

REVEALING PAIRWISE COUPLING IN LINEAR-NONLINEAR NETWORKS*

DUANE Q. NYKAMP†

Abstract. Through an asymptotic analysis of a simple network, we derive an estimate of the coupling between a pair of units when all other units are unobservable. The analysis is based on a model where the response of each unit is a linear-nonlinear function of a white noise stimulus. The results accurately determine the coupling when all unmeasured units respond to the stimulus differently than the measured pair. To account for the possibility of unmeasured units similar to the measured pair, we cast our results in the framework of “subpopulations,” which are defined as a group of units who respond to the stimulus similarly. We demonstrate that we can determine when correlations between two units are caused by a connection between their subpopulations, although the precise identity of the units involved in the connection may remain ambiguous. The result is rigorously valid only when the coupling is sufficiently weak to justify a second-order approximation in the coupling strength. We demonstrate through simulations that the results are still valid even with stronger coupling and in the presence of some deviations from the linear-nonlinear model. The analysis is presented in terms of neuronal networks, although the general framework is more widely applicable.

Key words. neural networks, correlations, Weiner analysis, white noise

AMS subject classification. 92C20

DOI. 10.1137/S0036139903437072

1. Introduction. This analysis of coupling within networks is motivated by neuroscience, and we use the vocabulary of neuroscience throughout. The measured response properties of a neuron arise from the structure of the neural network in which the neuron is embedded. To understand the relationship between these response properties and the neural network structure, one would like to simultaneously measure the response of neurons and estimate their connectivity. However, it has proven difficult to estimate the connectivity from measurements of neural activity because only a small subset of neurons can be monitored simultaneously.

In particular, a direct connection between two measured neurons is difficult to distinguish from a connection onto both neurons that originates from a third, unmeasured neuron. We refer to the latter configuration as the common input configuration. We address the case where one simultaneously measures two neurons in a network and attempts to distinguish the direct connection configuration from the common input configuration.

This distinction is especially difficult because when studying a network, one typically does not directly measure the internal state of neurons, but records only their discrete output events, called spikes. From simultaneous recordings of two neurons’ spike times, one can analyze the joint statistics of the two spike trains in hopes of detecting a direct connection. Two widely used tools are the joint peristimulus time histogram (JPSTH) and its integral, the shuffle-corrected correlogram [14, 1, 13]. Unfortunately, inferences from the JPSTH or correlogram about the connections between

*Received by the editors November 3, 2003; accepted for publication (in revised form) February 7, 2005; published electronically August 9, 2005. This research was supported in part by an NSF Mathematical Sciences Postdoctoral Research Fellowship and by NSF DMS-0415409.

<http://www.siam.org/journals/siap/65-6/43707.html>

†School of Mathematics, University of Minnesota, Minneapolis, MN 55455 (nykamp@math.umn.edu).

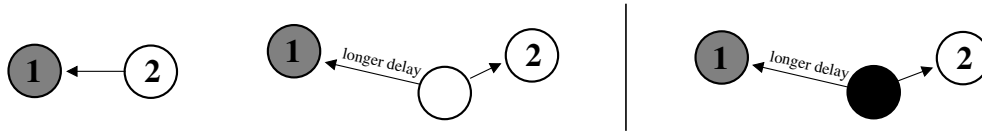


FIG. 1. To determine “subpopulation connectivity,” one needs to distinguish a direct connection from only certain kinds of common input. Three sample network configuration are shown, where neurons one and two are measured and the unlabeled neuron is not measured. The subpopulation of each neuron, which is defined within the context of a model, is indicated by the shading (white, gray, or black.) To determine subpopulation connectivity, as we have defined it, one must be able to distinguish the right configuration from the left two configurations. In both of the left configurations (but not in the right configuration), there is a connection from a neuron within neuron two’s subpopulation (white) onto a neuron within neuron one’s subpopulation (gray). Hence, we do not need to distinguish the left two configurations from each other in order to determine subpopulation connectivity.

the two measured neurons are ambiguous because these measures cannot distinguish a direct connection from common input.

The joint statistics of the two spike trains alone may be insufficient to distinguish a direct connection from common input. If one could measure the neurons inducing the common input effects, then the joint statistics of all the measured spike trains would be sufficient, and one could use analysis tools such as partial coherence [15] to distinguish a direct connection from common input. However, when one cannot measure all possible sources of common input, one cannot rule out common input through partial coherence.

Our approach is to analyze the joint statistics, not just of the measured spike trains, but also of an experimentally controlled stimulus. The idea motivating this approach is that the joint stimulus-spike statistics may be sufficient to distinguish the direct connection configuration from the common input configuration even if the neurons inducing the common input are unmeasured.

It turns out that we cannot distinguish a direct connection from all possible cases of common input. Instead, we can characterize connectivity only in terms of certain *subpopulations* of neurons, defined so that each neuron in a subpopulation responds to the stimulus in a similar manner (the definition of responding “similarly” is made in the context of a model). The concept of subpopulation connectivity is illustrated in Figure 1. Imagine that the spikes of neuron one are correlated with a delayed version of the spikes of neuron two, consistent with a direct connection from neuron two onto neuron one. Our central result is that we can distinguish between (A) a direct connection from neuron two onto neuron one and (B) common input that does not originate from neuron two’s subpopulation. On the other hand, if the common input does originate from neuron two’s subpopulation, the common input may not be distinguishable from a direct connection. However, in this latter case, the common input does contain a connection from neuron two’s subpopulation onto neuron one. Consequently, the identification of a direct connection from neuron two onto neuron one must be interpreted as the identification of a connection from neuron two’s subpopulation onto neuron one. The precise identity of the neuron originating the connection remains ambiguous (it could be neuron two or another neuron in neuron two’s subpopulation). To summarize this ambiguity, we say we can determine connectivity only at the level of subpopulations (and not at the level of individual neurons).

Our analysis is fundamentally model-driven. The structure imposed by an explicit

model gives the framework necessary for making the subtle distinction between a direct connection and most cases of common input. In this paper, we analyze a network modeled as interacting linear-nonlinear systems responding to a white noise stimulus. Clearly, this choice limits the applicability of this implementation to networks that can be approximated by this simple model. Our motivation for using this model is the ability to compute analytic expressions for necessary stimulus-spike statistics. We mention possible generalizations in the Discussion.

In section 2, we describe the model network and the assumptions required for the analysis. In section 3, we derive analytic expressions for measurable stimulus-spike statistics and solve the resulting system of equations for the coupling strength. We test our findings via simulations in section 4, and discuss the results in section 5.

2. The model.

2.1. The model network. We base our analysis on a model network of linear-nonlinear neurons that builds on the models we have presented previously [12, 10, 11, 9]. Let n be the (presumably unknown) number of neurons in the network. Let the random vector¹ \mathbf{X} denote the stimulus. The components of \mathbf{X} represent the spatio-temporal sequence of stimulus values, such as the pixel values for each refresh of a computer monitor.

The response of neuron $q = 1, 2, \dots, n$ will depend on the convolution of the stimulus with a spatio-temporal kernel² $\bar{\mathbf{h}}_q$, normalized so that $\|\bar{\mathbf{h}}_q\| = 1$. To make later notation simpler, we view the kernel $\bar{\mathbf{h}}_q$ as sliding along the stimulus with time, and denote by $\bar{\mathbf{h}}_q^i$ the kernel shifted for the discrete time point i . We implicitly view the temporal index of the stimulus as going backward in time, and write the convolution of the kernel with the stimulus as the dot product $\bar{\mathbf{h}}_q^i \cdot \mathbf{X}$.

Let the binary vector \mathbf{R} represent the spike times of neurons in the network. A component $R_q^i = 1$ indicates that neuron q spiked at time i ; otherwise, $R_q^i = 0$. When neuron p spikes, the probability that neuron q spikes j time steps later is modified by the connectivity factor \bar{W}_{pq}^j . The quantity \bar{W}_{pq}^j is simply added to the convolution $\bar{\mathbf{h}}_q^i \cdot \mathbf{X}$.

The only nonlinear part of the linear-nonlinear model is that the above linear sum is composed with a static monotonically increasing nonlinearity $\bar{g}_q(\cdot)$. This output nonlinearity represents, for example, the neuron’s spike generating mechanism and ensures that spiking probabilities stay between zero and one. The resulting linear-nonlinear network model is the following expression for the probability of a spike of neuron q at time i , conditioned on the stimulus and previous spikes (denoted $\mathbf{R}^{<i}$):

$$(2.1) \quad \Pr(R_q^i = 1 | \mathbf{X} = \mathbf{x}, \mathbf{R}^{<i} = \mathbf{r}^{<i}) = \bar{g}_q \left(\bar{\mathbf{h}}_q^i \cdot \mathbf{x} + \sum_{p=1}^n \sum_{j>0} \bar{W}_{pq}^j r^{i-j} \right).$$

We let the recent stimulus \mathbf{X} be a discrete approximation to temporal or spatio-temporal Gaussian white noise. For the analysis, we need to estimate stimulus-spike statistics conditioned on the stimulus. The estimation of these statistics implicitly assumes that we repeat each realization of the white noise stimulus multiple times.

¹With the exceptions of \bar{W} , we will use capital variables to denote random quantities.

²We use overbars (e.g., $\bar{\mathbf{h}}$) to indicate original model parameters, and will remove the bars (e.g., \mathbf{h}) to indicate their estimates from data. In addition, we use subscripts to denote neuron index, and superscripts to denote temporal indices.

2.2. The weak coupling assumption. To facilitate our analysis, we make a weak coupling assumption, which asserts that the coupling \bar{W}_{pq}^j is sufficiently small to justify a second-order approximation in \bar{W} . This assumption is really an assumption on how \bar{W} scales with the number of neurons n . As one expands equations such as (2.1) in powers of \bar{W} , one obtains terms that are k th-order in \bar{W} summed over the population k times. Hence, one obtains terms of the magnitude $(n\langle\bar{W}\rangle)^k$, where $\langle\bar{W}\rangle$ is an average of n values of \bar{W} . To truncate this series at finite k , one at minimum needs $n\langle\bar{W}\rangle < 1$. For a densely coupled large network, the coupling strength must, on average, scale at most like $1/n$. (Individual connections could be stronger, as long as the average scales like $1/n$.) We compute an approximation of order $k = 2$, and we ignore all terms that are third-order or higher in \bar{W} .

In our analysis, we go one step further. We ignore all second-order terms that are not summed over the population. Since, in this case, we are not summing over the population, it is no longer a scaling argument. This approximation simply asserts that any one connection cannot be too large. We will use \approx to indicate equality within this modified second-order approximation in \bar{W} .

We use this approximation out of necessity, not because we believe it is justified by the biology. However, we demonstrate with simulations that the results often still hold even for larger coupling than needed for the analytic results.

2.3. Effective uncoupled neuron model. Our first step is to fit the spikes of each neuron separately to an uncoupled linear-nonlinear model of the form³

$$(2.2) \quad \Pr(R_q^i = 1 | \mathbf{X} = \mathbf{x}) = g_q(\mathbf{h}_q^i \cdot \mathbf{x}),$$

where $\|\mathbf{h}_q^i\| = 1$. For the purpose of subpopulation definitions, below, we imagine we can do this for all neurons. In practice, of course, we can fit uncoupled models only to the two measured neurons. Fitting the uncoupled model (2.2) when the spikes were actually generated by the network model (2.1) defines the effective nonlinearities $g_q(\cdot)$ and kernels \mathbf{h}_q^i .

We derive expressions for the effective parameters in terms of the original parameters plus coupling effects. We simply need to calculate $\Pr(R_q^i = 1 | \mathbf{X} = \mathbf{x})$ from the network model (2.1). Because we assume a second-order approximation in coupling strength \bar{W} , it turns out that a first-order approximation in \bar{W} is sufficient for the effective single-neuron parameters.⁴

From a trivial generalization of the calculation in Appendix A.1 of [11], we can average the network model (2.1) over all spikes before time i to conclude that the probability of a spike at time i is

$$(2.3) \quad \Pr(R_q^i = 1 | \mathbf{X} = \mathbf{x}) = \bar{g}_q(\bar{\mathbf{h}}_1^i \cdot \mathbf{x}) + \sum_{p=1}^n \sum_{j>0} \bar{W}_{pq}^j \bar{g}'_q(\bar{\mathbf{h}}_1^i \cdot \mathbf{x}) \bar{g}_p(\bar{\mathbf{h}}_p^{i-j} \cdot \mathbf{x}) + O(\bar{W}^2).$$

Combining this expression with the uncoupled model (2.2), we obtain the following relationship between the effective kernels \mathbf{h}_q^i and nonlinearities $g_q(\cdot)$, on one hand,

³Note the absence of bars to indicate effective parameters that can be estimated from data (at least for the measured neurons).

⁴We will show that all terms for the spike-pair statistics will be first- or second-order in \bar{W} (all zero-order terms cancel out), and thus approximating the single-neuron parameters to first order is sufficient to retain a second-order approximation for the spike-pair statistics.

and the original model kernels $\bar{\mathbf{h}}_q^i$ and nonlinearities $\bar{g}_q(\cdot)$, on the other hand:

$$(2.4) \quad g_q(\mathbf{h}_q^i \cdot \mathbf{x}) = \bar{g}_q(\bar{\mathbf{h}}_q^i \cdot \mathbf{x}) + \sum_{p=1}^n \sum_{j>0} \bar{W}_{pq}^j \bar{g}'_q(\bar{\mathbf{h}}_q^i \cdot \mathbf{x}) \bar{g}_p(\bar{\mathbf{h}}_p^{i-j} \cdot \mathbf{x}) + O(\bar{W}^2).$$

Since $g_q(\mathbf{h}_q^i \cdot \mathbf{x})$ and $\bar{g}_q(\bar{\mathbf{h}}_q^i \cdot \mathbf{x})$ differ by only a first-order correction, and we are computing only to first order, we can simply erase the bars from \bar{g} and $\bar{\mathbf{h}}$ in the \bar{W}_{pq}^j term (creating a second-order error) to obtain

$$(2.5) \quad \bar{g}_q(\bar{\mathbf{h}}_q^i \cdot \mathbf{x}) = g_q(\mathbf{h}_q^i \cdot \mathbf{x}) - \sum_{p=1}^n \sum_{j>0} \bar{W}_{pq}^j g'_q(\mathbf{h}_q^i \cdot \mathbf{x}) g_p(\mathbf{h}_p^{i-j} \cdot \mathbf{x}) + O(\bar{W}^2).$$

This effective parameter relationship will be used in the following analysis to express all equations in terms of the effective parameters.

2.4. Subpopulation definition. A subpopulation is a group of neurons that respond to the stimulus in a similar manner. The effective kernel \mathbf{h} derived from fitting a neuron's spikes to the uncoupled model (2.2) describes the relationship of neuronal spikes to the stimulus. (In some contexts, this kernel would be referred to as the neuron's *receptive field*.) We base our subpopulation definitions on this effective kernel. We define the similarity between two neurons based on the correlation coefficient between the linear components from the uncoupled model (2.2):

$$(2.6) \quad cc_{pq}^k = \text{cor}(\mathbf{h}_p^i \cdot \mathbf{X}, \mathbf{h}_q^{i-k} \cdot \mathbf{X}),$$

where

$$\text{cor}(A, B) = \frac{\text{cov}(A, B)}{\sqrt{\text{var}(A) \text{var}(B)}}.$$

Note that $-1 \leq cc_{pq}^k \leq 1$. In fact, since each component of \mathbf{X} is a unit normal random variable, the correlation coefficient is simply the cosine of the angle between the kernels:

$$(2.7) \quad cc_{pq}^k = \frac{\mathbf{h}_p^i \cdot \mathbf{h}_q^{i-k}}{\|\mathbf{h}_p^i\| \|\mathbf{h}_q^{i-k}\|} = \mathbf{h}_p^i \cdot \mathbf{h}_q^{i-k}.$$

(The last equality results because the kernels are normalized to be unit vectors.)

Define the maximum correlation coefficient as

$$(2.8) \quad cc_{pq}^{\max} = \max_k cc_{pq}^k.$$

If cc_{pq}^{\max} is large, then neurons p and q respond to the stimulus similarly, and we consider the neurons as part of the same subpopulation. On the other hand, if cc_{pq}^{\max} is small, then we consider the neurons as parts of different subpopulations. For the analysis, when we assume that neurons p and q are from different subpopulations, we will effectively assume that each cc_{pq}^k is $O(\bar{W})$. We show via simulations that, in practice, we can relax this condition somewhat.

3. The analysis.

3.1. Overview of the analysis. We assume we have access to only the spikes of neuron one and two (R_1^i and R_2^i) as well as the discrete white noise stimulus \mathbf{X} . Given the stimulus, the probability of spikes from the network is specified by the network model (2.1). We initially assume that all unmeasured neurons (with index $p > 2$) are from different subpopulations than those of neurons one or two (in particular, that cc_{p1}^k and cc_{p2}^k for all k are $O(\bar{W})$). Using this assumption, we can solve for the direct connection (\bar{W}_{21}^j and \bar{W}_{12}^j) in terms of the joint statistics of the random variables R_1^i , R_2^i , and \mathbf{X} .

When we allow unmeasured neurons from the same subpopulations as the measured neurons, we do not change the algorithm to determine a direct connection. We demonstrate that, with this algorithm, common input from neuron one's subpopulation may be identified as a direct connection from neuron one onto neuron two. Similarly, common input from neuron two's subpopulation may be identified as a direct connection from neuron two onto neuron one. Although this results in a misidentification at the level of individual neurons, it still accurately identifies connectivity at the level of subpopulations (since, for example, common input from neuron two's subpopulation does contain a connection from neuron two's subpopulation onto neuron one's subpopulation).

For this analysis, we assume that we have an infinite dataset, so we can estimate the expected values of functions of the random variables. Since in practice, we will have much smaller datasets, we must reduce the bias in estimations from finite datasets using a procedure such as that outlined in [12, 11, 8]. We do not address such bias reduction here.

We give a brief overview of the analysis here and give more details of each step in the following sections. In the first step, we analyze the spikes of neuron one and neuron two separately. From their stimulus-spike statistics, we fit uncoupled linear-nonlinear models (2.2) as if we were using standard white noise analysis methods such as those outlined in [12]. This calculation is based on the mean spike rates⁵ ($E\{R_1^i\}$ and $E\{R_2^i\}$) and the stimulus-spike correlations ($E\{\mathbf{X}R_1^i\}$ and $E\{\mathbf{X}R_2^i\}$). Since the spike times are really given by the network model (2.1), the effective kernels (\mathbf{h}_q^i and \mathbf{h}_2^i) and nonlinearities ($g_1(\cdot)$ and $g_2(\cdot)$) are functions of network model parameters (including coupling and parameters from other neurons), as given by (2.4).

Next, we calculate the spike rates conditioned on a particular realization of the stimulus⁶ ($E\{R_1^i|\mathbf{X}\}$ and $E\{R_2^i|\mathbf{X}\}$). These statistics are equivalent to the peristimulus time histogram (PSTH) commonly used in the neuroscience literature.

We look at spike pairs, where neuron two spikes k units of time before neuron one (note that k could be positive or negative). We subtract off the product of the PSTHs from the rate of spike pairs conditioned on the stimulus, forming

$$E\{R_1^i R_2^{i-k}|\mathbf{X}\} - E\{R_1^i|\mathbf{X}\}E\{R_2^{i-k}|\mathbf{X}\}.$$

The result is the JPSTH cast into the notation of the model.

If we take the expected value of the JPSTH over all realizations of the stimulus,

⁵Note that, due to the stationarity of the stimulus and model, many stimulus-spike statistics do not depend on time, despite the notation. The mean rates $E\{R_q^i\}$, for example, do not depend on the time point i .

⁶Here we assume that each realization of the stimulus is repeated multiple times.

we obtain the shuffle-corrected correlogram or covariogram

$$(3.1) \quad C_{21}^k = E\{R_1^i R_2^{i-k}\} - E\{E\{R_1^i | \mathbf{X}\} E\{R_2^{i-k} | \mathbf{X}\}\}.$$

Here we have used the fact that $E\{E\{R_1^i R_2^{i-k} | \mathbf{X}\}\} = E\{R_1^i R_2^{i-k}\}$. For a given value of k , C_{21}^k is effectively a sum over the diagonal of the JPSTH corresponding to the delay k . From analysis of the network model (2.1), we derive an equation for C_{21}^k in terms of model parameters.

As we argued in the introduction, the covariogram alone is insufficient to distinguish common input from a direct connection. In terms of the model parameters, there are too many unknowns to solve for \bar{W}_{21}^k (or \bar{W}_{12}^{-k} if $k < 0$). To obtain more equations, we combine white noise analysis methods with the JPSTH.

The key of the approach is to calculate the correlation⁷ of the JPSTH with the stimulus:

$$(3.2) \quad D_{21}^{ki} = E\{\mathbf{X} R_1^i R_2^{i-k}\} - E\{\mathbf{X} E\{R_1^i | \mathbf{X}\} E\{R_2^{i-k} | \mathbf{X}\}\}.$$

Note that the stimulus-spike correlations (e.g., $E\{\mathbf{X} R_1^i\}$) were calculated by correlating the stimulus with a binary vector (e.g., the R_1^i). The above statistic D_{21}^{ki} is a correlation of the stimulus not with a spike vector, but with the vector composed of values from the diagonal of the JPSTH corresponding to delay k . This vector is, of course, not binary, but the correlation can be computed nearly identically. For a fixed k , the result D_{21}^{ki} will be a vector of the same dimension as the correlations $E\{\mathbf{X} R_1^i\}$ and $E\{\mathbf{X} R_2^{i-k}\}$ and hence the same dimension as the kernels \mathbf{h}_1^i and \mathbf{h}_2^{i-k} .

Consequently, for a given k , we can decompose D_{21}^{ki} into components parallel to the kernels \mathbf{h}_1^i and \mathbf{h}_2^{i-k} , calculating the coefficients A_1^k and A_2^k for which

$$(3.3) \quad D_{21}^{ki} = A_1^k \mathbf{h}_1^i + A_2^k \mathbf{h}_2^{i-k} + \mathbf{O}^{ki},$$

where \mathbf{O}^{ki} is perpendicular to \mathbf{h}_1^i and \mathbf{h}_2^{i-k} . By analyzing the network model (2.1), we calculate expressions for A_1^k and A_2^k in terms of model parameters. From C_{21}^k , A_1^k , and A_2^k , we have three equations for each delay k .

If one compares the number of unknown parameters with the number of equations, the situation still looks hopeless. Assume that we calculate the statistics for the delays $k = \pm 1, \pm 2, \dots, \pm M$, so that we have $2M \times 3 = 6M$ equations.⁸ Assume also that the coupling is zero for delays longer than M time units. Then the coupling parameters are \bar{W}_{pq}^j for $j \in \{1, 2, \dots, M\}$ and $p, q \in \{1, 2, \dots, n\}$, where n is the (presumably unknown) number of neurons, for a total of Mn^2 parameters.

If the number of neurons is more than two, the system appears vastly underdetermined. This limitation make sense. If we were sufficiently audacious as to claim that we could reconstruct the coupling of the entire network based on measures of just two neurons, our absurdity would be exposed by this reality check. Our goal is simply to estimate the direct connection \bar{W}_{21}^j and \bar{W}_{12}^j between the two measured neurons.

As demonstrated below, if we assume that unmeasured neurons are from different subpopulations than the measured neurons, all the coupling terms involving unmeasured neurons appear in the same combination⁹ in all three sets of equations. We denote this combination by \hat{U}^k , and refer to \hat{U}^k as the common input contribution¹⁰

⁷We use the term ‘‘correlation’’ loosely.

⁸We ignore the single-neuron statistics and single-neuron parameters for this rough calculation.

⁹For this overview, we ignore the presence of indirect connections (see section 3.3).

¹⁰It turns out that \hat{U}^k contains only combinations of coupling terms that correspond to common input.

to delay k . This notation makes it clear there are really only two unknowns per delay k (\hat{U}^k and either \bar{W}_{21}^k or \bar{W}_{12}^{-k} , depending on whether k is positive or negative, respectively). We have a total of $2M \times 2 = 4M$ parameters for $6M$ equations.¹¹ The system is actually overdetermined. Moreover, due to the weak coupling assumption of section 2.2, the system is linear in \hat{U}^k , \bar{W}_{21}^k , and \bar{W}_{12}^{-k} . We can easily solve it via least squares and estimate the direct connection between neurons one and two as well as the effective common input.

This estimate is, of course, valid only when unmeasured neurons are from different subpopulations than the measured neurons. We address the case of unmeasured neurons from the same subpopulations as the measured neurons in section 3.4. There we argue that our estimate of \bar{W}_{21}^k or \bar{W}_{12}^{-k} accurately reconstructs connectivity between the subpopulations of neuron one and neuron two.

3.2. Single-neuron statistics. In the following sections, we present more details of the analysis outlined above. As many of the calculations are long, we present only the key details, referring where possible to similar calculations from previous papers.

For each of the measured neurons $q = 1, 2$, we analyze its spikes R_q^i and the stimulus \mathbf{X} as though the uncoupled model (2.2) held. We view the parameters from the uncoupled model (2.2) as effective parameters that can be estimated from the stimulus-spike statistics. We have already calculated the effective parameter relationship (2.5) that relates effective parameters to the original model parameters.

In terms of the effective parameters, the stimulus-spike correlation is

$$\begin{aligned} E\{\mathbf{X}R_q^i\} &= E\{\mathbf{X}\Pr(R_q^i = 1|\mathbf{X})\} \\ &= E\{\mathbf{X}g_q(\mathbf{h}_q^i \cdot \mathbf{X})\} \\ (3.4) \quad &= E\{g'_q(\mathbf{h}_q^i \cdot \mathbf{X})\}\mathbf{h}_q^i, \end{aligned}$$

and the mean rate is

$$\begin{aligned} E\{R_q^i\} &= E\{\Pr(R_q^i = 1|\mathbf{X})\} \\ (3.5) \quad &= E\{g_q(\mathbf{h}_q^i \cdot \mathbf{X})\}, \end{aligned}$$

where we used the integration-by-parts formula (A.3) to obtain the final expression for $E\{\mathbf{X}R_q^i\}$.

Given (3.4) and the normalization $\|\mathbf{h}_q^i\| = 1$, the effective kernel can be calculated from the stimulus-spike correlation as

$$(3.6) \quad \mathbf{h}_q^i = \frac{E\{\mathbf{X}R_q^i\}}{\|E\{\mathbf{X}R_q^i\}\|}.$$

If we assume a two-parameter family of nonlinear functions for $g_q(\cdot)$, we can calculate those parameters from $E\{R_q^i\}$ and $\|E\{\mathbf{X}R_q^i\}\|$ (see [12]).

3.3. Neuron pair statistics. We repeat the calculations of Appendices A and B of [11], computing terms only up to the modified second-order approximation in \bar{W} , described above in section 2.2. To simplify the notation, we define $\bar{W}_{pq}^k = 0$ for $k \leq 0$.

¹¹Although we could, in principle, look for higher-order corrections by retaining higher-order terms in \bar{W} , the system would not collapse to $4M$ parameters, and we would have to look for more equations.

After long, tedious calculations and use of the effective parameter relationship (2.5), most of the terms cancel out, and we are left with

$$\begin{aligned}
& \Pr(R_1^i = 1 \& R_2^{i-k} = 1 | \mathbf{X} = \mathbf{x}) - \Pr(R_1^i = 1 | \mathbf{X} = \mathbf{x}) \Pr(R_2^{i-k} = 1 | \mathbf{X} = \mathbf{x}) \\
& \approx \left[\bar{W}_{21}^k + \sum_{p=3}^n \sum_{j>0} \bar{W}_{2p}^{k-j} \bar{W}_{p1}^j g'_p(\mathbf{h}_p^{i-j} \cdot \mathbf{x}) \right] g'_1(\mathbf{h}_1^i \cdot \mathbf{x}) g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{x}) [1 - g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{x})] \\
& + \left[\bar{W}_{12}^{-k} + \sum_{p=3}^n \sum_{j>0} \bar{W}_{1p}^j \bar{W}_{p2}^{-k-j} g'_p(\mathbf{h}_p^{i+j} \cdot \mathbf{x}) \right] g_1(\mathbf{h}_1^i \cdot \mathbf{x}) [1 - g_1(\mathbf{h}_1^i \cdot \mathbf{x})] g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{x}) \\
(3.7) \quad & + \sum_{p=3}^n \sum_{j>\max(0,k)} \bar{W}_{p1}^j \bar{W}_{p2}^{j-k} g'_1(\mathbf{h}_1^i \cdot \mathbf{x}) g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{x}) g_p(\mathbf{h}_p^{i-j} \cdot \mathbf{x}) [1 - g_p(\mathbf{h}_p^{i-j} \cdot \mathbf{x})],
\end{aligned}$$

where \approx indicates equality within our modified second-order approximation in \bar{W} . Note that if $k \leq 0$, then $\bar{W}_{21}^k = 0$ and $\bar{W}_{2p}^{k-j} = 0$, and the first expression in square brackets is zero. On the other hand, if $k \geq 0$, then $\bar{W}_{12}^{-k} = 0$ and $\bar{W}_{p2}^{-k-j} = 0$, and the second expression in square brackets is zero. Consequently, either the first or the second term is zero for any given k .

This expression is the expected value of the JPSTH, given that the stimulus $\mathbf{X} = \mathbf{x}$. Note that the first term is a direct connection with delay k from neuron two to neuron one, combined with an indirect connection through neuron p of total delay k . The second term is a direct connection, combined with an indirect connection, from neuron one to neuron two (of total delay $-k$, which is positive when this term is nonzero). The last term is due to common input from neuron p onto both neuron one and neuron two. (The expression $\bar{W}_{p1}^j \bar{W}_{p2}^{j-k}$ is nonzero only if neuron p is connected to both neuron one and neuron two.)

The covariogram (3.1) is the expected value of the JPSTH (3.7), and the statistic D (3.2) is the expected value of the JPSTH (3.7) times the stimulus \mathbf{X} . Without further assumptions on the unmeasured neurons, we cannot dissociate the contribution of unmeasured neurons from the contribution of measured neurons. In order to solve the equations, we assume that we can factor each expected value into (A) the expected value of an expression involving unmeasured neuron parameters multiplied by (B) the expected value of an expression involving measured neuron parameters. Note that unmeasured neuron parameters appear only in those terms that are second-order in \bar{W} . Given our second-order approximation in \bar{W} , this step assumes that, to zeroth order in \bar{W} , the $g_p(\mathbf{h}_p^{i-j} \cdot \mathbf{X})$ are independent of $g_1(\mathbf{h}_1^i \cdot \mathbf{X})$ and $g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})$ (i.e., the effective uncoupled models (2.2) for unmeasured neurons are independent of the effective uncoupled models for measured neurons). In particular, we are assuming that cc_{p1}^j and cc_{p2}^{j-k} are $O(\bar{W})$, which means that the unmeasured neurons are from different subpopulations than the measured neurons (as defined in section 2.4).

Under this assumption the covariogram (3.1) (i.e., the expected value of the JPSTH (3.7)) becomes¹²

$$\begin{aligned}
C_{21}^k & \approx \hat{W}_{21}^k E\{g'_1(\mathbf{h}_1^i \cdot \mathbf{X}) g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) [1 - g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\} \\
& + \hat{W}_{12}^{-k} E\{g_1(\mathbf{h}_1^i \cdot \mathbf{X}) [1 - g_1(\mathbf{h}_1^i \cdot \mathbf{X})] g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})\} \\
(3.8) \quad & + \hat{U}_{21}^k E\{g'_1(\mathbf{h}_1^i \cdot \mathbf{X}) g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})\},
\end{aligned}$$

¹²Recall that stationarity of the stimulus and model imply that the statistics in the equation for C_{21}^k (as well as A_1^k and A_2^k) do not depend on time point i , despite the notation.

where \approx indicates equality within our modified second-order approximation in \bar{W} and

$$\begin{aligned}
\hat{W}_{21}^k &= \bar{W}_{21}^k + \sum_{p=3}^n \sum_{j>0} \bar{W}_{2p}^{k-j} \bar{W}_{p1}^j E\{g'_p(\mathbf{h}_p^{i-j} \cdot \mathbf{X})\}, \\
\hat{W}_{12}^{-k} &= \bar{W}_{12}^{-k} + \sum_{p=3}^n \sum_{j>0} \bar{W}_{1p}^j \bar{W}_{p2}^{-k-j} E\{g'_p(\mathbf{h}_p^{i+j} \cdot \mathbf{X})\}, \\
(3.9) \quad \hat{U}_{21}^k &= \sum_{p=3}^n \sum_{j>\max(0,k)} \bar{W}_{p1}^j \bar{W}_{p2}^{j-k} E\{g_p(\mathbf{h}_p^{i-j} \cdot \mathbf{X})[1 - g_p(\mathbf{h}_p^{i-j} \cdot \mathbf{X})]\}.
\end{aligned}$$

The new parameters \hat{W}_{21}^k and \hat{W}_{12}^{-k} are the effective direct connections between neurons one and two. (By definition, $\hat{W}_{21}^k = 0$ for $k \leq 0$ and $\hat{W}_{12}^{-k} = 0$ for $k \geq 0$.) Note that this effective direct connection factor \hat{W} is a combination of both the direct connections and the indirect connections through any unmeasured neuron p . This fact indicates that we cannot distinguish between direct connections and indirect connections through unmeasured neurons. Our goal is to distinguish these effective direct connections \hat{W} from the effective common input \hat{U} , which is the sum total effect from all unmeasured neurons p that project to both neuron one and neuron two.

We multiply the JPSTH (3.7) by \mathbf{X} , take the expected value, and use the integration-by-parts formula (A.3) to obtain an expression for D_{21}^{ki} (given by (3.2))

$$(3.10) \quad D_{21}^{ki} \approx A_1^k \mathbf{h}_1^i + A_2^k \mathbf{h}_2^{i-k} + \mathbf{O}^{ki},$$

where \mathbf{O}^{ki} is an expression that is $O(\bar{W}^2)$ and involves the kernels of the unmeasured neurons. Since we are assuming that the unmeasured neurons are from different subpopulations than the measured neurons, \mathbf{O}^{ki} can be viewed as orthogonal to¹³ \mathbf{h}_1^i and \mathbf{h}_2^{i-k} . The components of D_{21}^{ki} parallel to \mathbf{h}_1^i and \mathbf{h}_2^{i-k} are

$$\begin{aligned}
A_1^k &= \hat{W}_{21}^k E\{g_1''(\mathbf{h}_1^i \cdot \mathbf{X})g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})[1 - g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\} \\
&\quad + \hat{W}_{12}^{-k} E\{g_1'(\mathbf{h}_1^i \cdot \mathbf{X})[1 - 2g_1(\mathbf{h}_1^i \cdot \mathbf{X})]g_2'(\mathbf{h}_2^{i-k} \cdot \mathbf{X})\} \\
&\quad + \hat{U}_{21}^k E\{g_1''(\mathbf{h}_1^i \cdot \mathbf{X})g_2'(\mathbf{h}_2^{i-k} \cdot \mathbf{X})\}, \\
A_2^k &= \hat{W}_{21}^k E\{g_1'(\mathbf{h}_1^i \cdot \mathbf{X})g_2'(\mathbf{h}_2^{i-k} \cdot \mathbf{X})[1 - 2g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\} \\
&\quad + \hat{W}_{12}^{-k} E\{g_1(\mathbf{h}_1^i \cdot \mathbf{X})[1 - g_1(\mathbf{h}_1^i \cdot \mathbf{X})]g_2''(\mathbf{h}_2^{i-k} \cdot \mathbf{X})\} \\
(3.11) \quad &\quad + \hat{U}_{21}^k E\{g_1'(\mathbf{h}_1^i \cdot \mathbf{X})g_2''(\mathbf{h}_2^{i-k} \cdot \mathbf{X})\}.
\end{aligned}$$

From measuring the spikes of neuron one and two (R_1^i and R_2^i) in response to the stimulus \mathbf{X} , we can calculate the effective uncoupled model parameters ($g_1(\cdot)$, $g_2(\cdot)$, \mathbf{h}_1^i , and \mathbf{h}_2^{i-k}), the covariogram C_{21}^k (via (3.1)), the statistic D_{21}^{ki} (via (3.2)), and consequently its components A_1^k and A_2^k . The nine expected values in (3.8) and (3.11) are simply Gaussian integrals of known functions that can be calculated. The only unknown quantities are the \hat{W}_{12}^{-k} , \hat{W}_{21}^k , and \hat{U}_{21}^k .

To emphasize that the direct connection is simply one variable per delay, we define

¹³Since, for any $p > 2$ and any j , the correlation coefficients cc_{p1}^j and cc_{p2}^{j-k} are assumed to be $O(\bar{W})$, the component of \mathbf{h}_p^{i-j} parallel to \mathbf{h}_1^i or \mathbf{h}_2^{i-k} is $O(\bar{W})$. Hence, the component of \mathbf{O}^{ki} parallel to these kernels must be $O(\bar{W}^3)$, which we can ignore.

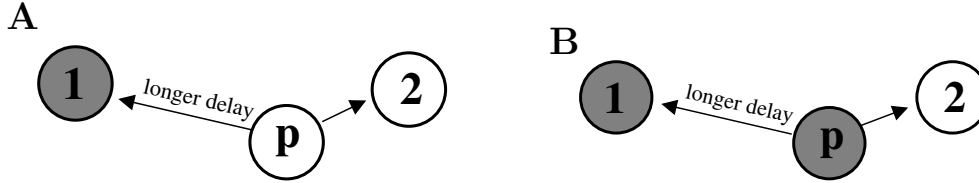


FIG. 2. Schematic of common input from unmeasured neuron p onto measured neurons one and two. The gray shading indicates which neurons are from the same subpopulation. The connection from neuron p to neuron one is delayed, so the common input introduces a correlation between the spikes of neuron one and two that is similar to the correlation induced by a connection from neuron two to neuron one. (A) Neuron p is within neuron two's subpopulation. In this case, the common input may appear as a direct connection from neuron two onto neuron one. (B) Neuron p is within neuron one's subpopulation. In this case, the common input will not be confused with a direct connection.

a new direct connection variable

$$(3.12) \quad \hat{W}^k = \begin{cases} \hat{W}_{12}^{-k} & \text{for } k < 0, \\ 0 & \text{for } k = 0, \\ \hat{W}_{21}^k & \text{for } k > 0. \end{cases}$$

Note that the equations for different delays k are uncoupled. For each $k \neq 0$, equations (3.8) and (3.11) are three linear equations for the two unknowns \hat{U}_{21}^k and \hat{W}^k . This system is easily solved for the two unknowns using linear least squares. For $k = 0$, the only unknown is \hat{U}_{21}^0 , for which we solve using (3.8).

3.4. Common input from the measured neurons' subpopulations. To complete the above analysis, we assumed that the unmeasured neurons were from different subpopulations than the measured neurons. In particular, we assumed that the correlation coefficients between the measured neurons and the unmeasured neurons (cc_{p1}^k and cc_{p2}^k for $p > 2$ and all k) were small.

In the brain, one typically finds groups of neurons that respond to a stimulus in a similar way and hence would be from the same subpopulation, as we defined them in section 2.4. Consequently, one would anticipate the presence of unmeasured neurons from the subpopulations of both neuron one and neuron two. Since such measured neurons could be the source of common input, we must address the case of common input from the neurons within the measured neurons' subpopulations. (With the exception of common input and indirect connection effects, all effects of unmeasured neurons had already been canceled in the analysis leading to the JPSTH (3.7), before we made any assumptions about subpopulations.)

We examine networks with a common input configuration where an unmeasured neuron p has a connection onto neuron two, and, with a longer delay, a connection onto neuron one, as schematized in Figure 2. To implement this, we let the connection onto neuron one have a delay of j time steps ($\bar{W}_{p1}^j \neq 0$) and the connection onto neuron two have a delay of $j - k$ time steps ($\bar{W}_{p2}^{j-k} \neq 0$), with $j > k > 0$. With this set of delays, the common input will introduce correlations between the measured neurons that mimic a direct connection from neuron two onto neuron one with a delay of k time steps.

We first show how common input from neuron two's subpopulation (Figure 2(A)) can be misidentified as a direct connection from neuron two onto neuron one. (Note

that this results in a correct identification of a connection from neuron two's subpopulation onto neuron one's subpopulation.) Since the connection from neuron p onto neuron two is due to the term $\bar{W}_{p2}^{j-k} \neq 0$, the correlation coefficient cc_{p2}^{j-k} will determine whether the neuron p acts as a member of neuron two's subpopulation (this inference comes from inspection of the last term in the JPSTH (3.7)). If cc_{p2}^{j-k} is large, neuron p will act as a member of neuron two's subpopulation.

We examine the extreme case where neuron p responds to nearly the same stimulus features as neuron two does $j - k$ time steps later; i.e., $\mathbf{h}_p^{i-j} = \mathbf{h}_2^{i-k} + O(\bar{W})$. Consequently, $cc_{p2}^{j-k} = 1 + O(\bar{W})$. Then, the contribution of this neuron to the JPSTH (3.7) is

$$\bar{W}_{p1}^j \bar{W}_{p2}^{j-k} g'_1(\mathbf{h}_1^i \cdot \mathbf{x}) g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{x}) g_p(\mathbf{h}_2^{i-k} \cdot \mathbf{x}) [1 - g_p(\mathbf{h}_2^{i-k} \cdot \mathbf{x})].$$

The contribution of neuron p to the covariogram is

$$(3.13) \quad C_{21}^k = \bar{W}_{p1}^j \bar{W}_{p2}^{j-k} E\{g'_1(\mathbf{h}_1^i \cdot \mathbf{X}) g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) g_p(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) [1 - g_p(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\}.$$

Since the kernel \mathbf{h}_p^{i-j} is identical to \mathbf{h}_2^{i-k} , an additional term will appear in A_2^k after the integration by parts, so that the contribution of neuron p to A_1^k and A_2^k is

$$(3.14) \quad \begin{aligned} A_1^k &= \bar{W}_{p1}^j \bar{W}_{p2}^{j-k} E\{g''_1(\mathbf{h}_1^i \cdot \mathbf{X}) g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) g_p(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) [1 - g_p(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\}, \\ A_2^k &= \bar{W}_{p1}^j \bar{W}_{p2}^{j-k} E\{g'_1(\mathbf{h}_1^i \cdot \mathbf{X}) g''_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) g_p(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) [1 - g_p(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\} \\ &\quad + \bar{W}_{p1}^j \bar{W}_{p2}^{j-k} E\{g'_1(\mathbf{h}_1^i \cdot \mathbf{X}) g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) g'_p(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) [1 - 2g_p(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\}. \end{aligned}$$

Compare this contribution to the effect of a direct connection from neuron two to neuron one at a delay of k units of time (the \bar{W}_{21}^k terms from (3.8) and (3.11)):

$$(3.15) \quad \begin{aligned} C_{21}^k &= \hat{W}_{21}^k E\{g'_1(\mathbf{h}_1^i \cdot \mathbf{X}) g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) [1 - g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\}, \\ A_1^k &= \hat{W}_{21}^k E\{g''_1(\mathbf{h}_1^i \cdot \mathbf{X}) g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) [1 - g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\}, \\ A_2^k &= \hat{W}_{21}^k E\{g'_1(\mathbf{h}_1^i \cdot \mathbf{X}) g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X}) [1 - 2g_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})]\}. \end{aligned}$$

Ignoring the first term of A_2^k in (3.14), we observe that the relationship among C_{21}^k , A_1^k , and A_2^k in (3.13) and (3.14) is nearly identical to their relationship in (3.15). For the case $g_p = g_2$, the only difference is the additional common factor of $g'_2(\mathbf{h}_2^{i-k} \cdot \mathbf{X})$ in the expected value.

If the second term of A_2^k in (3.14) does dominate the first term, then the common input from neuron two's subpopulation leads to a relationship among the statistics C_{21}^k , A_1^k , and A_2^k that mimics a direct connection from neuron two to neuron one. Consequently, we would expect that applying the results of section 3.3 would indicate the presence of a direct connection. Simulations confirm that the second term of A_2^k in (3.14) does indeed dominate, as network configurations such as Figure 2(A) are categorized as direct connection configurations.

We next consider the case where neuron p is from neuron one's subpopulation (Figure 2(B)). If neuron p responds to the stimulus almost exactly as neuron one does, j time steps later, the analysis does not give a clear answer. If we assume $\mathbf{h}_p^{i-j} = \mathbf{h}_1^i + O(\bar{W})$ so that $cc_{p1}^j = 1 + O(\bar{W})$, we do not obtain a relationship among C_{21}^k , A_1^k , and A_2^k that mimics their relationship in (3.15). In this case, simulations indicate that this configuration appears as common input. (Of course, if the connection from

neuron p to neuron two had the longer delay, the network would be equivalent to Figure 2(A) with the roles of neuron one and two reversed. In this case, the network would mimic a direct connection from neuron one to neuron two.)

We conclude that we cannot distinguish between a direct connection from neuron two onto neuron one and common input from neuron two's subpopulation. (An equivalent statement holds with the roles of neuron one and neuron two reversed.) Since common input from neuron two's subpopulation does contain a connection from neuron two's subpopulation onto neuron one, we conclude that the connectivity is correctly identified at the level of subpopulations. In applications where the distinction among particular neurons within a subpopulation is unimportant, the ambiguity in the precise identification of the connection is not problematic. See the Discussion for more details.

4. Tests via simulation.

4.1. Simulations of small linear-nonlinear networks. We tested our analytic results with simulations of networks of linear-nonlinear neurons given by (2.1). We used kernels $\bar{\mathbf{h}}_q^i$ that capture some features of the responses of neurons in visual cortex [6]. For spatial grid point $\mathbf{j} = (j_1, j_2)$ and time t , the kernels were of the form

$$(4.1) \quad \bar{h}_q(\mathbf{j}, t) = (t - b_q) \exp\left(-\frac{t - b_q}{\tau_h} - \frac{|\mathbf{j}|^2}{10}\right) \sin((j_1 \cos \psi_q + j_2 \sin \psi_q)f_q + \phi_q)$$

for $t > b_q$ and $\bar{h}_q(\mathbf{j}, t) = 0$ otherwise. We sampled $\bar{h}_q(\mathbf{j}, t)$ on a $10 \times 10 \times 10$ grid and normalized it to the unit vector $\bar{\mathbf{h}}_q^i$. For the analysis, the only important geometry is the inner product between the kernels, $\bar{\mathbf{h}}_q^i \cdot \bar{\mathbf{h}}_p^{i-j}$, which is the correlation between normal random variables $\bar{\mathbf{h}}_p^{i-j} \cdot \mathbf{X}$ and $\bar{\mathbf{h}}_q^i \cdot \mathbf{X}$.

For each example, we simulated the network response to 500,000 units of time, adjusting the nonlinearities $\bar{g}_q(\cdot)$ to obtain between 10,000 and 15,000 spikes from each neuron. Each simulation was composed of 100 trials, each lasting 5,000 units of time. For ten trials, the stimulus was independent realizations of the Gaussian white noise. We repeated each realization ten times to form the 100 trials. The repetitions allowed estimation of the spiking probabilities $\Pr(R_1^i | \mathbf{X} = \mathbf{x})$ and $\Pr(R_2^{i-k} | \mathbf{X} = \mathbf{x})$ needed for the JPSTH (3.7) by averaging over the ten repetitions (equivalent to a shuffle correction).

The analysis was based on expected values of stimulus-spike statistics. Naive estimates of products of these statistics, including the shuffle correction, from finite datasets can be highly biased. We reduced these biases using techniques described elsewhere [12, 11, 8]. From the independent trials, we estimated confidence intervals as described in Appendix C.

To compute the Gaussian integrals in (3.8) and (3.11), we needed to choose a form for the nonlinear functions $g_q(y)$. To allow us to compute the integrals analytically, we assume that the nonlinear functions are error functions of the form

$$(4.2) \quad \bar{g}_q(y) = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{y - \bar{y}_q}{\bar{\epsilon}_q \sqrt{2}}\right) \right],$$

where \bar{y}_q is the threshold, $\bar{\epsilon}_q$ defines the steepness of the nonlinearity, and the error function is $\operatorname{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y e^{-t^2} dt$. Note that $\lim_{y \rightarrow \infty} g_q(y) = 1$ and $\lim_{y \rightarrow -\infty} g_q(y) = 0$. The expressions for C_{21}^k , A_1^i , and A_2^{i-k} for the case of an error function nonlinearity are given in Appendix B. We demonstrate below that the results are not sensitive to

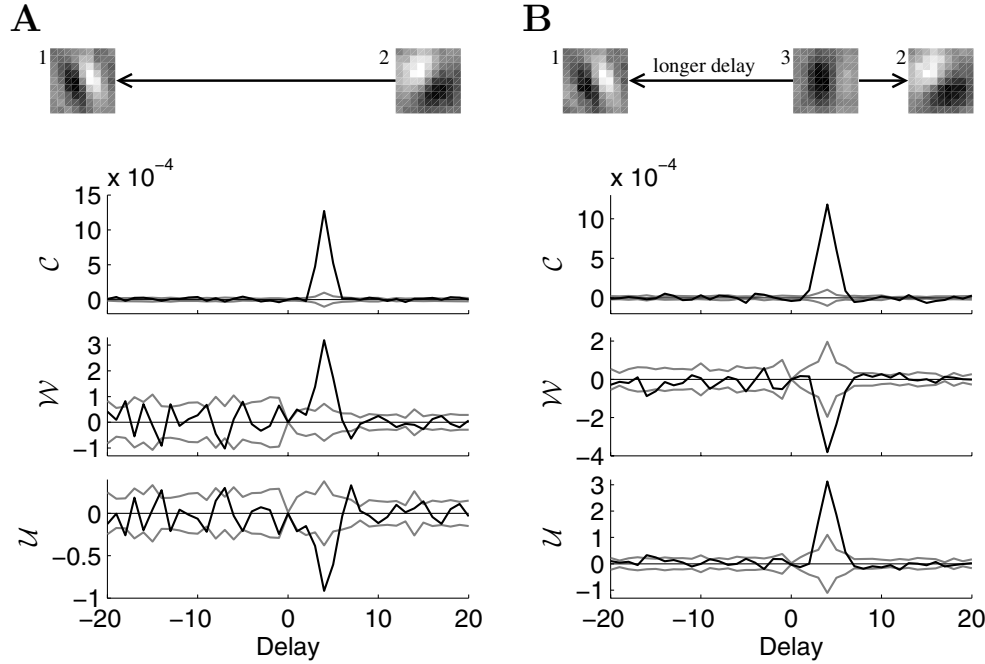


FIG. 3. Results from sample simulations showing successful distinction between (A) the direct connection configuration and (B) the common input configuration, using only the spike times from neurons one and two. The top panels are schematics of the network architecture, with grayscale plots of the time slice in which each kernel (\mathbf{h}_1 , \mathbf{h}_2 , or \mathbf{h}_3 , labeled by neuron number) reached its maximal value. (The spikes of neuron three were used to calculate \mathbf{h}_3 for this illustration. The remaining analysis used only the spikes of neuron one and two.) The bottom three panels plot with black lines the covariogram C , direct connection measure W , and common input measure U . The gray lines estimate confidence intervals of one standard error. Delay is the spike time of neuron one minus the spike time of neuron two. (A) The direct connection from neuron two onto neuron one creates a positive covariogram C around a delay of four units of time. The significantly positive direct connection measure W at that delay indicates that the correlation was due to a direct connection. The negative common input measure U , though indicating departure from the weak coupling assumption, does not confuse the direct connection interpretation. The direct connection was given by $\bar{W}_{21}^4 = 0.8$, $\bar{W}_{21}^3 = \bar{W}_{21}^5 = 0.4$. (All other \bar{W} were zero.) Parameters used: $\tau_h = 2$, $\psi_1 = \pi/8$, $\psi_2 = -\pi/4$, $\phi_1 = 0$, $\phi_2 = \pi$, $f_1 = 1.0$, $f_2 = 0.3$, $b_1 = b_2 = 0$, $\bar{T}_1 = 2.3$, $\bar{T}_2 = 2.8$, $\bar{\epsilon}_1 = 0.5$, $\bar{\epsilon}_2 = 1.0$. (B) For the network with common input from unmeasured neuron three, the covariogram C is nearly identical to the direct connection case from panel A. The fact that the correlation was due to common input is revealed by the positive U (and negative W). The common input to neuron one was delayed four more units of time compared with that to neuron two: $\bar{W}_{31}^6 = \bar{W}_{32}^2 = 1.8$, $\bar{W}_{31}^5 = \bar{W}_{31}^7 = \bar{W}_{32}^1 = \bar{W}_{32}^3 = 0.8$. Parameters as in (A) except $\psi_3 = 0$, $\phi_3 = -\pi/3$, $f_3 = 0.6$, $b_3 = 0$, $\bar{T}_1 = 2.6$, $\bar{T}_2 = 3.0$, $\bar{T}_3 = 2.4$, $\bar{\epsilon}_3 = 0.7$.

this particular choice of nonlinear function. One could perform a similar analysis for other nonlinear functions, although then one would presumably need to compute the integrals numerically.

We denote by C^k the covariogram C_{21}^k (3.1) estimated from a dataset. Similarly, we denote by W^k and U^k estimates of the direct connection \hat{W}^k (3.12) and common input \hat{U}_{21}^k (3.9), respectively.

To illustrate the method, we looked at minimal networks containing two or three neurons. First, we simulated a pair of neurons, where neuron two has a direct connection onto neuron one. The results are shown in Figure 3(A). The covariogram C shows

a peak at the delay corresponding to the connection. However, the covariogram does not indicate whether this spike correlation is due to a direct connection or common input from an unmeasured neuron.

This ambiguity is resolved by the measures \mathcal{W} and \mathcal{U} . The direct connection measure \mathcal{W} is significantly positive at the delay corresponding to the connection. On the other hand, the common input measure \mathcal{U} is negative at that delay. Hence, \mathcal{W} and \mathcal{U} indicate that the spike correlation was caused by a direct connection rather than by a common input.

Note that the noise in \mathcal{W} and \mathcal{U} is dramatically greater than in the covariogram \mathcal{C} . This increase is presumably due to the subtlety of the distinction we are attempting to make. For this reason, we required long simulations with up to 15,000 spikes to obtain good results.

The reciprocal behavior observed between \mathcal{W} and \mathcal{U} is not predicted by the analysis. According to the analysis, \mathcal{U} should be flat in the presence of a direct connection. The fact that \mathcal{U} is negative is surprising. Simulations indicate that this behavior is a result of a breakdown in the weak coupling assumption. For a weaker direct connection (and much longer simulation to compensate for noise), the reflection in \mathcal{U} disappears (not shown). The combination of a positive \mathcal{W} and a negative \mathcal{U} could theoretically be caused by either a positive direct connection or a negative common input. The ambiguity is removed by the positive \mathcal{C} , indicating that we indeed have a positive direct connection.

We next simulated three neurons, where the unmeasured neuron three was a source of common input to neurons one and two. In this example, neuron three was from a different subpopulation than neuron one or two, as defined in section 2.4. Figure 3(B) shows the results obtained from analyzing the spikes of neurons one and two. The covariogram \mathcal{C} is essentially identical to the direct connection case in Figure 3(A). The covariogram cannot be used to distinguish this common input configuration from the direct connection configuration. This distinction can be made from the measures \mathcal{W} and \mathcal{U} . In this case, the common input measure \mathcal{U} is significantly positive, while the direct connection measure \mathcal{W} is negative. Since \mathcal{C} is positive, this combination correctly indicates the common input configuration.

We demonstrate in Figure 4 two simulations to confirm our analysis, with common input from neurons within the subpopulation of neuron one or neuron two (section 3.4). We retain the same common input configuration of Figure 3(B), but change the kernel of the unmeasured neuron three to match the kernel of either neuron two or neuron one.

For the case when the unmeasured neuron three was in neuron two's subpopulation (Figure 4(A)), the common input appears as a direct connection from neuron two onto neuron one, because \mathcal{W} is significantly positive. The results fail to correctly identify that neuron two does not have a direct connection onto neuron one. If one cares about the distinction between neuron two and neuron three, then this result is unacceptable. If, on the other hand, the precise of identity of neurons within a subpopulation is unimportant, the results are adequate, as they correctly indicate that a neuron from neuron two's subpopulation has a direct connection onto neuron one.

For the case when the unmeasured neuron three was in neuron one's subpopulation (Figure 4(B)), the results correctly indicate the common input configuration. In this network, there was no direct connection from neuron two's subpopulation onto neuron one's subpopulation. Fortunately, the similarity between neuron three and neuron one does not affect the results.

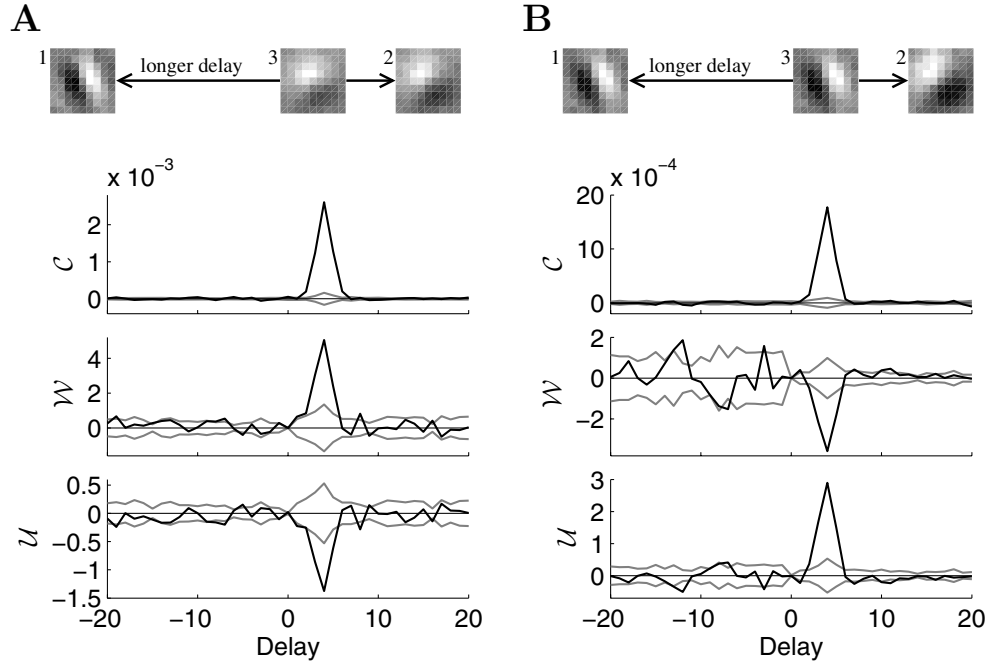


FIG. 4. Tests of the effect of common input from an unmeasured neuron within the subpopulation of neuron one or neuron two. Panels as in Figure 3. (A) The network configuration is identical to the common input of Figure 3(B) except that the kernel of neuron three is similar to that of neuron two ($cc_{32}^{\max} > 0.9$; see (2.8)), so that neuron three is in neuron two's subpopulation. In this case, the common input is misidentified as a direct connection (W is positive). W can be interpreted as indicating a direct connection from a neuron within neuron two's subpopulation onto neuron one. The connectivity \bar{W} and the parameters are identical to those of Figure 3(B) except $b_2 = 2$, $\psi_3 = -3\pi/8$, $\phi_3 = 7\pi/8$, $f_3 = 0.4$, $\bar{T}_2 = 3.4$. (B) When the kernel of neuron three is similar to that of neuron one ($cc_{31}^{\max} > 0.9$) so that neuron three is in neuron one's subpopulation, the common input is correctly identified (U is positive). In this case, there is no connection from neuron two's subpopulation to neuron one. The connectivity \bar{W} and the parameters are identical to those of Figure 3(B) except $b_1 = 6$, $\psi_3 = \pi/8$, $\phi_3 = 0.0$, $f_3 = 0.8$, $\bar{T}_1 = 3.0$.

As indicated by the analysis, our approach cannot distinguish a direct connection from an indirect connection via a third intermediate neuron. An example of an indirect connection is shown in Figure 5(A). Since the direct connection measure W is positive, the indirect connection is classified as a direct connection.

Although the analysis was based on an error function nonlinearity (4.2), the results are not sensitive to small changes in nonlinearity shape. In Figure 5, we demonstrate a simulation with a (truncated) power law nonlinearity: $\bar{g}_q(y) = \min\{A_q y^{\beta_q}, 1\}$ for $y > 0$, and $\bar{g}_q(y) = 0$ otherwise. This example includes both a direct connection from neuron one onto neuron two (we use the sign convention where this is a negative delay) and common input from an unmeasured neuron three (with a positive net delay, so that it mimics a connection from neuron two onto neuron one). We analyzed the spike responses from neurons one and two (R_1^i and R_2^i) and the stimulus \mathbf{X} just as we did in the previous examples; i.e., we used the results of Appendix B, in which the nonlinearities $\bar{g}_q(\cdot)$ are assumed to be error functions.

The covariogram contains two similar peaks at a positive and negative delay and therefore cannot distinguish the two type of connections. The measures W and U

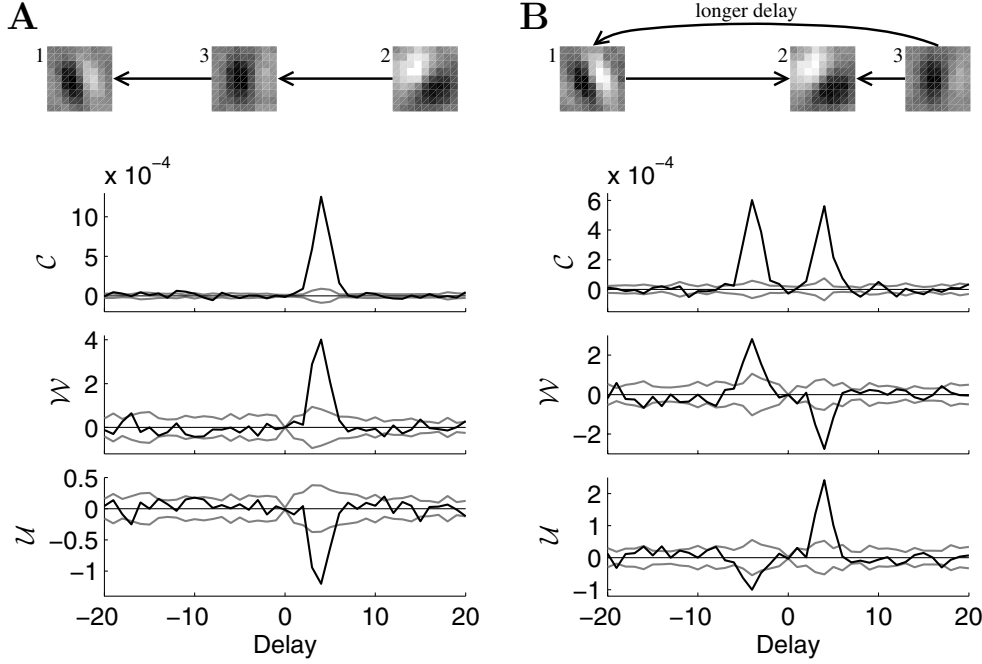


FIG. 5. Further demonstrations of the approach. Panels as in Figure 3. (A) An indirect connection from neuron two onto neuron one through an unmeasured neuron three. This connection appears as a direct connection (\mathcal{W} is positive). As shown by the analysis, we cannot distinguish such an indirect connection from a direct connection. The indirect connection is given by $\bar{W}_{23}^2 = \bar{W}_{31}^2 = 1.6$, $\bar{W}_{23}^1 = \bar{W}_{23}^3 = \bar{W}_{31}^1 = \bar{W}_{31}^3 = 0.8$. Parameters are identical to Figure 3(B) except $\bar{T}_1 = 2.5$, $\bar{T}_2 = 2.8$, $\bar{T}_3 = 2.6$. (B) A simulation with nonlinear functions $\bar{g}_q(\cdot)$ given by power laws. In this case, the network contains a direct connection from neuron one onto neuron two (corresponding to a negative delay) and common input from unmeasured neuron three onto neurons one and two (with a longer delay to neuron one to give a positive delay). The spikes of neuron one and two were analyzed as though the nonlinearities were error functions. Although the covariogram contains two virtually identical peaks, the measures \mathcal{W} and \mathcal{U} successfully identify the direct connection at negative delay and the common input at positive delay. The connections were given by $\bar{W}_{12}^4 = 0.6$, $\bar{W}_{12}^3 = \bar{W}_{12}^5 = 0.4$, $\bar{W}_{31}^6 = \bar{W}_{32}^6 = 1.8$, $\bar{W}_{31}^5 = \bar{W}_{31}^7 = \bar{W}_{31}^1 = \bar{W}_{32}^3 = 0.8$. Kernel parameters are as in Figure 3(B). Power law parameters: $A_1 = 0.02$, $A_2 = 0.035$, $A_3 = 0.05$, $\beta_1 = 2.6$, $\beta_2 = 2.0$, $\beta_3 = 2.3$.

differentiate between the origins of these peaks. Since \mathcal{W} is positive at a negative delay, it indicates a direct connection from neuron one onto neuron two. On the other hand, since \mathcal{U} is positive at a positive delay, it indicates a common input from a third neuron rather than any direct connection from neuron two onto neuron one. The method correctly identifies the circuitry of the model network even with a power law nonlinearity.

4.2. Simulation of integrate-and-fire networks. To test the robustness of the method to deviations from the linear-nonlinear model, we simulated a system of integrate-and-fire neurons. In this case, we viewed each time step as corresponding to a millisecond. The evolution of the voltage of neuron q in response to input $G_q(t)$ is given by $\tau_m \frac{dV_q}{dt} + V_q + G_q(t)(V_q - \mathcal{E}_s) = 0$. The spike times T_q^j of neuron q are those times when $V_q(T_q^j)$ reaches 1. After each spike, the voltage was reset to 0 and held there for an absolute refractory period of length τ_{ref} . Each neuron was driven by the

input conductance $G_q(t)$, which we specified by

$$G_q(t) = 0.05 \sum_{j>0} f(t - T_q^{\text{ext},j}) + \sum_{p=1}^n \sum_{j>0} W_{pq} f(t - T_p^j - d_{pq}),$$

where the first term is the response to external input events at times $T_p^{\text{ext},j}$ and the second term is due to internal coupling. The function $f(t) = \frac{e^2}{4} \left(\frac{t}{\tau_s}\right)^2 e^{-t/\tau_s}$ for $t > 0$, and $f(t) = 0$ otherwise. Here, W_{pq} specifies the strength of coupling from neuron p onto neuron q , and d_{pq} is the delay of that connection.

We set the external input to be a linear-nonlinear function of the stimulus. Accordingly, the $T_q^{\text{ext},j}$ were drawn from a modulated Poisson process with rate given by $\alpha_q [\mathbf{h}_q^i \cdot \mathbf{X}]_+$, where $[x]_+ = x$ if $x > 0$ and is zero otherwise.

We first simulated a network of three neurons that contained both a direct connection from neuron one onto neuron two and common input from neuron three onto neurons one and two (just as in Figure 5(B)). We used the same linear kernels (4.1) as before, sampling them on an $80 \times 10 \times 10$ grid in time and space. For realism, we sampled the white noise stimulus every ten units of time (i.e., every 10 ms). We simulated the network to 5,000 simulated seconds (nearly 1.4 simulated hours), recording 30,000 to 40,000 spikes per neuron. We needed such long simulations to obtain good results.

Figure 6(A) demonstrates that our analysis can distinguish common input from a direct connection even with integrate-and-fire neurons. The results are equivalent to Figure 5(B). The covariogram \mathcal{C} show peaks corresponding to the direct connection and the common input. The source of these correlations is distinguished by the measures \mathcal{W} and \mathcal{U} . The correlation at a negative delay is identified as a direct connection from neuron one onto neuron two; the correlation at a positive delay is identified as common input from a third neuron.

Since the integrate-and-fire neurons are driven by the stimulus in a fairly linear fashion, the basic relationship of neural response to the stimulus is similar to that assumed in the linear-nonlinear model (2.1). However, unlike model (2.1), the probability of a spike does depend strongly on previous spike times. The presence of a refractory period prevents the neuron from firing a spike immediately after spiking. Even after the refractory period, the voltage must integrate up to threshold, further increasing dependence among spike times. Figure 6(A) demonstrates that our approach can still work in the presence of these deviations from model assumptions.

As a final test of our approach, we simulated a slightly larger network of 20 integrate-and-fire neurons. The network included a direct connection from neuron one onto neuron two. In addition, four of the unmeasured neurons (neurons 3–20) were randomly selected to give common input onto both neurons one and two, where the connection onto neuron one had a delay that was 30 ms longer than the delay to neuron two. In this case, the measured spike trains had a correlation at a negative delay due to the direct connection and a correlation at a positive delay due to the common input, just as in the previous example. We randomly added additional connections to the network so that any given neuron had a 10% chance of connecting onto any given unmeasured neuron.

We simulated this network to 5,000 simulated seconds (nearly 1.4 simulated hours), measuring approximately 10,000–40,000 spikes per neuron. We discarded the spikes of all neurons except neurons one and two. The results from analyzing just these spikes are shown in Figure 6(B). In this case, the direct connection and

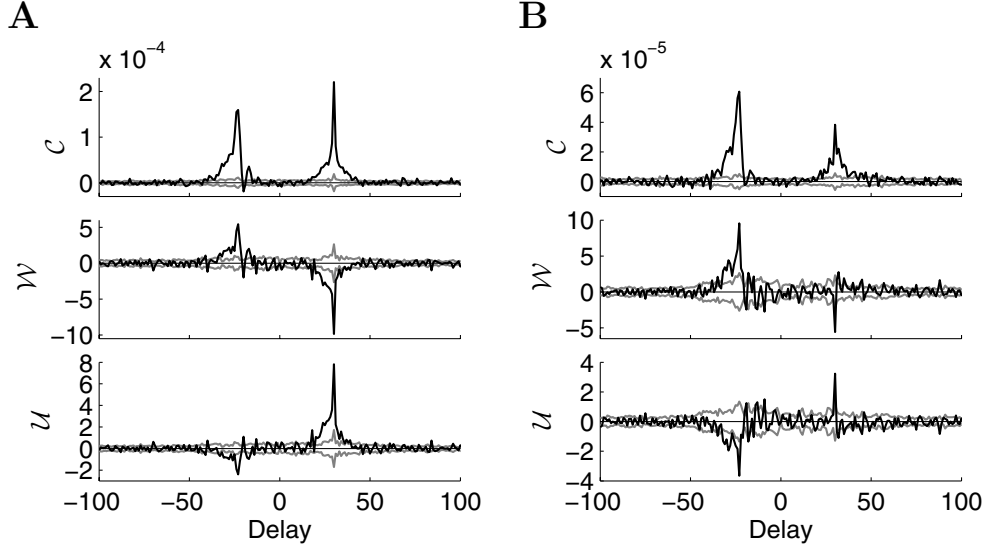


FIG. 6. Demonstration of the results applied to networks of integrate-and fire neurons. Panels as in the bottom panels of Figure 3. (A) Results from a simulation of three neurons with network architecture identical to that pictured at the top of Figure 5(B). (The network contained both a direct connection at negative delay and common input at positive delay.) The direct connection measure \mathcal{W} correctly identifies the direct connection from neuron one onto neuron two (appearing with negative delay). The common input measure \mathcal{U} correctly identifies the common input at positive delay. Parameters: $W_{12} = 0.1$, $W_{31} = W_{32} = 0.15$ (all other $W_{pq} = 0$), $d_{12} = 20$ ms, $d_{31} = 40$ ms, $d_{32} = 10$ ms, $\alpha_1 = \alpha_2 = 0.25$ ms $^{-1}$, $\alpha_3 = 0.3$ ms $^{-1}$, $\tau_m = 5$ ms, $\mathcal{E}_s = 6.5$, $\tau_s = 2$ ms, $\tau_{ref} = 2$ ms. Parameters for \mathbf{h} are the same as those in Figure 3(B) except that $\tau_h = 20$ ms. (B) Results from a simulation of a random network of twenty neurons. The measure \mathcal{W} correctly identified the direct connection from neuron one onto neuron two at negative delay (established by setting $W_{12} = 0.12$, $d_{12} = 20$ ms). The measure \mathcal{U} correctly identified the common input at positive delay. Four neurons with index $p > 2$ were randomly selected to give this common input. For these p , the connection strength was randomly selected from $W_{p1} \in (0.05, 0.15)$, and $W_{p2} = W_{p1}$. For these four neurons, the delays were coordinated so that the delay to neuron one was 30 ms longer: $d_{p2} = 2$ ms, $d_{p1} = 32$ ms. The remaining connections were randomly generated as follows. For any $p > 0$ and $q > 2$, $W_{pq} = 0$ with 90% probability; otherwise the parameters W_{pq} and d_{pq} were randomly generated with $W_{pq} \in (0.05, 0.15)$ and $d_{pq} \in (1, 40)$ ms. Parameters for \mathbf{h}_1 and \mathbf{h}_2 are as in Figure 3(A), except that $\tau_h = 20$ ms and $b_2 = 2$ ms. The remaining kernels with $p > 2$ were randomly generated with $\psi_p \in (0, 2\pi)$, $\phi_p \in (0, 2\pi)$, and $f_p \in (0.2, 1.0)$. We set $\alpha_1 = \alpha_2 = 0.25$ ms $^{-1}$ and, for $p > 2$, randomly generated $\alpha_p \in (0.15, 0.3)$ ms $^{-1}$. We set $\tau_m = 5$ ms, $\mathcal{E}_s = 6.5$, $\tau_s = 2$ ms, and $\tau_{ref} = 2$ ms.

common input are correctly identified by the measures \mathcal{W} and \mathcal{U} , respectively. We did not constrain the unmeasured neurons to be from different subpopulations than the measured neurons. For the four common input neurons p , the maximal correlation coefficient cc_{p2}^{\max} (see (2.8)) between neuron p and neuron two ranged from 0.0 to 0.7. Since the common input correlations mimicked a connection from neuron two to neuron one, these cc_{p2}^{\max} were the critical measures for determining whether the common input would be identified as a direct connection. The simulation indicates that the common input neurons (at least on average) were considered to be from subpopulations different from that of neuron two.

5. Discussion. The results demonstrate that we can correctly identify subpopulation connectivity when neural response can be captured by the linear-nonlinear model (2.1), the coupling is not too strong, and we have a lot of data. Before we focus

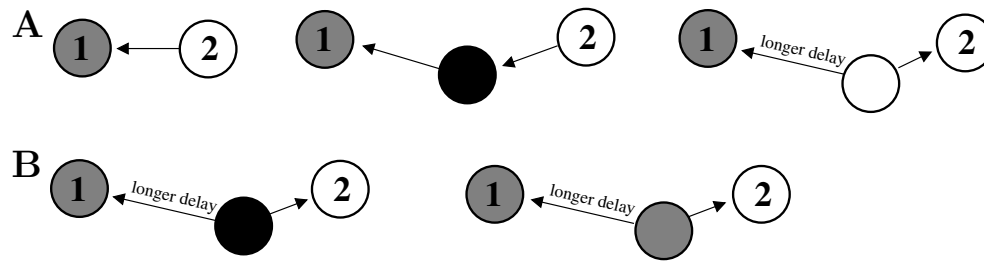


FIG. 7. Schematic summary of the determination of subpopulation connectivity that we are able to achieve. Each pictured network configuration leads to a correlation between the spikes of neuron one and a delayed version of the spikes of neuron two. If one analyzed the joint statistics of the spikes of neurons one and two (e.g., with a covariogram), each example would appear to involve a direct connection from neuron two onto neuron one. (The unlabeled neuron is not measured.) Our result is that we can distinguish (A) network configurations in the top row from (B) network configurations in the bottom row using our analysis of the joint statistics of the stimulus and the spikes of neurons one and two. The subpopulation of each neuron is indicated by the shading (white, gray, or black). (A) We consider the network configurations in the top row to belong to the “direct connection class,” as each configuration will be identified as a direct connection by our analysis. We cannot distinguish among the configurations in the direct connection class. Nonetheless, since each configuration contains a causal connection from neuron two’s subpopulation (white) onto neuron one’s subpopulation (gray), our analysis can still accurately determine connectivity at the level of subpopulations. From left to right, the configurations are a direct connection from neuron two onto neuron one, an indirect connection through an unmeasured neuron, and common input from neuron two’s subpopulation. (B) We consider network configurations in the bottom row to belong to the “common input class,” as each configuration will be identified as common input by our analysis. Since these configurations have no causal connection from neuron two’s subpopulation (white) onto neuron one’s subpopulation (gray), it is important that our analysis can distinguish them from the direct connection configurations (A). From left to right, the configurations are common input from a different subpopulation (black) and common input from neuron one’s subpopulation (gray).

on the limitations caused by these conditions, we discuss the significance of subpopulation connectivity and the relationship between this approach and other works.

5.1. Identification of subpopulation connectivity. Recall that two neurons are in the same subpopulations if their effective kernels \mathbf{h}_q (defined by fitting the uncoupled model (2.2)) are similar. In some contexts, neuroscientists would refer to these kernels as the neurons’ *receptive fields*; in this case, a subpopulation would be a group of neurons with similar receptive fields.

We have shown that common input originating from within one neuron’s subpopulation could appear like a direct connection from that neuron onto the other measured neuron. Hence, when we identify a direct connection between neurons, we can only conclude that there is a connection between those neurons’ subpopulations.

We summarize our conclusions in Figure 7. The direct connection measure \mathcal{W} and the common input measure \mathcal{U} effectively divide network configurations into two classes; we will call these a direct connection class and a common input class. The top row (A) shows network configurations that would be classified as having a direct connection from neuron two onto neuron one. Besides the actual direct connection, this direct connection class contains an indirect connection through an unmeasured neuron and common input from neuron two’s subpopulation. All three network configurations contain a causal connection from neuron two’s subpopulation (white) onto neuron one’s subpopulation (gray).

The bottom row (B) shows network configurations that would be classified as having common input. When the delay onto neuron one is longer (so that the correlations

mimic a direct connection from neuron two onto neuron one), the common input class contains networks with common input from different subpopulations and networks with common input from neuron one’s subpopulation. In neither of these cases is there a connection from neuron two’s subpopulation onto neuron one. Consequently, in order to accurately identify subpopulation connectivity, these configurations must be distinguished from the direct connection class of the first row. We have shown that, subject to the limitations mentioned above, we can make this distinction.

We argue that, in some experimental contexts, determining subpopulation connectivity is as informative as determining the actual connectivity between two measured neurons. In many experiments, electrodes are “blindly” inserted into the brain, and the precise identity of measured neurons remains unknown. In this situation, neurons are simply characterized by their response properties (e.g., their receptive fields), such as those captured by the effective kernels \mathbf{h}_q^i .

Since the precise identify of measured neurons is unknown, the best conclusion one can make about connectivity is that a neuron with response properties “A” is connected to a neuron with response properties “B.” In other words, the best one can say is that a neuron from the subpopulation characterized by response properties “A” is connected to a neuron from the subpopulation characterized by response properties “B.” This is the best possible conclusion even if we didn’t have to worry about ambiguity introduced by connections from unmeasured neurons. Our central result is that we have developed an approach to achieve this best possible conclusion even in the presence of common input from unmeasured neurons.

5.2. Precise identity of subpopulations. The above discussion assumes the presence of discrete subpopulations. If this were the case, the statement that two neurons are from the same subpopulation would be unambiguous. Of course, in general, this is not the case. The response properties of neurons across a large population may be better modeled as a continuum, where the correlation coefficients cc_{pq}^{\max} of (2.8) could be any value between 0 and 1. (Since cc_{pq}^k tends to zero for large $|k|$, the maximum is always nonnegative.) In order to make our subpopulation definition precise, we would like to have some cutoff value of cc_{pq}^{\max} , above which we could say that neurons p and q are from the same population and below which we could say they are from different subpopulations.

To explore this issue, we simulated the common input network of Figures 3(B) and 4 and the integrate-and-fire network of Figure 6(A), varying the model parameters to change¹⁴ cc_{32}^{\max} . Although there was no clean cutoff, the cutoff value was around $cc_{32}^{\max} = 0.6$. If cc_{32}^{\max} was near 0.6, the results were mixed, and the subpopulation of neuron three seemed to depend on model parameters. But for larger cc_{32}^{\max} , neuron three acted like a member of neuron two’s subpopulation because the common input appeared as a direct connection in measure \mathcal{W} (as in Figure 4(A)). Similarly, for cc_{32}^{\max} much smaller than 0.6, neuron three acted like a member of a different subpopulation because the common input was correctly identified as common input (as in Figures 3(B) and 4(B)). Hence, at least for these coupling strengths and roughly equivalent firing rates, neurons p and q were effectively in the same subpopulation when cc_{pq}^{\max} was well above 0.6. (See section 5.5 for examples of how strong coupling and disparity in firing rates can further complicate the picture.)

¹⁴To keep the discussion as simple as possible, we ensured that the maximum of cc_{32}^k occurred at the delay $k = 2$, since \bar{W}_{32}^j was maximal at $j = 2$. Section 3.4 shows why cc_{32}^2 is the important correlation coefficient for this case.

5.3. Heuristic explanation for results. To provide some intuition into how our approach successfully determines subpopulation connectivity, we give a heuristic explanation of why one should be able to distinguish subpopulation connectivity by analyzing the joint statistics of the measured spikes and the stimulus. We claim that one should expect that the relationship between the measured neurons' spikes and the stimulus will differ between the direct connection class of Figure 7(A) and the common input class of Figure 7(B).

For example, when the stimulus sequence happens to be optimal for neuron two and subsequently optimal for neuron one, the effectiveness of a connection from neuron two's subpopulation onto neuron one's subpopulation will be enhanced. (In this case, a spike from neuron two's subpopulation is likely to reach the neuron from neuron one's subpopulation when it is ready to fire.) Since in each network configuration in the direct connection class (Figure 7(A)) the correlation between neuron one and neuron two depends on a connection from neuron two's subpopulation onto neuron one, we expect the correlation to be especially strong for this particular stimulus sequence.

On the other hand, we would not expect the correlations in the common input class (Figure 7(B)) to be especially strong when the stimulus happens to be optimal for neuron two and subsequently optimal for neuron one. None of the connections leading to the correlation will be enhanced for this stimulus sequence, since no connection exists from neuron two's subpopulation onto neuron one's subpopulation. This example of an extreme stimulus sequence illustrates one case where the joint stimulus spike statistics will differ depending on subpopulation connectivity. One might expect the differences to be evident even with other stimuli. Our results show that, at least for a simple model, one can exploit this difference to determine subpopulation connectivity.

5.4. Comparison to other approaches. Our approach succeeds in reconstructing subpopulation connectivity by combining spike correlation analysis [14, 1, 13] with white noise analysis [7, 5, 4]. It builds on previous work [10, 11] that did not address the presence of unmeasured neurons. We have previously reported [9] on our early attempts to address the unmeasured neurons where, since we did not require stimulus repeats, we had to assume that all unmeasured neurons had dissimilar kernels (effectively, that every unmeasured neuron was in its own subpopulation).

Our approach differs from the partial coherence of Rosenberg et al. [15] because it does not require measurement of the neuron producing the common input. In cases where one monitors multiple neurons simultaneously, partial coherence can rule out common input from the other measured neurons without appealing to the model assumptions underlying our analysis. Although there is a large literature in which researchers have developed methods to reconstruct the connectivity among measured neurons, we are unaware of others that explicitly account for unmeasured neurons. Without accounting for unmeasured neurons, common input from unmeasured neurons would be erroneously identified as a direct connection.

5.5. The weak coupling assumption. The analysis underlying the measures \mathcal{W} and \mathcal{U} relied on the assumption that the coupling \bar{W} was small. The simulations demonstrate that one can obtain correct results even when the coupling is not weak. We used values of \bar{W}_{pq}^j as large as 1.8 and values of $\sum_j \bar{W}_{pq}^j$ as large as 3.4. For this parameter range, the weak coupling assumption is not justified, yet the results successfully determined subpopulation connectivity.

At this point, we lack an analysis of the effects of strong coupling. We have discovered through simulations that violations of the weak coupling assumption can cause invalid results when the firing rates of the two measured neurons are greatly different. For example, if $E\{R_1^i\} \gg E\{R_2^i\}$, strong coupling can cause a direct connection from neuron one to neuron two to appear as common input. The same situation can also cause common input to appear as a direct connection from neuron two to neuron one. In other words, there is a bias for a faster neuron appearing to have a connection from a slower neuron and a bias against a slower neuron appearing to have a connection from a faster neuron. The strength of coupling at which the misidentification occurs depends on the degree of inequality between the firing rates.

For example, we analyzed a sequence of simulations of the direct connection of Figure 3(A) where we increased the disparity between the firing rate of neurons one and two. By the time neuron two fired ten times faster than neuron one, the direct connection was misidentified as common input, and we failed to reconstruct the subpopulation connectivity. On the other hand, when we halved the strength of the direct connection (and ran very long simulations), a direct connection was still accurately identified even when neuron two fired more than fifty times faster than neuron one.

We also analyzed a similar sequence of simulations of the common input configuration of Figure 3(B). The common input appeared as a direct connection from neuron two onto neuron one when neuron one fired over 20 times faster than neuron two. Because neuron three was not in neuron two's subpopulation ($c_{32}^{\max} < 0.2$), this misidentification is a failure in reconstructing subpopulation connectivity. When we increased the connection strengths by 50% (adjusting kernel parameters to keep $c_{32}^{\max} < 0.2$), the misidentification began when neuron one fired only eight times faster than neuron two. (As one might infer from the observations of section 5.2, when we changed the kernels to increase cc_{32}^{\max} , the common input was identified as a direct connection with lower disparities in firing rate.)

5.6. Improving statistical efficiency. Our reconstruction is based on an analysis of just a few stimulus-spike moments. We employed this moment-based approach because our intuition on such moments' behavior could guide development of this initial implementation of our subpopulation connectivity approach. One important demerit of this choice was made clear in our simulations, where we needed long simulations to obtain good results. To apply this approach to realistic neuroscience data, we will presumably need more statistically efficient techniques, such as maximum likelihood estimators, which will yield reliable estimates of subpopulation connectivity with less data.

5.7. Validations. Clearly, the assumptions of the analysis are idealizations that will never be satisfied by biological neuronal networks. The approach is viable only because accurate results can be obtained outside the strict assumptions (as demonstrated throughout section 4). However, section 5.5 demonstrated some violations of the assumptions that do lead to inaccurate results. Another possible source of inaccuracies is covariation in latency or excitability, as discussed by Brody [2]. Since such covariation is not addressed by the network model (2.1), this covariation could invalidate our results. To address possible sources of error, we must develop validation methods that can identify critical violations of assumptions that may skew the results. Such validations will allow one to trust that the results are accurate.

Ideally, one would like to test the accuracy of these results with *in vitro* experiments, where the actual connectivity can be determined by other means (such as with electrodes that enter neurons). Unfortunately, our approach depends on having

an experimentally controlled stimulus \mathbf{X} , where the relationship between the firing probabilities is given by (2.1). Given that *in vitro* preparations are typically severed from sensory receptors, this requirement may be difficult to achieve. A more promising testbed may be a lower organism, where the connectivity is known and neurons can be driven by a stimulus.

5.8. Extensions to other models. The model (2.1) was made as simple as possible to facilitate the analysis. It assumes, for example, that the response is an approximately linear function of the stimulus, that the network is in an asynchronous state, and that the internal dynamics of the neuron can be neglected. The results are valid only when the network is stimulated by white noise. An extension to more general elliptically symmetric stimuli should be possible. Since in this case, a linear-nonlinear model can be reconstructed (see, for example, [3]), the results should be attainable if one replaces the integration-by-parts formula (A.3) with a more general version.

We view the implementation presented in this paper simply as an example of a new framework of network analysis. The principle of analyzing joint input-output statistics may be generalized to reveal pairwise coupling in other network models. The current version should have only limited applicability to neuroscience experiments because the relationship of neural response to a stimulus will in most cases be more fundamentally nonlinear than the linear-nonlinear model (2.1). Extension of the results to more complicated models and stimuli will increase the range of applicability, allowing the approach to evolve into a useful tool for analyzing neuronal networks and other stimulus-driven networks.

Appendix A. Integration-by-parts formula. In our notation, we do not explicitly distinguish spatial versus temporal components of the stimulus, but rather let time be represented only by the temporal index of the kernels \mathbf{h}_q^i and the spikes R_q^i . We let each of the m components of \mathbf{X} be independent standard normal variables, so that the probability density function of \mathbf{X} is

$$(A.1) \quad \rho_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{m/2}} e^{-\frac{\|\mathbf{x}\|^2}{2}}.$$

To assist the reader, we derive an integration-by-parts formula (A.3), although such a formula is not new. Let \mathbf{h}_k for $k = 1, 2, \dots, K$ be linearly independent unit vectors corresponding to K kernels. We wish to compute

$$E\{\mathbf{X}F(\mathbf{h}_1 \cdot \mathbf{X}, \dots, \mathbf{h}_K \cdot \mathbf{X})\},$$

where F is some smooth function with K arguments. Given the probability density function (A.1) for \mathbf{X} , this expected value is

$$\frac{1}{(2\pi)^{m/2}} \int \mathbf{x}F(\mathbf{h}_1 \cdot \mathbf{x}, \dots, \mathbf{h}_K \cdot \mathbf{x})e^{-\frac{\|\mathbf{x}\|^2}{2}} d\mathbf{x}.$$

Denote the standard unit vectors by \mathbf{e}_j , so that we can write the kernel \mathbf{h}_k and the dot product $\mathbf{h}_k \cdot \mathbf{x}$ in component form as

$$\mathbf{h}_k = \sum_j h_{kj} \mathbf{e}_j \quad \text{and} \quad \mathbf{h}_k \cdot \mathbf{x} = \sum_j h_{kj} x_j,$$

where h_{kj} is the j th component of \mathbf{h}_k .

We calculate the components of $E\{\mathbf{X}F(\mathbf{h}_1 \cdot \mathbf{X}, \dots, \mathbf{h}_K \cdot \mathbf{X})\}$. Through integration by parts with respect to x_j , the j th component is

$$\begin{aligned} E\{X_j F(\mathbf{h}_1 \cdot \mathbf{X}, \dots, \mathbf{h}_K \cdot \mathbf{X})\} &= \frac{1}{(2\pi)^{m/2}} \int x_j F \left(\sum_k h_{1k} x_k, \dots, \sum_k h_{Kk} x_k \right) e^{-\frac{\|\mathbf{x}\|^2}{2}} d\mathbf{x} \\ &= \frac{1}{(2\pi)^{m/2}} \sum_i h_{ij} \int F_i \left(\sum_k h_{1k} x_k, \dots, \sum_k h_{Kk} x_k \right) e^{-\frac{\|\mathbf{x}\|^2}{2}} d\mathbf{x} \\ &= \sum_i h_{ij} E\{F_i(\mathbf{h}_1 \cdot \mathbf{X}, \dots, \mathbf{h}_K \cdot \mathbf{X})\}, \end{aligned}$$

where F_i indicates the partial derivative of F with respect to the i th variable.

Putting the components together, we conclude that

$$\begin{aligned} E\{\mathbf{X}F(\mathbf{h}_1 \cdot \mathbf{X}, \dots, \mathbf{h}_K \cdot \mathbf{X})\} &= \sum_j E\{x_j F(\mathbf{h}_1 \cdot \mathbf{X}, \dots, \mathbf{h}_K \cdot \mathbf{X})\} \mathbf{e}_j \\ &= \sum_i E\{F_i(\mathbf{h}_1 \cdot \mathbf{X}, \dots, \mathbf{h}_K \cdot \mathbf{X})\} \left(\sum_j h_{ij} \mathbf{e}_j \right) \\ (A.2) \qquad \qquad \qquad &= \sum_i E\{F_i(\mathbf{h}_1 \cdot \mathbf{X}, \dots, \mathbf{h}_K \cdot \mathbf{X})\} \mathbf{h}_i. \end{aligned}$$

The special case we need for our derivation is

$$\begin{aligned} E\{\mathbf{X}g_p(\mathbf{h}_p^i \cdot \mathbf{X})g_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})g_r(\mathbf{h}_r^{i-k} \cdot \mathbf{X})\} &= E\{g'_p(\mathbf{h}_p^i \cdot \mathbf{X})g_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})g_r(\mathbf{h}_r^{i-k} \cdot \mathbf{X})\} \mathbf{h}_p^i \\ &\quad + E\{g_p(\mathbf{h}_p^i \cdot \mathbf{X})g'_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})g_r(\mathbf{h}_r^{i-k} \cdot \mathbf{X})\} \mathbf{h}_q^{i-j} \\ (A.3) \qquad \qquad \qquad &\quad + E\{g_p(\mathbf{h}_p^i \cdot \mathbf{X})g_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})g'_r(\mathbf{h}_r^{i-k} \cdot \mathbf{X})\} \mathbf{h}_r^{i-k} \end{aligned}$$

and the equivalent for fewer factors.

Appendix B. Equations for error function nonlinearity. The analysis for error function $\bar{g}_q(\cdot)$ (see (4.2)) mirrors the derivations outlined in [11]. In this appendix, we summarize the intermediate steps and then give the error function result for (3.8) and (3.11).

We define the effective error function parameters (ϵ_q, y_q) from the spikes of each neuron q by fitting to these spikes the uncoupled model (2.2) with nonlinearity,

$$g_q(y) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{y - y_q}{\epsilon_q \sqrt{2}} \right) \right].$$

Denote the inner product between kernels by $\cos \theta_{pq}^j = \mathbf{h}_p^i \cdot \mathbf{h}_q^{i-j}$. (Note that for Gaussian white noise $\cos \theta_{pq}^j = cc_{pq}^j$; see (2.7).) Define the following expressions as

functions of the parameters ϵ_q , y_q , and $\cos \theta_{pq}^j$:

$$\begin{aligned} \delta_q &= \frac{1}{\sqrt{1 + \epsilon_q^2}}, \\ \lambda_{qp}^j &= \frac{\delta_q y_q - \delta_p^2 \delta_q y_p \cos \theta_{qp}^j}{\sqrt{2(1 - \delta_p^2 \delta_q^2 \cos^2 \theta_{qp}^j)}}, \\ \mu_{qp}^j &= \frac{\delta_p \delta_q}{2\pi \sqrt{1 - \delta_p^2 \delta_q^2 \cos^2 \theta_{qp}^j}} \exp\left(-\frac{\delta_p^2 y_p^2 - 2\delta_p^2 \delta_q^2 y_p y_q \cos \theta_{qp}^j + \delta_q^2 y_q^2}{2(1 - \delta_p^2 \delta_q^2 \cos^2 \theta_{qp}^j)}\right), \\ \xi_{qp}^j &= \frac{\delta_q^2 (1 - \delta_p^2 \cos^2 \theta_{qp}^j)}{1 - \delta_p^2 \delta_q^2 \cos^2 \theta_{qp}^j}. \end{aligned}$$

Define a double complementary error function

$$(B.1) \quad \text{derfc}(a, b, c) = \frac{4}{\pi} \int_a^\infty dy e^{-y^2} \int_{\frac{b-cy}{\sqrt{1-c^2}}}^\infty dz e^{-z^2}.$$

The function derfc is a two-dimensional analogue of the complementary error function (see [10]).

Using the fact that $\mathbf{h}_p^i \cdot \mathbf{X}$ and $\mathbf{h}_q^{i-j} \cdot \mathbf{X}$ are joint unit normal random variables with correlation $\cos \theta_{qp}^j$, we compute the following expected values:

$$\begin{aligned} E\{g'_p(\mathbf{h}_p^i \cdot \mathbf{X})g_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})\} &= \frac{\delta_p}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_p^2 y_p^2}{2}\right) \text{erfc}(\lambda_{qp}^j), \\ E\{g'_p(\mathbf{h}_p^i \cdot \mathbf{X})(g_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X}))^2\} &= \frac{\delta_p}{4\sqrt{2\pi}} \exp\left(-\frac{\delta_p^2 y_p^2}{2}\right) \text{derfc}(\lambda_{qp}^j, \lambda_{qp}^j, \xi_{qp}^j), \\ E\{g'_p(\mathbf{h}_p^i \cdot \mathbf{X})g'_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})\} &= \mu_{qp}^j, \\ E\{g'_p(\mathbf{h}_p^i \cdot \mathbf{X})g'_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})g_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})\} &= \frac{\mu_{qp}^j}{2} \text{erfc}\left(\frac{\lambda_{qp}^j(1 - \xi_{qp}^j)}{\sqrt{1 - (\xi_{qp}^j)^2}}\right), \\ E\{g''_p(\mathbf{h}_p^i \cdot \mathbf{X})g_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})\} &= \frac{\delta_p^3 y_p}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_p^2 y_p^2}{2}\right) \text{erfc}(\lambda_{qp}^j) - \delta_p^2 \cos \theta_{qp}^j \mu_{qp}^j, \\ E\{g''_p(\mathbf{h}_p^i \cdot \mathbf{X})(g_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X}))^2\} &= \frac{\delta_p^3 y_p}{4\sqrt{2\pi}} \exp\left(-\frac{\delta_p^2 y_p^2}{2}\right) \text{derfc}(\lambda_{qp}^j, \lambda_{qp}^j, \xi_{qp}^j) \\ &\quad - \delta_p^2 \cos \theta_{qp}^j \mu_{qp}^j \text{erfc}\left(\frac{\lambda_{qp}^j(1 - \xi_{qp}^j)}{\sqrt{1 - (\xi_{qp}^j)^2}}\right), \\ E\{g''_p(\mathbf{h}_p^i \cdot \mathbf{X})g'_q(\mathbf{h}_q^{i-j} \cdot \mathbf{X})\} &= \frac{\delta_p^2 [y_p - \delta_q^2 y_q \cos \theta_{qp}^j] \mu_{qp}^j}{(1 - \delta_p^2 \delta_q^2 \cos^2 \theta_{qp}^j)}. \end{aligned}$$

We rewrite (3.8) and (3.11) in terms of the above quantities:

$$\begin{aligned} C_{21}^k &= \hat{W}_{21}^k \frac{\delta_1}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_1^2 y_1^2}{2}\right) \left[\text{erfc}(\lambda_{21}^k) - \frac{1}{2} \text{derfc}(\lambda_{21}^k, \lambda_{21}^k, \xi_{21}^k) \right] \\ &\quad + \hat{W}_{12}^{-k} \frac{\delta_2}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_2^2 y_2^2}{2}\right) \left[\text{erfc}(\lambda_{12}^{-k}) - \frac{1}{2} \text{derfc}(\lambda_{12}^{-k}, \lambda_{12}^{-k}, \xi_{12}^{-k}) \right] \\ (B.2) \quad &+ \hat{U}_{21}^k \mu_{21}^k, \end{aligned}$$

$$\begin{aligned}
 A_1^k = & \hat{W}_{21}^k \left\{ \frac{\delta_1^3 y_1}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_1^2 y_1^2}{2}\right) \left[\operatorname{erfc}(\lambda_{21}^k) - \frac{1}{2} \operatorname{derfc}(\lambda_{21}^k, \lambda_{21}^k, \xi_{21}^k) \right] \right. \\
 & \left. - \delta_1^2 \cos \theta_{21}^k \mu_{21}^k \left[1 - \operatorname{erfc}\left(\frac{\lambda_{21}^k(1 - \xi_{21}^k)}{\sqrt{1 - (\xi_{21}^k)^2}}\right) \right] \right\} \\
 & + \hat{W}_{12}^{-k} \mu_{12}^{-k} \left[1 - \operatorname{erfc}\left(\frac{\lambda_{12}^{-k}(1 - \xi_{12}^{-k})}{\sqrt{1 - (\xi_{12}^{-k})^2}}\right) \right] \\
 (B.3) \quad & + \hat{U}_{21}^k \frac{\delta_1^2 [y_1 - \delta_2^2 y_2 \cos \theta_{21}^k] \mu_{21}^k}{(1 - \delta_1^2 \delta_2^2 \cos^2 \theta_{21}^k)},
 \end{aligned}$$

$$\begin{aligned}
 A_2^k = & \hat{W}_{21}^k \mu_{21}^k \left[1 - \operatorname{erfc}\left(\frac{\lambda_{21}^k(1 - \xi_{21}^k)}{\sqrt{1 - (\xi_{21}^k)^2}}\right) \right] \\
 & + \hat{W}_{12}^{-k} \left\{ \frac{\delta_2^3 y_2}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_2^2 y_2^2}{2}\right) \left[\operatorname{erfc}(\lambda_{12}^{-k}) - \frac{1}{2} \operatorname{derfc}(\lambda_{12}^{-k}, \lambda_{12}^{-k}, \xi_{12}^{-k}) \right] \right. \\
 & \left. - \delta_2^2 \cos \theta_{12}^{-k} \mu_{12}^{-k} \left[1 - \operatorname{erfc}\left(\frac{\lambda_{12}^{-k}(1 - \xi_{12}^{-k})}{\sqrt{1 - (\xi_{12}^{-k})^2}}\right) \right] \right\} \\
 (B.4) \quad & + \hat{U}_{21}^k \frac{\delta_2^2 [y_2 - \delta_1^2 y_1 \cos \theta_{12}^{-k}] \mu_{12}^{-k}}{(1 - \delta_2^2 \delta_1^2 \cos^2 \theta_{12}^{-k})}.
 \end{aligned}$$

The key point of these long formulas is that, with the exception of \hat{W}_{21}^k , \hat{W}_{12}^{-k} , and \hat{U}_{21}^k , all expressions are functions of the error function parameters of the measured neurons (ϵ_1 , ϵ_2 , y_1 , and y_2) and $\cos \theta_{21}^k = \cos \theta_{12}^{-k}$. The kernels (and hence $\cos \theta_{21}^k$) are computed from (3.6). The parameters ϵ_q and y_q , for $q = 1, 2$, can be calculated from (3.4) and (3.5) with the use of the formulas

$$\begin{aligned}
 E\{g_q(\mathbf{h}_q^i \cdot \mathbf{X})\} &= \frac{1}{2} \operatorname{erfc}\left(\frac{\delta_q y_q}{\sqrt{2}}\right), \\
 E\{g'_q(\mathbf{h}_q^i \cdot \mathbf{X})\} &= \frac{\delta_q}{\sqrt{2\pi}} \exp\left(-\frac{\delta_q^2 y_q^2}{2}\right).
 \end{aligned}$$

Appendix C. Estimating confidence intervals. We estimate the confidence interval of our measures using essentially the procedure outlined in Appendix B of [10]. Besides changing the base variables to those needed for the current analysis, we make the following two minor changes. First, we calculate the covariances of inner products accurately using the covariances among all the factors in the product. Second, since the statistics from different delays are uncoupled in our equations, we ignore covariances among statistics from different delays.

Acknowledgment. The author thanks Dario Ringach for numerous helpful discussions throughout the development of this research.

REFERENCES

- [1] A. M. H. J. AERTSEN, G. L. GERSTEIN, M. K. HABIB, AND G. PALM, *Dynamics of neuronal firing correlation: Modulation of “effective connectivity,”* J. Neurophysiol., 61 (1989), pp. 900–917.

- [2] C. D. BRODY, *Correlations without synchrony*, *Neural Comput.*, 11 (1999), pp. 1537–1551.
- [3] E. J. CHICHILNISKY, *A simple white noise analysis of neural light responses*, *Network: Comput. Neural Syst.*, 12 (2001), pp. 199–213.
- [4] Y. DAN, J.-M. ALONSO, W. M. USREY, AND R. C. REID, *Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus*, *Nature Neurosci.*, 1 (1998), pp. 501–507.
- [5] E. DEBOER AND P. KUYPER, *Triggered correlation*, *IEEE Trans. Biomed. Eng.*, 15 (1968), pp. 169–179.
- [6] S. MARCELJA, *Mathematical description of the responses of simple cortical cells*, *J. Opt. Soc. Am.*, 70 (1980), pp. 1297–1300.
- [7] P. N. MARMARELIS AND V. Z. MARMARELIS, *Analysis of Physiological Systems: The White Noise Approach*, Plenum Press, New York, 1978.
- [8] D. Q. NYKAMP, *Measuring linear and quadratic contributions to neuronal response*, *Network: Comput. Neural Syst.*, 14 (2003), pp. 673–702.
- [9] D. Q. NYKAMP, *Reconstructing stimulus-driven neural networks from spike times*, in *Advances in Neural Information Processing Systems 15*, S. Becker, S. Thrun, and K. Obermayer, eds., MIT Press, Cambridge, MA, 2003, pp. 309–316.
- [10] D. Q. NYKAMP, *Spike correlation measures that eliminate stimulus effects in response to white noise*, *J. Comp. Neurosci.*, 14 (2003), pp. 193–209.
- [11] D. Q. NYKAMP, *White noise analysis of coupled linear-nonlinear systems*, *SIAM J. Appl. Math.*, 63 (2003), pp. 1208–1230.
- [12] D. Q. NYKAMP AND D. L. RINGACH, *Full identification of a linear-nonlinear system via cross-correlation analysis*, *J. Vision*, 2 (2002), pp. 1–11.
- [13] G. PALM, A. M. H. J. AERTSEN, AND G. L. GERSTEIN, *On the significance of correlations among neuronal spike trains*, *Biol. Cybern.*, 59 (1988), pp. 1–11.
- [14] D. H. PERKEL, G. L. GERSTEIN, AND G. P. MOORE, *Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains*, *Biophys. J.*, 7 (1967), pp. 419–440.
- [15] J. R. ROSENBERG, A. M. AMJAD, P. BREEZE, D. R. BRILLINGER, AND D. M. HALLIDAY, *The Fourier approach to the identification of functional coupling between neuronal spike trains*, *Prog. Biophys. Mol. Biol.*, 53 (1989), pp. 1–31.