J. Math. Biol. (2009) 59:147–173
DOI 10.1007/s00285-008-0224-9

Mathematical Biology

A stimulus-dependent connectivity analysis of neuronal networks

Duane Q. Nykamp

Received: 24 March 2008 / Revised: 15 August 2008 / Published online: 2 October 2008

Abstract We present an analysis of interactions among neurons in stimulus-driven networks that is designed to control for effects from unmeasured neurons. This work builds on previous connectivity analyses that assumed connectivity strength to be constant with respect to the stimulus. Since unmeasured neuron activity can modulate with the stimulus, the effective strength of common input connections from such hidden neurons can also modulate with the stimulus. By explicitly accounting for the resulting stimulus-dependence of effective interactions among measured neurons, we are able to remove ambiguity in the classification of causal interactions that resulted from classification errors in the previous analyses. In this way, we can more reliably distinguish causal connections among measured neurons from common input connections that arise from hidden network nodes. The approach is derived in a general mathematical framework that can be applied to other types of networks. We illustrate the effects of stimulus-dependent connectivity estimates with simulations of neurons responding to a visual stimulus.

Keywords Neural networks · Correlations · Causality · Penalized likelihood

Mathematics Subject Classification (2000) 92C20

1 Introduction

We have recently developed a framework for estimating the connectivity among nodes in a network [12–15]. The key feature of this approach is its ability to control for

This research was supported by the National Science Foundation grants DMS-0415409 and DMS-0748417.

D. Q. Nykamp (✉)
School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA
e-mail: nykamp@math.umn.edu

the effects of connections from hidden nodes. Because of the hidden nodes, we frame our goal in terms of identifying *causal connections* among measured nodes. A causal connection is a direct connection between measured nodes or an indirect connection consisting of a chain of connections via hidden nodes through which one measured node influences the other. We do not attempt to distinguish a direct connection from an indirect connection but refer to either as simply a causal connection.

The presence of hidden nodes can corrupt estimates of causal connections among the measured nodes. For example, if two measured nodes receive common input connections from a hidden node, the activity of those nodes may be correlated in a way that mimics a causal connection, even though neither measured node causally influences the other. We were able to successfully distinguish between causal connections and such unmeasured common input. However, the success relied on making assumptions about the hidden nodes [13,14].

The most significant assumption was about the relationship between the activity of the hidden nodes and the measured external variables, such as a stimulus. Since by definition, we have no data from the hidden nodes, we cannot determine any of their properties. It turns out that we need not estimate many of these properties, such as how the activity of a hidden node is influenced by its own history. However, in the analysis to determine network connectivity, we obtain equations containing factors representing the relationship between the stimulus and the activity of hidden nodes. In previous work, we effectively ignored [13,14] any such relationship in order to solve for common input from hidden nodes.

Despite such a seemingly drastic assumption, the method worked surprisingly well. The analysis was able to accurately distinguish common input from causal connection as long as the hidden common input nodes responded differently to the stimulus than did the measured nodes. If a hidden common input node happened to respond to the stimulus in a similar manner, then this common input might be misidentified as a causal connection between the measured nodes. Since one cannot verify the presence or absence of such a similar common input node, the results contain a degree of ambiguity in the identification of causal connections. We will illustrate this ambiguity more precisely in the context of simulation results (see Sect. 3).

In this paper, we demonstrate that one can eliminate this ambiguity if one repeats a stimulus many times and allows the connectivity strength to vary in time with the stimulus. In Sect. 2, we develop the analysis underlying these stimulus-dependent connectivity measures. In Sect. 3, we test the results with simulations of neurons responding to a visual stimulus, and we discuss the results in Sect. 4.

2 The analysis

2.1 The model used in the analysis

Since we are interested in the case where a stimulus is repeated many times, we can avoid postulating a model of how nodes respond to the stimulus. Instead, we use history and histogram (HAH) models [15] that allow an arbitrary dependence of node

activity on stimulus time (though not on stimulus repeat). Our primary application of these methods is in neuroscience, and we use the name HAH because the model can be viewed as based on a neuron's peristimulus time histogram (PSTH) combined with the dependence of the neuron's activity on its own spiking history.

Let $R_s^{k,i}$ be the activity of node s at stimulus time i (we use discrete time) and repeat k. Eventually, we will let $R_s^{k,i}$ be a Bernoulli random variable indicating the presence or absence of a spike of a neuron in that time bin. For now, we will not specify the probability distribution as the analysis applies to more general networks, allowing the activity of a node be a continuous or discrete random variable. (For simplicity, we use the notation of a discrete random variable.)

We allow the probability distribution of $R_s^{k,i}$ to depend in an arbitrary manner on stimulus time i. In addition, it could depend on the activity of all (measured and hidden) nodes for all times before i during stimulus repeat k, which we denote by[1] $\mathbf{R}^{k,<i}$. We separate out the history of node s itself, which we denote by the vector $\mathbf{R}_s^{k,<i}$. At this point, we will allow an arbitrary dependence on the node's own history but assume that inter-nodal coupling adds linearly. The resulting HAH model for the probability distribution of the activity of node s can be written as

$$\Pr(R_s^{k,i} = r_s^{k,i} \mid \mathbf{R}^{k,<i} = \mathbf{r}^{k,<i}) = \bar{P}_s \left(r_s^{k,i}; i, \mathbf{r}_s^{k,<i}, \sum_{\tilde{s} \neq s} \sum_{j>0} \bar{W}_{\tilde{s},s}^{j,i} r_{\tilde{s}}^{k,i-j} \right), \quad (1)$$

where $\bar{P}_s$ is a probability distribution in its first argument. The remaining three arguments determine how $R_s^{k,i}$ depends on stimulus time, the history of node s, and the total coupling from all other nodes. To implement this approach, we will need to specify how these arguments influence the probability distribution; for now, we leave the formulation generic.

Note that the coupling kernel $\bar{W}_{\tilde{s},s}^{j,i}$ is allowed to depend not only on delay j but also on stimulus time point i. $\bar{W}_{\tilde{s},s}^{j,i}$ determines how the activity of node s at stimulus time i is influenced by the activity of node $\tilde{s}$ from j time steps previous. We have assumed that the coupling strength is independent of stimulus repeat. We refer to this variable connectivity strength as stimulus-dependent connectivity.

We assume that the activity of each node was given by an equation of the form (1). Since in this model, all interactions among nodes are delayed by at least one time step, the random variables for the activity of all nodes in a single time step are independent, conditioned on the stimulus and the history of all nodes. Hence, we can repeatedly use Bayes' rule to expand the probability distribution of the activity of all nodes, denoted $\mathbf{R}$, into factors of the form of (1). The result is

[1] If the stimulus is repeated with no break in between presentations, we can allow $R_s^{k,i}$ to depend on the $R_{\tilde{s}}^{\tilde{k},\tilde{i}}$ for previous stimulus repeats $\tilde{k} < k$. However, for simplicity of notation, we ignore any such "wrapping" of the time from the end of one stimulus repeat to the beginning of the next.

$$\Pr(\mathbf{R} = \mathbf{r}) = \prod_{s,k,i} \Pr(R_s^{k,i} = r_s^{k,i} \mid \mathbf{R}^{k,<i} = \mathbf{r}^{k,<i})$$

$$= \prod_{s,k,i} \bar{P}_s \left(r_s^{k,i}; i, \mathbf{r}_s^{k,<i}, \sum_{\tilde{s} \neq s} \sum_{j>0} \bar{W}_{\tilde{s},s}^{j,i} r_{\tilde{s}}^{k,i-j} \right). \tag{2}$$

In the following analysis, we assume that the activity of all nodes was generated by (2).

2.2 Sketch of the analysis

Our goal is to use (2) to estimate the connectivity $\bar{W}$ among measured nodes in the network. The only data we have at our disposal is the stimulus and the activity of measured nodes, which we denote by $\mathbf{R}_{\mathcal{Q}} = \{R_q^{k,i} \mid q \in \mathcal{Q}\}$, where $\mathcal{Q}$ is the set of indices corresponding to measured nodes. However, the $\mathbf{R}$ of (2) also includes the activity of hidden nodes, which we denote by $\mathbf{R}_{\mathcal{P}} = \{R_p^{k,i} \mid p \in \mathcal{P}\}$, where $\mathcal{P}$ is the set of indices corresponding to hidden nodes. Clearly, we cannot hope to determine the parameters of (2), as it is underconstrained by the available data.

Our strategy to estimate connectivity is to average out the activity of hidden nodes from (2) and fit the resulting equation to the activity of the measured nodes. In other words, we marginalize the full probability $\Pr(\mathbf{R})$ to just the probability distribution $\Pr(\mathbf{R}_{\mathcal{Q}})$ of measured nodes.

The analysis follows that of Ref. [14], so we just briefly summarize the steps here. We assume that the coupling $\bar{W}$ is weak (i.e., the $\bar{W}_{\tilde{s},}^{j,i}$ are small parameters) so that we can expand (2) into a second-order Taylor series in the $\bar{W}_{\tilde{s},}^{j,i}$. Each term of the resulting expression is simply a polynomial in the r_s^i times the probability distributions $\bar{P}_s$ (or derivatives of the $\bar{P}_s$). Since each term is so simple, we can analytically compute the averages over any of the r_s^i. In other words, we can derive an analytic expression for the marginal probability distribution $\Pr(\mathbf{R}_{\mathcal{Q}})$ in terms of the quantities from (2).

To write the probability distribution $\Pr(\mathbf{R}_{\mathcal{Q}})$ in terms of known parameters, we marginalize (2) over the activity of all nodes except one. Hence, we obtain an expression for $\Pr(\mathbf{R}_s)$, the probability distribution for the activity $\mathbf{R}_s$ of node s.

Note that the probability distributions in (2) have a bar over them. This bar indicates that these are the original probability distributions that we assume generated the activity of the nodes. We cannot determine those probability distributions as they depend on the hidden variables of the hidden nodes. We can, however, separately fit effective probability distributions to the activity $\mathbf{R}_s$ of each measured node, which we write without a bar,

$$\Pr(\mathbf{R}_s = \mathbf{r}_s) = \prod_{k,i} P_s \left(r_s^{k,i}; i, \mathbf{r}_s^{k,<i}, 0 \right). \tag{3}$$

When fitting the effective probability distributions (3), we ignore the fact that the nodal activity $\mathbf{R}_s$ depends on the activity of other nodes due to coupling (this is just a restatement of the fact that the left hand size of (3) is the marginal distribution of

$\mathbf{R}_s$). To reflect the assumption in this fit, we let the coupling argument be zero at this stage. Since the activity $\mathbf{R}_s$ used to fit the effective probability distributions (3) were influenced by the coupling, (3) includes the average effect of the coupling. (Fitting a model to the activity of just one node corresponds to effectively averaging over the activity of the other nodes.) Hence, when we reintroduce coupling terms back into (3), below, such coupling terms will reflect deviations from the average effect of the coupling.

The assumptions on the P_s required for the analysis are detailed in Ref. [14]. The main assumptions are that the underlying model is identifiable (meaning we can determine all parameters from measurements of $\mathbf{R}_s$ and the stimulus) and that the coupling argument is chosen so that we can determine $P_s(r_s^{k,i}; i, \mathbf{r}_s^{k,<i}, w)$ for any w once we fit model (3) from $\mathbf{R}_s$ alone.

By equating (3) to the marginalization of (2) for $\Pr(\mathbf{R}_s)$, we derive an expression for the original probability distributions $\bar{P}_s$ in terms of the effective probability distribution P_s. Then, we can rewrite our expression for the probability distribution $\Pr(\mathbf{R}_\mathcal{Q})$ for measured activity $\mathbf{R}_\mathcal{Q}$ in terms of the effective probability distributions. We end up with the equation[2]

$$\Pr(\mathbf{R}_\mathcal{Q} = \mathbf{r}_\mathcal{Q}) \approx \prod_{q\in\mathcal{Q}} \prod_{k,i} P_q \left(r_q^{k,i}; i, \mathbf{r}_q^{k,<i}, \sum_{\substack{\tilde{q}\in\mathcal{Q}\\ \tilde{q}\neq q}} \sum_{j>0} W_{\tilde{q},q}^{j,i} \left[r_{\tilde{q}}^{k,i-j} - E_0(R_{\tilde{q}}^{k,i-j}) \right] \right.$$
$$\left. + \sum_{\substack{\tilde{q}\in\mathcal{Q}\\ \tilde{q}\neq q}} \sum_{j>0} U_{\tilde{q},q}^{j,i} \frac{\partial P_{\tilde{q}}^{k,i-j}}{\partial w} \frac{1}{P_{\tilde{q}}^{k,i-j}} \right). \tag{4}$$

Model (4) is in the same form as the original network equations (2) except that it is in terms of the measured node activity $\mathbf{R}_\mathcal{Q}$, the effective models (3), and additional quantities which we now define.

We define the effective expected value $E_0(\cdot)$ to be the expected value given the effective model (3). Given any function $g(\mathbf{R}_s)$ of the activity $\mathbf{R}_s$ of node s, we define

$$E_0(g(\mathbf{R}_s)) = \sum_{\mathbf{r}_s} g(\mathbf{r}_s) \prod_{k,i} P_s \left(r_s^{k,i}; i, \mathbf{r}_s^{k,<i}, 0 \right), \tag{5}$$

where the sum is over all possible values of the activity $\mathbf{r}_s$ of node s. In particular $E_0(R_s^{k,i})$ is equal to the expected value of the activity of node s at stimulus time point i. (By assumption, it is independent of stimulus repeat k.) Recall that we assume the stimulus is repeated many time. Hence, if $\mathbf{R}_s$ indicates the spiking activity of neuron s, then $E_0(R_s^{k,i})$ is proportional to the PSTH of neuron s at stimulus time point i,

[2] C.f. Eqs. (3.10a) and (3.13) of Ref. [14]. Note that we have ignored terms quadratic in W to facilitate computations.

assuming that model (3) adequately captures the average activity of neuron s. We also use the shorthand notation

$$\begin{aligned} P_s^{k,i} &= P_s\left(r_s^{k,i}; i, \mathbf{r}_s^{k,<i}, 0\right), \\ \frac{\partial P_s^{k,i}}{\partial w} &= \frac{\partial}{\partial w} P_s\left(r_s^{k,i}; i, \mathbf{r}_s^{k,<i}, w\right)\Big|_{w=0}. \end{aligned} \tag{6}$$

to represent the effective model (3) and its derivative.

Equation (4) contains two new effective connectivity terms. The first is $W_{\tilde{q},q}^{j,i}$, which is the effective causal connection onto measured node q at stimulus time i coming from the activity of measured node $\tilde{q}$ that occurred j time steps previous. It includes not only the direct connection $\bar{W}_{\tilde{q},q}^{j,i}$, but also indirect connections from node $\tilde{q}$ onto node q through an intermediate hidden node p. (Since we used a second-order Taylor approximation, we do not directly account for chains of connections longer than two.) In terms of the original connectivity, W can be written as

$$W_{\tilde{q},q}^{j,i} = \bar{W}_{\tilde{q},q}^{j,i} + \sum_{p\in\mathcal{P}} \sum_{\substack{\tilde{j}_1,\tilde{j}_2 \\ j>\tilde{j}_1\geq\tilde{j}_2>0}} \bar{W}_{\tilde{q},p}^{j-\tilde{j}_1,i-\tilde{j}_1} C_p^{i-\tilde{j}_1,i-\tilde{j}_2} \bar{W}_{p,q}^{\tilde{j}_2,i} \tag{7}$$

where $C_p^{i-\tilde{j}_1,i-\tilde{j}_2}$ is an expression involving the model of hidden node p; it indicates the influence of an input at stimulus time $i-\tilde{j}_1$ on its spiking probability at stimulus time $i-\tilde{j}_2$ (of the same stimulus repeat). Clearly, we cannot determine any of the parameters of the sum in $W_{\tilde{q},q}^{j,i}$; our goal is only to compute the effective causal connection $W_{\tilde{q},q}^{j,i}$.

The second connectivity term is $U_{\tilde{q},q}^{j,i}$, which is the effective common input from hidden nodes that reaches measured node q at stimulus time i and reaches measured node $\tilde{q}$ at the time j steps previous. In terms of the original connectivity, U can be written

$$\begin{aligned} U_{\tilde{q},q}^{j,i} = \sum_{p\in\mathcal{P}} \sum_{\substack{\tilde{j}_1,\tilde{j}_2 \\ \tilde{j}_1>0,\tilde{j}_2>j}} \bar{W}_{p,q}^{\tilde{j}_1,i} \bar{W}_{p,\tilde{q}}^{\tilde{j}_2-j,i-j} \Big[& E_0(R_p^{k,i-\tilde{j}_1} R_p^{k,i-\tilde{j}_2}) \\ & - E_0\,(R_p^{k,i-\tilde{j}_1}) E_0(R_p^{k,i-\tilde{j}_2}) \Big]. \end{aligned} \tag{8}$$

The bracketed expression is the covariance between the activity of hidden node p at stimulus times $i-\tilde{j}_1$ and $i-\tilde{j}_2$. (Recall the effective expected values do not depend on stimulus repeat k.) Again, we cannot determine the individual quantities in the sum defining $U_{\tilde{q},q}^{j,i}$, but seek only to determine the total effective common input $U_{\tilde{q},q}^{j,i}$.

A key point is that once we fit effective models (3) to each of the measured nodes, the only unknown factors in (4) are the effective connectivity factors W and U. All other quantities are functions of the effective models of measured nodes. All

quantities involving hidden nodes parameters are contained within the factors W and U as specified in Eqs. (7) and (8).

2.3 Avoiding additional ambiguity

The above analysis parallels that of Ref. [14]. However, since the analysis of Ref. [14] did not exploit stimulus repeats, it could not allow W and U to depend on time points (otherwise, there would have more unknowns than data points). In Ref. [14], $W^{j,i}_{\tilde{q},q}$ and $U^{j,i}_{\tilde{q},q}$ were allowed to be functions only of delay j and not of time point i.

Note from Eqs. (7) and (8) that the effective coupling terms $W^{j,i}_{\tilde{q},q}$ and $U^{j,i}_{\tilde{q},q}$ will still depend on stimulus time point i even if the original coupling terms $\bar{W}^{j,i}_{\tilde{q},q}$ and $\bar{U}^{j,i}_{\tilde{q},q}$ were independent of i. Both Eqs. (7) and (8) contain factors involving hidden node activity that depend explicitly on i. Hence, the approximation made in Ref. [14] neglects how hidden node activity can depend on the stimulus. As discussed in Ref. [14], this approximation leads to some ambiguity in the identification of causal connections. We will return to this ambiguity when demonstrating the results of our new method in Sect. 3. In this paper, we develop a modification to remove this ambiguity for the case when a short stimulus is repeated many times.

The essential change from the analysis of Ref. [14] was already made in the previous subsection: we allowed the connectivity strength to depend on the stimulus time point but not on the stimulus repeat. If we have many stimulus repeats, then we can allow $W^{j,i}_{\tilde{q},q}$ and $U^{j,i}_{\tilde{q},q}$ to depend on both stimulus time point i as well as the delay j. We do not need to make any additional assumptions about the hidden nodes that introduce ambiguity in the identification of the causal connections.

By allowing the coupling terms to depend on stimulus time point, we have greatly increased the number of unknown parameters we must determine. To allow for tractable computations with feasible amounts of data, we reduce the degrees of freedom by splining. We expect that the connectivity terms could depend on the delay j with a fine temporal resolution, since that captures the relative timing between the activity of the nodes. However, we expect the dependence on absolute stimulus time point to change much more slowly, on a time scale determined by the temporal structure of the stimulus. Consequently, we approximate $W^{j,i}_{\tilde{q},q}$ and $U^{j,i}_{\tilde{q},q}$ with quadratic splines in delay j using a fine grid spacing of Δt_d and with linear splines in stimulus time point i using a larger grid spacing of Δt_s.

In previous approaches, we used the logarithm of (4) to compute maximum likelihood estimates of W and U. Since allowing stimulus-dependent connectivity greatly increases the number of unknowns, maximum likelihood estimates would be prone to overfit the data. To avoid this, we penalize the likelihood in two ways. First, we use the standard procedure of penalizing large values of any of the spline coefficients for W and U.[3] Second, we penalize large deviations of $W^{j,i}_{q,\tilde{q}}$ or $U^{j,i}_{q,\tilde{q}}$ with respect to stimulus

[3] We subtract from the logarithm of (4) the sum of the squares of the spline coefficients, multiplied by a factor $\lambda_1 = 0.001$.

time point i.[4] In the end, we calculate maximum a posteriori (MAP) estimates of W and U using the resulting penalized log-likelihood.

3 Tests via neuronal network simulations

To test the performance of this approach, we simulated small networks of neurons with known circuitry and tested how well we could reconstruct the connections. We simulated networks with both direct connections and common input connections in order determine under what conditions allowing stimulus-dependent connectivity improved our connectivity estimates.

3.1 The format of the tests

3.1.1 The simulated networks

Our primary goal is to determine under which conditions we can distinguish between causal connections and common input connections. For simplicity, our causal connection network consisted of two neurons where neuron 2 had a direct connection onto neuron 1. Our common input network consisted of three neurons where neuron 3 was connected to both neurons 1 and 2; in the analysis, the spikes of neuron 3 were ignored so that it was an unmeasured neuron.

We let the response $R_s^{k,i}$ of neuron s be a Bernoulli random variable where $R_s^{k,i} = 1$ corresponds to the presence of a spike in time bin i during stimulus repeat k. We used time bins of width $\Delta t = 0.5$ ms. We let $\mathbf{X}^i$ represent a 2D visual stimulus at time i and simulated the network activity as generalized linear models (GLMs) of the form

$$\begin{aligned}\Pr(R_s^{k,i} = 1 \mid \mathbf{R}^{k,<i} = \mathbf{r}^{k,<i}) \\ = \bar{A}_s \left[\sum_{j>0} \left(\bar{\mathbf{h}}_{1,s}^j \cdot \mathbf{X}^{i-j} + \bar{h}_{2,s}^j r_s^{k,i-j} + \sum_{\tilde{s}\neq s} \bar{W}_{\tilde{s},s}^j r_{\tilde{s}}^{k,i-j} \right) + \bar{y}_s \right]_+^2 \end{aligned} \tag{9}$$

where $[y]_+ = \max(y, 0)$. The threshold quadratic nonlinearity for each neuron was determined by a gain parameter $\bar{A}_s$ and offset parameter $\bar{y}_s$. (We truncated to 1 any values of the probability that exceeded 1.)

Our GLM model had three sets of linear kernels. The first kernels $\bar{\mathbf{h}}_{1,s}^j$ specified how the neurons responded to the stimulus. We used Gabor functions in space multiplied by an alpha function in time. At each position $\mathbf{z} = (z_1, z_2)$ and delay of j, the kernel $\bar{\mathbf{h}}_{1,s}^j$ was proportional to

[4] For each node pair and delay index, we calculate the difference between each pair of spline coefficients with adjacent stimulus time indices. Then, we subtract from the logarithm of (4) the sum of the squares of these differences, multiplied by a factor $\lambda_2 = 0.1$.

$$\bar{h}^j_{1,s}(\mathbf{z}) \propto j \exp\left(-\frac{j\Delta t}{\tau_s^{\text{stim}}} - \frac{|\mathbf{z}-\mathbf{z}_0|^2}{2\sigma_s^2}\right)\cos(\mathbf{k}_s \cdot (\mathbf{z}-\mathbf{z}_0) + \phi_s) \tag{10}$$

with $\mathbf{k}_s = 2\pi(\cos\psi_s, \sin\psi_s) f_s$. The constant of proportionality was chosen so that $\sum_j \bar{\mathbf{h}}^j_{1,s} \cdot \mathbf{X}^{i-j}$ had unit variance. The offset $\mathbf{z}_0$ was chosen to center the kernel on the stimulus ($\mathbf{z}_0 = (N_0, N_0)/2$, where N_0 was the length of one side of the visual image as described below).

The second kernels $\bar{h}^j_{2,s}$ specified how each neuron was influenced by its own spiking history. We included an absolute refractory period of length τ_s^{ref} by setting $\bar{h}^j_{2,s} = -100$ for $j\Delta t \le \tau_s^{\text{ref}}$, which insured that $\Pr(R_s^{k,i} = 1)$ was zero for an interval of τ_s^{ref} after each spike. After the absolute refractory period, we included a relative refractory period by setting

$$\bar{h}^j_{2,s} = -a_s e^{-j\Delta t/\tau_s^{\text{hist}}} \qquad \text{for } j\Delta t > \tau_s^{\text{ref}}. \tag{11}$$

The third kernels $\bar{W}^j_{\tilde{s},s}$ specified the connectivity. We used the form

$$\bar{W}^j_{\tilde{s},s} = b_{\tilde{s},s} \frac{j\Delta t - d_{\tilde{s},s}}{\tau_w^2} \exp\left(-\frac{j\Delta t - d_{\tilde{s},s}}{\tau_w}\right)$$

for $j\Delta t > d_{\tilde{s},s}$ and $\bar{W}^j_{\tilde{s},s} = 0$ otherwise. The alpha function form of $\bar{W}^j_{\tilde{s},s}$ could be viewed as reflecting the time course of postsynaptic potential in neuron s in response to an input from neuron $\tilde{s}$. The parameter $d_{\tilde{s},s}$ represented the delay and $b_{\tilde{s},s}$ the strength of the connection. For all connections, we set the time scale to $\tau_w = 0.5$ ms. Note that we did not let the coupling strength $\bar{W}$ depend explicitly on time.

3.1.2 The stimulus

In the first set of experiments, we stimulated the networks with a simple stimulus: a grating drifting at 10 Hz. We let the stimulus at position $\mathbf{z} = (z_1, z_2)$ and time bin i be

$$X^i(\mathbf{z}) = \cos(2\pi(\mathbf{k} \cdot \mathbf{z} + \omega i \Delta t)),$$

where $\omega = 0.01$ ms^{-1} and $\mathbf{k} = 0.033(\cos\pi/4, \sin\pi/4)$. We recorded spikes from the neurons for 10 simulated minutes. We adjusted the nonlinearity parameters $\bar{A}_s$ and $\bar{y}_s$ so that each neuron spiked between 15 and 20 Hz, obtaining between 9,000 and 12,000 spikes per neuron.

In the second set of experiments, we stimulated the networks with a five second movie consisting of a sequence of sinusoidal gratings. For a given grating $\mathbf{I}^{\mathbf{k}}$ with wave vector $\mathbf{k} = (k_1, k_2)$, the image at position $\mathbf{z} = (z_1, z_2)$ was $I^{\mathbf{k}}(\mathbf{z}) = \text{cas}(2\pi\mathbf{k} \cdot \mathbf{z}/N_0)$ where $\text{cas}\, x = \cos x + \sin x$ and $N_0 = 100$ was the length of one side of the image ($0 \le z_1, z_2 \le N_0 - 1$). Every 50 simulated milliseconds, a new image was selected, with replacement, from the set composed of the $\mathbf{I}^{\mathbf{k}}$ and $-\mathbf{I}^{\mathbf{k}}$, for $k_1, k_2 \in \{-5, -4, \ldots, 4, 5\}$. We repeated this movie for one simulated hour. We adjusted the

nonlinearity parameters $\bar{A}_s$ and $\bar{y}_s$ so that each neuron spiked around 10 Hz, obtaining around 35,000–38,000 spikes per neuron.

3.1.3 The HAH models used to estimate connectivity

The first step in determining the connectivity is fitting an effective probability distribution (3) to the spikes of each measured neuron. We do not attempt to reconstruct the GLM models (9) that were used in the simulations. (Neither the drifting grating stimulus nor the short five second movie sampled the stimulus space sufficiently well to obtain good estimates of the kernels $\bar{\mathbf{h}}_{1,s}$.) Instead, we fit HAH models of the GLM form[5]

$$\Pr(R_s^{k,i} = 1 \mid \mathbf{R}_s^{k,<i} = \mathbf{r}_s^{k,<i}) = P_s\left(1; i, \mathbf{r}_s^{k,<i}, 0\right) = g_s\left(P_s^i + \sum_{j>0} h_s^j r_s^{k,i-j}\right) \tag{12}$$

where $g_s(y) = A_s \log(1 + \exp(y + y_0))$. The procedure for fitting the parameters is described in the Appendix.

After we have determined all parameters for the effective models of all measured neurons, we use (4) to fit the effective causal connection parameters $W^{j,i}$ and effective common input parameters $U^{j,i}$, where we simply add the coupling term underneath the nonlinearity, using

$$P_s\left(1; i, \mathbf{r}_s^{k,<i}, w\right) = g_s\left(P_s^i + \sum_{j>0} h_s^j r_s^{k,i-j} + c_s w\right). \tag{13}$$

See the Appendix for how we determined the scaling parameter c_s.

Recall that we splined the $W^{j,i}$ and the $U^{j,i}$ to reduce degrees of freedom. We used a fine grid spacing of $\Delta t_d = 2$ ms for the splines in delay j. For the splines in stimulus time point i, we used the grid spacing $\Delta t_s = 10$ ms for the drifting grating stimulus and the grid spacing $\Delta t_s = 50$ ms for the random grating stimulus.

Since in all the examples, we will have only two measured neurons, we combine for display purposes the connectivity parameters as

$$W^{j,i} = \begin{cases} W_{12}^{-j,i} & \text{for } j < 0, \\ 0 & \text{for } j = 0, \\ W_{21}^{j,i} & \text{for } j > 0, \end{cases}$$

[5] Since $R_s^{k,i}$ is a Bernoulli random variable, it is clear that the probability that $R_s^{k,i}$ is zero is $P_s\left(0; i, \mathbf{r}_s^{k,<i}, 0\right) = 1 - P_s\left(1; i, \mathbf{r}_s^{k,<i}, 0\right)$.

$$U^{j,i} = \begin{cases} U_{12}^{-j,i} & \text{for } j < 0, \\ 0 & \text{for } j = 0, \\ U_{21}^{j,i} & \text{for } j > 0. \end{cases} \tag{14}$$

so that connections from neuron 2 onto neuron 1 would be observed at positive delays and connections from neuron 1 onto neuron 2 would be observed at negative delays.

To further facilitate display of the results, we will also graph the average W and U across all stimulus time points

$$W_{\text{ave}}^{j} = \frac{1}{N_t} \sum_{i=1}^{N_t} W^{j,i} \quad \text{and} \quad U_{\text{ave}}^{j} = \frac{1}{N_t} \sum_{i=1}^{N_t} U^{j,i}, \tag{15}$$

where N_t is the number of time points in the stimulus.

Lastly, we are interested in comparing results based on stimulus-dependent connectivity with those based on fixed connectivity strength as in Ref. [14]. To calculate these latter estimates, we use (4) with (13) to obtain maximum likelihood estimates of the parameters $W_{\tilde{q},q}^{j}$ and $U_{\tilde{q},q}^{j}$ that are not allowed to depend on stimulus time i. We denote these constant connectivity estimates of the connectivity between neurons 1 and 2 as W_{const}^{j} and U_{const}^{j}. We will compare W_{ave}^{j} and U_{ave}^{j} with W_{const}^{j} and U_{const}^{j} to assess the performance of the stimulus-dependent connectivity estimates.

3.2 Result with drifting grating stimulus

In the first set of experiments, we used a drifting grating stimulus to stimulate both a network with a direct connection and a network with a common input connection from an unmeasured neuron. In the direct connection network, neuron 2 had a direct connection onto neuron 1. In the common input network, an unmeasured neuron had a connection onto both neurons 1 and 2, and the connection onto neuron 1 had a longer delay. With this delay, the spikes of neuron 1 were correlated with a delayed version of the spikes of neuron 2, mimicking the direct connection network.

To demonstrate this correlation, we plot the joint peristimulus time histogram (JPSTH) [1,20] for these two networks in Fig. 1a, b. In our notation, the JPSTH can be written as $\langle R_1^{k,i_1} R_2^{k,i_2} \rangle - \langle R_1^{k,i_1} \rangle \langle R_2^{k,i_2} \rangle$ where $\langle \cdot \rangle$ indicates average over stimulus repeats. Note that in both cases, the JPSTH has large values (indicated by light color) just below the diagonal, indicating the correlation between neuron 1's spikes and a delayed version of neuron 2's spikes. The JPSTHs do not distinguish between the direct connection network (Fig. 1a) and the common input network (Fig. 1b).

To distinguish between the direct connection network and the common input network, we calculate our causal connection factor W and common input factor U, as described above. The results are shown in Fig. 2a, b. For comparison, we also replot each JPSTH using the delay j and stimulus time i convention that we used for $W^{j,i}$ and $U^{j,i}$ in (14). In the direct connection network of Fig. 2a, we see a band of large JPSTH values (light color) around the delay of $j = 5$ ms, corresponding to the delay of the direct connection. A similar band of large values at delay $j = 5$ ms is also visible in

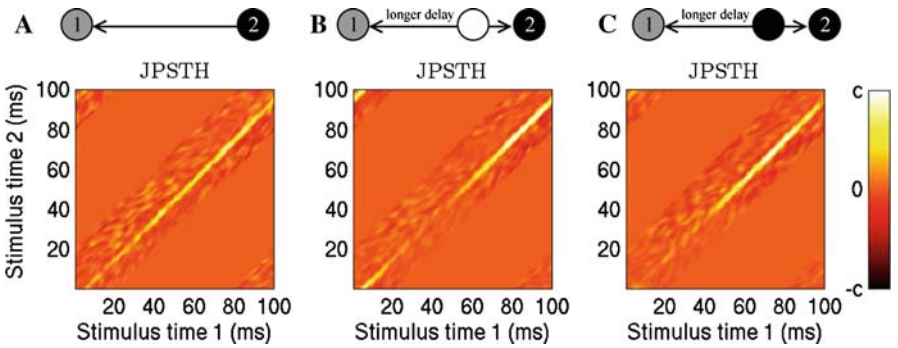

Fig. 1 JPSTHs calculated from networks driven by a drifting grating. The x-axis indicates the time of a spike in neuron 1 relative to the stimulus, and the y-axis indicates the time of a spike in neuron 2 relative to the stimulus. For each network, the JPSTH has large values just below the diagonal, indicating that neuron 1's spikes are correlated with a delayed version of neuron 2's spikes. The JPSTH does not distinguish between the different networks. **a** A network where neuron 2 is connected to neuron 1, as schematized by the top diagram. Simulation parameters: $\bar{A}_1 = 0.06$, $\bar{A}_2 = 0.05$, $\bar{y}_1 = \bar{y}_2 = -0.18$, $\sigma_1 = 15$, $\sigma_2 = 10$, $\psi_1 = 0$, $\psi_2 = \pi/4$, $f_1 = 0.4/\sigma_1$, $f_2 = 0.5/\sigma_2$, $\phi_1 = 0$, $\phi_2 = \pi$, $\tau_1^{\text{stim}} = \tau_2^{\text{stim}} = 40$ ms, $\tau_1^{\text{ref}} = 2$ ms, $\tau_2^{\text{ref}} = 1$ ms, $\tau_1^{\text{hist}} = 15$ ms, $\tau_2^{\text{hist}} = 20$ ms, $a_1 = 4$, $a_2 = 3$, $d_{2,1} = 4$ ms, $b_{2,1} = 2$, and all other $b_{\tilde{s},s} = 0$. **b** A network where an unmeasured neuron has a common input connection onto both neurons 1 and 2, with a longer delay to neuron 1. Simulation parameters as in **a**. Exceptions and additional parameters (index 3 refers to unmeasured neuron): $A_3 = 0.06$, $\bar{y}_1 = -0.13$, $\bar{y}_2 = -0.14$, $\bar{y}_3 = -0.15$, $\sigma_3 = 20$, $\psi_3 = \pi/2$, $f_3 = 0.3/\sigma_3$, $\phi_3 = 3\pi/2$, $\tau_3^{\text{stim}} = 40$ ms, $\tau_3^{\text{ref}} = 1.5$ ms, $\tau_3^{\text{hist}} = 25$ ms, $a_3 = 2$, $d_{3,1} = 5$ ms, $d_{3,2} = 0$ ms, $b_{3,1} = b_{3,2} = 7$, and all other $b_{\tilde{s},s} = 0$. **c** Another network with common input connections onto neurons 1 and 2. The only difference from the network in **b** is that the unmeasured neuron responds to similar stimulus features as does neuron 2, as indicated by the identical shading of both neurons in the top diagram. (Simulation parameters as in **b** except that $\phi_3 = 1.2\pi$.) In all three panels, the color scale as shown at the far right ranges from $-c = -0.0002$ up to $c = 0.0002$. JPSTHs were smoothed with a Gaussian filter with standard deviation of 1 ms. Note that in order to facilitate comparison with Fig. 2, we set to zero all values of each JPSTH where the stimulus times differ by more than 20 ms, modulo the stimulus period of 100 ms. (There was no structure in the JPSTHs outside the region that is shown)

the causal connection factor $W^{j,i}$ but not in the common input factor $U^{j,i}$, indicating that the correlation was due to a causal connection from neuron 2 onto neuron 1. Conversely, in the common input network of Fig. 2b, $U^{j,i}$, but not $W^{j,i}$, has a band of large values around the delay $j = 5$ ms, correctly indicating that the correlation observed in the JPSTH was not due to a causal connection between the measured neurons, but rather arises from common input that originates from an unmeasured neuron.

Note, however, that the estimates of W and U do not appear as clean as the JPSTHs. In Fig. 2a, for example, there is a region around delay $j = -8$ ms[6] and stimulus time 70 ms where the values of $W^{j,i}$ are even larger than those along the $j = 5$ ms band. Given that the JPSTH contains no structure in that region, one may suspect that those large values of W may be due to noise rather than indicating connectivity. (Indeed, we know it is unrelated to the connectivity since we know the underlying network.) Such noisy behavior of W and U is a consequence of allowing the connectivity parameters to depend on stimulus time, greatly increasing the degrees of freedom in our connectivity measures.

[6] A delay of $j = -8$ ms corresponds to neuron 2 firing 8 ms after neuron 1.

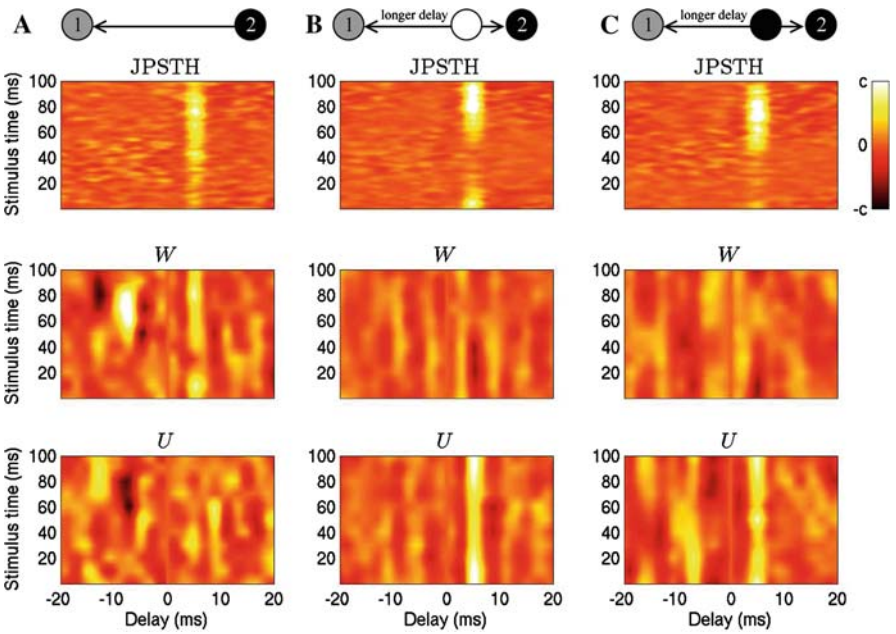

Fig. 2 Determination of the connectivity patterns underlying the correlations shown in Fig. 1. Network diagrams are repeated from Fig. 1. In the *top panels*, the JPSTHs of Fig. 1 are replotted using the delay j and stimulus time i convention that we used for $W^{j,i}$ and $U^{j,i}$ in (14). The $W^{j,i}$ and $U^{j,i}$ are shown in the *middle* and *bottom panels*, respectively. For all networks, the JPSTH indicates that the spikes of neuron 1 are correlated with the spikes of neuron 2 delayed by about 5 ms. For JPSTH plots, the color scale (see far right) ranges from $-c = -0.0002$ up to $c = 0.0002$. **a** Since $W^{j,i}$, but not $U^{j,i}$, has a band of large values around the delay $j = 5$ ms, the results are evidence that the correlations were due to a causal connection from neuron 2 onto neuron 1. However, especially since W and U are also large at other points, this raw data is difficult to interpret. For W and U, the color scale ranges from $-c = -2.5$ to $c = 2.5$. **b** Since $U^{j,i}$, but not $W^{j,i}$, has a band of large values around the delay $j = 5$ ms, the results correctly indicate that the correlation was due to common input from a measured neuron. For W and U, the color scale ranges from $-c = -4.5$ to $c = 4.5$. **c** Even if neuron two and the unmeasured neuron respond to the stimulus similarly (as schematized by both *circles being black* in the top diagram), the connectivity measures W and U still correctly indicate the presence of the common input connection, as only $U^{j,i}$ has a band of large values around the delay $j = 5$ ms. For W and U, the color scale ranges from $-c = -4$ to $c = 4$. Delay is spike time of neuron 1 minus spike time of neuron 2. Stimulus time is spike time (relative to stimulus) of the later firing neuron

To average out this noise, we average $W^{j,i}$ and $U^{j,i}$ over stimulus time i, obtaining W^j_{ave} and U^j_{ave} of Eq. (15). In Fig. 3a, b, we compare these average connectivity measures with the JPSTH similarly averaged over stimulus times, i.e., the shuffle corrected correlogram or covariogram [1,17,20]. For both the direct connection network (Fig. 3a) and the common input network (Fig. 3b), the covariogram has a peak at delay $j = 5$ ms. The presence of a causal connection cannot be deduced from the covariogram. On the other hand, the average connectivity measures W_{ave} and U_{ave} clearly distinguish the circuitry, as only W_{ave} has a significant peak for the direct connection network and only U_{ave} has a significant peak for the common input network, and these peaks occur around the delay of $j = 5$ ms that corresponds to the peaks in the covariograms.

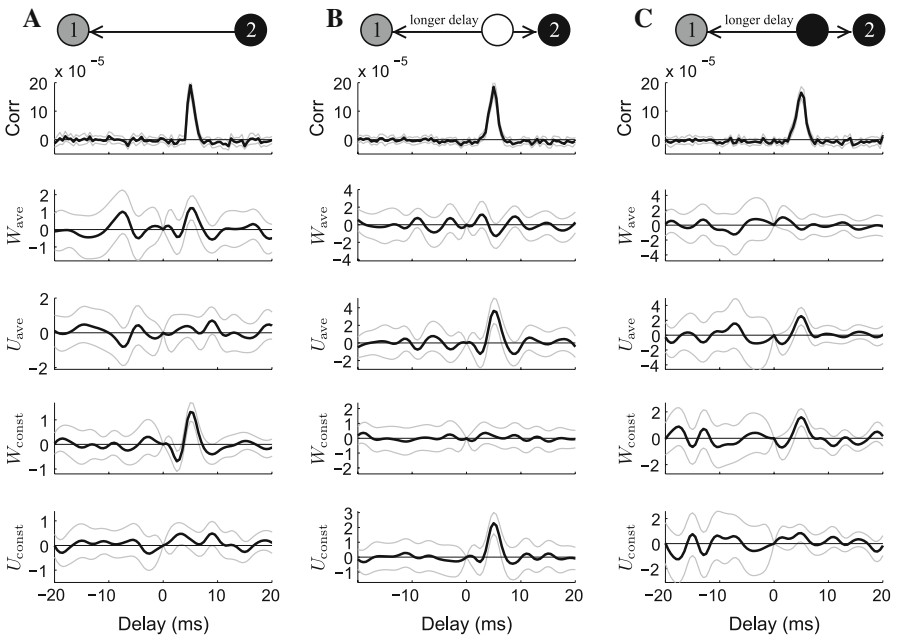

Fig. 3 The averaged connectivity factors successfully determine the circuitry of the networks from Fig. 1. The *top three panels* are the corresponding measures from Fig. 2 averaged over all stimulus times. The *two bottom panels* are the connectivity measures calculated based on the assumption that the connectivity is constant with respect to stimulus time. The covariogram (top plot, average of JPSTH) show a peak in the correlation at the delay of 5 ms for all networks. **a** The averaged causal connection measure W_{ave} correctly determines that the correlation at a delay of 5 ms was due to a causal connection from neuron 1 onto neuron 2. W_{ave} has a peak that is 3.5 standard errors above zero while U_{ave} is within a standard error from zero. The constant causal connection measure W_{const} even more reliability indicates the direct connection, as it reaches seven standard errors above zero. *Gray lines* indicate a bootstrap estimate of two standard errors, calculated from 50 resamples. **b** The averaged common input measure U_{ave} correctly determines that the correlation at a delay of 5 ms was due to common input from an unmeasured neuron. U_{ave} has a peak that is five standard errors above zero while W_{ave} is negative. The constant common input measure U_{const} also detects the common input, rising six standard errors above zero. **c** Even when the unmeasured neuron and neuron 2 are from the same subpopulation (as indicated by *black circles*), the averaged common input measure U_{ave} correctly determines that the correlation at a delay of 5 ms was due to common input from an unmeasured neuron. U_{ave} has a peak that is four standard errors above zero while W_{ave} is negative. On the other hand, the constant causal connection measure W_{const} incorrectly indicates that the correlation was caused by a causal connection. W_{const} rises nearly five standard errors above zero, while U_{const} barely exceeds two standard errors above zero. In this way, the constant connectivity measures W_{const} and U_{const} must be interpreted with *subpopulation ambiguity*, as described in the text. Network diagrams (*top*) are repeated from Fig. 1

We compare these results with the analogous connectivity measures W^j_{const} and U^j_{const} that are estimated based on the assumption that they are constant with respect to stimulus time i [14]. These measures are shown in the bottom two panels of Fig. 3. We see that W_{const} and U_{const} also correctly identify the circuitry, as W_{const} has a significant peak for the causal connection and U_{const} has a significant peak for common input network. In fact, it appears that W_{const} and U_{const} do a better job at distinguishing the circuitry, as the peak in W_{const} is much more significant than the peak in W_{ave} for the direct connection network of Fig. 3a.

The fact that W_{const} performed better than W_{ave} is understandable for two reasons. First, W_{const} contains many fewer degrees of freedom than W_{ave}. And second, for the direct connection network, $W^{j,i}$ should indeed be independent of time because the simulated network (9) contained connectivity $\bar{W}$ independent of stimulus time (and, as shown by Eq. (7), W should be equal to $\bar{W}$ if the network just has a direct connection).

The natural question then is why would one prefer W_{ave} and U_{ave} over W_{const} and U_{const} if the former result in noisier estimates of the connectivity. The answer is revealed by examination of Eq. (8) that defines U. Assume that the underlying connectivity $\bar{W}^{j,i}_{\tilde{s},s}$ does not depend on stimulus time point i. If ones assumes that $U^{j,i}_{\tilde{s},s}$ does not depend on i, then one is neglecting how the variance in the firing of unmeasured common input neurons may change with stimulus time point. As discussed in Ref. [14], this assumption does not affect the estimation of W_{const} and U_{const} except in one case. And that case is when the activity of the unmeasured common input neuron modulates with time in the same way as does the measured neuron that receives the connection with shorter delay, i.e., neuron 2 in our example network. If the unmeasured common input neuron of Fig. 3b happened to respond to the stimulus in a similar way as neuron 2, then W_{const} and U_{const} would misidentify the connectivity pattern as arising from a casual connection from neuron 2 onto neuron 1.

In the common input network of Fig. 3b, we chose the phase ϕ_3 of the stimulus kernel (10) of the unmeasured neuron to be different from the phase ϕ_2 of neuron 2 ($\phi_2 = \pi$, $\phi_3 = 3\pi/2$). In this way, the phase of the modulation of the unmeasured neuron's activity with the sinusoidal stimulus was different than that of neuron 2. This difference was enough for W_{const} and U_{const} to correctly identify the correlations as arising from common input.

Of course, the entire motivation for this analysis is to address the case of common input arising from *unmeasured* neurons. If the common input neuron is unmeasured, one would have no way of discerning whether or not the neuron responded to the stimulus in a similar way as neuron 2. Any interpretation of the results must include the possibility that an unmeasured common input neuron responds to the stimulus in a way similar to neuron 2. As mentioned above, such common input could be misidentified by the constant connectivity measures. Indeed, if we change the phase ϕ_3 of the unmeasured common input neuron to be similar to that of neuron 2 (setting $\phi_3 = 1.2\pi$), we find that W_{const} and U_{const} misidentify the common input as a causal connection (bottom two panels of Fig. 3c).

This observation leads to what we term *subpopulation ambiguity* in the interpretation of W_{const} and U_{const}. If W_{const} indicates that a correlation was due to a causal connection from neuron 2 onto neuron 1, one can conclude that either there is such a causal connection, or there is a causal connection onto neuron 1 from an unmeasured neuron that responds to the stimulus in a similar manner as neuron 2 (i.e., is in the same *subpopulation,* where a subpopulation is a group of neurons that respond similarly to the stimulus). The drawback of this subpopulation ambiguity is mitigated by the fact that, if one is measuring neurons' activity only by their spike times, those neurons can be identified only by how they respond to the stimulus, i.e., they are identified only by their subpopulation. Hence, any conclusions one makes based on such data already inherently contain subpopulation ambiguity [14].

Nonetheless, with a simple stimulus such as a drifting grating, many neurons responds similarly to the stimulus, with the main distinguishing factors being the amount by which their activity is modulated by the stimulus and the phase of that modulation. In this sense, the subpopulations defined by such a stimulus are large, so that the subpopulation ambiguity significantly detracts from one's ability to understand connectivity using W_{const} and U_{const}. For instance, the unmeasured neuron in Fig. 3c has a stimulus kernel (10) that is quite different than that of neuron 2. But it was effectively part of the subpopulation of neuron 2 because both neurons had similar phase of modulation with the drifting grating stimulus.

Our motivation for allowing stimulus-dependent connectivity in our estimates of W and U (and hence W_{ave} and U_{ave}) was to overcome this severe subpopulation ambiguity with simple stimuli such as a drifting grating. In Figs. 2c and 3c, we demonstrate that these stimulus-dependent connectivity measures can correctly identify the circuitry even if the unmeasured common input neuron responds to the stimulus in a way similar to the neuron that receives the connection with shorter delay (i.e., neuron 2 in our example network). In Fig. 2c, only $U^{i,j}$ has a band of large values around the delay of $j = 5$ ms. Similarly, in Fig. 3c, only U^j_{ave} shows a significant peak at the delay of $j = 5$ ms. The key point is that only W_{ave} and U_{ave} (and not W_{const} and U_{const}) correctly identify the circuitry that underlies the correlation in the covariogram. With the stimulus-dependent connectivity measures, one does not have subpopulation ambiguity in the interpretation of the results.

To further explore the relative performance of the stimulus-dependent and constant connectivity measures, we tested how well we could use these measures to classify networks as causal connection versus common input networks. We simulated 100 direct connection networks and 100 common input networks similar to those from Figs. 1, 2 and 3 with randomly chosen parameters. We held the delays in the connectivity fixed so that in all cases, the correlation due to the connectivity occurred around a delay of 5 ms. For each of the 200 simulations, we calculated W_{ave} and U_{ave}, and attempted to classify the network structure based on the values of these measures at the delay of 5 ms, which we denoted W_5 and U_5, respectively. We used a simple linear classifier, classifying the network as a causal connection network if $W_5 - U_5 > \theta$ for a threshold θ. We then repeated this classification using W_{const} and U_{const}.

To evaluate each classification, we perform receiver operating characteristic (ROC) analyses [26]. We let *true positives* correspond to direct connection networks that were correctly classified as causal connection networks and *false negatives* be those that were incorrectly classified as common input networks. Similarly, we let *true negatives* be correctly classified common input networks and *false positives* be those incorrectly classified as causal connection networks. With these definitions, we would expect networks like those of Fig. 3c to lead to extra false positives classified from W_{const} and U_{const}, as such common input networks would be likely misclassified as causal connection networks.

For stimulus-dependent and constant connectivity measures, we calculated ROC curves by classifying the results using a range of thresholds θ. For each θ, we calculated the total number of true positives (denoted *TP*), false negatives (*FN*), true negatives (*TN*), and false positives (*FP*) based on applying the connectivity analysis to the

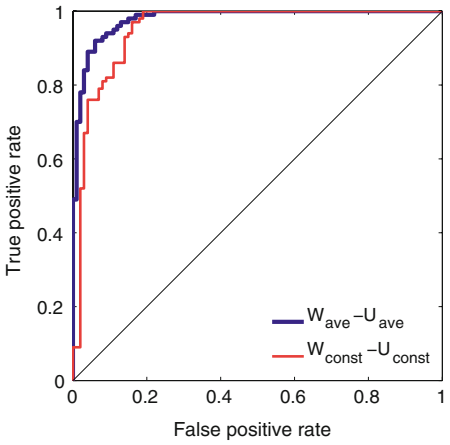

Fig. 4 Receiver operating characteristic (ROC) curves for simulated networks driven by a drifting grating stimulus. The *curves* are based on a simple linear classifier, a thresholded value of $W - U$ at a delay of 5 ms, where $W = W_{\text{ave}}$ and $U = U_{\text{ave}}$ (*thick dark line*) and where $W = W_{\text{const}}$ and $U = U_{\text{const}}$ (*thin light line*). True positive rate is the fraction of direct connection networks correctly classified, and false positive rate is the fraction of common input networks misclassified as causal connection networks. The subpopulation ambiguity in W_{const} and U_{const} increased the false positive rate of the resulting classifier so that the classifier based on W_{ave} and U_{ave} performed better. Results were based on simulations of 100 direct connection networks similar to that of Fig. 3a and 100 common input networks similar to those of Fig. 3b, c. For all neurons $s \in \{1, 2, 3\}$, parameters were chosen independently from uniform distributions over the following intervals: $\bar{A}_s \in (0.04, 0.08)$, $\bar{y}_s \in (0.1, 0.2)$, $\sigma_s \in (5, 20)$ $\psi_s \in (0, 2\pi)$, $f_s \in (0, 0.6/\sigma_s)$, $\phi_s \in (0, 2\pi)$, $\tau_s^{\text{stim}} \in (10, 50)$ ms, $\tau_s^{\text{ref}} \in (0, 2)$ ms, $\tau_s^{\text{hist}} \in (10, 30)$ ms, $a_s \in (2, 4)$. For direct connection networks, we set $d_{2,1} = 4$ ms, sampled the direct connection strength uniformly from the interval $b_{2,1} \in (0.4, 4)$, and set all other $b_{\tilde{s},s} = 0$. For common input networks, we set $d_{3,1} = 5$ ms and $d_{3,2} = 0$ ms, sampled the common input connection strengths uniformly and independently from the intervals $b_{3,1} \in (2, 10)$ and $b_{3,2} \in (2, 10)$, and set all other $b_{\tilde{s},s} = 0$

spikes times of neurons 1 and 2 in each of the 200 simulated networks. Then, we calculated the true positive rate as $TPR = TP/(TP + FN)$ and the false positive rate as $FPR = FP/(FP + TN)$. As θ varies, the points (FPR, TPR) trace out the ROC curve for a given classifier.

The ROC curves for the classifiers based on stimulus-dependent and constant connectivity measures are shown in Fig. 4. As a perfect classifier corresponds to $(FPR, TPR) = (0, 1)$, a better classifier would have an ROC curve that stays closer to the upper ($TPR = 1$) and left ($FPR = 0$) boundaries of the plot. Conversely, random chance corresponds to the diagonal. Figure 4 shows that, at least for the drifting grating stimulus and the parameter ranges simulated, the stimulus-dependent connectivity measures lead to a better classification. The false positives of the constant connectivity measures that are caused by the subpopulation ambiguity tend to shift the ROC curve for $W_{\text{const}} - U_{\text{const}}$ to the right.

Only the relative position of the ROC curves, but not their absolute shape, is a relevant indication of the classification ability of the connectivity measures. We could move the ROC curves closer to the diagonal or the upper left corner by, for example, changing the lower limits of the connectivity strengths or the simulation length. Both classifiers will perform worse as the connectivity strength becomes small (for a given

simulation length). Figure 4 demonstrates that, for these simulations, the elimination of the subpopulation ambiguity in the stimulus-dependent connectivity measures outweighs the cost of the additional parameters and leads to a better classification of the circuitry.

3.3 Result with random grating stimulus

In our second set of experiments, we stimulated the same networks with the rich random grating stimulus. Because the stimulus was longer and elicited much more temporal structure in the response of the neurons, W and U had many more degrees of freedom than with the drifting grating stimulus. Moreover, the rich stimulus would tend to expose differences among neurons' response properties, making it less likely that two neurons would have similar responses. In this way, the size of effective subpopulations would be reduced compared to the drifting grating stimulus, and the constant connectivity factors W_{const} and U_{const} would perform with less ambiguity. Hence, the random grating stimulus would test the limits of the stimulus-dependent connectivity factors and explore their value for experiments where the stimulus has rich structure.

We first simulate the response of the direct connection network of Fig. 2a in response to the random grating stimulus. The results are shown in Figs. 5a and 6a. The JPSTH (Fig. 5a) clearly shows the correlation at a delay of 5 ms. The raw connectivity measures W and U are difficult to interpret because, due to the many degrees of freedom in W and U, they appear quite noisy. Upon close inspection, a band of large values (light colors) at a delay of 5 ms is noticeable in W. The connectivity becomes evident upon averaging over all stimulus time. The average connectivity measures W_{ave} and U_{ave} clearly indicate that the correlation was due to a causal connection (Fig. 6a), as W_{ave} has a significant peak at the delay of 5 ms. The constant connectivity measures W_{const} and U_{const} also correctly identify the circuitry, and the peak in W_{const} is even more significant than the peak in W_{ave}.

Next, we simulate the response of the common input network of Fig. 2c. Note that for the drifting grating stimulus, the unmeasured common input neuron was considered part of neuron 2's effective subpopulation, as the phase of its modulation with the drifting grating was similar to that of neuron 2. Since most parameters of the unmeasured neuron's stimulus kernel (10) differed substantially from those of neuron 2, we expect the rich random grating stimulus to drive the the unmeasured neuron differently than neuron 2. Hence, the unmeasured neuron would no longer be grouped as part of neuron 2's subpopulation and even the constant connectivity measures W_{const} and U_{const} should correctly detect that the correlation was due to common input. These results are borne out in Figs. 5b and 6b. The stimulus-dependent common input factor U has a band of large values around the delay of 5 ms (Fig. 5b). Both the averaged U_{ave} and the constant U_{const} common input factors have significant peaks around 5 ms.

The advantage of the stimulus-dependent connectivity factors is revealed only if we choose the parameters of the unmeasured common input so that its stimulus kernel (10) closely matches that of neuron 2. As shown in Figs. 5c and 6c, only then is the unmeasured neuron considered part of neuron 2's subpopulation and the subpopula-

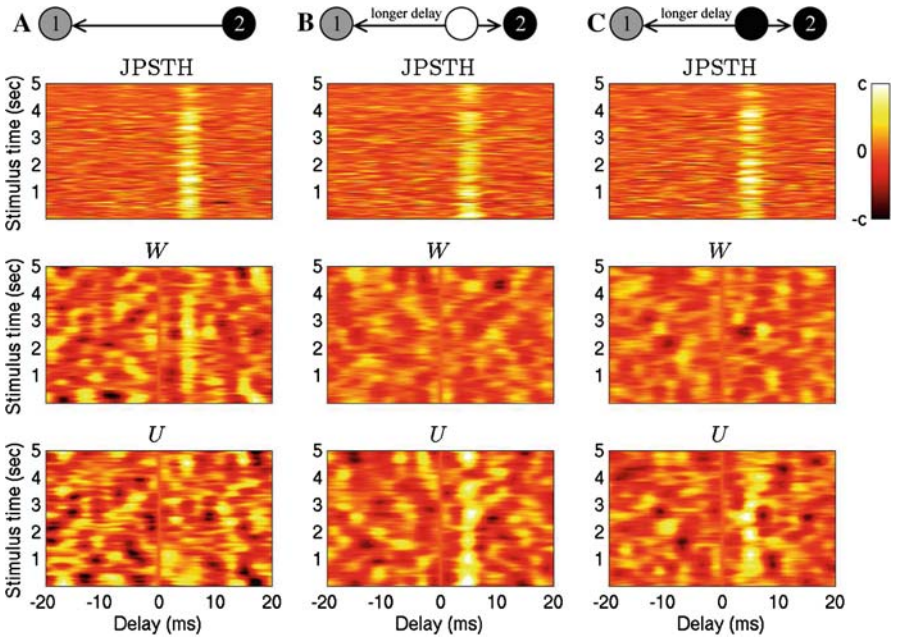

Fig. 5 Determination of the connectivity patterns for networks driven by random grating stimulus. Panels as in Fig. 2, but note different scale for stimulus time. For all networks, the JPSTH indicates that the spikes of neuron 1 are correlated with the spikes of neuron 2 delayed by about 5 ms. For JPSTH plots, the color scale (see far right) ranges from $-c = -5 \times 10^{-5}$ to $c = 5 \times 10^{-5}$. **a** Results from the direct connection network of Fig. 1a. A band of large values (*light colors*) of W is barely noticeable, giving a hint that the correlation was due to a causal connection from neuron 2 onto neuron 1. For W and U, the color scale ranges from $-c = -3.5$ to $c = 3.5$. Parameters as in Fig. 1a, except that $\bar{A}_1 = 0.04$, $\bar{A}_2 = 0.03$, and $\bar{y}_1 = \bar{y}_2 = -0.12$. **b** Results from the common input network of Fig. 1c. Note that the common input neuron in the top diagram is now shaded differently to indicate that it is no longer in neuron 2's subpopulation when the stimulus is the random grating sequence. The common input measure U shows a band of large values around a delay of 5 ms, indicating that the correlation was due to common input from an unmeasured neuron. For W and U, the color scale ranges from $-c = -5$ to $c = 5$. Parameters as in Fig. 1c, except that $\bar{A}_1 = \bar{A}_3 = 0.04$, $\bar{A}_2 = 0.03$, $\bar{y}_1 = \bar{y}_2 = \bar{y}_3 = -0.11$, and $b_{3,1} = b_{3,2} = 8$. **c** Results from a common input network, where the parameters of the unmeasured common input neuron were chosen to be nearly identical to those of neuron 2. The common input neuron is shaded the same color as neuron 2 in the top diagram to indicate it is in neuron 2's subpopulation. A band of large values in U at the delay of 5 ms indicates the correlation was due to common input. For W and U, the color scale ranges from $-c = -5$ to $c = 5$. Parameters as in **b** except that $\bar{y}_3 = -0.12$, $\sigma_3 = 10$, $\psi_3 = \pi/4$, $f_3 = 0.5/\sigma_3$, and $\phi_3 = 1.1\pi$

tion ambiguity in W_{const} and U_{const} is revealed. The stimulus-dependent connectivity factors W and U (especially their averages W_{ave} and U_{ave}) correctly identity the common input while the constant connectivity factors W_{const} and U_{const} misclassify the correlation as arising from a causal connection.

We repeat the ROC analysis for the networks driven by the random grating stimulus. As shown in Fig. 7, we get a different result than we did for networks driven by the drifting grating stimulus (Fig. 4). In this case, the classifier based on the constant connectivity measures performs as good as or better than the classifier based on the stimulus dependent connectivity measures. Given the number of

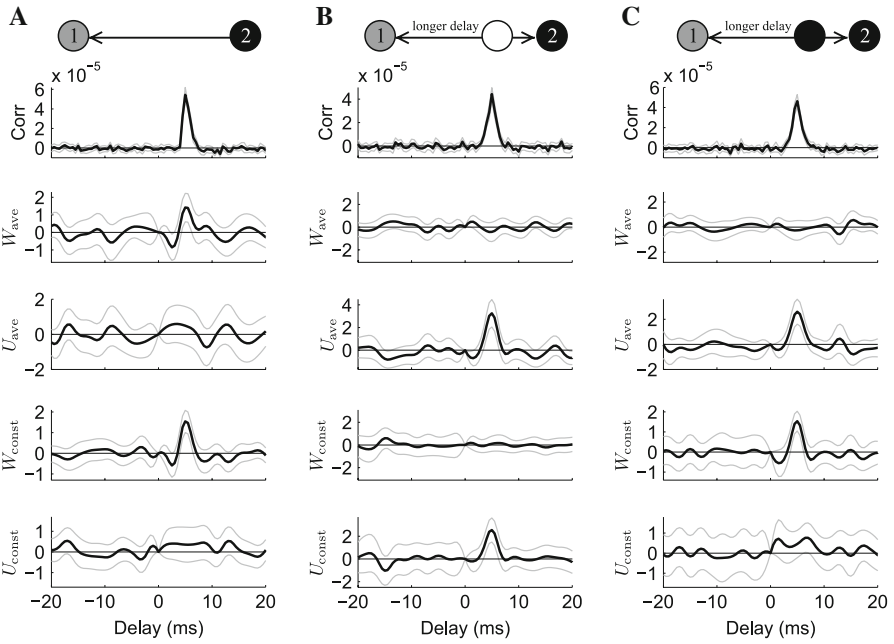

Fig. 6 The averaged connectivity factors successfully determine the circuitry for the networks of Fig. 5. The panels are as in Fig. 3. The covariogram has a peak at a delay of 5 ms for all networks. **a** The averaged causal connection measure W_{ave} correctly indicates that the correlation was due to a causal connection from neuron 1 onto neuron 2. W_{ave} has a peak that is 3.5 standard errors above zero and U_{ave} is within a standard error of zero at a delay of 5 ms. The constant causal connection measure W_{const} also correctly identifies the circuity, as it has an even more significant peak, rising over five standard errors above zero. **b** The averaged common input measure U_{ave} correctly identifies the source of the correlation as being due to common input from an unmeasured neuron. U_{ave} has a peak that is five standard errors above zero, while W_{ave} is negative. Although the network was the same as in Fig. 3c, the rich random grating stimulus brings out the differences between neuron 2 and the unmeasured common input neuron so that they are no longer in the same subpopulation. Hence, the constant connectivity measures also correctly identify the common input, with a peak in U_{const} that rises nearly five standard errors above zero. **c** Even when the common input neuron responds to almost the same stimulus features as neuron 2, the averaged connectivity measures correctly identify the common input. The peak in U_{ave} is five standard errors above zero and W_{ave} is negative. On the other hand, the constant connectivity measures classify the correlation as arising from a causal connection from neuron 2 onto neuron 1, as W_{const} has a peak that is over six standard errors above zero while U_{const} remains within about two standard errors from zero. Hence, for the constant connectivity measures W_{const} and U_{const}, the unmeasured neuron is in neuron 2's subpopulation

kernel parameters that we randomly selected, it was highly unlikely that the kernels for neurons 2 and 3 would be as similar as they were for Fig. 6c. Since we were unlikely to randomly create a common input network where the subpopulation ambiguity in W_{const} and U_{const} would cause them to misidentify the correlation as due to a causal connection, the additional false positives in the classifier based on $W_{\text{const}} - U_{\text{const}}$ were rare. On the other hand, the additional variability due to more parameters in W_{ave} and U_{ave} still affects the the classifier based on $W_{\text{ave}} - U_{\text{ave}}$. For these reasons, adding stimulus dependence did not improve the classification of the networks.

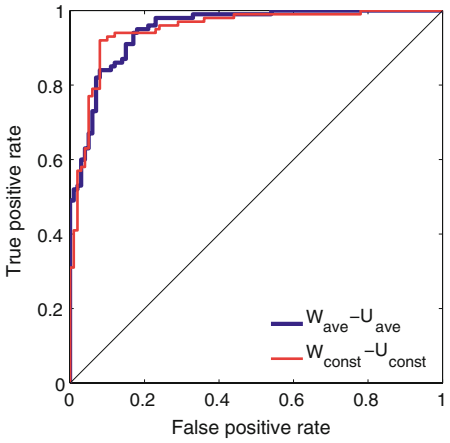

Fig. 7 Receiver operating characteristic (ROC) curves for simulated networks driven by a random grating stimulus. Figure is identical to Fig. 4 except for the different stimulus. The rich stimulus decreased the subpopulation ambiguity in W_{const} and U_{const} so that the additional the false positive rate is negligible. In this case, the classifier based on the constant connectivity measures performs as good as or better than the classifier based on the stimulus-dependent connectivity measures, as the corresponding ROC curve is primarily closer to the upper left corner. Parameters identical to Fig. 4 except that the intervals for the nonlinearity parameters were chosen to decrease the firing rate: $\bar{A}_s \in (0.02, 0.06)$, $\bar{y}_s \in (0.05, 0.15)$

4 Discussion

4.1 Eliminating subpopulation ambiguity

The present work is an extension of earlier analyses designed to control for the effects of hidden nodes that can corrupt estimates of causal connections among measured nodes of the network [12–15]. These analyses were designed to distinguish causal connections among measured nodes from common input connections arising from hidden nodes. Since the earlier treatments did not allow the resulting connectivity estimates to depend on stimulus time, they effectively assumed that hidden node activity was not modulated by the stimulus. This assumption does not detract from their ability to distinguish circuitry with one exception. A particular form of common input from a hidden node would be misclassified as a causal connection: if a hidden node happened to be (a) connected to a measured node that was modulated by the stimulus in a similar manner and (b) connected to another measured node with a longer delay, then this common input configuration would be misidentified as a causal connection from the first measured node onto the second.

Since one cannot rule out the possibility that such a common input configuration was really underlying any identification of a causal connection, the resulting estimate of causal connection must be interpreted with *subpopulation ambiguity*. We use the term subpopulation to describe a group of nodes that respond to the stimulus in a similar manner. Hence, a subpopulation is not an absolute grouping but one that depends on the chosen stimulus. The subpopulation ambiguity means that the identification of a causal connection from node 1 onto node 2 must be interpreted as the identification

of a causal connection from a node within node 1's subpopulation onto node 2. This interpretation acknowledges the possibility that the correlation could have been due to a hidden node from node 1's subpopulation having common input connections onto both node 1 and node 2 (with a longer delay onto node 2). Since this common input configuration does include a connection from a node in node 1's subpopulation onto node 2, the subpopulation ambiguity precisely describes the degree of ambiguity in the identification of connections. We stress that the subpopulation ambiguity does not refer to connections among populations of nodes but simply ambiguity in the identification of the individual nodes involved in the connections.

We have demonstrated how one can eliminate subpopulation ambiguity in connectivity estimates by allowing these estimates to vary with the stimulus. With the stimulus-dependent connectivity measures developed in this paper, we make no assumption about how hidden node activity is modulated by the stimulus. Even if a hidden common input node happened to respond to the stimulus in exactly the same manner as one of the connected measured nodes, the resulting correlation between the measured nodes would still be correctly identified as arising from common input. Hence, the identification of a causal connection between measured nodes can be interpreted as a causal connection between the actual measured nodes. We have eliminated the misidentification that lead to the subpopulation ambiguity.

4.2 Stimulus features and subpopulation ambiguity

The cost of the extension to stimulus-dependent connectivity measures was greater degrees of freedom in the connectivity parameters that must be estimated, requiring more data to obtain significant results. For this reason, the examples explored were all based on long neuronal network simulations, especially in the case of the longer random grating stimulus. Because one is limited in the length of experiment over which one can record the same neurons, this stimulus-dependent connectivity analysis can be practically applied only in those experiments where one has a short stimulus so that the stimulus can be repeated many times. However, a short stimulus is exactly the type of stimulus that leads to the most substantial subpopulation ambiguity in the constant connectivity measures W_{const} and U_{const} of previous versions of the analysis.

With a short or simple stimulus, the effective subpopulations of the constant connectivity measures would be large. If the range of responses elicited by the stimulus were limited, then many nodes would respond to the stimulus in virtually identical manners. These nodes would be effectively in the same subpopulation, and the subpopulation ambiguity of the constant connectivity measures would limit their usefulness. Any identification of causal connectivity inferred from those measures must be interpreted with a large degree of ambiguity. Thus, with such stimuli, the stimulus-dependent connectivity measures developed in this paper would allow a better determination of the connectivity, as these measures are not subject to subpopulation ambiguity. It is fortunate that it is precisely with such short stimuli that the stimulus-dependent connectivity measures can be practically computed.

On the other hand, with a longer and richer stimulus, the effective subpopulations of the constant connectivity measures are much smaller. A rich stimulus with many

different features would more rigorously probe the properties of the nodes. A much larger range of responses would be elicited by the stimulus, and fewer nodes would respond identically. The subpopulation ambiguity would add only a little ambiguity to the interpretation of the identified causal connectivity. If two nodes respond to the stimulus in a similar manner, then it is likely that those nodes really do have similar properties. In those cases, one may be content with the subpopulation ambiguity of the constant connectivity measures and may not need to attempt to use the stimulus-dependent connectivity measures to eliminate the subpopulation ambiguity. Indeed, the results indicate that adding stimulus dependence to the connectivity measures may have limited utility for properly classifying networks in cases where one expects subpopulation ambiguity to be rare.

4.3 History-dependence and subpopulation ambiguity

As discussed in [14], the connectivity measures W and U (W_{const} and U_{const} as well) distinguish common input from causal connections by exploiting two differences in their influence upon the activity patterns of measured nodes. First, the connectivity measures exploit how models such as (3) predict a different relationship between activity patterns and the stimulus depending on the nature of the connectivity structure. The subpopulation ambiguity results from exploiting these differences while keeping the connectivity constant with respect to stimulus time.

The second difference between common input and causal connections is how they interact with history-dependent properties of nodes. In the context of neuron spiking patterns, such history dependent properties would be refractory periods or bursting behavior. The connectivity measures exploit how models such as (3) predict a different relationship between intra-neuron spiking patterns (i.e., history-dependent effects such as those observed in an autocorrelation) and inter-neuron spiking patterns (i.e., correlations such as observed in the covariogram) depending on network circuitry [14]. Since such history-dependent effects do not involve the stimulus, they lead to a distinction between common input and causal connections that is not subject to subpopulation ambiguity, even if one assumes the connectivity is constant with respect to stimulus time.

Both differences are contained in (4) that we use to calculate W and U. Normally, these differences reinforce each other to produce the best estimate of the connectivity patterns. However, when we use the constant connectivity measures W_{const} and U_{const} and the actual underlying connectivity mirrors that of Figs. 3c or 6c, these two differences oppose each other. Because of the subpopulation ambiguity, the exploitation of the stimulus differences would tend to increase W_{const} as the connectivity should be classified as a causal connection subject to subpopulation ambiguity. On the other hand, the exploitation of history-dependent differences would tend to increase U_{const} since, without subpopulation ambiguity, the connectivity should be classified as common input.

In the examples shown in Figs. 3c and 6c, W_{const} identified the correlations as arising from a causal connection, indicating that the history-dependent effects were not strong enough to overcome the subpopulation ambiguity. We can change the balance

between the history-dependent effects and the subpopulation ambiguity of the stimulus effects by increasing the history-dependence in the model neurons. For example, we can increase the relative refractory period of the neurons by increasing the parameters a_s in (10). Doing so, increases U_{const} and decreases W_{const} at the delay where the correlation is observed (i.e., at a 5 ms delay). In the examples of Fig. 3c and 6c, the peaks in W_{const} are essentially replaced by peaks in U_{const} by the time we quadruple all the a_s (not shown). Hence, adding strong history-dependence to the model neurons eliminated the subpopulation ambiguity in these examples, enabling even the constant connectivity measures W_{const} and U_{const} to correctly identify the common input. These results suggest that if one is measuring from nodes that have strong history dependence[7] (e.g., neurons whose spiking patterns are far from Poisson), then the subpopulation ambiguity may have a smaller effect on the constant connectivity measures, and one may not need to use the stimulus-dependent connectivity measures.

4.4 Relationship to other approaches

The question of determining causal interactions among networks has spawned many analysis methods, such as Granger causality [6], partial coherence [3], partial directed coherence [2,22], and transfer entropy [23]. There is also a large literature focused on analyzing interactions among neurons [1,4,5,7,9–11,16,17,19–21,25,27]. Some of these methods are designed to address common input, but, with one exception, they can only control for common input arising from other measured nodes.

The one other approach that we are aware of that addresses common input arising from hidden neurons is that of Kulkarni and Paninski [9]. Their analysis framework contains a latent noise source that could model hidden common input. Their model is a doubly-stochastic process or Cox process [24] that allows the latent noise source to evolve with time. Although the approaches differ significantly, one similarity is that our stimulus-dependent connectivity was derived from an analysis of how the hidden common input node activity varies with time. A future task will be to compare the two approaches.

4.5 Model dependence

Our approach is model based so that its success depends on choosing a model (3) that captures how the activity of each measured node is modulated by its own history. In the examples, we have used a GLM model of history-dependence (12), but any other identifiable model of history-dependence could be used in the analysis. Since we assume the stimulus has been repeated many times, we do not need to assume a model for how the nodal activity is modulated by the stimulus. The advantage of the HAH model is that is a model-independent description of the stimulus-dependence of the nodal activity [15].

[7] The unknown history-dependence of the hidden nodes does not play a role.

The model (3) also requires that one postulate how the coupling from other nodes could influence each measured node's activity (i.e., one must choose a functional form for the effect of the last argument of P_s). In our GLM models, the natural way to include connectivity was underneath the nonlinearity (13). For the analysis to succeed, the model must capture the essence of how other nodes could modulate the response of the measured nodes.

As demonstrated by the examples, allowing stimulus-dependent connectivity estimate does require that one has a lot of data and can repeat any stimulus many times. Nonetheless, especially in cases where the stimulus is short and simple and where the nodes do not exhibit a high degree of history-dependence, the stimulus-dependent connectivity measures can be used to eliminate the significant subpopulation ambiguity that would otherwise degrade the connectivity estimates, thus allowing better estimation of the underlying connectivity.

Appendix: Estimating model parameters

We sketch our algorithm for fitting the parameters of (12) for each measured neuron. To reduce the dimension of the parameter space, we restricted both the P_s and h_s to live in subspaces as follows. We splined the stimulus parameters P_s with a linear spline with grid spacing of $\Delta t = 5$ ms for the drifting grating stimulus and $\Delta t = 20$ ms for the random grating stimulus.

To estimate the history parameters h_s of each measured neuron, we first calculated the absolute refractory period τ_s^{absref} as the minimum number of time bins observed between spikes. Then, so that our model predicts absolutely no firing for τ_s^{absref} time steps after each spike, we set $h_s^j = -10^{100}$ for $j \leq \tau_s^{\text{absref}}$. We restricted the remainder of the history kernel h_s to be in the subspace spanned by the vectors

$$B_s^k(i) = \begin{cases} \sin\left(\pi k \left[2\frac{i-\tau_s^{\text{absref}}}{\tau_s} - \left(\frac{i-\tau_s^{\text{absref}}}{\tau_s}\right)^2\right]\right) & \text{for } 0 < i - \tau_s^{\text{absref}} < \tau_s \\ 0 & \text{otherwise,} \end{cases}$$

for $1 \leq k \leq 39$. We set $\tau_s = 200 - \tau_s^{\text{absref}}$ time bins. We obtained basis vectors through Gram-Schmidt orthonormalization of the B_s. These basis vectors are analogous to those used in Ref. [8]; they can represent fine temporal structure for the time immediately after the spike but are smoother for longer time scales.

For each fixed value of the scale factor A_s of the nonlinearity $g_s(\cdot)$, we computed MAP estimates of the nonlinearity parameter y_s and the coefficients for the P_s and the h_s. We penalized the log-likelihood by adding the sum of the squares of the parameters, multiplied by $\lambda = 0.1$. This penalty term prevented any parameters from becoming too large, especially important when segments of the stimulus produced no spikes from a neuron. In those cases, attempting to find maximum likelihood estimates would pull some of the P_s^i toward minus infinity so that the probability $g_s(P_s^i)$ would tend to zero.

For a fixed value of A_s, calculating the MAP estimates was computationally tractable, as we chose our $g_s(\cdot)$ to be convex and log-concave. This choice assured that

maximization problem to determine all parameters except A_s would be a free of non-global local maxima [18]. The MAP estimation of the parameters defined all parameters as a function of A_s. We then searched for a value of A_s that maximized the log-likelihood while keeping the other parameters set at this function of A_s.

The scale parameter A_s allows the nonlinearity $g_s(\cdot)$ to interpolate between a exponential function and a threshold-linear function. If A_s is a very large value, then the MAP estimates of the parameters underneath the nonlinearity in (12) will be small values (so that the variance of $g_s(\cdot)$ will match the data). In the extreme case, the exponential in $g_s(y) = A_s \log(1 + \exp(y + y_0))$ will always be small compared to 1, and the nonlinearity will be nearly exponential over its entire domain of sampled values $g_s(y) \approx A_s \exp(y + y_0)$. On the other hand, if A_s is a very small value, then the MAP estimates of the other parameters will be relatively large. For positive arguments of the exponential, the exponential will be large compared to 1 and $g_s(y) \approx A_s(y + y_0)$, and for negative values the exponential will be very small so that $g_s(y) \approx A_s \exp(y + y_0) \approx 0$. In the extreme case, the nonlinearity is nearly a threshold linear function $g_s(y) \approx \max(A_s(y + y_0), 0)$.

Despite the extra work to compute a value of A_s (because the log-likelihood is not convex in A_s), we needed to allow A_s to be a free parameter fit from the data in order to get good connectivity estimates. Equation (4) which we use to fit W and U depends on the derivative of the probability function $P_s(\cdot)$. By allowing A_s to vary in our fit of the parameters of $P_s(\cdot)$, we do not prescribe the relationship between values of $P_s(\cdot)$ and its derivative. Rather, allowing A_s to be fit from the data gives a degree of freedom in this relationship which, in our simulation tests, was important to achieving good results.

A side effect of allowing A_s to vary among neurons means that the scale of the parameters under the nonlinearity could vary substantially (inversely with A_s, as described above). In an attempt to maintain a similar scale in the connectivity measures W and U, we introduced the factor c_s in (13). We determined c_s as follows. Let $Y = P_s^i + \sum_{j>0} h_s^j R_s^{k,i-j}$ (the argument of the nonlinearity) and let μ_Y equal its mean over the data set, ignoring those points where $Y < -10^{100}$ because of the refractory period. We let c_s be the standard deviation of a normal random variable Z with mean μ_Y that would make the variance of $g_s(Z)$ equal the variance of $g_s(Y)$. In this way, c_s scaled the values of W comparable to the effective scaling of the other parameters inside the nonlinearity. We matched the variance of $g_s(Y)$ rather than Y because large negative values of Y (from the relative refractory period) contribute little to the actual variance of the spiking probability $g_s(Y)$ even though they have a large effect on the variance of Y itself.

References

1. Aertsen AMHJ, Gerstein GL, Habib MK, Palm G (1989) Dynamics of neuronal firing correlation: Modulation of "effective connectivity". J Neurophysiol 61:900–917
2. Baccalá LA, Sameshima K (2001) Partial directed coherence: a new concept in neural structure determination. Biol Cybern 84:463–474
3. Brillinger DR (1981) Time series: data analysis and theory. Holden Day, San Francisco

4. Brown EN, Kass RE, Mitra PP (2004) Multiple neural spike train data analysis: state-of-the-art and future challenges. Nat Neurosci 7:456–461
5. Chornoboy ES, Schramm LP, Karr AF (1988) Maximum likelihood identification of neural point process systems. Biol Cybern 59:265–275
6. Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37:424–438
7. Harris KD, Csicsvari J, Hirase H, Dragoi G, Buzsáki G (2003) Organization of cell assemblies in the hippocampus. Nature 424:552–556
8. Keat J, Reinagel P, Reid RC, Meister M (2001) Predicting every spike: a model for the responses of visual neurons. Neuron 30:803–817
9. Kulkarni JE, Paninski L (2007) Common-input models for multiple neural spike-train data. Network Comput Neural Syst 18:375–407
10. Martignon L, Deco G, Laskey K, Diamond M, Freiwald W, Vaadia E (2000) Neural coding: higher-order temporal patterns in the neurostatistics of cell assemblies. Neural Comp 12:2621–2653
11. Nicolelis MAL, Dimitrov D, Carmena JM, Crist R, Lehew G, Kralik JD, Wise SP (2003) Chronic, multisite, multielectrode recordings in macaque monkeys. Proc Natl Acad Sci USA 100:11041–11046
12. Nykamp DQ (2005) Revealing pairwise coupling in linear-nonlinear networks. SIAM J Appl Math 65:2005–2032
13. Nykamp DQ (2007) A mathematical framework for inferring connectivity in probabilistic neuronal networks. Math Biosci 205:204–251
14. Nykamp DQ (2007) Exploiting history-dependent effects to infer network connectivity. SIAM J Appl Math 68:354–391
15. Nykamp DQ (2008) Pinpointing connectivity despite hidden nodes within stimulus-driven networks. Phys Rev E 78:021902
16. Okatan M, Wilson MA, Brown EN (2005) Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. Neural Comp 17:1927–1961
17. Palm G, Aertsen AMHJ, Gerstein GL (1988) On the significance of correlations among neuronal spike trains. Biol Cybern 59:1–11
18. Paninski L (2004) Maximum likelihood estimation of cascade point-process neural encoding models. Network Comput Neural Syst 15:243–262
19. Paninski L, Pillow JW, Simoncelli EP (2004) Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model. Neural Comp 16:2533–2561
20. Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. Biophys J 7:419–440
21. Rosenberg JR, Amjad AM, Breeze P, Brillinger DR, Halliday DM (1989) The Fourier approach to the identification of functional coupling between neuronal spike trains. Prog Biophys Mol Biol 53:1–31
22. Sameshima K, Baccalá LA (1999) Using partial directed coherence to describe neuronal ensemble interactions. J Neurosci Methods 94:93–103
23. Schreiber T (2000) Schreiber. Phys Rev Lett 85:461–464
24. Snyder D, Miller M (1991) Random point processes in time and space. Springer, Heidelberg
25. Stuart L, Walter M, Borisyuk R (2005) The correlation grid: analysis of synchronous spiking in multi-dimensional spike train data and identification of feasible connection architectures. Biosystems 79:223–234
26. Swets JA (1996) Signal detection theory and ROC analysis in psychology and diagnostics: collected papers. Lawrence Erlbaum Associates, Mahwah
27. Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN (2005) A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. J Neurophysiol 93:1074–1089