FINITE-STATE SUBNETWORK EXPANSION

For the purposes of discussion, we will define a pulse-coupled network as follows. A pulse-coupled network consists of a directed graph amongst several nodes, with each node possessing its own internal dynamics. The nodes evolve independently from the rest of the network, except for instantaneous pulses or "firing events," which are generated when any one node reaches a specified set of states—at which point there is an effect on other nodes in the network to which the firing node is connected. We will assume that the state space of each node has been appropriately expanded so that the effects of a pulse are instantaneous. Specifically, we assume that the dynamics of each node within the network differs from the uncoupled dynamics associated with that node only at times which correspond to firing events. This restriction is not severe, and many biologically relevant pulse-coupled networks fall into this framework.

Some previous work on pulse-coupled systems, such as [5,6], assumes that the system is deterministic, and that the dynamics reach a fixed point (which may depend on the system's initial conditions). With these assumptions, their work then analyzes the basins of attraction (with respect to initial conditions) of each of these fixed points. However, there are many pulse-coupled systems, such as neuronal networks within the mammalian visual cortex, which never attain a steady state, and which perpetually drift between a variety of states even when driven by constant input (or while in background) [7–9]. In this paper we will focus on this second type of system. We assume that the dynamics produced by the network (even under constant input) does not settle to a steady state, but instead involves an ongoing sequence of different system states. We also assume that there are some statistical features of the dynamics (such as firing rates or correlations in the activity of different nodes in the network), which are relevant either for observers, or for controlling the dynamics of downstream networks.

Many previous attempts to link the architecture and dynamics of pulse-coupled networks, such as the kinetic theory of neuronal populations [10–16] and recent field-theoretic techniques [17,18], typically assume that (a) the number of network nodes is large, (b) the uncoupled dynamics of each node is the same, and (c) the coupling between nodes is statistically homogeneous. However, many real pulse-coupled networks are not very large, and possess particular heterogeneous composition and connectivity (e.g., the network may be composed of various types of nodes with different internal properties, which are sparsely coupled together with nonuniform coupling strength). The framework we introduce is capable of addressing these network-specific details.

In what follows we will derive the diagrammatic expansion used later in the paper. For ease of presentation we will discuss a discrete-time, finite-state version of a conductance-based integrate-and-fire neuronal network [19–27], for which it is clear where each term in the expansion originates. Note that the extension to continuous time and infinitely many states is relatively straightforward (although more difficult to numerically implement and test), and will be briefly discussed later. Within this finite-state network, each node represents a neuron, and each firing event represents a "spike" or neuronal action potential. Assume that a given neuron, say Z, can be in one of m voltage states, and one of n conductance states. If the voltage reaches the final mth voltage state (the threshold voltage), we say the neuron "fires" (affecting other neurons in the network), and the voltage is forced to reset to the first voltage state (the reset voltage) in the next time step. Thus at each point in time the neuron is in one (and only one) of $s=mn$ possible states, which we enumerate with the indices $i_Z^V \in \{1,\ldots,m\}$ and $i_Z^G \in \{1,\ldots,n\}$ [collectively referred to as $i_Z \in \{(1,1),\ldots,(m,n)\}$]. We also assume that each neuron is driven by some input $\eta_Z \in \mathcal{R}$, which could depend on Z, but which does not depend on time (many types of time-varying input can be accounted for by appropriately increasing the state space of each individual neuron).

A. State transition matrix

In the absence of input from other neurons in the network, the discrete-time evolution of neuron Z is governed by some $s \times s$ state-transition matrix $L(\eta_Z)$. The entry L_{ij} equals the probability that the neuron Z transitions from state j to state i in one time step, given an input η_Z at the start of that time step. In other words, if $\mathbb{S}(t)$ is the state of the neuron Z at time t, then $L_{ij}(\eta_Z)=P(\mathbb{S}(t+1)=i|\mathbb{S}(t)=j$, and input $\eta_Z)$. We will place two restrictions on the transition matrix. The first restriction ensures that the system behaves like an integrate-and-fire neuron (i.e., when the voltage reaches a threshold it is forced to reset to a lower value [22,26]). We enforce this restriction by requiring that $L_{ij}=0$ whenever $j^V=m$ and $i^V \neq 1$. The second restriction we place on the state-transition matrix is that L should "mix" (i.e., have a unique eigenvector ρ of eigenvalue 1). This second restriction ensures that our subnetwork expansion for ρ^F (detailed below) will be asymptotically accurate in the weak-coupling limit [3].

We now use the state-transition matrix for individual neurons to build the state-transition matrix for a network of N neurons. As mentioned earlier, we assume that the network of neurons is pulse coupled, meaning that neurons only affect

one another when they fire. This assumption is critical for this particular framework, but is a very reasonable one. We use the multi-index $\vec{i}=\{i_{Z_1},i_{Z_2},\ldots,i_{Z_N}\}$ to reflect the state of the network, implying that Z_l is in state $i_{Z_l}=(i^V_{Z_l},i^G_{Z_l})$. The $s^N \times s^N$ transition matrix for the network is given by the product of N single-neuron state-transition operators (of the form $L_{i_{z'}j_z}$), each dependent on the state of the other neurons in the network

$$L^F_{\vec{i}\vec{j}} = P(\mathbb{S}(t+1)=\vec{i}\,|\mathbb{S}(t)=\vec{j}) = \prod_z L_{i_{z'}j_z}\left(\eta_z + \sum_w \Delta_{zw}[\delta_{i^V_w,1}\delta_{j^V_w,m}]\right), \tag{5}$$

where $\mathbb{S}(t)$ is the state of the system at time t, and the entry Δ_{zw} of the connectivity matrix $\Delta \in \mathcal{R}^{N\times N}$ specifies the effect of a spike from neuron w on neuron z. The pulse coupling is encapsulated by the product of Kronecker deltas $\delta_{i^V_w,1}\delta_{j^V_w,m}$, which reflects the fact that neuron w only affects neuron z during a step in which neuron w is transitioning from threshold to reset (i.e., a step in which w fires). At this point, if we could simply find the equilibrium distribution ρ^F associated with L^F, we would have "solved" for the system's dynamics, and we could then measure a variety of quantities, such as firing rates, at our leisure. Unfortunately, even though we may be capable of solving the $s\times s$ eigenvalue problem associated with L, the $s^N\times s^N$ eigenvalue problem associated with L^F is typically impossible, even for moderate values of m,n,N. In what follows, we will form an expansion for L^F (in terms of the coupling strength Δ), which will allow us to approximate ρ^F without solving the eigenvalue problem associated with an N-neuron eigensystem.

We can Taylor-expand Eq. (5) in terms of the coupling strength Δ as follows:

$$L^F_{\vec{i}\vec{j}} = \prod_z \left[L_{i_{z'}j_z}(\eta_z) + \sum_{p=1}^{p=\infty} \frac{1}{p!}\partial^p_\eta L_{i_{z'}j_z}(\eta_z)\left(\sum_w \Delta_{zw}[\delta_{i^V_w,1}\delta_{j^V_w,m}]\right)^p \right].$$

Now, we introduce the following notation:

$$L_z = L_{i_{z'}j_z}(\eta_z), \quad L^{(p)}_z = \partial^p_\eta L_{i_{z'}j_z}(\eta_z),$$

$$F_w = F_{i_w j_w} = [\delta_{i^V_w,1}\delta_{j^V_w,m}],$$

and note the equality

$$\left(\sum_w F_w \Delta_{zw}\right)^p = \tilde{\sum} \mathcal{M}(p,\{q_\beta\}) \prod_\beta F_{w_\beta}(\Delta_{zw_\beta})^{q_\beta}, \tag{6}$$

where the sum $\tilde{\Sigma}$ is taken over all subsets of unique neurons $\{w_\beta\}$ and sets of positive integers $\{q_\beta\}$, such that $\Sigma_\beta q_\beta = p$. With this notation one can think of z as a "receiver" and consider each term of Eq. (6) as involving $|\{q_\beta\}|$ "transmitter" neurons $\{w_\beta\}$ (each considered with multiplicity q_β). The symbol $\mathcal{M}(p,\{q_\beta\})$ refers to the number of ways to choose $|\{q_\beta\}|$ uniquely defined elements with multiplicities $q_1,q_2,\ldots$ out of p identical sets of N elements. Using Eq. (6) we can write $L^F = L^{[0]} + L^{[1]} + \cdots$, where the Mth order term $L^{[M]}$ involves a sum over subnetworks,

$$L^{[M]}_{\vec{i}\vec{j}} = \tilde{\sum}^{[1]} \left[\prod_{\alpha,\beta} \mathcal{M}(p_\alpha,\{q^\alpha_\beta\}_\beta) \Delta^{q^\alpha_\beta}_{z_\alpha w^\alpha_\beta} \right] \times \left[\prod_\alpha \frac{1}{p_\alpha!} L^{(p_\alpha)}_{z_\alpha} \right] \left[\prod_{\alpha,\beta} F_{w^\alpha_\beta} \right] \left[\prod_{y\notin\{z_\alpha\}} L_y \right], \tag{7}$$

where the sum $\tilde{\Sigma}^{[1]}$ is taken over all subsets of unique neurons $\{z_\alpha\}$, $\{w^\alpha_\beta\}$ and sets of positive integers $\{p_\alpha\}$, $\{q^\alpha_\beta\}$ such that $\Sigma_\alpha p_\alpha = M$, and $\Sigma_\beta q^\alpha_\beta = p_\alpha$. Each term in this Mth-order sum corresponds to some set of $|\{p_\alpha\}| \le M$ unique receiver neurons $\{z_\alpha\}$ each considered with multiplicity p_α. The index α is used to label the receiver neurons. Each receiver z_α corresponds to $|\{q^\alpha_\beta\}_\beta|$ unique transmitter neurons $\{w^\alpha_\beta\}_\beta$, each considered with multiplicity q^α_β (the subscript on $\{q^\alpha_\beta\}_\beta$ and $\{w^\alpha_\beta\}_\beta$ is added for clarity—to imply that the sets are to be considered with index β varying, and index α fixed). The index β is used to label the transmitter neurons associated with any particular receiver z_α. Note that in this particular form of the expansion, multiplicity amongst receivers is disallowed (i.e., each z_α is unique); however multiplicity amongst transmitters is allowed (i.e., the same w can be in set $\{w^1_\beta\}$ as well as in set $\{w^2_\beta\}$, thus implying that the connections $\Delta_{z_1 w}$ and $\Delta_{z_2 w}$ are being simultaneously considered). Note also that, as $L^{[M]}$ is associated with only a single time step, the effects of each transmitter are considered simultaneously. The sets $\{z_\alpha\}$, $\{w^\alpha_\beta\}$ along with their multiplicities define a subnetwork of the original network of N neurons, and the corresponding term of $L^{[M]}$ describes the Mth-order effect of this subnetwork on the state-transition matrix of the full system (see Figs. 1 and 2 for examples of various subnetworks and their graphical representations).

To clarify the rest of the presentation, we will let "$\cdot$" denote the standard matrix product (whenever there is confusion), and we will use the symbol "$\otimes$" to denote a matrix direct product. The direct product involves a "nestling" or "interlacing" of matrices—given a $p_A\times q_A$ matrix A and an $p_B\times q_B$ matrix B, the $p_A p_B \times q_A q_B$ matrix $C=A\otimes B$ is given by $C_{(i,k),(j,l)}=A_{ij}B_{kl}$, where (i,k) and (j,l) are both multi-indices. Note that we could choose some index ordering, such as $\iota = i + p_A(k-1)$ and $\varphi = j + q_A(l-1)$ and define $C_{\iota\varphi}$ with respect to standard indices, but this requires an essentially arbitrary ordering of dimensions, and will have no impact on our results. Note that $[A\otimes B]\otimes C = A\otimes[B\otimes C]$, $A\otimes(B+C) = A\otimes B + A\otimes C$, and (if A and B are both invertible) $(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$. Note, however, that $[I\otimes I - A\otimes B]^{-1} \neq (I-A)^{-1}\otimes(I-B)^{-1}$, which will become important later [see Eq. (10)]. With this notation we present some special cases of Eq. (7), such as

$$L^{[0]}_{\vec{i}\vec{j}} = \prod_z L_{i_{z'}j_z} = \bigotimes_z L_z,$$

which is the simple direct product of individual uncoupled state matrices, and (writing $L' = \partial_\eta L$)

$$L^{[1]} = \sum_{z,w \text{ distinct}} \Delta_{zw} L'_z \otimes (F^R \cdot L_w \cdot F^T) \bigotimes_{y\neq z,w} L_y + \sum_z \Delta_{zz} (F^R \cdot L'_z \cdot F^T) \bigotimes_{y\neq z} L_y, \tag{8}$$

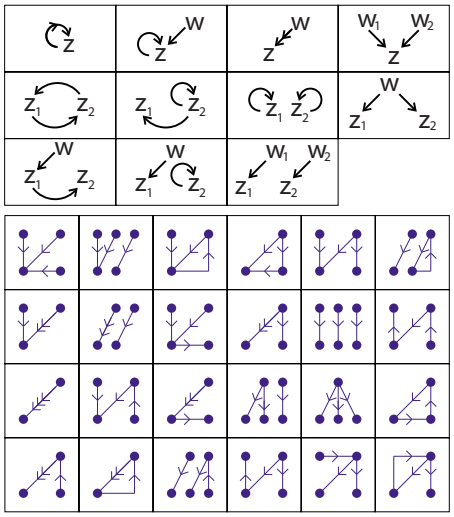

FIG. 1. (Color online) Examples of subnetworks used for expanding $L^{[M]}$, as described in Sec. I A. (Upper three rows) Displayed are 11 topologically distinct subnetworks associated with different terms in the subnetwork expansion of $L^{[2]}$. Each subnetwork is composed of up to two receiver neurons (indicated by "Z") and up to two transmitter neurons (indicated by "W" if they are distinct from the receivers). The degree p_α associated with receiver Z_α is equal to the number of arrowheads pointing at Z_α. The degree q_β^α associated with the transmitter-receiver pair W_β^α, Z_α is equal to the number of arrowheads on the arrow connecting W_β^α to receiver Z_α. Starting from the upper left corner and reading from left to right, the subnetworks correspond to $\{p_1=2, q_1^1=2, Z_1=W_1^1\}$, $\{p_1=2, q_1^1=q_2^1=1, Z_1=W_1^1\}$, $\{p_1=2, q_1^1=2\}$, $\{p_1=2, q_1^1=q_2^1=1\}$, $\{p_1=p_2=1, Z_1=W_1^2, Z_2=W_1^1\}$, $\{p_1=p_2=1, Z_2=W_1^2, Z_2=W_1^1\}$, $\{p_1=p_2=1, Z_2=W_1^2, Z_1=W_1^1\}$, $\{p_1=p_2=1, W_1^2=W_1^1\}$, $\{p_1=p_2=1, Z_1=W_1^2\}$, $\{p_1=p_2=1, Z_2=W_1^2\}$, and $\{p_1=p_2=1\}$, respectively, where neurons with distinct labels are assumed distinct unless otherwise specified. (Lower four rows) Displayed are 20 topologically distinct subnetworks associated with different terms in the subnetwork expansion of $L^{[3]}$. Only subnetworks excluding autapses are shown. The direction and number of arrowheads are sufficient to specify the role of each of the neurons (indicated by circles).

FIG. 2. (Color online) Examples of multistage subnetworks used for expanding $\rho^{[3]}$, as described in Sec. I A. (Left) Displayed is the two-stage subnetwork associated with the terms $\tilde{\rho}$ and $\widetilde{[b]}$ described in Sec. I C. The first stage is colored dark gray (blue online), and the second stage is colored light gray (red online). Each stage γ of this subnetwork involves a set of receivers and transmitters (denoted by $\{Z_\alpha^\gamma\}_\alpha$ and $\{W_\beta^{\gamma,\alpha}\}_{\alpha,\beta}$, respectively). The degree p_α^γ of every receiver Z_α^γ is equal to the number of arrowheads pointing into Z_α^γ, and the degree $q_\beta^{\gamma,\alpha}$ of every transmitter-receiver pair $W_\beta^{\gamma,\alpha}, Z_\alpha^\gamma$ is equal to the number of arrowheads on the arrow connecting the pair. The stages span time in the sense of "causality" since the contribution to $\tilde{\rho}$ coming from stage 1 involves the correction associated with stage 2. (Right) Displayed is the three-stage "loop" subnetwork described in Sec. I C. The first, second and third stages are different shades of gray, and are colored online blue, red, and green, respectively.

which consists of first-order contributions from w to z in addition to the effect of autapses (i.e., direct connections from a neuron to itself). In Eq. (8) we have introduced the notation

$$F_{ij}^R = \delta_{ij}\delta_{i^V,1}, \quad F_{ij}^T = \delta_{ij}\delta_{j^V,m}$$

to disambiguate the element-wise product $L_{i_{z^j z}} F_{i_{z^j z}}$ from a matrix product (note that the only nonzero entries of the $s \times s$ matrices $F^R L F^T$ and $F^R L' F^T$ are the "firing blocks" of the matrices L and L', respectively). As another example, the term $L^{[2]}$ includes 11 topologically different types of subnetworks (see Fig. 1) ranging from pairs of distinct first-order contributions to second-order contributions from autapses.

B. Equilibrium distribution

The eigenvector ρ^F of L^F with eigenvalue 1 corresponds to the equilibrium distribution of the network—$\rho_{\vec{j}}^F$ can be interpreted as the probability that the system is in state $\vec{j}$ at a randomly chosen time, or that a particular system chosen out of an infinite ensemble is in state $\vec{j}$ [3]. Determining ρ^F will allow us to compute several useful quantities, such as, say, the firing rates of individual neurons. Once we have an expression for $L^{[M]}$, we can determine the distribution ρ^F order by order (i.e., $\rho^F = \rho^{[0]} + \rho^{[1]} + \ldots$). The 0th order term $\rho^{[0]}$ is given by

$$\rho_{\vec{j}}^{[0]} = \bigotimes \rho_z,$$

where $\rho_z = \rho_{j_z}$ is the single-neuron distribution corresponding to L_z (i.e., the eigenvector of L_z with eigenvalue 1). We know that $L^{[0]}\rho^{[0]} = \rho^{[0]}$, but determining $\rho^{[M]}$ is slightly more difficult. Here we write out a formal expansion for $\rho^{[M]}$ in terms of $\rho^{[0]}$ and the elements of Eq. (7),

$$L^F \cdot \rho^F = \rho^F \Rightarrow (I - L^F) \cdot (\Sigma_{M>0}\rho^{[M]}) = \Sigma_{M>0} L^{[M]} \cdot \rho^{[0]}. \tag{9}$$

We note that $(I - L^F)$ is not invertible (as ρ^F is in the null space). However, every other eigenvector of L^F has an eigenvalue less than 1, and so $(I - L^F)^{-1}$ is well defined on the right eigenspace of L^F, which excludes ρ^F. It is easy to show that this eigenspace is $e_\perp$, where e is the vector of all 1's (indeed, since the column sums of L^F are equal to 1, e^{T} is a left eigenvector of L^F, and we have that $e^{\mathsf{T}}\psi = e^{\mathsf{T}} L^F \psi = \lambda e^{\mathsf{T}}\psi$ for any right eigenvector $L^F\psi = \lambda\psi$). Thus, it is reasonable to treat $(I - L^F)^{-1}$ as an operator from $e_\perp \to e_\perp$. An appropriate expansion of $(I - L^F)^{-1}$ is given by

$$(I - L^F)^{-1} = \left[\sum_{M>0} \tilde{\sum}^{[2]} \prod_\gamma G \cdot L^{[r_\gamma]} \right] \cdot G, \tag{10}$$

where the sum $\tilde{\Sigma}^{[2]}$ is taken over all sets of positive integers $\{r_\gamma\}$ such that $\Sigma_\gamma r_\gamma = M$, and the operator G is given by

$$G = (I - L^{[0]})^{-1} : e_\perp \to e_\perp,$$

[note that Span $(L^{[r]}) \in e_\perp$ for $r > 0$]. Combining Eqs. (9) and (10) yields

$$\rho^{[M]} = \left[\tilde{\sum}^{[2]} G \cdot L^{[r_\gamma]} \right] \cdot \rho^{[0]}, \tag{11}$$

where again the sum $\tilde{\Sigma}^{[2]}$ is taken over all sets of positive integers $\{r_\gamma\}$ such that $\Sigma_\gamma r_\gamma = M$. Each term in the sum can be interpreted as consisting of $t = |\{r_\gamma\}|$ stages, with each stage γ involving r_γ transmitter-receiver pairs (counted with multiplicity). Because the effects of each transmitter associated with stage γ are considered simultaneously, one can consider a t-stage term as consisting of t unique firing times. Once we finish discussing a representation for G, we will return to Eq. (11) and interpret the terms that arise [see Eq. (13)].

Before we write down the operator $G \in \mathcal{C}^{s^N \times s^N}$, we first solve for $G_z = (I - L_z)^{-1} : e_\perp \to e_\perp$ (where $G_z \in \mathcal{C}^{s \times s}$ acts only on the neuron z). We diagonalize $L_z = \Psi_z \cdot \Sigma_z \cdot \Psi_z^{-1}$ [where the diagonal matrix $\Sigma_z = \mathrm{diag}(\sigma^z)$, and $\sigma_1^z = 1$ and $|\mathrm{real}(\sigma_j^z)| < 1$ for $j > 1$], and write

$$G_z = \Psi_z \cdot \mathrm{diag}(\Lambda^z) \cdot \Psi_z^{-1},$$

where $\mathrm{diag}(\Lambda^z) = (I - \Sigma_z)^{-1}$ is the standard pseudoinverse of $(I - \Sigma_z)$, given by

$$\Lambda^z = \left[0, \frac{1}{1-\sigma_2^z}, \frac{1}{1-\sigma_3^z}, \ldots, \frac{1}{1-\sigma_s^z} \right].$$

In a similar fashion we can now define

$$G_{\{z_\alpha\}} = (I - \bigotimes_\alpha L_{z_\alpha})^{-1} = [\bigotimes_\alpha \Psi_{z_\alpha}] \cdot \mathrm{diag}(\Lambda^{\{z_\alpha\}}) \cdot [\bigotimes_\alpha \Psi_{z_\alpha}^{-1}],$$

for any set of neurons $\{z_\alpha\}$, where, $\Lambda^{\{z_\alpha\}} = (I - \bigotimes_\alpha \Sigma_{z_\alpha})^{-1}$ is the appropriate pseudoinverse (note that, with this notation, $G = G_{\{Z_1, Z_2, \ldots, Z_N\}}$). While not all that practical for actual computation, the previous representation of G can be used to derive two important equalities, which we will use later to reduce the dimensionality of expressions involving G,

$$G_{\{z_\alpha\}} \cdot [\rho_{z_1} \otimes \zeta] = \rho_{z_1} \otimes [G_{\{z_\alpha\}_{\alpha>1}} \cdot \zeta],$$

$$[e^{\mathbf{T}} \otimes \zeta^{\mathbf{T}}] \cdot G_{\{z_\alpha\}} = e^{\mathbf{T}} \otimes [\zeta^{\mathbf{T}} \cdot G_{\{z_\alpha\}_{\alpha>1}}], \tag{12}$$

both of which are true for any $s^{|\{z_\alpha\}|-1}$ element vector ζ, as long as $|\{z_\alpha\}| \geq 2$.

Now we can rewrite Eq. (11) in a more descriptive manner,

$$\rho^{[M]} = \tilde{\sum}^{[3]} \left[\prod_{\gamma,\alpha} \mathcal{M}(p_\alpha^\gamma, \{q_\beta^{\gamma,\alpha}\}_\beta) \prod_\beta \Delta_{z_\alpha^\gamma w_\beta^{\gamma,\alpha}}^{q_\beta^{\gamma,\alpha}} \right] \cdot \prod_\gamma \left\{ G \cdot [I \otimes_{\alpha,\beta} F^R_{w_\beta^{\gamma,\alpha}}] \cdot \left[\bigotimes_\alpha \frac{1}{p_\alpha^\gamma!} L_{z_\alpha^\gamma}^{(p_\alpha^\gamma)} \bigotimes_{a \notin \{z_\alpha^\gamma\}_\alpha} L_a \right] \cdot [I \otimes_{\alpha,\beta} F^T_{w_\beta^{\gamma,\alpha}}] \rho^{[0]} \right\}, \tag{13}$$

where the sum $\tilde{\Sigma}^{[3]}$ is taken over all sets of positive integers $\{r_\gamma\}$, $\{p_\alpha^\gamma\}$, and $\{q_\beta^{\gamma,\alpha}\}$, and all sets of neurons $\{z_\alpha^\gamma\}$, $\{w_\beta^{\gamma,\alpha}\}$ such that $\Sigma_\gamma r_\gamma = M$, $\Sigma_\alpha p_\alpha^\gamma = r_\gamma$, and $\Sigma_\beta q_\beta^{\gamma,\alpha} = p_\alpha^\gamma$. To recapitulate, each of the terms in Eq. (13) can be interpreted as a subnetwork of $t = |\{r_\gamma\}|$ stages, spanning both space and "time" (in the sense that a t-stage subnetwork involves t distinct spike times). The index γ is used to label the distinct spike times. Each stage γ corresponds to some set of $|\{p_\alpha^\gamma\}| \leq r_\gamma$ unique receiver neurons $\{z_\alpha^\gamma\}_\alpha$ each considered with multiplicity p_α^γ. The index α is used to label the receivers associated with any particular stage γ. Each receiver z_α^γ corresponds to $|\{q_\beta^{\gamma,\alpha}\}_\beta|$ unique transmitter neurons $\{w_\beta^{\gamma,\alpha}\}_\beta$ each considered with multiplicity $q_\beta^{\gamma,\alpha}$. The index β is used to label the transmitters associated with any particular receiver z_α^γ. The sets $\{z_\alpha^\gamma\}$, $\{w_\beta^{\gamma,\alpha}\}$ along with their multiplicities define a subnetwork of the original network of N neurons, and the corresponding term of $\rho^{[M]}$ describes the Mth-order effect of this subnetwork on the eigendistribution of the full system. It is important to note that no homogeneity has been assumed in the derivation of Eq. (13). The internal dynamics L_z associated with each neuron, as well as every element of the connectivity matrix Δ, are treated independently.

As a specific case, let us examine a particular term within the expression for $\rho^{[3]}$ which corresponds to a two-stage 3rd-order subnetwork with $r_1 = 2$, $r_2 = 1$, $p_1^1 = 2$, $q_1^{1,1} = 1$, and $q_2^{1,1} = 1$ (and consequently $p_1^2 = 1$, $q_1^{1,2} = 1$), for which the set $\{z_1^1, w_1^{1,1}, w_2^{1,1}, z_1^2, w_1^{2,1}\}$ contains distinct elements, with the exception that $w_2^{1,1} = z_1^2$. To simplify notation, let us rename the transmitters and receivers $\{z_1^1, w_1^{1,1}, w_2^{1,1}, z_1^2, w_1^{2,1}\}$ to be $\{z, a, y, y, c\}$. This term (say, $\tilde{\rho}$) corresponds to

$$\tilde{\rho} = 2\Delta_{za}\Delta_{zy}\Delta_{yc} G \cdot \left[\frac{1}{2} L_z'' \otimes F^R L_a F^T \otimes F^R L_y F^T \bigotimes_{x \neq z,a,y} L_x \right] \cdot G \cdot [L_y' \otimes F^R L_c F^T \bigotimes_{x \neq y,c} L_x] \cdot \rho^{[0]},$$

which [using Eq. (12)] can be reduced to

$$\tilde{\rho} = \Delta_{za}\Delta_{zy}\Delta_{yc} G_{\{z,a,y,c\}} \cdot [L_z'' \rho_z \otimes F^R L_a F^T \rho_a \otimes [[F^R L_y F^T \otimes L_c] \cdot G_{\{y,c\}} \cdot [L_y' \rho_y \otimes F^R L_c F^T \rho_c]]] \bigotimes_{x \neq z,a,y,c} \rho_x. \tag{14}$$

The interior term $G_{\{y,c\}}[L_y' \rho_y \otimes F^R L_c F^T \rho_c]$ in Eq. (14) represents the first-order correction to ρ^F, resulting from incorporating the effects of neuron c firing on the dynamics of neuron y (as if c were uncoupled from the network). Note that the term $F^R L_c F^T \rho_c$ picks out the component of ρ_c that flows over threshold in one time step. The full term shown in Eq. (14) represents the third order correction to ρ^F resulting from correcting the dynamics of neuron z by incorporating the effects of neuron a (firing as if uncoupled) as well as the corrected effects of neuron y (whose firing is corrected by incorporating the effects of neuron c). We can interpret the term $\tilde{\rho}$ as a correction involving two subnetworks—the first involving c affecting y, and the second involving a and y affecting z (see Fig. 2). It is worth noting at this point that the utility of the subnetwork expansion comes from (a) the fact that neurons only influence each other when in a select few states (so that the coupling term $\delta_{i_{w'}^\nu,1}\delta_{j_{w'}^\nu,m}$ is quite simple), and (b) the equalities given in Eq. (12) allow us to solve for terms involving various subnetworks without considering the entire system. Indeed, the expression for $\tilde{\rho}$ given in Eq. (14) is effectively quite low dimensional—involving a four-neuron subsystem $\{z, a, y, c\}$ interlaced with $N-4$ uncoupled neurons $\{x\}_{x \neq z,a,y,c}$. Note that terms involving $G_{\{z_\alpha\}}$ with

$|\{z_a\}|>1$ can be further reduced to a series of single-neuron operators, as described in Appendix B.

It is important to point out that the validity of this expansion hinges on the assumption that the spectrum of $L_{\Gamma}(\Delta)$ admits a single unique eigendistribution of eigenvalue 1, which varies smoothly as the coupling strength Δ is increased away from 0. Validating this assumption (which is stronger than merely requiring ergodicity of L_{Γ} for a particular Δ) is beyond the scope of this paper. Nevertheless, it should be noted that if this assumption is violated (e.g., when a network exhibits bistability), then the results computed using the subnetwork expansion may not correspond to an observable solution, and should be interpreted with care.

C. Projections of the equilibrium distribution

Once we have an approximation to ρ^{Γ}, we can approximate several important features of the system by projecting ρ^{Γ} in various ways. For example, we can determine the firing rate of any neuron b (denoted by $[b]$) by computing

$$[b] = [e_b^{\mathsf{T}} F_b^T \bigotimes_{x \neq b} e_x^{\mathsf{T}}] \cdot \rho^{\Gamma}, \tag{15}$$

where, for clarity, we use the subscript b on e_b^{T}, F_b^T and x on e_x^{T} to indicate that the operator acts on the indices associated with neuron b, x, respectively (when the indices are obvious we will omit these subscripts). Obviously, since $e_x^{\mathsf{T}} \rho_x = 1$ for every x, we have that

$$[b]^{[0]} = e_b^{\mathsf{T}} F_b^T \rho_b = f_b^{\mathsf{T}} \rho_b$$

is the uncoupled firing rate of neuron b (where we have introduced $f^{\mathsf{T}} = e^{\mathsf{T}} F^T = \delta_{i^V,m}$). In addition, since Span (L'), span $(F^R L' F^T) \in e_{\perp}$, $e^{\mathsf{T}} F^R L F^T = f^{\mathsf{T}}$, and $f^{\mathsf{T}} F^R L' F^T = 0$, we have that

$$[b]^{[1]} = \sum_{z \neq b} \Delta_{bz} (f_b^{\mathsf{T}} G_b L_b' \rho_b)(f_z^{\mathsf{T}} \rho_z) + \Delta_{bb} (f_b^{\mathsf{T}} G_b F^R L_b' F^T \rho_b), \tag{16}$$

implying that, to first order, only direct synapses onto b affect the firing rate of b. With the possible exception of the autapse term, these two results can be obtained from a mean-field theory. Notably, however, our framework allows us to systematically expand the projection $[b]$ to higher order. As an example, the contribution $\widetilde{[b]}$ to $[b]^{[3]}$ arising from the two-stage 3rd-order term $\tilde{\rho}$ [shown in Eq. (14)] is 0 unless $b=z$, in which case we have

$$\widetilde{[b]} = [f_z^{\mathsf{T}} \bigotimes_{x \neq z} e_x^{\mathsf{T}}] \tilde{\rho} = \Delta_{za} \Delta_{zy} \Delta_{yc} (f_z^{\mathsf{T}} G_z L_z'' \rho_z) \times (f_a^{\mathsf{T}} \rho_a)(f_y^{\mathsf{T}} G_y L_y' \rho_y)(f^{\mathsf{T}} \rho_c),$$

which is a product of four scalars, each of which can be formed by considering an uncoupled single neuron. It is important to note that, while typical, $\widetilde{[b]}$ is not representative of every subnetwork contributing to $[b]^{[3]}$. For example, the three-stage 3rd-order "three-element loop" subnetwork (see Fig. 3) corresponding to $r_1=r_2=r_3=1$, for which the set $\{z_1^1, w_1^{1,1}, z_1^2, w_1^{2,1}, z_1^3, w_1^{3,1}\}$ contains three distinct elements,

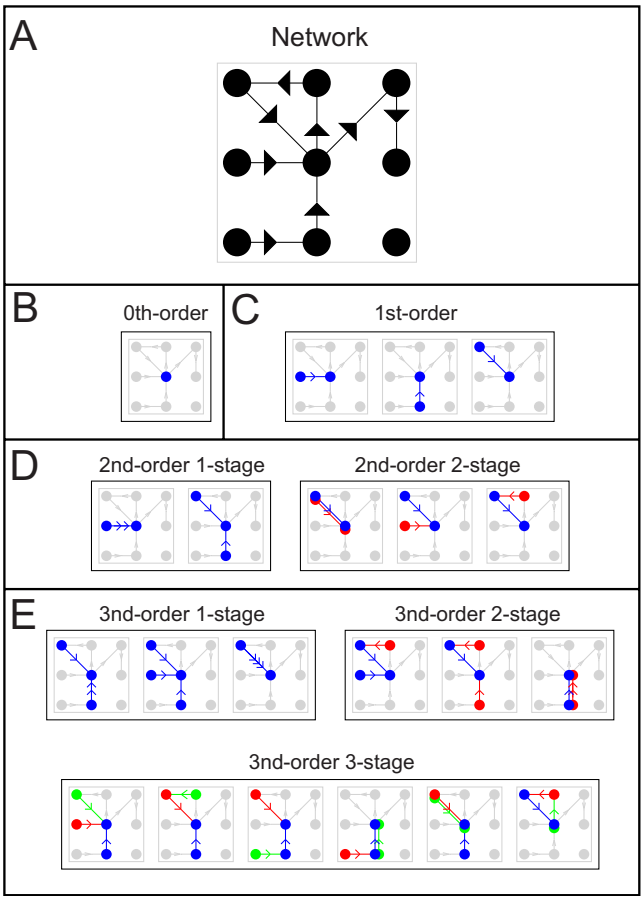

FIG. 3. (Color online) Examples of multistage subnetworks used for expanding $[b]^{[0]}$, $[b]^{[1]}$, $[b]^{[2]}$, and $[b]^{[3]}$, as described in Sec. I C. (a) Displayed is a specific network. Neurons are indicated by circles and connections are indicated by solid black arrows. In the rest of this figure we demonstrate the subnetwork expansion associated with the firing rate of the central neuron in this network (denoted by $[b]$). The diagrams used to denote subnetworks are described in Fig. 2. The first, second, and third stages are shown in different shades of gray, which are colored online blue, red, and green, respectively. (b) The subnetworks contributing to $[b]^{[0]}$. Note that only the neuron b itself is included. (c) The subnetworks contributing to $[b]^{[1]}$. Note that b must be the receiver. (d) Sample subnetworks contributing to $[b]^{[2]}$, grouped into one-stage subnetworks and two-stage subnetworks. (e) Sample one-stage, two-stage, and three-stage subnetworks contributing to $[b]^{[3]}$. The three-stage subnetwork topologically equivalent to Fig. 2(b) is in the bottom right corner. Note that each t-stage subnetwork only includes neurons that can be connected to b in t steps or less. Thus, no t-stage subnetwork will include the neurons on the right side of the network.

and $w_1^{1,1}=z_1^2$, $w_1^{2,1}=z_1^3$, and $w_1^{3,1}=z_1^1$, corresponds to the contribution

$$[([f_y^{\mathsf{T}} \otimes f_z^{\mathsf{T}} G_z L_z'] G_{yz} [L_y' \otimes L_z]) \otimes f_x^{\mathsf{T}}],$$

$$[\rho_y \otimes G_{zx} [F^R L_z F^T \rho_z \otimes L_x' \rho_x]],$$

where we have relabeled $\{z_1^1, w_1^{1,1}, z_1^2, w_1^{2,1}, z_1^3, w_1^{3,1}\} = \{z, y, y, x, x, z\}$, and assumed that $b=z$ (otherwise the con-

tribution is 0). This contribution is the inner product of two vectors, each of dimension s^3 (note, however, that the highest dimensional linear operators involved in this expression are G_{yz}, G_{zx}, which each correspond to systems of two uncoupled neurons).

There are many other useful projections of ρ^F, other than firing rate. For example, we may try and determine the distribution of states associated with a particular neuron b in the presence of the network (denoted by ρ_b^F), by computing

$$\rho_b^F = [I_b \underset{x\neq b}{\otimes} e^{\mathsf{T}}]\rho^F = \rho_b + \sum_{w\neq b} \Delta_{bw}(G_b L_b' \rho_b)(f_w^{\mathsf{T}}\rho_w) + \Delta_{bb}(G_b F^R L_b' F^T \rho_b) + O(\Delta^2).$$

We may also be interested in the distribution of a particular neuron b, conditioned on another neuron a firing (denoted by $\rho_{b|a \text{ fires}}^F$). Such a projection can be computed via

$$\rho_{b|a \text{ fires}}^F = \frac{1}{[a]}[f_a^{\mathsf{T}} \otimes I_b \underset{x\neq b,a}{\otimes} e^{\mathsf{T}}]\rho^F,$$

which immediately allows us to write

$$[a](\rho_{b|a \text{ fires}}^F - \rho_b^F) = +\Delta_{ba}[I_b \otimes f_a^{\mathsf{T}}]G_{ba}[L_b'\rho_b \otimes F^R L_a F^T \rho_a] - \Delta_{ba}(G_b L_b'\rho_b)(f_a^{\mathsf{T}}\rho_a)(f_a^{\mathsf{T}}\rho_a) + \Delta_{ab}[f_a^{\mathsf{T}} \otimes I_b]G_{ab}[L_a'\rho_a \otimes F^R L_b F^T \rho_b] - \Delta_{ab}(f_a^{\mathsf{T}} G_a L_a' \rho_a)(f_b^{\mathsf{T}}\rho_b)\rho_b + O(\Delta^2).$$

This expression implies that even though (to first order) the distribution ρ_b^F depends on the entire connectivity matrix Δ, the conditional distribution $\rho_{b|a \text{ fires}}^F$ is shifted away from ρ_b^F only through the two neuron subnetworks involving a and b. Subnetworks involving a neuron other than a, b do come into play within the higher-order terms.

D. Projections of the filtration

Given ρ^F and L^F, we can write out the distribution associated with any sequence of states. For example, if we let $\mathbb{S}(t)$ denote the state of the system at time t, then the probability that the system is in state $\vec{j}_0$ at a randomly chosen time, and in state $\vec{j}_1$ during the next time step (i.e., a sequence of two states) can be written as $P(\mathbb{S}(t)=\vec{j}_0, \mathbb{S}(t+1)=\vec{j}_1)=L_{\vec{j}_1\vec{j}_0}^F \rho_{\vec{j}_0}^F$, noting that there is no sum over $\vec{j}_0$. Similarly, the probability that the sequence of states $\{\vec{j}_0, \vec{j}_1, \ldots, \vec{j}_\theta\}$ is observed is $P(\mathbb{S}(t+\omega)=\vec{j}_\omega \text{ for } \omega=0,\ldots,\theta) = L_{\vec{j}_\theta\vec{j}_{\theta-1}}^F \cdots L_{\vec{j}_1\vec{j}_0}^F \rho_{\vec{j}_0}^F$. The collection of distributions associated with all state-sequences of arbitrary length is termed the "filtration" of the network [28]. This filtration can be formally expanded order by order as necessary using Eqs. (7) and (13).

By itself, the filtration is a cumbersome object. However, elements of the filtration can be projected in various ways in order to analyze higher-order correlations between neuronal firing events in the network. Within this framework one of the most straightforward projections of the filtration is the projection onto "event-chain rates," which are probabilities that given sequences of firing events occur [4]. For example, the event that a particular neuron a fires and a particular neuron b fires θ time steps later is referred to as a "two-event-chain," and the occurrence rate of this two-event-chain [denoted by $[a\rightarrow_\theta b]=P$ (b fires at time $t+\theta | a$ fires at time t)] is given by

$$[a\rightarrow_\theta b] = \sum_{\vec{j}_0,\vec{j}_1,\ldots,\vec{j}_\theta} \delta_{[j_\theta]_b^V,m} L_{\vec{j}_\theta\vec{j}_{\theta-1}}^F \cdots L_{\vec{j}_1\vec{j}_0}^F \delta_{[j_0]_a^V,m} \rho_{\vec{j}_0}^F,$$

which can otherwise be written as

$$[a\rightarrow_\theta b] = [f_b^{\mathsf{T}} \underset{x\neq b}{\otimes} e^{\mathsf{T}}] \cdot [L^F]^\theta \cdot [F_a^T \underset{x\neq a}{\otimes} I_x] \cdot \rho^F. \quad (17)$$

The occurrence rate of longer event chains, such as the probability that a particular neuron a fires, and a particular neuron b fires θ time steps later, and a particular neuron c fires θ' time steps afterwards (referred to as a three-event-chain, with rate denoted by $[a\rightarrow_\theta b\rightarrow_{\theta'} c]$), can also be projected out of the filtration easily,

$$[a\rightarrow_\theta b\rightarrow_{\theta'} c] = [f_c^{\mathsf{T}} \underset{x\neq c}{\otimes} e^{\mathsf{T}}] \cdot [L^F]^{\theta'} \cdot [F_b^T \underset{x\neq b}{\otimes} I_x] \cdot [L^F]^\theta \cdot [F_a^T \underset{x\neq a}{\otimes} I_x] \cdot \rho^F. \quad (18)$$

More standard measurements of higher-order statistics, such as two-point correlations and the occurrence rates of nearly synchronous firing events, can be obtained by combining/summing over various event-chain rates as necessary (e.g., a two-point correlation between neurons a and b can be constructed via $[a\rightarrow_\theta b]-[a][b]$).

II. DISCRETE STATE EXAMPLES

There are several ways in which our formalism can be used to analyze neuronal networks. By means of illustration, we will use the subnetwork expansion to analyze the event-chain rates (and, as a consequence, the neuronal correlations) in a network of neurons similar to "current-based" integrate-and-fire neurons [22]. We will also use the subnetwork expansion to address network-induced activity in networks of neurons similar to "conductance-based" integrate-and-fire neurons.

A. Analysis of a "current-based integrate-and-fire" network model

For the purposes of this example we will consider a specific discrete-time finite-state neuronal model, in which each neuron has only m states corresponding to the m voltages $V_j=j/(m-1)$ linearly spaced between 0 and 1 [with no notion of conductance built in, equivalent to $n=1$ in Eq. (5)]. We construct this finite-state neuronal model so that it captures many of the qualitative features of the standard current-based integrate-and-fire equation. Given a fixed voltage state j, we define matrices

$$A_{ij} \propto \exp[-(V_j - V_j e^{-\tau})^2/(2\sigma)],$$

$$B_{ij} \propto \exp\{-[V_j - (V_j e^{-\tau} + S)]^2/(2\sigma)\},$$

where τ, σ, S are fixed, and both A and B are normalized to have column sum 1. We set the state-transition matrix for a single uncoupled neuron to be equal to

$$L_{ij}(\eta) = (1-\eta)A_{ij} + \eta B_{ij}, \quad \text{for } j < m,$$

$$L_{ij} = \delta_{i,1}, \quad \text{for } j = m,$$

for inputs $\eta \in [0,1]$. Thus, in the absence of any input a neuron's voltage more or less "decays" (with some probabilistic spread σ) by a factor of $e^{-\tau}$. However, an input of $\eta > 0$ implies that, with probability η, the neuron's voltage will receive a kick of size S in addition to decaying by a factor of $e^{-\tau}$. With this individual neuronal model, we construct a network of N neurons by choosing a coupling matrix Δ_{ab}, a common input $\eta_Z = \eta$ for all neurons, and setting up the full state-transition matrix via Eq. (5). Thus, when any one neuron in the network fires (i.e., transitions from state m to state 1), the other neurons to which it is connected receive a transient increase in their input (over a single time step), which corresponds to a higher chance to receive a voltage kick of size S. Note that when numerically simulating such a system, one never needs to directly apply L^F to the full state of the system (indeed, as L^F is $m^N \times m^N$ dimensional, this operation cannot be carried out on most computers even for moderate values of m and N). Instead, each neuron can be evolved step by step independently of the other neurons—with the exception that, if a neuron starts out in state m, the neurons to which it is connected are evolved (for that time step) with the appropriately corrected input.

1. Analysis of firing rates

In order to demonstrate the subnetwork expansion's application to firing rates, we will first consider a randomly connected network of such neurons (i.e., assume we are given a fixed network for which Δ_{ab} and η_z have both been randomly chosen [29]). Some neurons will fire more than others, and we can use the subnetwork expansion to quantify how often the various neurons will fire. For this situation we can use Eq. (15) to reveal that the firing rate $[a]$ of any neuron a is given (up to second order) by

$$\begin{aligned}[a] = {} & f_a^{\mathsf{T}}\rho_a + \sum_y \Delta_{ay}(f_y^{\mathsf{T}}\rho_y)(f_a^{\mathsf{T}}G_a L_a'\rho_a) \\ & + \sum_{x,y,a \text{ distinct}} \Delta_{ax}\Delta_{xy}(f_y^{\mathsf{T}}\rho_y)(f_a^{\mathsf{T}}G_a L_a'\rho_a)(f_x^{\mathsf{T}}G_x L_x'\rho_x) \\ & + \sum_{x,y,a \text{ distinct}} \Delta_{ax}\Delta_{ay}(f_x^{\mathsf{T}}\rho_x)(f_y^{\mathsf{T}}\rho_y)(f_a^{\mathsf{T}}G_a L_a' G_a L_a'\rho_a) \\ & + \sum_{x,a \text{ distinct}} \Delta_{ax}\Delta_{xa}(f_a^{\mathsf{T}}G_a[[L_a' \otimes f_x^{\mathsf{T}}]G_{ax}[F^R L_a F^T \rho_a \\ & \otimes L_x'\rho_x]]) + \sum_{a,y \text{ distinct}} \Delta_{ay}\Delta_{ay}(f_a^{\mathsf{T}}G_a[[f_y^{\mathsf{T}} \\ & \otimes L_a']G_{ya}[F^R L_y F^T \rho_y \otimes L_a'\rho_a]]) + O(\Delta^3). \end{aligned} \tag{19}$$

Note that $L_z'' = 0$, and so only the first derivative L' appears in this expression. Also note that, as $f^{\mathsf{T}}L' = 0$ for this particular neuronal model, there are no autapse terms in Eq. (19) (indeed, autapses have no effect on the dynamics of the system). We can plug in the connectivity matrix Δ and external inputs η associated with any given network directly into Eq. (19) to approximate the firing rates of each different neuron in that network [note that the dependence on Δ is indicated explicitly in Eq. (19), whereas the terms L_z, L_z', ρ_z, and G_z implicitly depend on η_z]. As a numerical test of this expansion, we set $N=8$, $m=12$, $\tau=0.1$, $\sigma=0.025$, $S=0.5$, and randomly select a set of networks by randomly choosing matrices Δ (each Δ_{ab} drawn from a uniform distribution on the interval [0.0,0.2]) and input η (each η_a drawn from a uniform distribution on the interval [0.225,0.275]). For each of these networks we measure the 2-norm of the difference between the analytically computed firing-rate vector [see Eq. (19)], and the firing-rate vector measured from the full network simulation (lasting 2×10^6 time steps). In Fig. 4(c) we plot the mean ± 1 standard deviation of this error on a log

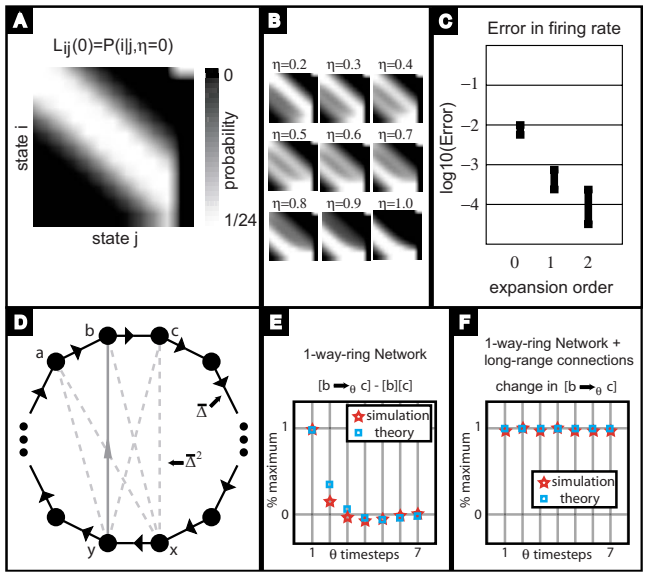

FIG. 4. (Color online) Application of the subnetwork expansion to a current-based integrate-and-fire model. (a) The single-neuron state-transition matrix $L_{ij}(\eta)$ for the neuronal model described in Sec. II A 1, with parameters $m=12$, $\tau=0.1$, $\sigma=0.025$, and $S=0.5$. The shade of the square in the ith row and jth column corresponds to the probability of state j transitioning to state i in one time step (given input $\eta=0$). Note that if a neuron is in the final 12th state, then that neuron fires, and is reset to the first state in the next time step ($L_{i,m}=\delta_{i,1}$). (b) The state- transition matrix L_{ij} for values of $\eta \in [0.2, 1.0]$. The higher the input η, the more likely it is that a neuron will transition to the firing state $i=m$. (c) Error between calculated and simulated firing rates. (d) We consider the "one-way-ring" network (see Sec. II A 2). Neurons are indicated by circles, and short-range connections of strength $\bar{\Delta}$ are indicated by solid black arrows. Weaker long-range connections of strength $\bar{\Delta}^2$ are indicated by gray lines/arrows. The network can be of arbitrary size, as indicated by the ellipses. (e) The correlation between adjacent neurons b and c in the ring can be calculated to first order, which is sufficient to capture the true correlation obtained by numerical simulation. (f) The contribution to $[b \to_\theta c]$ coming from weak $O(\bar{\Delta}^2)$ long- range connections.

scale. The error decreases as the order of the subnetwork expansion increases. Moreover, the second-order approximations to the firing rates calculated via the subnetwork expansion are very close to the measured values, even though the individual Δ_{ab} can be quite large (note that $\Delta_{ab}=1$ is the maximum possible connectivity strength).

2. Analysis of two-point correlations in a one-way-ring network

Next, in order to demonstrate the subnetwork expansion's application to correlated firing events, we will consider a "one-way-ring" network with uniform drive to each of the neurons in the network. Specifically, given N neurons, let neuron z_k be coupled to neuron z_{k+1} (i.e., $\Delta_{Z_{k'}Z_k}=0$ unless $k'=k+1$, in which case $\Delta_{Z_{k+1}Z_k}=\bar{\Delta}$, where $Z_{N+1}=Z_1$), and let η_{z_k} be constant independent of k [see Fig. 4(d)]. If the coupling coefficient $\bar{\Delta}$ is positive, one might expect that any given firing event would subsequently give rise to more firing events "downstream." More specifically, given that neuron c is directly downstream from b in the one-way ring, then one might expect that if $\Delta_{cb}=\bar{\Delta}>0$, then the two-event-chain rate $[b\rightarrow_\theta c]$ would be higher than $[b][c]$ for some values of θ. Using Eq. (17) we can justify (and quantify) this prediction, showing that for this class of networks

$$\begin{aligned}([b\rightarrow_\theta c]-[b][c])/\bar{\Delta} &= [f^{\mathsf{T}}\otimes f^{\mathsf{T}}L^\theta]G_{bc}[F^RLF^T\rho\otimes L'\rho] \\ &\quad -(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}G_zL'\rho)+(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}L^{\theta-1}L'\rho) \\ &\quad +\sum_{\omega=1}^{\omega=\theta-1}(f^{\mathsf{T}}L^{\theta-\omega}F^RLF^T\rho)(f^{\mathsf{T}}L^{\omega-1}L'\rho)+O(\bar{\Delta}).\end{aligned} \tag{20}$$

This correlation is proportional to $\bar{\Delta}$ to first order. Note that since the inputs η_z are independent of z, the terms $L_z, L'_z, \rho_z, G_z, G_{zw}$ are independent of z, w as well (as such, the subscripts have been dropped when unnecessary). By evaluating Eq. (20) (using $m=12$, $\tau=0.1$, $\sigma=0.025$, $S=0.5$, $\eta=0.25$, and assuming $\bar{\Delta}>0$), we can reveal that $[b\rightarrow_\theta c]$ is indeed markedly higher than $[b][c]$, but only for $\theta=1,2$ [see Fig. 4(e)]. For higher values of $\theta\geq 3$, the correlation coefficient $([b\rightarrow_\theta c]-[b][c])/\bar{\Delta}$ decays quite quickly, becoming negative for values of $\theta=4,5$, and falling within a few percent of $[b][c]$ for $\theta>6$. This first-order calculation is quite representative of the correlation obtained by numerical simulation (with $N=8$, and $\bar{\Delta}=0.1$), as shown in Fig. 4(e).

The subnetwork expansion can also be used to compute the change in $[b\rightarrow_\theta c]$ that would result if we were to introduce weak 'long-range' connections into the ring network described above. To clarify the presentation, assume that multiple elements of the connectivity matrix $\Delta_{Z_{k'}Z_k}$ (with $k'\neq k+1$) are set to various nonzero values, but are bounded in magnitude by $\bar{\Delta}^2$ (thus these additional connections are weak in comparison to the connections which form the ring). By expanding Eq. (17) to $O(\bar{\Delta}^3)$, we find that the lowest-order terms that involve the weak subnetwork are those which contain a single weak connection. The influence of any particular weak connection Δ_{by} on the event-chain rate $[b\rightarrow_\theta c]$ involves collections of terms with prefactors Δ_{by}, $\Delta_{cb}\Delta_{by}$, $\Delta_{ba}\Delta_{by}$, and $\Delta_{by}\Delta_{yx}$ (where we have assumed that x,y,a,b,c,d are distinct; x connects strongly to y; a connects strongly to b; and b connects strongly to c). For example, the $O(\bar{\Delta}^2)$ term corresponding to Δ_{by} (for $y\neq b$) has the form $\Delta_{by}(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}GL'\rho)$, which is constant as a function of θ, and the $O(\bar{\Delta}^3)$ term corresponding to $\Delta_{cb}\Delta_{by}$ is given in Appendix A. When evaluated for this particular neuronal model (with values of m, τ, σ, and S given above), these collections of third-order terms reveal that the influence of Δ_{by} on the event-chain rate $[b\rightarrow_\theta c]$ is nearly constant as a function of θ, implying that $O(1/\bar{\Delta})$ multiple positive weak long-range connections to b can substantially increase $[b\rightarrow_\theta c]$, but will not substantially increase the correlation $[b\rightarrow_\theta c]-[b][c]$. This fact can be verified by numerical simulation (with $N=8$, $\bar{\Delta}=0.1$, and the probability of weak connections $p=0.25$), see Fig. 4(f). In addition, by analyzing both $[a]$, $[b]$, $[c]$, $[a\rightarrow_\theta b]$, and $[b\rightarrow_\theta c]$ [via Eq. (19) and Eq. (20)], it can also be shown that multiple positive weak long-range connections to b will (to third order in $\bar{\Delta}$) increase $[a\rightarrow_\theta b]$ and $[b\rightarrow_\theta c]$ by about the same amount (more or less independent of θ), while increasing $[b]$ substantially, and increasing $[c]$ only marginally (again, this fact can be verified by simulation).

B. Analysis of a "conductance-based integrate-and-fire" network model

For the purposes of this example we will consider a specific discrete-time finite-state neuronal model, which is qualitatively similar to the conductance-based integrate-and-fire equations often used in neuroscience [22,26]. Given a network of neurons, the integrate-and-fire equations are a system of ODEs which prescribe the evolution of each model "point-neuron," with state variables $V_a(t)$ and $g_a(t)$,

$$\begin{aligned}\frac{d}{dt}V_a &= -g_LV_a-g_a(V_a-V_{EX})=-(g_a+g_L)[V_a-V_S(g_a)], \\ \frac{d}{dt}g_a &= -g_a/\tau_G+I_a^{in}+\sum_b\Delta_{ab}\delta(t-t_k^b),\end{aligned} \tag{21}$$

where g_L, V_{EX}, and τ_G are constants, the slaving voltage $V_S(g)=(gV_{EX})/(g+g_L)$, and the input I_a^{in} is fixed. The voltage V_a and conductance g_a of any neuron a evolve continuously under Eq. (21), until $V_a=V_T$, at which point neuron a fires, and V_a is reset to $V_R=0$. The time t_k^b is the kth spiketime of neuron b in the network.

To mimic this system of ODEs, we choose a fixed constant for dt, and design a discrete-time neuronal model for which each neuron has m voltage states corresponding to the voltages $V_{j^V}=j^V/(m-1)$ linearly spaced between 0 and 1, and n conductance states corresponding to the n conductances $g_{j^G}=g_{\max}j^G/(n-1)$ linearly spaced between 0 and $g_{\max}$. For each state $j=(j^V,j^G)$, we define V_j^{next} and g_j^{next} as follows:

$$g_j^{\text{next}}(\eta) = g_{jG}\exp(-dt/\tau_G) + \eta\frac{\tau_G}{dt}[1-\exp(-dt/\tau_G)],$$

$$V_j^{\text{next}}(\eta) = V_S(g_j^{\text{next}}) + [V_{jV} - V_S(g_j^{\text{next}})]e^{-dt(g_j^{\text{next}}+g_L)},$$

for any given input η. We set the state-transition probability P (i at time $t+1 \mid j$ at time t) for a single uncoupled neuron to be equal to

$$L_{ij}(\eta) \propto e^{-|V_iV - V_j^{\text{next}}|/\sigma_V} e^{-|g_{iG}-g_j^{\text{next}}|/\sigma_G}, \quad \text{for } j^V < m,$$

$$L_{ij}(\eta) \propto \delta_{i^V,1}\exp(-|g_{iG} - g_j^{\text{next}}|/\sigma_G), \quad \text{for } j^V = m,$$

where L is normalized to have column sums equal to 1. Using this single neuronal model, we define the full state-transition matrix (for any network) via Eq. (5). Note that for the correct ordering of limits $m,n\to\infty$, $\sigma_V,\sigma_G,dt\to 0$, this discrete system is equivalent to the conductance-based integrate-and-fire equations with $I^{in}=0$ (in this discrete system I^{in} is not explicitly modeled). In practice, the discrete system described above with $m\sim 8$ and $n\sim 8$ can be used to construct networks capable of reproducing many qualitative features associated with integrate-and-fire-type dynamics, such as nonlinear gain, bistability, hysteresis, spontaneous synchronization, and oscillations [20,23,24,27].

1. Analysis of autocorrelations

Because of the internal conductance states of the neuronal model, the autapse terms in the associated subnetwork expansion are nonzero. These autapse terms influence the autocorrelations associated with any particular network. As an example, we consider randomly constructed networks for which $g_L=1/20$, $\tau_g=5$, $V_{EX}=14/3$, $m=8$, $n=8$, $g_{\max}=\frac{1}{8}$, $dt=1$, $\sigma_V=\frac{1}{4m}$, $\sigma_G=\frac{g_{\max}}{4n}$, and $\eta_k=\eta^{bkg}=0.001$ is uniform. With this fixed input, it is easy to calculate [via Eq. (15)] that the firing rate $[a]$ of an uncoupled neuron a is $[a]^{[0]}\sim 0.0045$ (i.e., an uncoupled neuron fires about once every 200 steps on average). It can also be shown [via Eq. (17)] that the event-chain rate $[a\to_\theta a]$ is given by

$$\begin{aligned}[a\to_\theta a] &= (f_a^{\mathsf{T}} L_a^\theta F^T \rho_a) + \Delta_{aa}(f_a^{\mathsf{T}} L_a^\theta F^T G_a F^R L_a' F^T \rho_a) \\ &\quad + \sum_{\omega=1}^{\omega=\theta} \Delta_{aa}(f_a^{\mathsf{T}} L_a^{\theta-\omega} F^R L_a' F_a^T L_a^{\omega-1} F^T \rho_a) \\ &\quad + \sum_{d,a\ \text{distinct}} \sum_{\omega=1}^{\omega=\theta} \Delta_{ad}(f_a^{\mathsf{T}} L_a^{\theta-\omega} L_a' L_a^{\omega-1} F^T \rho_a)(f_d^{\mathsf{T}} \rho_d) \\ &\quad + \sum_{d,a\ \text{distinct}} \Delta_{ad}(f_a^{\mathsf{T}} L_a^\theta F^T G_a L_a' \rho_a)(f_d^{\mathsf{T}} \rho_d) + O(\Delta^2).\end{aligned} \tag{22}$$

Notably, even these first-order terms reveal that the effect of any coupling on the event-chain rate $[a\to_\theta a]$ is quite large.

Let us first consider the autapse terms in Eq. (22) associated with the coupling parameter Δ_{aa}. These terms are negligible ($\sim 10^{-7}$) for values of $\theta<5$, but are very large (~ 0.014) for values of $\theta=8,9$. This term (combined with the negligible term $f_a^{\mathsf{T}} L_a^\theta F^T \rho_a$) can be interpreted to imply that the presence of a weak autapse will give rise to an autocorrelation that peaks at around eight time steps. This prediction is consistent with numerical simulations of this system for $\Delta_{aa}=0.0015$ (see Fig. 5). Furthermore, the first-order terms in Eq. (22) corresponding to connections Δ_{ad} (with d,a distinct) also contribute most to the rate $[a\to_\theta a]$ when $\theta=8,9$, as is consistent with numerical simulations (data not shown). Thus, for this example, the first-order subnetwork expansion is sufficient to capture the autocorrelation structure of the network dynamics.

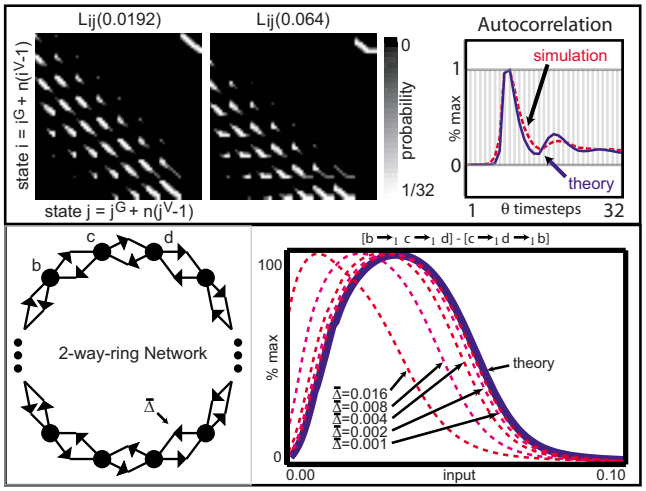

FIG. 5. (Color online) Application of the subnetwork expansion to a conductance-based integrate-and-fire model. (Upper left) The single-neuron state-transition matrix $L_{ij}(\eta)$ for the neuronal model described in Sec. II B 1, with $\eta=0.0192$. The shade of the square in the ith row and jth column corresponds to the probability of state j transitioning to state i in one time step (given input $\eta=0$). Note that if a neuron is in state (m,j^G), then that neuron fires, and is reset in the next time step. The only nonzero entries of $F^R L F^T$ are given by the upper right $n\times n$ block of L, which corresponds to neurons transitioning across the voltage threshold. (Upper center) The state-transition matrix L_{ij} for $\eta=0.064$. (Upper right) We simulate a single neuron with an autapse of strength $\Delta_{aa}=0.0125$, and measure the autocorrelation (dashed line). The structure of this autocorrelation can be predicted (solid line) by the first-order contribution to $[a\to_\theta a]$ coming from the autapse Δ_{aa}. (Lower Left) In order to illustrate an application of the subnetwork expansion to network dynamics, we consider the "two-way-ring" network, composed of neurons with uniform input η and parameters given in Sec. II B 2. Neurons are indicated by circles, and short-range connections of strength $\bar{\Delta}$ are indicated by solid black arrows. The network can be of arbitrary size, as indicated by the ellipses. (Lower right) One way to quantify the wavelike nature of the network's dynamics is to measure the quantity $[b\to_1 c\to_1 d]-[c\to_1 d\to_1 b]$, which is proportional to $\bar{\Delta}$ up to first order, and is a function of the input η. As shown in the graph, this particular indicator of wavelike dynamics has a maximum for $\eta\sim 0.03$. Numerical simulation of this two-way-ring network (for $N=128$ and various values of $\bar{\Delta}\le 0.0175$) confirm this analysis for $\bar{\Delta}\le 0.004$.

2. Analysis of waves in a two-way-ring network

Next we will demonstrate an application of the subnetwork expansion to the analysis of wavelike activity. We will

consider a "two-way-ring" network. Specifically, given N neurons, let neuron z_k be coupled to neuron $z_{k\pm1}$ (i.e., $\Delta_{Z_{k'}Z_k}=0$ unless $k'=k\pm1$, in which case $\Delta_{Z_{k\pm1}Z_k}=\bar{\Delta}$, where $Z_{N+1}=Z_1$ and $Z_0=Z_N$), and let η_{z_k} be constant (independent of k). One important feature of such networks is that certain dynamic regimes give rise to wavelike phenomena (note that the dynamic regime will be a function of the architectural parameters $\bar{\Delta}$, η, as well as the neuronal parameters g_L, τ_g, V_{EX}, m, n, $g_{\max}$, dt, σ_V, and σ_G).

The concept of waves is not that well defined for pulse-coupled systems since there is not generally a straightforward notion of spatial proximity for nodes in a graph [30]. However, it is possible to use event-chains to posit a reasonable definition of waves for this particular class of networks. For clarity, let us label neurons $\{Z_1,Z_2,Z_3,Z_4,Z_5\}$ as $\{a,b,c,d,e\}$. If the event-chain rate $[b\to_\theta c\to_\theta d]$ is significantly higher than the event-chain rate $[c\to_\theta d\to_\theta b]$ for a range of θ, then it is reasonable to assume that sequences of firing events propagate through the network preferentially in the direction of increasing neuronal index or decreasing neuronal index. Note that, due to the symmetry of this simple network, the event-chain rate $[b\to_\theta c\to_\theta d]$ must equal the rate $[d\to_\theta c\to_\theta b]$, and so any long-time observation of the system should reveal equal rates of wave propagation in either direction. Obviously, if $\bar{\Delta}=0$, then both $[b\to_\theta c\to_\theta d]$ and $[c\to_\theta d\to_\theta b]$ are equal to $(f^{\mathsf{T}}\rho)^3$, and there will be no obvious wavefronts within this uncoupled system (since the neuronal model and input η are the same for each of the neurons, we omit subscripts). Using Eq. (18) we can compute the difference $[b\to_\theta c\to_\theta d]-[c\to_\theta d\to_\theta b]$. This difference is proportional to $\bar{\Delta}$ to first order, and is a function of θ, as well as η and the various neuronal parameters. As a specific case (see Appendix A), we can consider the difference $([b\to_1 c\to_1 d]-[c\to_1 d\to_1 b])/\bar{\Delta}$ as a function of η, with uniform drive to each of the neurons in the network, and the rest of the neuronal parameters fixed as: $g_L=1/20$, $\tau_g=15$, $V_{EX}=14/3$, $m=2$, $n=3$, $g_{\max}=\frac{1}{4}$, $dt=1$, $\sigma_V=\frac{1}{4m}$, and $\sigma_G=\frac{g_{\max}}{4n}$. As shown in Fig. 5, there is a critical $\eta_{crit}\sim0.03$ for which this difference is maximal. Numerical simulations (with $\bar{\Delta}\leq0.0175$ and $N=128$) confirm that this analysis is quantitatively accurate for $\bar{\Delta}\leq0.004$.

III. CONTINUOUS-TIME INFINITE-STATE NEURONAL MODELS

This subnetwork expansion can be formally extended to infinite-state continuous-time systems. For purposes of illustration, consider the term $(f^{\mathsf{T}}G_zL'\rho)$ in Eq. (16). In the infinite-state continuous-time limit, the vector ρ_{jVjG} becomes a continuous distribution $\rho(V,g)$, the state-transition matrix L becomes an infinitesimal state-transition operator $\mathcal{L}$, the derivative L' becomes a differential operator $\mathcal{L}'$, the operator $G=(I-L)^{-1}$ becomes a boundary-value problem $\mathcal{G}$, and the operator f^{T} becomes an integral of the probability flux over the $V=V_T$ boundary of the domain. For example, the infinitesimal state-transition operator $\mathcal{L}$ associated with an uncoupled neuron obeying the integrate-and-fire equations [Eq. (21)] is given by

$$\mathcal{L}_{(V_2,g_2),(V_1,g_1)}(\eta,I^{in}) = \lim_{h\to0}\delta([1-\delta_{V_1,V_T}](V_1+h[-g_LV_1-g_1(V_1-V_{EX})]) + \delta_{V_1,V_T}V_R-V_2)\cdot\delta(g_1+h[-g_1/\tau_G+I^{in}]+\eta-g_2),$$

where the parameter η accounts for the pulse coupling. The operator $[I-\mathcal{L}(0,I^{in})]$ is proportional to the differential operator

$$[I-\mathcal{L}(0,I^{in})]\rho(V,g)\propto\frac{d}{dt}\rho(V,g) = -\partial_V J^V(V,g,\rho)-\partial_g J^g(V,g,I^{in},\rho) + \delta(V-V_R)J^V(V_T,g,\rho),$$

where the fluxes J^V,J^g are given by

$$J^V(V,g,\rho)=[-g_LV-g(V-V_{EX})]\rho,\quad \text{for } V>V_R,$$

$$J^V(V,g,\rho)=0,\quad \text{for } V=V_R,$$

$$J^g(V,g,I^{in},\rho)=[-g/\tau_G+I^{in}]\rho.$$

The derivative $\mathcal{L}'$ can be interpreted as

$$\frac{d}{d\eta}\mathcal{L}_{(V_2,g_2),(V_1,g_1)}(0,I^{in}) = \lim_{h\to0}\delta\{[1-\delta_{V_1,V_T}]\{V_1+h[-g_LV_1-g_1(V_1-V_{EX})]\} + \delta_{V_1,V_T}V_R-V_2\}\cdot\lim_{h'\to0}\frac{1}{h'}\{\delta[g_1+h(-g_1/\tau_G+I^{in}) + h'-g_2]-\delta[g_1+h(-g_1/\tau_G+I^{in})-g_2]\}.$$

With this interpretation, the distribution $\mathcal{L}'\psi$ is proportional to $-\frac{d}{dg}\psi$. Note that, since the distribution $\rho(V,g)$ tends to 0 as $g\to0$ and as $g\to\infty$, the integral $\int_{V_R}^{V_T}\int_0^\infty\mathcal{L}'\rho dgdV=0$ as required. The distribution $\hat{\rho}=\mathcal{G}\mathcal{L}'\rho$ satisfies the relationships

$$[I-\mathcal{L}(0,I^{in})]\hat{\rho}(V,g)=\mathcal{L}'\rho(V,g),$$

$$\int_{V_R}^{V_T}\int_0^\infty\hat{\rho}dgdV=0,$$

and the boundary integral $f^{\mathsf{T}}\hat{\rho}$ is given by

$$f^{\mathsf{T}}\hat{\rho}=\int_0^\infty J^V(V_T,g,\hat{\rho})dg.$$

Infinite state examples

Here we analyze two different types of autapse-free random networks: (type-A) randomly constructed networks with sparsity coefficient p and (type-B) all-to-all connected networks with stochastic post-synaptic connection probability p per firing event [i.e., a "failure coefficient" of $(1-p)$]. These networks function differently, yet many current methods for analyzing statistically homogeneous random networks cannot be used to differentiate the dynamics exhibited by these two types of networks [10–12,16,31]. Nevertheless, these dy-

namical differences can be readily analyzed by appealing to our subnetwork expansion. A network of type-A will be determined by setting the elements of the connectivity matrix Δ_{ab} (with $a \neq b$) equal to 0 or $\bar{\Delta}$, with probability $(1-p)$ and p, respectively. A network of type-B will be all-to-all connected (with no autapses), and will have a modified state-transition matrix to account for the fact that each neuron either ignores or receives each incoming network-generated firing event with probability $(1-p)$ and p, respectively. Note that, obviously, autapses would strongly differentiate these two networks—a type-A network neuron which has an autapse will consistently affect itself, whereas a type-B network neuron with an autapse will only affect itself p of the time. Thus, we assume for now that there are no autapses $(\Delta_{aa}=0, \forall a)$.

If we consider the type-B network, one can show that the operators $\mathcal{L}', \mathcal{L}'', \ldots$ are each proportional to p, and are equal to the corresponding operators in the type-A network only when $p=1$. There will be neurons in the type-A network which receive more synaptic input than others. Those neurons will in fact fire more often. In general there will be a distribution of firing rates across the neuronal population in a type-A network. Note, however, that every neuron in a type-B network has exactly the same firing rate. Up to first order, the mean of the firing-rate distribution of type-A networks (sampled across many random networks) will be equal to the mean firing rate exhibited by the type-B network. Indeed, for any neuron b one can show

$$[b]^{[0]} + [b]^{[1]} = (f^{\mathsf{T}}\rho) + \sum_{b,w \text{ distinct}} \Delta_{bw}(f^{\mathsf{T}}\mathcal{G}\mathcal{L}'\rho)(f^{\mathsf{T}}\rho).$$

Thus, on average, a typical type-A neuron receives only $p(N-1)$ of the Δ_{bw} terms, each at full strength, whereas a type-B neuron receives all $N-1$ coupling terms, each at p strength (due to the fact that $\mathcal{L}'$ is proportional to p in a type-B network). Notably however, the second-order contribution to $[b]$, given by

$$\begin{aligned}
[b]^{[2]} = & \sum_{z,w,y \text{ distinct}} \Delta_{zw}\Delta_{zy}(f^{\mathsf{T}}\mathcal{G}\mathcal{L}'\mathcal{G}\mathcal{L}'\rho)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}\rho) \\
& + \sum_{z,w,y \text{ distinct}} \Delta_{zw}\Delta_{zy}\left(\frac{1}{2!}f^{\mathsf{T}}\mathcal{G}\mathcal{L}''\rho\right)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}\rho) \\
& + \sum_{z,w,y \text{ distinct}} \Delta_{zw}\Delta_{wy}(f^{\mathsf{T}}\mathcal{G}\mathcal{L}'\rho)(f^{\mathsf{T}}\mathcal{G}\mathcal{L}'\rho)(f^{\mathsf{T}}\rho) \\
& + \sum_{z,w \text{ distinct}} \Delta_{zw}\Delta_{zw}\left(\frac{1}{2!}f^{\mathsf{T}}\mathcal{G}\mathcal{L}''\rho\right)(f^{\mathsf{T}}\rho) \\
& + \sum_{z,w \text{ distinct}} \Delta_{zw}\Delta_{zw}(f^{\mathsf{T}}\mathcal{G}_z[\mathcal{L}' \otimes f^{\mathsf{T}}]\mathcal{G}_{zw}[\mathcal{L}'\rho \\
& \otimes F^R\mathcal{L}F^T\rho]) + \sum_{z,w \text{ distinct}} \Delta_{zw}\Delta_{wz}(f^{\mathsf{T}}\mathcal{G}_z[\mathcal{L}' \\
& \otimes f^{\mathsf{T}}]\mathcal{G}_{zw}[F^R\mathcal{L}F^T\rho \otimes \mathcal{L}'\rho]),
\end{aligned}$$

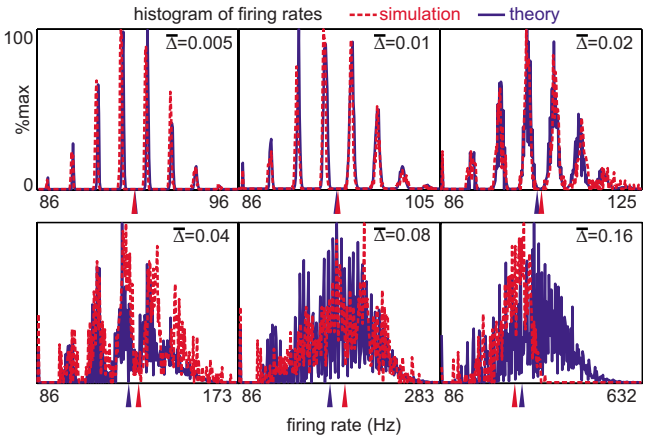

FIG. 6. (Color online) The various panels show the histograms of time-averaged neuronal firing rates produced by type-A random networks with $p=0.5$ (see text) as calculated to second-order using our subnetwork expansion (solid) and via simulation (dotted). The shaded arrowheads (blue and red online) below the firing-rate axis indicate the calculated and simulated time-averaged neuronal firing rates produced by type-B networks with $p=0.5$. Each random network is composed of $N=8$ conductance-based integrate-and-fire neurons obeying Eq. (21), with $G_L=0.05$ ms^{-1}, $V_{EX}=14/3$, $\tau_G=2.0$ ms, with each I_a^{in} given by $I_a^{in}(t)=f\Sigma_k\delta(t-t_{a,k}^{in})$, with strength $f=0.0365$ and each set of feedforward input spikes $\{t_{a,k}^{in}\}_k$ drawn from an independent Poisson process with rate 500 Hz. Note that the theory matches the simulations very well for $\bar{\Delta} \leq 0.04$.

is different for type-A and type-B networks. Specifically, the typical type-A neuron receives $p^2(N-1)(N-2)$ of the following terms:

$$\sum_{z,w,y \text{ distinct}} \Delta_{zw}\Delta_{zy}\left(\frac{1}{2!}f^{\mathsf{T}}\mathcal{G}\mathcal{L}''\rho\right)_z (f^{\mathsf{T}}\rho)_w(f^{\mathsf{T}}\rho)_y,$$

each at full strength, whereas a type-B neuron receives $(N-1)(N-2)$ of these terms at strength p. Similarly, a type-A neuron receives $p(N-1)$ of the following terms:

$$\sum_{z,w \text{ distinct}} \Delta_{zw}\Delta_{zw}(f_z^{\mathsf{T}}\mathcal{G}_z[\mathcal{L}'_z \otimes f_w^{\mathsf{T}}]\mathcal{G}_{zw}[\mathcal{L}'_z\rho_z \otimes F\mathcal{L}_w\rho_w]),$$

each at full strength, whereas a type-B neuron receives $(N-1)$ such terms at strength p^2. Thus, the subnetwork expansion can be used to show that the mean of the firing-rate distribution of type-A networks is not equal to the mean firing rate of type-B networks. Moreover, by computing the appropriate terms, we can predict (to second order) the firing-rate distribution of neurons within type-A networks, as well as the firing rate of neurons within type-B networks (see Fig. 6). This same type of analysis can also be used to investigate higher-order correlations within these networks.

IV. CONCLUSIONS

This subnetwork expansion provides a systematic step towards an understanding of the relationships between the architecture of a network and the dynamics of that network.

This framework can provide substantial insight into the dynamics of many commonly studied network architectures. Ultimately, this framework provides a map (of chosen order) between the specific architectural properties of a network, and various projections of the network dynamics. If this map can be approximately inverted (e.g., when the expansion order is low), then the architecture of the network can be deduced by observing/measuring various statistical features of the dynamics. For example, by expanding the firing rates and two-event chains out to second order, we can construct a quadratic map between the connectivity matrix Δ and the correlation matrix $C_{ab}(\theta)=[a\rightarrow_\theta b]-[a][b]$ for fixed θ. The inversion of this map corresponds to a "best guess" for the connectivity matrix given the observed correlations. Note that this procedure does not require perturbation of the system, and may potentially provide a useful analytical tool which can complement noninvasive experiments.

We note that computation of many of the higher-order terms involved in any diagrammatic expansion may be difficult, and may require either a numerical solution of the associated population-dynamics equations [32], or a numerical simulation of smaller networks with similar high-order terms. However, once the terms in the expansion are computed, they can be used to dissect the full network, and associate causality with each particular class of subnetwork. In addition, the diagrammatic terms can be used to analyze a range of network connectivities by recomputing the appropriate sum derived from Eq. (13).

Finally, while the subnetwork expansion detailed within this paper involves an expansion around the uncoupled dynamic regime (i.e., a weak-coupling expansion in terms of Δ around $\Delta=0$), the same formal structure can be applied towards an expansion around another dynamic regime for which the equilibrium distribution and state-transition matrix are known. For example, if the network contains many neurons, and we make the assumption that the coupling matrix is all-to-all, and that the coupling is weak, the equilibrium state can be approximated via kinetic equations derived in the mean-field limit [13,15]. The full network dynamics can then be approximated using an expansion around the state-transition operator associated with this mean-field limit.

ACKNOWLEDGMENTS

This work was supported by a grant from the Schwartz Foundation and NSF Grants.

APPENDIX A: VARIOUS EXPANSIONS

The contribution to $[b\rightarrow_\theta c]$ from the term $\Delta_{cb}\Delta_{by}$ in the current-based one-way ring network described in Sec. II A 2 is given by (up to third order in $\bar{\Delta}$)

$$\Delta_{cb}\Delta_{by}\Bigg\{(f^{\mathsf{T}}GL'\rho)(f^{\mathsf{T}}L^{\theta-2}L'\rho)(f^{\mathsf{T}}\rho)+\sum_{\substack{\omega_1+\omega_2+\omega_3=\theta-3,\\ \omega_3>0}}(f^{\mathsf{T}}L^{\omega_1}L'\rho)(f^{\mathsf{T}}L^{\omega_2}L'L^{\omega_3}F^RLF^T\rho)(f^{\mathsf{T}}\rho)+\sum_{\substack{\omega_1+\omega_2=\theta-2,\\ \omega_2>0}}(f^{\mathsf{T}}L^{\omega_2}F^RLF^TGL'\rho)$$
$$\times(f^{\mathsf{T}}L^{\omega_1}L'\rho)(f^{\mathsf{T}}\rho)+(f^{\mathsf{T}}\rho)([f^{\mathsf{T}}L^{\alpha}\otimes f^{\mathsf{T}}]G_{ba}[L'\rho\otimes F^RLF^TGL'\rho])+(f^{\mathsf{T}}\rho)([f^{\mathsf{T}}\otimes f^{\mathsf{T}}L^{\alpha}]G_{ab}[L'\otimes L]G_{ab}[F^RLF^T\rho\otimes L'\rho])\Bigg\}.$$

The two-event-chain rates $[a\rightarrow_\theta b]$ for the conductance-based system described in Sec. II B are given (up to first-order) by

$$[a\rightarrow_\theta b]=[a][b]+\Delta_{ba}\sum_{\omega=1}^{\omega=\theta}(f_b^{\mathsf{T}}L_b^{\theta-\omega}L_b'\rho_b)(f_a^{\mathsf{T}}L_a^{\omega-1}F^T\rho_a)+\Delta_{ab}\cdot[f_a^{\mathsf{T}}\otimes f_b^{\mathsf{T}}L^{\theta}]G_{ab}[L_a'\rho_a\otimes F^RL_bF^T\rho_b]$$
$$+\Delta_{ba}\cdot[f_b^{\mathsf{T}}L^{\theta}\otimes f_a^{\mathsf{T}}]G_{ba}[L_b'\rho_b\otimes F^RL_aF^T\rho_a],$$

in the case that a and b are distinct.

The difference between the three-event-chain rates $[b\rightarrow_1 c\rightarrow_1 d]$ and $[c\rightarrow_1 d\rightarrow_1 b]$ for the two-way-ring network described in Sec. II is given (up to first-order) by

$$([b\rightarrow_1 c\rightarrow_1 d]-[c\rightarrow_1 d\rightarrow_1 b])/\bar{\Delta}=(f^{\mathsf{T}}L'\rho)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}\rho)-(f^{\mathsf{T}}L'\rho)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}LF^T\rho)-(f^{\mathsf{T}}LL'\rho)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}\rho)$$
$$+(f^{\mathsf{T}}LF^RLF^T\rho)(f^{\mathsf{T}}L'\rho)(f^{\mathsf{T}}\rho)-(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}L'\rho)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}\rho)-(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}LG_zL'\rho)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}\rho)$$
$$+(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}\rho)(f^{\mathsf{T}}G_zL'\rho)(f^{\mathsf{T}}\rho)+[f^{\mathsf{T}}LL\otimes f^{\mathsf{T}}L]G_{zw}[L'\rho\otimes F^RLF^T\rho](f^{\mathsf{T}}\rho)$$
$$-[f^{\mathsf{T}}LL\otimes f^{\mathsf{T}}]G_{zw}[L'\rho\otimes F^RLF^T\rho](f^{\mathsf{T}}\rho)+[f^{\mathsf{T}}L\otimes f^{\mathsf{T}}LL]G_{zw}[L'\rho\otimes F^RLF^T\rho](f^{\mathsf{T}}\rho)$$
$$-[f^{\mathsf{T}}\otimes f^{\mathsf{T}}LL]G_{zw}[L'\rho\otimes F^RLF^T\rho](f^{\mathsf{T}}\rho),$$

where we have omitted subscripts since the internal dynamics associated with every neuron in the network is identical (z and w are dummy variables indicating the dimension of the G operator).

APPENDIX B: A NOTE ON THE CALCULATION OF $G_{\{z_j\}}$ for $|\{z_j\}|>1$

Here we address one important technical point relating to the actual computation of the terms in the subnetwork expansion. The operator $G_{\{z_j\}}=(I-\bigotimes_{z_j} L_{z_j})^{-1}$ is $s^{|\{z_j\}|}\times s^{|\{z_j\}|}$ dimensional, which raises numerical computation issues even for low values of $|\{z_j\}|\gtrsim 2$. For example, consider the operator G_{ab} appearing in the terms $G_{ab}[L'\rho\otimes F^R L F^T\rho]$ associated with the second-order expansion of firing rate. Even for moderate values of $s\sim 10\text{^}3$ (or, equivalently, a discretization of a continuous-time problem involving $10\text{^}3$ gridpoints), direct computation of G_{ab} requires manipulating a $10\text{^}6\times 10\text{^}6$ matrix, and direct computation of higher-order terms (such as G_{abc}) becomes practically impossible. Fortunately, the terms arising in the subnetwork expansion have a specific structure—namely that every term involving G only incorporates the operators G, F^R, F^T, and $L^{[M]}$ (for various M). Thus, if one were able to express G_{ab} as the direct product of two single-neuron operators (say, $G_{ab}\sim A\otimes B$), then we could successfully reduce every operation involving G_{ab} down to the product of single-neuron operations (e.g., one could represent $G_{ab}[L'\otimes L]G_{ab}[F^R L F^T\rho\otimes L'\rho]$ as $(AL'AF^R L F^T\rho)(BLBL'\rho)$, which only involves operators of size $s\times s$). Unfortunately, this is not quite possible since G_{ab} has no representation as the direct product of two single-neuron operators. However, we can express G_{ab} as a rapidly converging series of direct products of single-neuron operators.

Using the notation of Sec. I B, we first apply the coordinate transformation

$$(I-\Sigma_a\otimes\Sigma_b)^{-1}=[\Psi_a^{-1}\otimes\Psi_b^{-1}]G_{ab}[\Psi_a\otimes\Psi_b].$$

We have that $\Lambda^{ab}=\mathrm{diag}((I-\Sigma_a\otimes\Sigma_b)^{-1})$ is an s^2-element vector such that $\Lambda^{ab}_{(1,1)}=0$, and $\Lambda^{ab}_{(i,j)}=(1-\sigma_i^a\sigma_j^b)^{-1}$ as long as either $i\neq 1$ or $j\neq 1$. Letting $\Lambda=\Lambda^{ab}_{(i,j)}-\Lambda_i^a\mathbf{1}_j-\mathbf{1}_i\Lambda_j^b$ (where $\mathbf{1}$ is the s-element unit vector with first entry 1), we note that $\Lambda_{(i,j)}=(1-\delta_{i0})(1-\delta_{j0})(1-\sigma_i^a\sigma_j^b)^{-1}$. Note that $\Lambda\in\mathbb{C}^{s^2}$ is an s^2-element vector, $\mathrm{diag}(\Lambda)\in\mathbb{C}^{s^2\times s^2}$ is an $s^2\times s^2$ matrix, and

$$G_{ab}=G_a\otimes(\rho_b e_b^{\mathsf{T}})+(\rho_a e_a^{\mathsf{T}})\otimes G_b+[\Psi_a\otimes\Psi_b]\mathrm{diag}(\Lambda)[\Psi_a^{-1}\otimes\Psi_b^{-1}].$$

As the diagonal matrix $\mathrm{diag}(\Lambda)$ only has s^2 relevant entries, there is another natural way to view $\mathrm{diag}(\Lambda)$—namely as an $s\times s$ matrix $\Lambda'\in\mathbb{C}^{s\times s}$, where $\Lambda'_{i,j}=\Lambda_{(i,j)}$. A key observation is that any representation of $\mathrm{diag}(\Lambda)$ in terms of a sum of direct products of $s\times s$ diagonal matrices is entirely equivalent to a representation of Λ in terms of a sum of direct products of $s\times 1$ vectors, which is again entirely equivalent to a representation of Λ' in terms of a sum of outer products of $s\times 1$ vectors. More specifically, given any approximation of Λ' in terms of a sum of outer products of s-element vectors $\theta^{[\mu]}$, $\phi^{[\mu]}\in\mathbb{C}^s$,

$$\Lambda'_{i,j}=\theta_i^{[1]}[\phi_j^{[1]}]^*+\theta_i^{[2]}[\phi_j^{[2]}]^*+\cdots,$$

we can immediately write

$$\Lambda_{(i,j)}=\theta_i^{[1]}\bar{\phi}_j^{[1]}+\theta_i^{[2]}\bar{\phi}_j^{[2]}+\cdots,$$

which corresponds to a representation of G_{ab},

$$G_{ab}=G_a\otimes(\rho_b e_b^{\mathsf{T}})+(\rho_a e_a^{\mathsf{T}})\otimes G_b+\sum_\mu[\Psi_a\,\mathrm{diag}(\theta^{[\mu]})\Psi_a^{-1}]\otimes[\Psi_b\,\mathrm{diag}(\bar{\phi}^{[\mu]})\Psi_b^{-1}].$$

A straightforward approach is to choose $\theta^{[\mu]},\phi^{[\mu]}$ for each μ so that

$$\left|\Lambda-\sum_{l=1}^{l=\mu}\theta^{[l]}\otimes\phi^{[l]}\right|_2\sim\left|\Lambda'-\sum_{l=1}^{l=\mu}\theta^{[l]}[\phi^{[l]}]^*\right|_2$$

is as small as possible. One way to do this is to iteratively solve the following eigenvalue problems:

$$\tilde{\Lambda}_{i,j}=\Lambda'_{i,j}-\sum_{l=1}^{l=\mu-1}\theta_i^{[l]}[\phi_j^{[l]}]^*,$$

$$\lambda^{[\mu]}\begin{bmatrix}\theta^{[\mu]}\\ \phi^{[\mu]}\end{bmatrix}=\begin{bmatrix}&\tilde{\Lambda}\\ \tilde{\Lambda}^*&\end{bmatrix}\begin{bmatrix}\theta^{[\mu]}\\ \phi^{[\mu]}\end{bmatrix},\tag{B1}$$

where, at each step, $\tilde{\Lambda}$ is an $s\times s$ matrix formed from the elements of $\Lambda'-\Sigma_{l=1}^{l=\mu-1}\theta^{[l]}[\phi^{[l]}]^*$. With this formulation the largest eigenvalue $\lambda^{[\mu]}=\|\theta^{[\mu]}\|^2=\|\phi^{[\mu]}\|^2$. This procedure is identical to constructing a singular value decomposition of Λ' [33]—indeed, $\theta^{[\mu]}$ and $\phi^{[\mu]}$ are parallel to the left and right singular vectors of Λ', with squared norms equal to the corresponding singular values of Λ'. If $L_z=L_w$ (i.e., the neurons z and w are identical), then $\theta^{[\mu]}=\phi^{[\mu]}$, and the eigenvalue problem in Eq. (B1) reduces to $\lambda^{[\mu]}\theta^{[\mu]}=\tilde{\Lambda}\theta^{[\mu]}$. In practice this process converges very quickly, and $|\tilde{\Lambda}|$ typically decreases by an order of magnitude with every iteration. For instance, for the discrete-state examples given in this paper, G_{ab} can be represented to double-precision accuracy by a sum of $\lesssim 4$ such direct products. Similarly, for a continuous-time infinite-state current-based integrate-and-fire system, G_{ab} can be represented to double-precision accuracy by a sum of ≤ 16 such direct products (even for numerical discretizations which incorporate far more than 16 voltage states per neuron).

[1] W. McCulloch and W. Pitts, Bull. Math. Biophys. **7**, 115 (1943).
[2] B. Quenet, R. Dubois, S. Sirapian, G. Dreyfus, and D. Horn, Biosystems **67**, 203 (2002).
[3] D. Stroock, *An Introduction to Markov Processes* (Springer, New York, 2005).
[4] A. Rangan, D. Cai, and D. McLaughlin, Proc. Natl. Acad. Sci. U.S.A. **105**, 10990 (2008).
[5] J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).
[6] J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **81**, 3088 (1984).
[7] A. Arieli, D. Shoham, R. Hildesheim, and A. Grinvald, J. Neurophysiol. **73**, 2072 (1995).
[8] A. Arieli, A. Sterkin, A. Grinvald, and A. Aertsen, Science **273**, 1868 (1996).
[9] M. Tsodyks, T. Kenet, A. Grinvald, and A. Arieli, Science **286**, 1943 (1999).
[10] B. Knight, J. Gen. Physiol. **59**, 734 (1972).
[11] L. F. Abbott and C. van Vreeswijk, Phys. Rev. E **48**, 1483 (1993).
[12] B. Knight, D. Manin, and L. Sirovich, in *Symposium on Robotics and Cybernetics: Computational Engineering in Systems Applications*, edited by E. Gerf (Cite Scientifique, Lille, France, 1996).
[13] D. Nykamp and D. Tranchina, J. Comput. Neurosci. **8**, 19 (2000).
[14] L. Sirovich, A. Omurtag, and B. Knight, SIAM J. Appl. Math. **60**, 2009 (2000).
[15] D. Cai, L. Tao, M. Shelley, and D. Mclaughlin, Proc. Natl. Acad. Sci. U.S.A. **101**, 7757 (2004).
[16] W. Gerstner, Phys. Rev. E **51**, 738 (1995).
[17] M. A. Buice and J. D. Cowan, Phys. Rev. E **75**, 051919 (2007).
[18] M. A. Buice and C. C. Chow, Phys. Rev. E **76**, 031118 (2007).
[19] D. Hansel, G. Mato, C. Meunier, and L. Neltner, Neural Comput. **10**, 467 (1998).
[20] N. Brunel and V. Hakim, Neural Comput. **11**, 1621 (1999).
[21] G. D. Smith, C. Cox, S. Sherman, and J. Rinzel, J. Neurophysiol. **83**, 588 (2000).
[22] P. Dayan and L. Abbott, *Theoretical Neuroscience* (MIT press, Cambridge, MA, 2001).
[23] N. Fourcaud and N. Brunel, Neural Comput. **14**, 2057 (2002).
[24] A. Casti, A. Omurtag, A. Sornborger, E. Kaplan, B. Knight, J. Victor, and L. Sirovich, Neural Comput. **14**, 957 (2002).
[25] A. Rauch, G. LaCamera, H. Luscher, W. Senn, and S. Fusi, J. Neurophysiol. **90**, 1598 (2003).
[26] T. Vogels and L. Abbott, J. Neurosci. **25**, 10786 (2005).
[27] D. French and E. Gruenstein, J. Comput. Neurosci. **21**, 227 (2006).
[28] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications* (Springer, Berlin, 2003).
[29] P. Erdos and A. Renyi, Publ. Math. (Debrecen) **6**, 290 (1959).
[30] J. Bondy and U. Murty, *Graph Theory with Applications* (Elsevier, New York, 1976).
[31] A. V. Rangan and D. Cai, Phys. Rev. Lett. **96**, 178101 (2006).
[32] A. Rangan and D. Cai, J. Comput. Neurosci. **22**, 81 (2007).
[33] L. Trefethen and D. Bau, *Numerical Linear Algebra* (SIAM, Philadelphia, 1997).