Mean-Field Approximations for Coupled Populations of Generalized Linear Model Spiking Neurons with Markov Refractoriness

Taro Toyoizumi
taro@neurotheory.columbia.edu
Department of Neuroscience, Columbia University, New York, NY 10032, U.S.A.

Kamiar Rahnama Rad
kamiar@stat.columbia.edu
Department of Statistics, Columbia University, New York, NY 10027, U.S.A.

Liam Paninski
liam@stat.columbia.edu
Department of Neuroscience and Department of Statistics, Columbia University,
New York, NY 10027, U.S.A.

There has recently been a great deal of interest in inferring network connectivity from the spike trains in populations of neurons. One class of useful models that can be fit easily to spiking data is based on generalized linear point process models from statistics. Once the parameters for these models are fit, the analyst is left with a nonlinear spiking network model with delays, which in general may be very difficult to understand analytically. Here we develop mean-field methods for approximating the stimulus-driven firing rates (in both the time-varying and steady-state cases), auto- and cross-correlations, and stimulus-dependent filtering properties of these networks. These approximations are valid when the contributions of individual network coupling terms are small and, hence, the total input to a neuron is approximately gaussian. These approximations lead to deterministic ordinary differential equations that are much easier to solve and analyze than direct Monte Carlo simulation of the network activity. These approximations also provide an analytical way to evaluate the linear input-output filter of neurons and how the filters are modulated by network interactions and some stimulus feature. Finally, in the case of strong refractory effects, the mean-field approximations in the generalized linear model become inaccurate; therefore, we introduce a model that captures strong refractoriness, retains all of the easy fitting properties of the standard generalized linear model, and leads to much more accurate approximations of mean firing rates and cross-correlations that retain fine temporal behaviors.

1 Introduction

One of the fundamental problems in the statistical analysis of neural data is to infer the connectivity and stimulus dependence of a network of neurons given an observed sequence of extracellular spike times in the network (Brown, Kass, & Mitra, 2004). Estimating the parameters of network models from data is in general a computationally difficult problem because of the high dimensionality of the parameter space and the possible complexity of the model's objective function given an observed data set of network spike times.

One class of models that has proven quite useful is known in the statistics literature as the *generalized linear model* (GLM; Brillinger, 1988; Chornoboy, Schramm, & Karr, 1988; see also Martignon et al., 2000; Kulkarni & Paninski, 2007; Nykamp, 2007), variants of which have been applied successfully in the hippocampus (Harris, Csicsvari, Hirase, Dragoi, & Buzsaki, 2003; Okatan, Wilson, & Brown, 2005), motor cortex (Paninski, Fellows, Shoham, Hatsopoulos, & Donoghue, 2004; Truccolo, Eden, Fellows, Donoghue, & Brown, 2005), retina (Pillow et al., 2008), and cultured cortical slice (Rigat, de Gunst, & van Pelt, 2006). In its simplest form, this model incorporates both stimulus-dependence terms and direct coupling terms between each observed neuron in the network; fitting the model parameters leads to an inhomogeneous (stimulus-driven), nonlinear coupled spiking model with delay terms. Statistically speaking, the model is attractive due to its explicitly probabilistic nature and because fitting the model parameters is surprisingly simple: under certain simple conditions, the log-likelihood function with respect to the model parameters is concave, and the maximum likelihood estimation of the parameters can be easily performed using standard ascent methods (Paninski, 2004; Paninski, Pillow, & Lewi, 2007).

Once the model parameters are obtained, we are left with an obvious question: What do we do next? One of the key applications of such a network model is to better understand the input-output properties of the network. For example, we would like to be able to predict the mean firing response of the network given a novel input and to carve out the impact of the network coupling terms on this stimulus-response relationship (e.g., what is local inhibition's impact on the stimulus filtering properties of the network?). We would also like to know how the correlation properties of spike trains in the network might depend on the stimulus, and in general how correlations might encode stimulus information (Riehle, Grün, Diesmann, & Aertsen, 1997; Hatsopoulos, Ojakangas, Paninski, & Donoghue, 1998; Oram, Hatsopoulos, Richmond, & Donoghue, 2001; Nirenberg, Carcieri, Jacobs, & Latham, 2002; Schnitzer & Meister, 2003; Kohn & Smith, 2005; Pillow et al., 2008).

We can in general study these questions with direct Monte Carlo simulations. However, simulation of a large-scale probabilistic spiking network is computationally expensive, since we often need to draw many samples

(i.e., run many simulated trials) in order to compute the quantities of interest to the desired precision. More importantly direct numerical simulation often provides limited analytical insight into the mechanisms underlying the observed phenomena.

The goal of this study is to investigate how much of the behaviors of these GLM networks can be understood using standard analytical mean-field approximations (Renart, Brunel, & Wang, 2003; Wilson & Cowan, 1972; Amit & Tsodyks, 1991; Ginzburg & Sompolinsky, 1994; Hertz, Krogh, & Palmer, 1991; Kappen & Spanjers, 2000; Gerstner & Kistler, 2002; Meyer & van Vreeswijk, 2002). In particular, we develop approximations for the mean firing rates of the network given novel stimuli, as well as the auto- and cross-correlation and input-output filtering properties of these networks. These approximations are valid when the contributions of individual network coupling terms are small and lead to deterministic ordinary differential equations that are much easier to solve and analyze than direct Monte Carlo simulation of the network activity.

However, in the case of strong refractory effects, these mean-field approximations become inaccurate, since the spike history terms in the generalized linear model must be large to induce strong refractoriness, and this pushes our approximations beyond their region of accuracy. Therefore we introduce a modified model: a generalized linear model with Markovian refractoriness. This model has several advantages in this setting: it captures strong refractoriness, retains all of the easy-fitting properties of the standard generalized linear model, and leads to much more accurate mean-field approximations.

2 The Generalized Linear Point Process Model

We begin by precisely defining the generalized linear point process model (see Figure 1). We consider N recurrently connected spiking neurons, modeled as a multivariate point process (Snyder & Miller, 1991). The firing rate (conditional intensity function) of neuron i at time t is a function of the total input, $u_i(t)$, the neuron receives:

$$\lambda_i(t) = f(u_i(t)), \tag{2.1}$$

where $f(.)$ is a smooth, nonnegative, and monotonically increasing function. For numerical simulations, we use exponential nonlinarity, $f(u) = e^u$, but the analysis in this article is not limited to this exponential nonlinearity. The total input to the neuron is expressed as the sum of external input I_i and recurrent input H_i:

$$u_i(t) = I_i(t) + H_i(t). \tag{2.2}$$
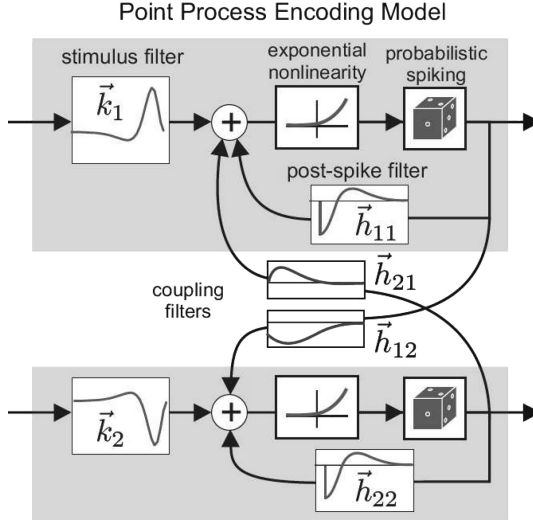
Point Process Encoding Model

Figure 1: A schematic illustration of the generalized linear point process model.

The recurrent input is modeled as a sum of identical responses caused by each presynaptic spike and written as

$$H_i(t) = \sum_{j=1}^{N} \sum_{t_{j,n} < t} w_i^j(t - t_{j,n}),$$ (2.3)

where $w_i^j(t)$ is a coupling term from neuron j (upper index) to i (lower index), and $t_{j,n}$ is the nth spike time of neuron j. Without loss of generality, we may express each coupling term $w_i^j(t)$ as a weighted sum of exponential functions of different time constants. As we will see below, this sum-of-exponentials representation is extremely useful: we will take full advantage of the Markovian nature of the decaying exponential function. To keep notation under control, we will restrict our attention to the case of a single exponential function with time constant τ_i^j,

$$w_i^j(t) = J_i^j \exp(-t/\tau_i^j)\Theta(t),$$ (2.4)

where Θ is the Heaviside step function, which takes one for positive arguments and zero otherwise; generalizations to the case where $w_i^j(t)$ are given by a sum of exponentials with different time constants will be left implicit and will be straightforward in all cases. If we write the output spike train of neuron i as

$$S_i(t) = \sum_n \delta(t - t_{i,n}),$$ (2.5)

the recurrent input of equation 2.3 is rewritten as the convolution of the synaptic filter and the presynaptic spike train:

$$H_i(t) = \sum_{j=1}^{N} \int ds \, w_i^j(t-s) S_j(s)$$

$$= \sum_{j=1}^{N} J_i^j h_i^j(t) \qquad (2.6)$$

with

$$h_i^j(t) = \int_{-\infty}^{t} e^{-(t-s)/\tau_i^j} S_j(s) \, ds. \qquad (2.7)$$

This model is a generalization of the inhomogeneous Poisson process, since past output spikes modulate the firing intensity function. If we set all the coupling weights J_i^j to zero, we recover the the inhomogeneous Poisson process model. The model is also closely related to the spike-response model of Gerstner and Kistler (2002) and the inhomogeneous Markov interval process of Kass and Ventura (2001), in which the firing rate depends on only the last observed spike. Note that in the GLM, the firing intensity of a neuron may depend not only on the last spike of the neuron but on the past several spikes of this neuron (see, e.g., Paninski, 2004; Paninski et al., 2007, for further discussion). As mentioned in section 1, this model is attractive because it is fairly easy to fit to data and seems to capture the firing statistics of neural populations in a variety of experimental settings. However, in this article, we will not discuss the estimation problem but rather limit our attention to the problem of determining the network firing rate, auto- and cross-correlation functions, and input-output properties, given fixed, known model parameters.

3 Master Equation

The main object we would like to compute here is $P(\boldsymbol{h}, t)$: this is the distribution of the state variable $\boldsymbol{h} = (h_1^1, \ldots, h_N^N)^T$ at time t. These h_i^j are the solutions of the multivariate coupled Markov process,

$$\frac{dh_i^j(t)}{dt} = -\frac{h_i^j(t)}{\tau_i^j} + S_j(t), \qquad (3.1)$$

and therefore solving for $P(\boldsymbol{h}, t)$ will allow us to compute a number of important quantities. For example, to compute the mean firing intensity

$v_i(t) = E[S_i(t)]$ of neuron i given some input $\{I_i(t)\}$, we may compute

$$v_i(t) = \int P(\boldsymbol{h}, t)\lambda_i(t)d\boldsymbol{h}$$

$$= \int P(\boldsymbol{h}, t)f\left(I_i(t) + \sum_j J_i^j h_i^j\right)d\boldsymbol{h},$$

where $E[.]$ is the average over the spike history $\{S_i(t)|i = 1, \ldots, N, 0 \leq t < T\}$ in an entire duration T given external input $\{I_i(t)\}$. Hence, we can read the mean firing rate off directly from $P(\boldsymbol{h}, t)$.

We begin by writing down the master equation—the Kolmogorov forward equation that governs the time evolution of this Markov process (Oksendal, 2002):

$$\frac{\partial P(\boldsymbol{h}, t)}{\partial t} = \sum_{i,j}\left(\frac{1}{\tau_i^j}\frac{\partial[h_i^j P(\boldsymbol{h}, t)]}{\partial h_i^j}\right) + \sum_j \lambda_j(\boldsymbol{h} - \boldsymbol{e}_j, t)P(\boldsymbol{h} - \boldsymbol{e}_j, t)$$

$$- \sum_j \lambda_j(\boldsymbol{h}, t)P(\boldsymbol{h}, t), \tag{3.2}$$

with $\lambda_i(\boldsymbol{h}, t) = f(I_i(t) + \sum_j J_i^j h_i^j)$ as in equation 2.1 and the term $\boldsymbol{e}_j$ denoting the vector with its $\{i'j'\}$ slot $\delta_{jj'}$. Note that there are no diffusion terms here; the random nature of the stochastic differential equation enters instead through the jump terms.

While this jump-advection partial differential equation (PDE) in principle completely describes the behavior of this system, unfortunately this equation remains difficult to handle analytically (due to the nonlinearity in the λ terms) and numerically (due to the large dimensionality of the state variable $\boldsymbol{h}$). We will pursue a direct analysis of this PDE elsewhere; here, we will attack this system using a different approach, based on a self-consistent expansion of the first and second moments of $S_i(t)$.

In the one-dimensional case (i.e., a single neuron, with one $h(.)$), however, we may easily solve the PDE numerically to obtain an exact solution for the mean firing rate. Figure 2 compares this exact solution versus the mean-field approximations derived in the following section. The mean-field approximation is reasonably good here but not perfect, because of the fluctuation in the recurrent input ($J = -1$, $\tau = 10$ ms).

4 Mean-Field Approximation of GLM Statistics

Now we turn to the main topic of this article. As emphasized above, we would like to calculate the mean-firing intensities of neurons, $v_i(t) = E[S_i(t)]$, or the spike-cross-correlation function between neurons, $\phi_{ij}(t, t') = E[S_i(t)S_j(t')]$. The calculation of these statistics is difficult, however, because
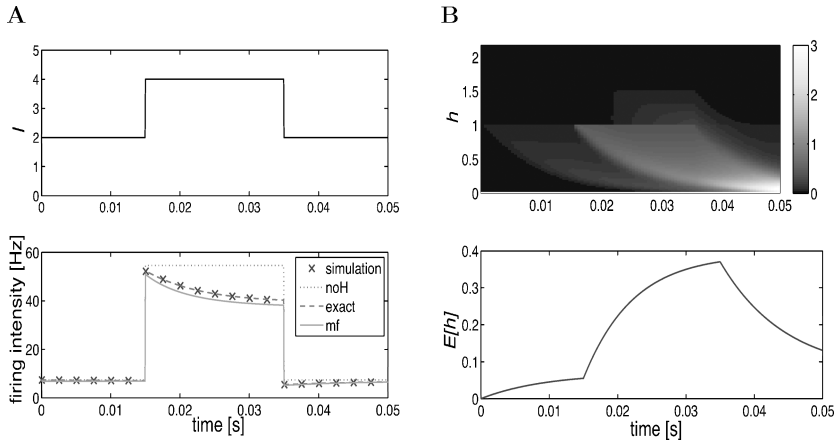
Figure 2: Direct solution of the master equation in the one-dimensional case. (A) Top: A single neuron is simulated with a pulse input $I(t)$ shown. Bottom: Comparison of the firing rates computed via direct Monte Carlo sampling (simulation); the inhomogeneous Poisson case, with no history dependence, that is, $w = 0$ (noH); the master equation of equation 3.2 (exact); and the mean-field approximation methods in the limit of $N_c \to \infty$ introduced in section 4. The mean-field approximation gives good approximations of the output firing intensity. (B) Top: The probability distribution $P(h, t)$; grayscale axis indicates the height of the probability density. The self-history term $h(t)$ was the convolution of the output spike train and a single exponential filter with time constant $\tau = 10$ ms and weight $J = -1$; thus, the neuron was inhibited for a brief period following each spike. Note the discontinuous behavior visible at stimulus onset, where probability mass jumps significantly toward $h = 1$, then decays back down toward zero, according to equation 3.1. Bottom: The mean of h calculated from the probability distribution $P(h, t)$. This $E[h]$ also increases following the input pulse, which has the effect of slowing the firing rate. Since the weight J is negative, the history term is inhibitory here. (A color version of this figure is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/neco.2008.04-08-757.)

of the nonlinearity f and the recurrent input term H in the model. A typical way of approximating these recurrent effects is to replace the interaction term H by its mean field (Renart et al., 2003; Wilson & Cowan, 1972; Amit & Tsodyks, 1991; Ginzburg & Sompolinsky, 1994; Hertz et al., 1991; Kappen & Spanjers, 2000; Gerstner & Kistler, 2002; Meyer & van Vreeswijk, 2002).

Let us assume that each neuron receives input from some subset, N_c, of the total population of neurons N. If the synaptic strength scales $J \sim 1/N_c$ with N_c, and assuming that the population is in the "asynchronous" state, then it is known that the cross-covariance function between two neurons $\Delta\phi_{ij}(t, t') = \text{Cov}[S_i(t), S_j(t')]$ scales with $1/N_c$ for large N_c (Ginzburg & Sompolinsky, 1994). This means that the contribution of each network

interaction term is negligible and all the neuron activities are independent in the $N_c \to \infty$ limit. Hence, calculating the mean-firing intensities and cross-correlations is easy in this limit.

We first derive the mean-firing intensity in the limit of $N_c \to \infty$ and then evaluate the finite-size effect of order $1/N_c$. Note that this mean-field solution is a good approximation not only when N_c is very large; equivalently, this is a good approximation when the synaptic coupling J is small for fixed N_c, where the contribution of an individual synapse is small.

Now our first key approximation is to appeal to the central limit theorem applied to the sum of equation 2.6, and to assume that the recurrent input H is roughly gaussian. We can express the mean firing intensities of the neurons in our network as a function of two sufficient state variables: the mean recurrent input $E[H]$ and the variance of the recurrent input $\mathrm{Var}[H]$,

$$
\begin{aligned}
\nu_i(t) &= E[\lambda_i(t)] \\
&= E[f(I_i(t) + H_i(t))] \\
&\approx F(I_i(t) + \mu_i(t), \sigma_i(t)),
\end{aligned}
\tag{4.1}
$$

where, in the third line, the distribution of the recurrent input, $H_i(t)$, is assumed to be gaussian with mean $\mu_i(t) = E[H_i(t)]$ and variance $\sigma_i^2(t) = \mathrm{Var}[H_i(t)]$, and the expectation was replaced by a gaussian integral:

$$
F(\mu, \sigma) = \int \frac{f(u)}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-\mu)^2}{2\sigma^2}} \, du.
\tag{4.2}
$$

Note for an exponential nonlinearity $f(u) = e^u$, the gaussian integral is given by $F(\mu, \sigma) = \exp(\mu + \sigma^2/2)$.

Next we calculate the cross-correlation function,

$$
\begin{aligned}
\phi_{ij}(t, t') &= E[S_i(t)S_j(t')] \\
&= \delta_{ij}\delta(t - t')E[S_i(t)] + \Theta(t - t')E[\lambda_i(t)S_j(t')] \\
&\quad + \Theta(t' - t)E[S_i(t)\lambda_j(t')] \\
&= \delta_{ij}\delta(t - t')\nu_i(t) + \Theta(t - t')\nu_{ij}(t \mid t')\nu_j(t') \\
&\quad + \Theta(t' - t)\nu_{ji}(t' \mid t)\nu_i(t),
\end{aligned}
\tag{4.3}
$$

where in the second line, we decomposed the correlation function into the three terms: the first term shows the simultaneous correlation with its amplitude proportional to the firing intensity due to the point-process nature of the spike train, the second term shows the correlation when $t > t'$ (the spike variable $S_i(t)$ is averaged given past spike history to yield $\lambda_i(t)$), and the third term is the correlation for $t' > t$. In the third line, we introduced conditional firing intensity,

$$
\nu_{ij}(t \mid t') = E[S_i(t) \mid \text{spike of } j \text{ at } t'] = \frac{E[S_i(t)S_j(t')]}{E[S_j(t')]},
\tag{4.4}
$$

which describes the probability of neuron i firing at t given the spike of j at t'. We also introduce an abbreviation in the following and write $E[.|S_j(t')]$ instead of $E[.|\text{spike of } j \text{ at } t']$. Similar to the approximation in equation 4.1, we assume a gaussian distribution of H given a spike of neuron j at time t' to find

$$
\begin{aligned}
v_{ij}(t \mid t') &= E[\lambda_i(t) \mid S_j(t')] \\
&= E[f(I_i(t) + H_i(t)) \mid S_j(t')] \\
&\approx F(I_i(t) + \mu_{ij}(t \mid t'), \sigma_{ij}(t \mid t'))
\end{aligned} \tag{4.5}
$$

for $t > t'$ with conditional mean, $\mu_{ij}(t \mid t') = E[H_i(t) \mid S_j(t')]$, and conditional variance, $\sigma_{ij}^2(t \mid t') = \mathrm{Var}[H_i(t) \mid S_j(t')]$. Note that equation 4.5 is valid only for $t > t'$; if $t < t'$ instead, we evaluate $v_{ji}(t' \mid t)$ and use the Bayes theorem to calculate $v_{ij}(t \mid t') = v_{ji}(t' \mid t)v_i(t)/v_j(t')$. The unconditional and conditional mean and variance of H in equations 4.1 and 4.5 are evaluated as

$$
\mu_i(t) = \sum_k J_i^k \alpha_i^k(t),
$$

$$
\mu_{ij}(t \mid t') = \sum_k J_i^k \alpha_{ij}^k(t \mid t'),
$$

$$
\sigma_i^2(t) = \sum_{k,l} J_i^k J_i^l A_i^{kl}(t),
$$

$$
\sigma_{ij}^2(t \mid t') = \sum_{k,l} J_i^k J_i^l A_{ij}^{kl}(t \mid t'), \tag{4.6}
$$

where we defined

$$
\alpha_i^k(t) = E[h_i^k(t)] = \int_{-\infty}^t ds \, e^{-(t-s)/\tau_i^k} v_k(s),
$$

$$
\alpha_{ij}^k(t \mid t') = E[h_i^k(t) \mid S_j(t')] = \int_{-\infty}^t ds \, e^{-(t-s)/\tau_i^k} v_{kj}(s \mid t'),
$$

$$
A_i^{kl}(t) = \mathrm{Cov}[h_i^k(t), h_i^l(t)]
$$
$$
= \int_{-\infty}^t ds \int_{-\infty}^t ds' \, e^{-(t-s)/\tau_i^k - (t-s')/\tau_i^l} \Delta\phi_{kl}(s, s'),
$$

$$
A_{ij}^{kl}(t \mid t') = \mathrm{Cov}[h_i^k(t), h_i^l(t) \mid S_j(t')]
$$
$$
= \int_{-\infty}^t ds \int_{-\infty}^t ds' \, e^{-(t-s)/\tau_i^k - (t-s')/\tau_i^l} \mathrm{Cov}[S_k(s), S_l(s') \mid S_j(t')],
$$

$$
\tag{4.7}
$$

with $\Delta\phi_{kl}(s, s') = \phi_{kl}(s, s') - v_k(s)v_l(s')$. The evolution of α_i^k, α_{ij}^k, and A_i^{kl} in equation 4.7 follows simple ordinary differential equations (ODEs) and is easy to simulate,

$$\frac{d}{dt}\alpha_i^k(t) = -\frac{\alpha_i^k(t)}{\tau_i^k} + v_k(t),$$

$$\frac{d}{dt}\alpha_{ij}^k(t \mid t') = -\frac{\alpha_{ij}^k(t \mid t')}{\tau_i^k} + v_{kj}(t \mid t'),$$

$$\frac{d}{dt}A_i^{kl}(t) = -\left(\frac{1}{\tau_i^k} + \frac{1}{\tau_i^l}\right)A_i^{kl}(t) + \lim_{\epsilon \to 0}\left[\int_{-\infty}^{t+|\epsilon|}e^{-(t-s)/\tau_i^k}\Delta\phi_{kl}(s, t)\,ds\right.$$

$$+ \left.\int_{-\infty}^{t-|\epsilon|}e^{-(t-s')/\tau_i^l}\Delta\phi_{kl}(t, s')\,ds'\right]$$

$$= -\frac{A_i^{kl}(t)}{\tau_i^{kl}} + \lim_{\epsilon \to 0}\left[\int_{-\infty}^{t+|\epsilon|}e^{-(t-s)/\tau_i^k}\Delta v_{kl}(s \mid t)v_l(t)\,ds\right.$$

$$+ \left.\int_{-\infty}^{t-|\epsilon|}e^{-(t-s')/\tau_i^l}\Delta v_{lk}(s' \mid t)v_k(t)\,ds'\right]$$

$$= -\frac{A_i^{kl}(t)}{\tau_i^{kl}} + \Delta\alpha_{il}^k(t^+ \mid t)v_l(t) + \Delta\alpha_{ik}^l(t^- \mid t)v_k(t), \qquad (4.8)$$

where $1/\tau_i^{kl} = 1/\tau_i^k + 1/\tau_i^l$, $\Delta v_{ij}(t \mid t') = v_{ij}(t \mid t') - v_i(t)$, $\Delta\phi_{kl}(s, s') = \Delta v_{kl}$ $(s \mid s')v_l(s') = \Delta v_{lk}(s' \mid s)v_k(s)$, $\Delta\alpha_{ij}^k(t \mid t') = \alpha_{ij}^k(t \mid t') - \alpha_i^k(t)$, and $\Delta\alpha_{il}^k(t^\pm \mid t) = \lim_{\epsilon \to 0}\Delta\alpha_{il}^k(t\pm \mid \epsilon \mid\mid t)$. Note that the differentiation of $A_i^{kl}(t)$ for $k = l$ is a little tricky because $\Delta\phi_{kl}(s, s')$ has a delta peak at $s = s'$. While the definition of $A_i^{kl}(t)$ in equation 4.7 is clear, we have to make sure in the evaluation of $\frac{d\,A_i^{kl}(t)}{dt}$ not to double-count the integration of the delta peak. Hence, in the third equation of equation 4.8, we included the delta peak along the integration over s (including the peak at $s = t$) but excluded this peak from the integration along s' (excluding the peak at $s' = t$).

4.1 Solution in the Mean-Field Limit.

In this section we evaluate mean firing intensities, v_i, and cross-correlation functions, ϕ_{ij}, in the mean-field limit of $N_c \to \infty$ (which is also valid in the noncoupled limit $J \to 0$ for finite N_c), that is, synaptic strengths are scaled $J \sim 1/N_c$ so that the individual contribution of a spike is negligible for large N_c; and we also assume, as discussed above, that the network is in the asynchronous state. Under these conditions, the cross-covariances between neurons are known to scale as $\Delta\phi_{ij} \sim 1/N_c$ for $i \neq j$ and $\Delta\phi_{ij} \sim 1$ for $i = j$ (Ginzburg & Sompolinsky, 1994). Then we can easily see $\sigma_i^2(t) \sim 1/N_c$ because $A_i^{kl} \sim 1$ for $k = l$ and $A_i^{kl} \sim 1/N_c$ for $k \neq l$ from equation 4.7. Hence, in the limit of $N_c \to \infty$, we

obtain from equations 4.1 and 4.8,

$$v_i(t) \to f\left(I_i(t) + \mu_i(t)\right),$$

$$\mu_i(t) = \sum_k J_i^k \alpha_i^k(t),$$

$$\frac{d}{dt}\alpha_i^k(t) = -\frac{\alpha_i^k(t)}{\tau_i^k} + v_k(t). \tag{4.9}$$

This self-consistent update rule corresponds to a well-known mean-field dynamics of the mean-firing intensity (Ginzburg & Sompolinsky, 1994). Similarly, the scaling of cross-covariance $\Delta\phi_{ij} \sim 1/N_c$ for $i \neq j$ implies $\Delta\alpha_{ij}^k(t \mid t') \sim 1/N_c$ for $k \neq j$ and $\Delta\alpha_{ij}^k(t \mid t') \sim 1$ for $k = j$ from equation 4.7. Moreover the conditional covariance $\mathrm{Cov}[S_k(s), S_l(s') \mid S_j(t')]$ scales with $1/N_c$ for $k \neq l \neq j$ and thus $\sigma_{ij}^2(t \mid t') \sim 1/N_c$. Hence, in the limit of $N_c \to \infty$, we find $v_{ij}(t \mid t') \to v_i(t)$ for $t > t'$ and, from equation 4.3,

$$\Delta\phi_{ij}(t, t') \to \delta_{ij}\delta(t - t')v_i(t). \tag{4.10}$$

These equations are very easy to simulate, and the approximation of the mean-firing intensity is good even for a small N_c if J is small (see Figure 2A). On the other hand, the mean-field approximation is not good for large J (see Figures 3 and 4B). This is because, for fixed N_c, the fluctuation of the recurrent input increases with J, making the mean-field assumption (which neglects these fluctuations) invalid. Figure 4 shows the cross-correlation functions between two neurons. Because the self-interaction term $J \sim 1/N_c$ vanishes for the large N_c limit, the autocorrelation function in this limit can capture the Poisson peak only at $t = t'$; it does not capture any nontrivial correlation caused by the interaction term J for finite N_c (see Figure 4).

4.2 Estimating the Finite Size Effect. In the previous section, we calculated the mean-firing intensity, $v_i(t)$, and the cross-covariance function, $\phi_{ij}(t, t')$, in the $N_c \to \infty$ limit. Ideally, once provided the true $\sigma_{ij}^2(t \mid t')$ term, which is of order $1/N_c^2$ different from $\sigma_i^2(t)$, we can iteratively evaluate equations 4.1 to 4.8, starting from the mean-field limit solution of equations 4.9 and 4.10 to find the precise approximation of $v_i(t)$ and $\phi_{ij}(t, t')$ as long as the gaussian H assumption is valid. However, we need the third-order correlations to evaluate $\sigma_{ij}^2(t \mid t')$, and the equations are not closed. Here, we give up evaluating $O(1/N_c^2)$ terms and just evaluate the finite size corrections of $v_i(t)$ and $\phi_{ij}(t, t')$ up to $O(1/N_c)$ terms. Up to this order, it can be argued that $\sigma_{ij}^2(t \mid t') \approx \sigma_i^2(t)$, and under this approximation, we obtain a self-consistent expression for $v_i(t)$ and $\phi_{ij}(t, t')$. Equations 4.1
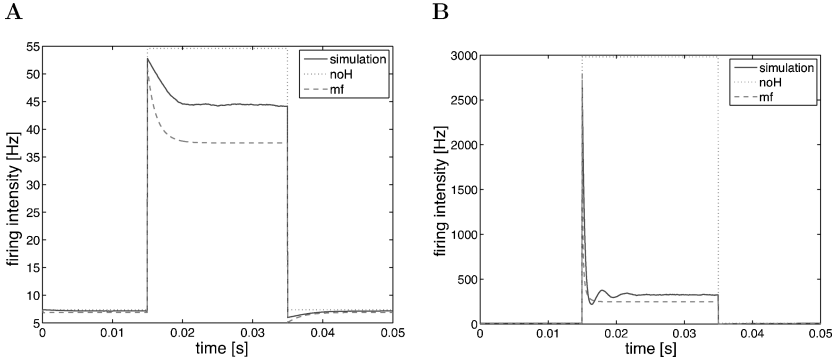
A

B

Figure 3: Comparison of the direct Monte Carlo simulation of a single neuron (solid), the firing intensity without recurrent input H (dotted), and the mean-field approximation (dashed). (A) Low step input (baseline $I = 2$, peak $I = 4$) is applied. The self-inhibition term is set to $J = -5$ and $\tau = 2$ ms. (B) High step input (baseline $I = 2$, peak $I = 8$) is applied. The self-inhibition term is set to $J = -5$ and $\tau = 2$ ms. Because of the large nongaussian fluctuations of input caused by the strong self-inhibition, the mean-field approximation loses accuracy here. (A color version of this figure is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/neco.2008.04-08-757.)

and 4.5 yield, up to the first order of $1/N_c$,

$$
\begin{aligned}
v_i(t) &= F(I_i(t) + \mu_i(t), \sigma_i(t)), \\
v_{ij}(t \mid t') &= F(I_i(t) + \mu_{ij}(t \mid t'), \sigma_i(t)),
\end{aligned}
\tag{4.11}
$$

where $\sigma_{ij}^2(t \mid t') \approx \sigma_i^2(t)$ to the first order of $1/N_c$. Equation 4.11 can be evaluated using equations 4.6 to 4.8, which we repeat here for convenience:

$$
\mu_i(t) = \sum_k J_i^k \alpha_i^k(t),
$$

$$
\mu_{ij}(t \mid t') = \sum_k J_i^k \alpha_{ij}^k(t \mid t'),
$$

$$
(\sigma_i(t))^2 = \sum_{k,l} J_i^k J_i^l A_i^{kl}(t),
\tag{4.12}
$$

$$
\frac{d}{dt} \alpha_i^k(t) = -\frac{\alpha_i^k(t)}{\tau_i^k} + v_k(t),
$$

$$
\frac{d}{dt} \alpha_{ij}^k(t \mid t') = -\frac{\alpha_{ij}^k(t \mid t')}{\tau_i^k} + v_{kj}(t \mid t'),
$$

$$
\frac{d}{dt} A_i^{kl}(t) = -\frac{A_i^{kl}(t)}{\tau_i^{kl}} + \Delta \alpha_{il}^k(t^+ \mid t) v_l(t) + \Delta \alpha_{ik}^l(t^- \mid t) v_k(t).
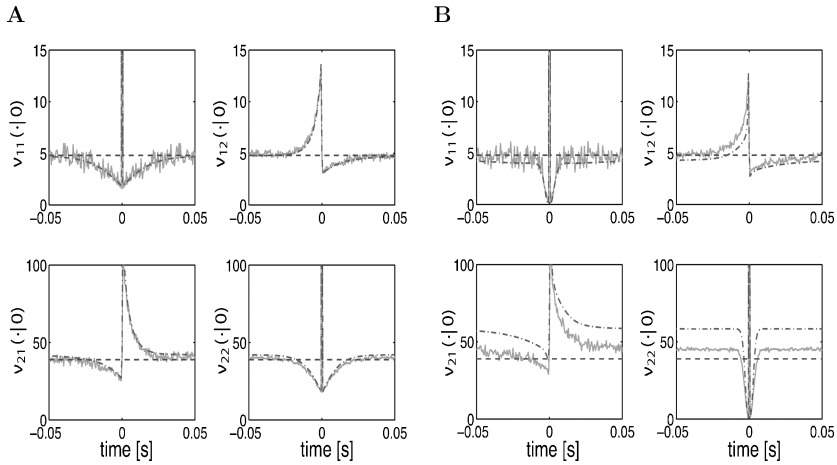$$

Figure 4: Comparison of cross-correlation functions of two neurons calculated by direct simulation (solid), calculated from the mean-field approximation (see equations 4.9 and 4.10; dashed), and the mean-field approximation with finite size correction (see equations 4.11 and 4.12; dot-dashed). (A) Weak self-interaction. Parameters are set to $J_{11} = -1$, $J_{22} = -1$, $J_{21} = 1$, $J_{12} = -0.5$, $\tau_{11} = 10$ ms, $\tau_{22} = 10$ ms, $\tau_{21} = 10$ ms, $\tau_{12} = 20$ ms, and $I_1 = 2$, $I_2 = 4$. The result with finite size correction is obtained after five iterations of the self-consistent equations. The calculation in the mean-field limit captures only the steady-state firing intensity, not any temporal cross-correlation function. Cross-correlation functions are well approximated with the finite-size correction terms. (B) Strong self-interaction: parameters are set to $J_{11} = -5$, $J_{22} = -5$, $J_{21} = 1$, $J_{12} = -0.5$, $\tau_{11} = 2$ ms, $\tau_{22} = 2$ ms, $\tau_{21} = 10$ ms, $\tau_{12} = 20$ ms, and $I_1 = 2$, $I_2 = 4$. Because of the large nongaussian fluctuation of the input caused by the strong self-inhibition term, the mean-field approximation does not work very well. In this case the finite size solution converges rather slowly. Hence, the result is after 10 iterations of the self-consistent equations. (A color version of this figure is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/neco.2008.04-08-757.)

In principle, the above equations are valid even with time-dependent input, and several iterations of the above set of equations yield time-dependent cross-correlation functions. However, evaluation of the above equations is computationally expensive with time-dependent input. If one neuron is connected to N_c surrounding neurons and time is discretized into T time bins, it takes $O(N^2 \cdot N_c \cdot T^2)$ operations and memory to evaluate $\alpha_{ij}^k(t \mid t')$. If the input is constant in time, $\alpha_{ij}^k(t \mid t')$ is a function of $t - t'$ only, and just $O(N^2 \cdot N_c \cdot T)$ operations are required to solve the equations. Therefore, we focus on calculating the finite size effect for constant input in this

article.[1] Generally several iterations of the above set of equations are required for the convergence of mean firing intensities and cross-correlations. However, by initially setting those variables to the solution in the limit of $N_c \to \infty$, we found good approximations after a couple of iterations in many cases.

The mean firing rate and cross-correlation functions calculated from equations 4.11 and 4.12 are precise for small J and large N_c because the recurrent input is close to gaussian in this case (see Figure 4A). On the other hand, for large J and small N_c, the gaussian input approximation is poor, and therefore the approximation becomes worse (see Figure 4B). One key result here is that both the autocorrelations and the cross-correlations in these networks depend on the inputs $I(t)$ in general, as can be seen from equations 4.6 to 4.8. Changing just the baseline (mean) of the input $I(t)$ can significantly change the correlation structure in the network as seen in a number of physiological preparations (e.g., Hatsopoulos et al., 1998; Kohn & Smith, 2005), even if the coupling terms J and w are unchanged.

4.3 Linear Input-Output Filter. The ODE derived in equation 4.9 is general and may be applied given arbitrary time-dependent input $\{I_i(t)\}$. In this section, we exploit these equations to explore the linear response properties of these spiking networks. A standard physiological technique is white noise analysis (Marmarelis & Marmarelis, 1978; Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997), in which we inject white noise input in order to study a network's input-output filtering properties. When this analysis is applied to our GLM system, the external input is given by $I_i(t) = \int K_i(t-s)\xi_i(s)\,ds$, where K_i is neuron i's linear stimulus filter and $\xi_i(t)$ is the white noise stimulus with mean $\langle \xi(t) \rangle = \xi_i^{(0)}$ and small variance $\langle (\xi_i(t) - \xi_i^{(0)})(\xi_j(t') - \xi_j^{(0)}) \rangle = \sigma_\xi^2 \delta_{ij} \delta(t-t')$. We keep the input variance small here in order to treat the input as almost constant and develop a consistent linear response theory. Note that $\langle . \rangle$ describes the average over the fluctuating stimulus $\xi_i(t)$ here. We use $\delta X = X - \langle X \rangle$ for any function X of input stimulus and neglect $O(\delta I^2)$ terms in the following calculations.

Let us summarize the calculations in the mean-field ($N_c \to \infty$) limit. The mean-firing intensity of a neuron is given by the following self-consistent equations:

$$\nu_i(t) = f(I_i(t) + \mu_i(t)),$$

$$\mu_i(t) = \sum_k \int w_i^k(t-s)\nu_k(s)\,ds. \tag{4.13}$$

---

[1]Note that the computational complexity to calculate the finite size effect is not N^2 but $N^2 \cdot N_c$ because each synaptic time constant, τ_i^k, is distinct. Hence, we needed to evaluate each $\alpha_{ij}^k = E[h_i^k \mid S_j]$ term separately. However, if all the synaptic constants onto each neuron are constant, that is, $\tau_i^k = \tau_i$, we can directly evaluate $\mu_{ij} = \sum_k J_i^k \alpha_{ij}^k$ without evaluating α_{ij}^k separately, and the computational complexity to evaluate the finite size effect reduces to $O(N^2 \cdot T)$.
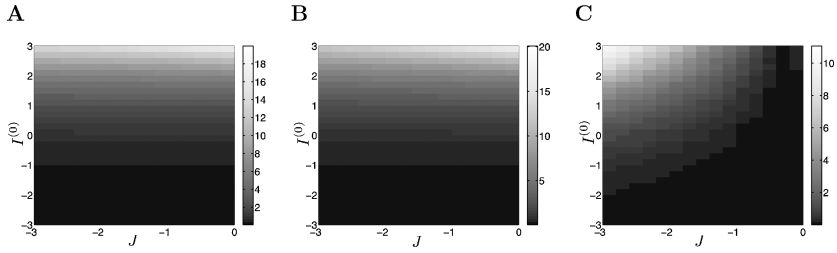
A **B** **C**

Figure 5: Comparison of (A) a simulated firing rate of a single neuron and (B) the self-consistent solutions of equation 4.13 for different input levels, $I^{(0)}$, and self-inhibition strengths, J. (C) The percentage errors of the mean-field approximation are shown with respect to the simulated firing intensity. We set $\tau = 10$ ms. (A color version of this figure is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/neco.2008.04-08-757.)

In particular, for a constant stimulus, $\xi_i(t) = \xi_i^{(0)}$, we obtain

$$v_i^{(0)} = f\left(I_i^{(0)} + \mu_i^{(0)}\right),$$
$$\mu_i^{(0)} = \sum_k \tau_i^k J_i^k v_k^{(0)} \tag{4.14}$$

from equation 2.4. Note that $X^{(0)}$ describes responses at constant stimulus $\xi_i^{(0)}$, for example, $I_i^{(0)} = \xi_i^{(0)} \int K(t)\,dt$. We can solve the zeroth-order terms in equation 4.14 self-consistently by finding the roots of the equations $v_i^{(0)} = f(I_i^{(0)} + \sum_k \tau_i^k J_i^k v_k^{(0)})$. In Figure 5, we compare this self-consistent mean-field solution with numerical simulation of a single neuron as a function of constant input, $I^{(0)}$, and self-inhibition strength J. We see that the approximation is good for small values of self-interaction J, but that the error increases with $|J|$ and $I^{(0)}$, as expected given the analysis in the previous section.

To perform the linear response analysis, we require the first-order terms. Differentiation of equation 4.12 with respect to stimulus perturbation, $\delta\xi$, yields

$$\delta v_i(t) = f'\left(I_i^{(0)} + \mu_i^{(0)}\right)(\delta I_i(t) + \delta\mu_i(t))$$

$$= f'\left(I_i^{(0)} + \mu_i^{(0)}\right) \int K_i(t-s)\delta\xi_i(s)\,ds + f'\left(I_i^{(0)} + \mu_i^{(0)}\right)$$

$$\times \sum_k \int w_i^k(t-s)\delta v_k(s)\,ds. \tag{4.15}$$

Now we can rewrite equation 4.15 by introducing vector notation as

$$\delta\boldsymbol{v}(t) = \int \left(\boldsymbol{K}(t-s)\delta\boldsymbol{\xi}(s) + \boldsymbol{w}(t-s)\delta\boldsymbol{v}(s)\right)\,ds, \tag{4.16}$$
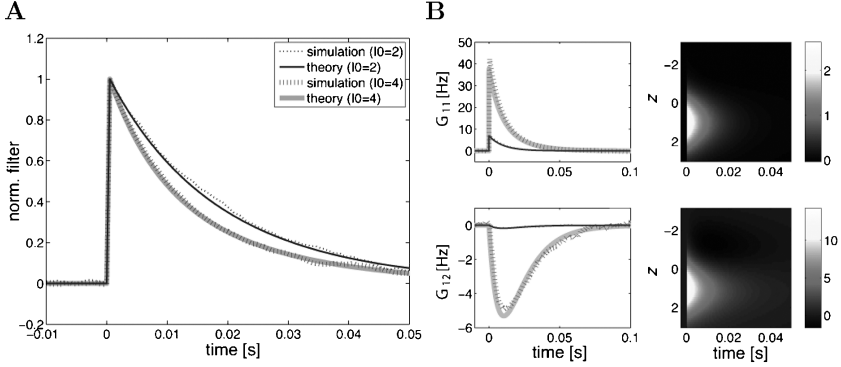
Figure 6: Comparison of the analytically derived and simulated input-output filters. (A) A simulated normalized linear input-output filter for $I^{(0)} = 2$ (thin-dotted) and $I^{(0)} = 4$ (thick-dotted) is compared with analytical results at $I^{(0)} = 2$ (thin-solid) and $I^{(0)} = 4$ (thick-solid) of equation 4.18. Because the strength of the self-inhibition component changes with the baseline input through $f'(I^{(0)} + \mu^{(0)})$, the input filter is sharpened for high baseline input. Other parameters are set to $J = 1$, $\tau = 10$ ms, $K(t) = e^{-t/\tau_I}\Theta(t)$ with $\tau_I = 20$ ms and $\sigma_\xi = 1$ Hz. All filter amplitudes are normalized by their maximum amplitudes. (B) Left panels show the comparison of analytically derived input-output filters $G_{11}(t)$ and $G_{12}(t)$ of equation 4.20 (solid lines) and simulated input-output filters (dotted lines) for $I^{(0)} = 2$ (thin lines) and $I^{(0)} = 4$ (thick lines). Right panels show neuron 1's spatiotemporal input-output filter $\tilde{G}_1(t, z)$ for $I^{(0)} = 2$ (top) and $I^{(0)} = 4$ (bottom). See the text for parameter values. (A color version of this figure is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/neco.2008.04-08-757.)

where we defined $[\boldsymbol{K}(s)]_{ij} = \delta_{ij} f'(I_i^{(0)} + \mu_i^{(0)})K_i(s)$ and $[\boldsymbol{w}(s)]_{ij} = f'(I_i^{(0)} + \mu_i^{(0)})w_i^j(s)$. Therefore, the Fourier transformation of equation 4.16 yields

$$\delta\hat{\boldsymbol{v}}(\omega) = \hat{\boldsymbol{K}}(\omega)\delta\hat{\boldsymbol{\xi}}(\omega) + \hat{\boldsymbol{w}}(\omega)\delta\hat{\boldsymbol{v}}(\omega)$$

$$= \hat{\boldsymbol{G}}(\omega)\delta\hat{\boldsymbol{\xi}}(\omega), \tag{4.17}$$

with a gain function $\boldsymbol{G}(t) = \int \frac{d\omega}{2\pi} e^{i\omega t}[\boldsymbol{I} - \hat{\boldsymbol{w}}(\omega)]^{-1}\hat{\boldsymbol{K}}(\omega)$. Note that $\boldsymbol{I}$ is an identity matrix here (not the input vector). We should note that a mathematically identical linear response filter was derived for a related model in Ginzburg and Sompolinsky (1994). One difference is that in their model, spinlike binary variables flip depending on input level with a fixed time constant. This corresponds, in our model, to the case that all the synaptic time constants τ_i^j are identical. Hence, the linear response filter in equation 4.17 is a natural generalization of their result.

We show in Figure 6A a simulated normalized receptive field of a single neuron and the analytical result of equation 4.17. In this single-neuron case, where K is set to be a single exponential function with time constant τ_I, for simplicity, the Fourier transformations of the synaptic filter and the input filter are $\hat{w}(\omega) = \tau/(1 + i\omega\tau)$ and $\hat{K}(\omega) = \tau_I/(1 + i\omega\tau_I)$, respectively. In this case, we obtain, after a Fourier inverse transformation by a Cauchy integral,

$$G(t) = \frac{f'}{1 + Jf'\tau_-} e^{-t/\tau_I} + \frac{Jf'^2\tau_-}{1 + Jf'\tau_-} e^{-(1/\tau - Jf')t}, \tag{4.18}$$

where $\tau_- = (1/\tau_I - 1/\tau)^{-1}$ and $f' = f'(I^{(0)} + \mu^{(0)})$. We see that the linear response may be described as a combination of the stimulus-filter term K along with an additional term due to the recurrent filter w. Of course, in the limit $J \to 0$, we recover $G \to f'K$. Note also that the input-output filter changes with the baseline input through f'. In the case of the exponential nonlinearity, f, and $J < 0$, the larger the input baseline, the sharper the linear filter. This is because the effective strength of of the spike history effect is modulated by the input-output nonlinearity and the baseline input level (see, e.g., Bryant & Segundo, 1976, for some similar effects in an in vitro physiological preparation).

Next, we calculate how interactions between neurons can change the input-output filter. For concreteness, we consider two neurons that receive a spatiotemporal white noise stimulus $\xi(z, t)$ with unit variance. Input to the neuron i is described by $\xi_i(t) = \int_{-\infty}^{\infty} dz L_i(z)\xi(t, z)$, where $L_i(z) = \frac{1}{\sqrt{2\pi}} e^{-(z-z_i)^2/2}$ are spatial filters with means $z_1 = 1$ and $z_2 = -1$. For simplicity, we assume that two neurons receive the same baseline of input, $I_1^{(0)} = I_2^{(0)} = I^{(0)}$ with an identical temporal stimulus filter, $K(t) = e^{-t/\tau_I}\Theta(t)$, and have symmetric synaptic interactions $J_{12} = J_{21} = J = -1$, $J_{11} = J_{22} = 0$, and $\tau_i^k = \tau_I = \tau = 10$ ms. The derivative of the nonlinear function is described by $f_i' = f(I_i^{(0)} + \mu_i^{(0)})$ for $i = 1, 2$. Proceeding as above, we can calculate the spatiotemporal input-output filter of neuron i as $\tilde{G}_i(t, z) = \sum_j G_{ij}(t)L_j(z)$. Similar to the previous single-neuron case, the Fourier transformation of the temporal stimulus filter is $[\hat{K}(\omega)]_{ij} = \delta_{ij}\frac{f_i'}{1/\tau + i\omega}$, while the synaptic filter is $[\hat{w}(\omega)]_{ij} = \frac{f_i' J_{ij}}{1/\tau + i\omega}$, which yield the Fourier transformation of the temporal input-output filter,

$$\hat{G}(\omega) = [I - \hat{w}(\omega)]^{-1}\hat{K}(\omega)$$

$$= \frac{1/(1/\tau + i\omega)}{1 - (f_1'f_2'J^2)/(1/\tau + i\omega)^2}$$

$$\times \begin{pmatrix} f_1' & f_1'f_2'J/(1/\tau + i\omega) \\ f_1'f_2'J/(1/\tau + i\omega) & f_2' \end{pmatrix}. \tag{4.19}$$

After a Fourier inverse transformation using Cauchy integration, we find

$$G(t) = e^{-t/\tau}\Theta(t)\begin{pmatrix} f_1'\cosh(\sqrt{f_1'f_2'}Jt) & \sqrt{f_1'f_2'}\sinh(\sqrt{f_1'f_2'}Jt) \\ \sqrt{f_1'f_2'}\sinh(\sqrt{f_1'f_2'}Jt) & f_2'\cosh(\sqrt{f_1'f_2'}Jt) \end{pmatrix}.$$

(4.20)

Figure 6B compares the analytically derived input-output filter of equation 4.20 with numerical simulation for different input baseline $I^{(0)} = 2$ and $I^{(0)} = 4$. Figure 6B, also shows the spatiotemporal input-output filter of neuron 1, $\tilde{G}_1(t, z) = G_{11}(t)L_1(z) + G_{12}(t)L_2(z)$, for the two input baselines. Under large input baseline conditions, we see stronger interactions between the two neurons and, hence, significant differences in the effective input-output filtering properties of the neuron in the two conditions. Thus, as emphasized above, the filtering properties of the network can change in a stimulus-dependent way, without any changes in the system parameters K, J, or w. (See, e.g., Pillow et al., 2008, for some related results concerning the effects of interneuronal coupling terms on the filtering properties of populations of primate retinal ganglion neurons.)

5 The Generalized Linear Point Process Model with Markov Refractoriness

In section 4, we saw that we could derive simple approximate expressions for many important quantities (firing rate, correlations) in the basic generalized linear point process model as long as the individual synaptic coupling terms $| J |$ are not so large that the accuracy of the mean-field expansion is compromised (see Figures 3 and 4).

In particular, the small $| J |$ condition is likely acceptable for the multineuronal coupling terms, since in many cases, the inferred coupling parameters—in motor cortex (Paninski et al., 2004; Truccolo et al., 2005) and retina (Pillow et al., 2008), for example—have been empirically found to be small (though larger network effects are found in hippocampus; see Harris et al., 2003; Okatan et al., 2005). Similarly, we might expect the "slow" self-inhibition terms in the GLM (the terms responsible for adaptation in the firing rate, for example) to be relatively small as well (Pillow, Paninski, Uzzell, Simoncelli, & Chichilnisky, 2005; Truccolo et al., 2005).

However, it is clear that the generalized linear model with small weights $| J |$ will not be able to account for the strong refractoriness that is a fundamental feature of neural point-process dynamics at fine timescales: in the GL model, strong, brief inhibition is produced by large, negatively weighted, and sharply decaying history effects $w_i^i(t)$, and these large $w_i^i(t)$ terms spoil

our mean-field approximation. Thus, it is clear that we need to adapt the methods introduced above in order to handle strong refractory effects.

One possible solution would be to take a discrete-time approach: instead of modeling neural responses as point processes in continuous time, we could simply model each cell's spike count within a discrete-time bin as a Bernoulli random variable whose rate $v_i(t)$ is an increasing function of $u_i(t)$, very much as in the continuous-time model described above (Chornoboy et al., 1988; Okatan et al., 2005). This discrete-time Bernoulli model shares much of the ease of fitting (including concavity of the log likelihood; Escola & Paninski, 2007) as the continuous-time model. In fact, the discrete-time model converges in a natural sense to the continuous-time model as the time binwidth dt becomes small. The advantage of this discrete-time formalism is that the firing rate $v_i(t)$ can never exceed $1/dt$ (since the Bernoulli probabilities $v_i(t)\,dt$ cannot exceed one), and therefore, in a crude sense, the discrete-time model has a refractory effect of length dt. A mean-field analysis can be developed for this discrete-time model, which exactly mirrors that developed above in the continuous-time model. We simply need to exchange our ordinary differential equations for discrete-time difference equations.

Of course, this discrete-time approach is unsatisfactory in at least one respect, since we would like, if possible, to model the firing behavior of neurons down to a millisecond timescale (Berry & Meister, 1998; Keat, Reinagel, Reid, & Meister, 2001; Pillow et al., 2005; Paninski et al., 2007), and the discrete-time approach, by construction, ignores these fine timescales. Thus, we introduce a model that allows us to incorporate strong refractory effects directly. We set the rates,

$$\lambda_i(t) = f(u_i(t))I(x_i(t) = M), \tag{5.1}$$

where $f(.)$ and $u_i(t)$ are defined as in section 4 and $I(.)$ is the indicator function that takes the value 1 if the argument is true. We have introduced an auxiliary refractory variable $x_i(t)$, which takes a discrete state from 1 to M. We assume that this variable $x_i(t)$ is itself Markovian, with transition rates

$$W_i(t) = \begin{pmatrix} -1/\tau_r & 0 & 0 & 0 & & S_i(t) \\ 1/\tau_r & -1/\tau_r & 0 & 0 & \cdots & 0 \\ 0 & 1/\tau_r & -1/\tau_r & 0 & & 0 \\ & \vdots & & & \ddots & -S_i(t) \end{pmatrix}; \tag{5.2}$$

that is, when $x_i(t)$ is in state M, it will transition to state 1 with each spike, and then transitions occur from state m to state $m+1$ with rate $1/\tau_r$ until $x_i(t)$ has reached the spiking state M once again. Refractoriness in this

model is enforced because the neuron is silent whenever $x_i(t)$ is in one of the postspike quiescent states $x_i(t) < M$. It is easy to see that this is a kind of inhomogeneous renewal model (the delay τ_d required for the state variable x_i to move from state 1 to first reach the active state M is a sum of $(M-1)$ independent and identically (i.i.d.) distributed exponential random variables of mean τ_r, and so τ_d is an i.i.d. gamma variable with parameters $(M-1, 1/\tau_r)$) and may be considered a special case of the "inhomogeneous Markov interval" model introduced by Kass and Ventura (2001). By adjusting the number of states M and the rate $1/\tau_r$, we may adjust the refractory properties of the model: for example, $M = 1$ implies that we have an inhomogeneous Poisson model, while $M = 2$ provides an exponentially decaying relative refractory effect, and if we let M be large, with τ_r scaling like $\tau_r \sim \tau/M$, we obtain an absolute refractory effect of length τ.

Finally, it is easy to show that the additional Markov term in the definition of the rate $\lambda_i(t)$ does not have a negative impact on the estimation of the GL model parameters J given spiking data; as discussed in Paninski (2004), maximizing the likelihood in this model (or constructing an EM algorithm, as in Escola & Paninski, 2007) requires that we maximize a nonnegatively weighted version of the standard point-process log likelihood, and this weighted log likelihood retains all of the concavity properties of the original GL model. Thus, this new model requires just a single concave optimization and is as easy to fit as the standard GL point process model. The advantage is that a strong relative refractory effect is intrinsic to the model and does not need to be enforced by a large, brief, negative $w_i^i(t)$ term. Instead, we may use small adjustments to the J_i^i terms to fine-tune the model to match the short-time details of the observed interspike interval density.[2] As we will see below, this leads to much more accurate mean-field approximations of the firing rates and correlation functions in these networks.

Let us define the probability $\boldsymbol{p}_i(t) = (P(x_i(t) = 1), \ldots, P(x_i(t) = M))^T$ that the neuron i is in state $x_i(t) = 1, \ldots, M$. Note that the boldface characters denote vectors. The dynamics of $\boldsymbol{p}_i$ is described by

$$\frac{d\,\boldsymbol{p}_i(t)}{dt} = \overline{W}_i(t)\,\boldsymbol{p}_i(t), \tag{5.3}$$

---

[2]In this article, we mostly use $M = 3$ (except for Figure 7B, where the effect of abrupt refractoriness is studied), because this is the smallest value of M that requires the following mean-field formulation (we do not need a vector formulation for an $M = 2$ case because there is only one degree of freedom that corresponds to the probability of being in the active state). It is worth noting that Escola and Paninski (2007) discuss methods for additionally estimating an optimal transition matrix $W_i(t)$ via an EM algorithm. This provides another method for adjusting the short-time details of the model's responses. In addition, we may extend many of the mean-field methods developed below to the case of more general rate matrices $W_i(t)$. However, for simplicity, we will not pursue these extensions here.

where $\overline{W}_i(t)$ has the same elements as $W_i(t)$ of equation 5.2 except for

$$
\begin{aligned}
{[\overline{W}_i(t)]}_{1M} = -[\overline{W}_i(t)]_{MM} &= E[S_i(t) \mid x_i(t) = M] \\
&= \frac{E[S_i(t)I(x_i(t) = M)]}{E[I(x_i(t) = M)]} \\
&= \frac{v_i(t)}{p_i(t)},
\end{aligned}
\tag{5.4}
$$

which describes the probability of neuron i firing at time t given $x_i(t) = M$. Note that we used $v_i(t) = E[S_i(t)] = E[S_i(t)I(x_i(t) = M)]$ and wrote the last component of $\boldsymbol{p}_i(t)$ as $[\boldsymbol{p}_i(t)]_M = p_i(t)$.

We also define $[\boldsymbol{p}_{ij}(t \mid t')]_m = E[I(x_i(t) = m) \mid S_j(t')]$ as the probability of neuron i being in state m given a spike of neuron j at time t'. In particular, the last component is written as $p_{ij}(t \mid t') = [\boldsymbol{p}_{ij}(t \mid t')]_M$. The evolution of $\boldsymbol{p}_{ij}(t \mid t')$ is described by a similar equation,

$$
\frac{d\,\boldsymbol{p}_{ij}(t \mid t')}{dt} = \overline{W}_{ij}(t \mid t')\,\boldsymbol{p}_{ij}(t \mid t'),
\tag{5.5}
$$

where the transition matrix $\overline{W}_{ij}(t \mid t')$ has the same elements as $W_i(t)$ of equation 5.2 except for

$$
\begin{aligned}
{[\overline{W}_{ij}(t \mid t')]}_{1M} = -[\overline{W}_{ij}(t \mid t')]_{MM} &= E[S_i(t) \mid x_i(t) = M, S_j(t')] \\
&= \frac{E[S_i(t)I(x_i(t) = M) \mid S_j(t')]}{E[I(x_i(t) = M) \mid S_j(t')]} \\
&= \frac{v_{ij}(t \mid t')}{p_{ij}(t \mid t')}.
\end{aligned}
\tag{5.6}
$$

Before turning to "mean-field" approximation, it is worth noting that we may solve for the firing rates and correlations exactly in the special case of no synaptic couplings, $J = 0$. (The analogous case in the standard GL model is the inhomogeneous Poisson case, which is of course trivial.) We can proceed by exploiting either the renewal nature of the model (Gerstner & Kistler, 2002), which leads to convolution or infinite-sum formulas for the firing rate, or the Markovian structure, which leads to somewhat more intuitive ordinary differential equations. We pursue the second approach here.

In this case, the mean firing intensity is given by

$$
v_i(t) = E[f(I_i(t))I(x_i(t) = M)] = f(I_i(t))p_i(t).
\tag{5.7}
$$

This in turn gives, together with equation 5.4, $[\overline{W}_i(t)]_{1M} = -[\overline{W}_i(t)]_{MM} = f(I_i(t))$. Hence, we can find the dynamics of mean firing intensity by solving equation 5.3. In the special case of constant input I_i, it is easy to solve the fixed point of $p_i(t)$; we find for the last component that $p_i = [1 + (M - 1)\tau_r f(I_i)]^{-1}$.

The autocorrelation function in this case is similarly easy to derive. (Note that the cross-covariance functions are zero in this $J = 0$ case, since the cross-coupling terms are set to zero.) In case of no couplings, $J = 0$, the conditional firing intensity is given by

$$v_{ij}(t \mid t') = E[f(I_i(t))I(x_i(t) = M) \mid S_j(t')] = f(I_i(t))p_{ij}(t \mid t') \qquad (5.8)$$

for $t > t'$. We find those corner elements of $\overline{W}_{ij}(t \mid t')$ as

$$[\overline{W}_{ij}(t \mid t')]_{1M} = -[\overline{W}_{ij}(t \mid t')]_{MM} = f(I_i(t)) \qquad (5.9)$$

for $t > t'$. This means that $\overline{W}_{ij}(t \mid t') = \overline{W}_i(t)$ for $t > t'$. Hence, $p_{ij}(t \mid t')$ follows the same differential equation as $p_i(t)$ in this $J = 0$ case but starting from the initial condition

$$p_{ij}(t \mid t^-) = \lim_{\epsilon \to 0} p_{ij}(t \mid t - \mid \epsilon \mid)$$

$$= \delta_{ij}(1, 0, \ldots, 0)^T + (1 - \delta_{ij})p_i(t) \qquad (5.10)$$

because the state x_i is reset to state 1 just after a given spike of neuron i.

In the following sections, we apply mean-field approximations for the firing rates and correlations in the nonzero J case.

5.1 Mean-Field Approximation of the GL Model with Markov Refractoriness. Let us approximate the mean firing intensity and the cross-correlations of GL model with Markov refractoriness assuming that the contribution of individual synaptic coupling is small. Similar to the calculation without the Markov refractoriness, we assume a gaussian distribution of recurrent input $H_i(t)$ given $x_i(t) = M$ to find

$$v_i(t) = E[f(u_i(t))I(x_i(t) = M)]$$

$$= E[f(u_i(t)) \mid x_i(t) = M]p_i(t)$$

$$\approx F(I_i(t) + \mu_i(t), \sigma_i(t))p_i(t), \qquad (5.11)$$

with conditional mean $\mu_i(t) = E[H_i(t) \mid x_i(t) = M]$ and conditional variance $\sigma_i^2(t) = \text{Var}[H_i(t) \mid x_i(t) = M]$. Also, assuming a gaussian distribution of recurrent input $H_i(t)$ given $x_i(t) = M$ and a spike of neuron j at time t',

we find

$$
\begin{aligned}
v_{ij}(t \mid t') &= E[f(u_i(t))I(x_i(t) = M) \mid S_j(t')] \\
&= E[f(u_i(t)) \mid x_i(t) = M, S_j(t')]E[I(x_i(t) = M) \mid S_j(t')] \\
&\approx F(I_i(t) + \mu_{ij}(t \mid t'), \sigma_{ij}(t \mid t'))p_{ij}(t \mid t')
\end{aligned}
\tag{5.12}
$$

for $t > t'$ with conditional mean $\mu_{ij}(t \mid t') = E[H_i(t) \mid x_i(t) = M, S_j(t')]$ and variance $\sigma_{ij}^2(t \mid t') = \mathrm{Var}[H_i(t) \mid x_i(t) = M, S_j(t')]$. The dynamics of $p_i(t)$ and $p_{ij}(t \mid t')$ are described by equations 5.3 and 5.5, respectively.

Next we calculate the conditional means and variances of the recurrent input. We find

$$
\mu_i(t) = \sum_k J_i^k \frac{\beta_i^k(t)}{p_i(t)},
$$

$$
\mu_{ij}(t \mid t') = \sum_k J_i^k \frac{\beta_{ij}^k(t \mid t')}{p_{ij}(t \mid t')},
$$

$$
\sigma_i^2(t) = \sum_{k,l} J_i^k J_i^l \frac{B_i^{kl}(t)}{p_i(t)},
$$

$$
\sigma_{ij}^2(t \mid t') = \sum_{k,l} J_i^k J_i^l \frac{B_{ij}^{kl}(t \mid t')}{p_{ij}(t \mid t')}
\tag{5.13}
$$

with auxiliary variables defined as

$$
\begin{aligned}
\left[\boldsymbol{\beta}_i^k(t)\right]_m &= E[h_i^k(t) \mid x_i(t) = m] \cdot [\boldsymbol{p}_i(t)]_m, \\
\left[\boldsymbol{\beta}_{ij}^k(t \mid t')\right]_m &= E[h_i^k(t) \mid x_i(t) = m, S_j(t')] \cdot [\boldsymbol{p}_{ij}(t \mid t')]_m, \\
\left[\boldsymbol{B}_i^{kl}(t)\right]_m &= \mathrm{Cov}[h_i^k(t), h_i^l(t) \mid x_i(t) = m] \cdot [\boldsymbol{p}_i(t)]_m, \\
\left[\boldsymbol{B}_{ij}^{kl}(t \mid t')\right]_m &= \mathrm{Cov}[h_i^k(t), h_i^l(t) \mid x_i(t) = m, S_j(t')] \cdot [\boldsymbol{p}_{ij}(t \mid t')]_m.
\end{aligned}
\tag{5.14}
$$

Note that we use the final component of the above quantities without an index (e.g., $\beta_i^k(t) = \left[\boldsymbol{\beta}_i^k(t)\right]_M$). After some calculation (see the appendix), we obtain

$$
\beta_i^k(t) = \int_{-\infty}^t ds\, e^{-(t-s)/\tau_i^k}\, p_{ik}(t \mid s)v_k(s),
$$

$$
\beta_{ij}^k(t \mid t') = \int_{-\infty}^t ds\, e^{-(t-s)/\tau_i^k}\, p_{ikj}(t \mid s, t')v_{kj}(s \mid t'),
$$

$$\left[\boldsymbol{B}_i^{kl}(t)\right]_m = \int_{-\infty}^t ds \int_{-\infty}^t ds' \, e^{-(t-s)/\tau_i^k - (t-s')/\tau_i^l} \left\{ \left[\boldsymbol{p}_{ikl}(t \mid s, s')\right]_m \phi_{kl}(s, s') \right.$$

$$\left. - \frac{\left[\boldsymbol{p}_{ik}(t \mid s)\right]_m \left[\boldsymbol{p}_{il}(t \mid s')\right]_m}{\left[\boldsymbol{p}_i(t)\right]_m} v_k(s) v_l(s') \right\}, \tag{5.15}$$

where $\left[\boldsymbol{p}_{ikl}(t \mid s, s')\right]_m = E[I(x_i(t) = m) \mid S_k(s), S_l(s')]$ and, particularly, the last component is written as $p_{ikl}(t \mid s, s') = \left[\boldsymbol{p}_{ikl}(t \mid s, s')\right]_M$.

5.2 Solution in the Mean-Field Limit. Similar to the case without Markov refractoriness, we first derive a self-consistent equation for the mean firing intensity and cross-covariances in the mean-field limit $N_c \to \infty$, where $J \sim 1/N_c$ and assuming an asynchronous state so that $\Delta\phi_{ij} \sim 1/N_c$. Because all the cross-covariance functions vanish and $\sigma_i^2(t) \to 0$ in the limit, we find, from equations 5.11 to 5.13, for $t > t'$,

$$v_i(t) \to f(I_i(t) + \mu_i(t)) p_i(t),$$

$$v_{ij}(t \mid t') \to f(I_i(t) + \mu_i(t)) p_{ij}(t \mid t'),$$

$$\mu_i(t) = \sum_k J_i^k \frac{\beta_i^k(t)}{p_i(t)}, \tag{5.16}$$

while the cross-correlation function is given by equation 4.3. For the corner terms of the transition matrices, we find $[\overline{W}_i(t)]_{1M} \to f(I_i(t) + \mu_i(t))$ from equation 5.4 and $[\overline{W}_{ij}(t \mid t')]_{1M} \to f(I_i(t) + \mu_i(t))$ for $t > t'$ from equation 5.6. Hence, the evolution of $\boldsymbol{p}_i(t)$ and $\boldsymbol{p}_{ij}(t \mid t')$ is described by

$$\frac{d\boldsymbol{p}_i(t)}{dt} = \overline{W}_i(t) \boldsymbol{p}_i(t),$$

$$\frac{d\boldsymbol{p}_{ij}(t \mid t')}{dt} = \overline{W}_i(t) \boldsymbol{p}_{ij}(t \mid t') \tag{5.17}$$

for $t > t'$, where the initial condition of $\boldsymbol{p}_{ij}(t \mid t')$ is given by equation 5.10. Finally, from equation 5.15, the evolution of $\boldsymbol{\beta}_i(t)$ is written as an ODE,

$$\frac{d}{dt} \boldsymbol{\beta}_i^k(t) = \left(-\frac{1}{\tau_i^k} + \overline{W}_i(t)\right) \boldsymbol{\beta}_i^k(t) + \boldsymbol{p}_{ik}(t \mid t^-) v_k(t), \tag{5.18}$$

where we can apply equation 5.10 to evaluate $p_{ik}(t \mid t^-)$ in this limit. Figure 7 plots the mean firing intensity calculated from the above equations for various input, and the mean-field approximation provides good approximation of the time dependent firing intensities. Figure 8 compares the above approximation ($N_c \to \infty$ limit) with the mean-field approximation, including the finite size effect that we discuss in the next section. In the mean-field limit, all the cross-correlations vanish and the autocorrelation functions
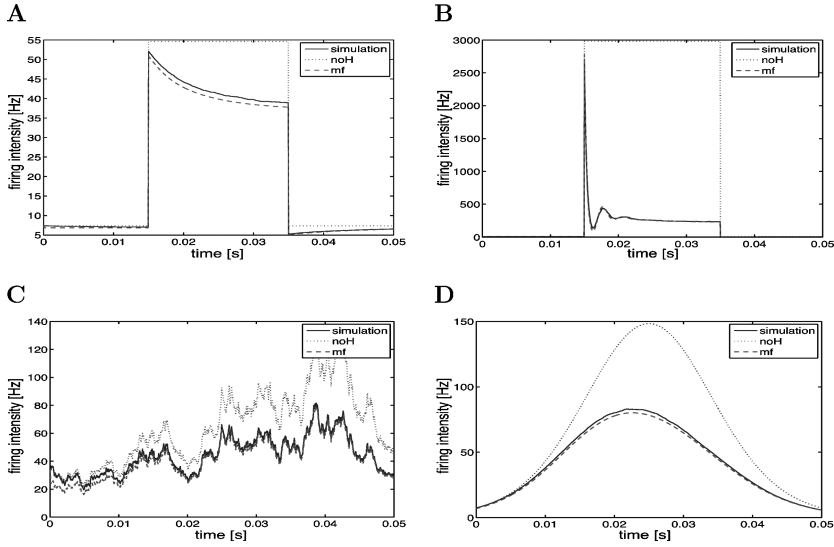
Figure 7: Comparison of the mean firing intensity obtained from direct simulation of a single neuron (solid), without recurrent input H and no Markov refractory effect (dotted), and the mean-field equation in the limit of $N_c \rightarrow \infty$ (dashed). In all cases, the Markov refractory effect with $\tau_r = 2$ ms$/(M - 1)$ was included. (A) Weak step input (baseline $I = 2$, peak $I = 4$) and refractoriness $M = 3$; $J = -1$, $\tau = 10$ ms. Although the model includes strong Markov refractory effect, the mean-field approximation is not abolished. (B) Strong step input (baseline $I = 2$, peak $I = 8$) and refractoriness $M = 10$; $J = -1$, $\tau = 10$ ms. The mean-field approximation also approximates very well the transient oscillation caused by the the strong refractory effect and steep current. (C) The mean-field approximation works for colored noise input: $\tau_I \frac{dI(t)}{dt} = -I(t) + \xi(t)$ with $E[\xi(t)] = 4$, $\text{Cov}[\xi(t), \xi(t')] = 0.1\delta(t - t')$, and $\tau_I = 20$ ms; $M = 3$, $J = -1$, $\tau = 10$ ms. The initial error is due to the preset initial values that deviate from the true ones. (D) Sinusoidal input input: $I(t) = 2 + \sin(2\pi \cdot 10 \text{ Hz} \cdot t)$; $M = 3$, $J = -1$, $\tau = 10$ ms. The mean-field results give good approximations. (A color version of this figure is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/neco.2008.04-08-757.)

are approximated by including the Markov refractory effect but not the self-interaction terms J_i^i. We can also calculate the linar input-output filter discussed in section 4.3 with this Markov refractoriness (see the appendix).

5.3 Estimating the Finite Size Effect. Similar to the case without Markov refractoriness, we evaluate the mean firing intensity and cross-covariance functions up to $1/N_c$ terms. We find, as in the case without Markov refractoriness, $\Delta\phi_{ij}(t, t') \sim 1/N_c$ for $i \neq j$, $\Delta p_{ik}(t \mid s) \sim 1/N_c$ for $i \neq k$, and the third-order covariances, such as $\Delta p_{ikl}(t \mid s, s')$, are of order
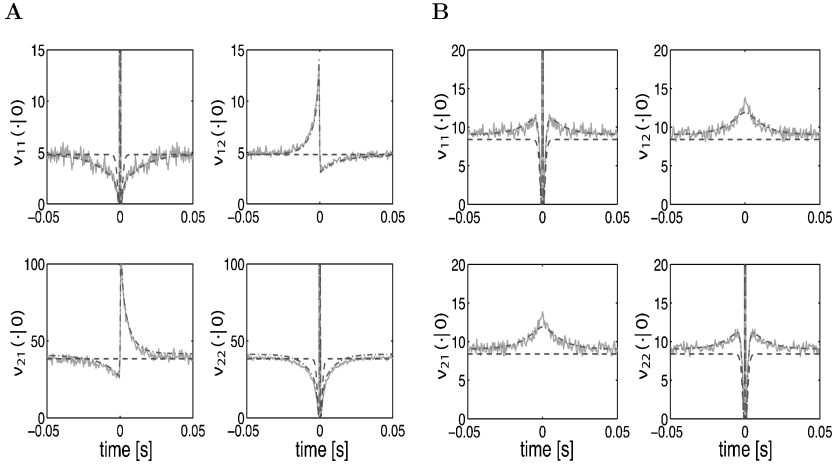
A

B

Figure 8: Comparison of spike cross-correlation functions of two neurons obtained by direct numerical simulation (solid), mean-field approximation in the limit of $N_c \to \infty$ (dashed), and with finite size correction terms (dot-dashed). While the approximation in the limit of $N_c \to \infty$ captures only the strong Markov refractory effect, the weaker refractory effect caused by the self-interaction terms is well captured with the finite size correction. The finite size correction provides a good approximation despite the strong refractoriness because of an explicit evaluation of the Markov refractory effect. Mean-field equations are iteratively solved over five iterations. (A) Two neurons are directly coupled: parameters are set to $J_{11} = -1$, $J_{22} = -1$, $J_{12} = -0.5$, $J_{21} = 1$, $\tau_{11} = 10$ ms, $\tau_{22} = 10$ ms, $\tau_{12} = 20$ ms, $\tau_{21} = 10$ ms, $M = 3$, $I_1 = 2$, $I_2 = 4$. (B) Two neurons are not directly connected but receive common input from a third neuron: parameters are set to $J_{13} = J_{23} = 2$, $\tau_{13} = \tau_{23} = 10$ ms, $M = 3$, $I_1 = I_2 = I_3 = 2$. (A color version of this figure is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/neco.2008.04-08-757.)

$1/N_c^2$ under the asynchronous state if the contributions of individual synaptic weights are small, $J \sim 1/N_c$. This implies $\sigma_{ij}^2(t \mid t') - \sigma_i^2(t) \sim 1/N_c^2$.

The dynamics of $p_i(t)$ and $p_{ij}(t \mid t')$ are described by equations 5.3 and 5.5, respectively. We now want to calculate $\mu_i(t)$, $\mu_{ij}(t \mid t')$, and $\sigma_i^2(t)$ to the first order of $1/N_c$. First, direct differentiation of $\beta_i^k(t)$ in equation 5.15 yields an ODE update equation (see the appendix):

$$\frac{d}{dt}\beta_i^k(t) = \left(-\frac{1}{\tau_i^k} + \overline{W}_i(t)\right)\beta_i^k(t) + p_{ik}(t \mid t^-)\nu_k(t)$$

$$-m\left[\beta_i^k(t) - \alpha_{ii}^k(t^- \mid t)p_i(t)\right]\frac{\nu_i(t)}{p_i(t)}, \quad (5.19)$$

with $m = (1, 0, \ldots, 0, -1)^T$. The origin of the second term on the right-hand side of equation 5.19 is due to the correlation of spike variable $S_k(s)$ and the

refractory variable $x_i(t)$. Next, we need to evaluate $\boldsymbol{\beta}_{ij}^k(t \mid t')$ to the zeroth order of $1/N_c$ for $k = i$ or $k = j$ because $J \sim 1/N_c$ makes the contribution of them $\sim 1/N_c$, and we need to evaluate $\boldsymbol{\beta}_{ij}^k(t \mid t')$ to the first order of $1/N_c$ for $k \neq i$ and $k \neq j$ case. Except for the case $i = j = k$, we find (see the appendix)

$$\boldsymbol{\beta}_{ij}^k(t \mid t') \approx \boldsymbol{\beta}_i^k(t) + \alpha_{ij}^k(t \mid t')\, p_{ij}(t \mid t') - \alpha_i^k(t)\, p_i(t) \tag{5.20}$$

to the order described above. The precise evaluation of this term for $i = j = k$ is much harder but, to a good approximation, equation 5.20 holds (see the appendix). Note that the approximation of the $i = j = k$ term affects the order $1/N_c$ term of the autocovariance function, $\Delta\phi_{ii}(t, t')$, but affects only the $O(1/N_c^2)$ terms of the mean firing intensities and cross-covariance functions.

Finally, dropping higher-order terms that result as $O(1/N_c^2)$ on the quantities in equation 5.13, we find (see the appendix)

$$B_i^{kl}(t) \approx p_i(t) A_i^{kl}(t). \tag{5.21}$$

Altogether, using equation 4.12, we obtain the following update equations to evaluate the finite size effect up to order $1/N_c$:

$$v_i(t) = F(I_i(t) + \mu_i(t), \sigma_i(t))\, p_i(t),$$

$$v_{ij}(t \mid t') = F(I_i(t) + \mu_{ij}(t \mid t'), \sigma_i(t))\, p_{ij}(t \mid t'),$$

$$\mu_i(t) = \sum_k J_i^k \frac{\beta_i^k(t)}{p_i(t)},$$

$$\mu_{ij}(t \mid t') = \sum_k J_i^k \left[\alpha_{ij}^k(t \mid t') + \frac{\beta_i^k(t) - \alpha_i^k(t)p_i(t)}{p_{ij}(t \mid t')} \right],$$

$$\sigma_i^2(t) = \sum_{k,l} J_i^k J_i^l A_i^{kl}(t),$$

$$\frac{d}{dt}\alpha_i^k(t) = -\frac{\alpha_i^k(t)}{\tau_i^k} + v_k(t),$$

$$\frac{d}{dt}\alpha_{ij}^k(t \mid t') = -\frac{\alpha_{ij}^k(t \mid t')}{\tau_i^k} + v_{kj}(t \mid t'),$$

$$\frac{d}{dt}A_i^{kl}(t) = -\frac{A_i^{kl}(t)}{\tau_i^{kl}} + \Delta\alpha_{il}^k(t^+ \mid t)v_l(t) + \Delta\alpha_{ik}^l(t^- \mid t)v_k(t),$$

$$\frac{d}{dt}\boldsymbol{\beta}_i^k(t) = \left(-\frac{1}{\tau_i^k} + \overline{W}_i(t) \right) \boldsymbol{\beta}_i^k(t) + p_{ik}(t \mid t^-)v_k(t)$$

$$- m\left[\beta_i^k(t) - \alpha_{ii}^k(t^- \mid t)p_i(t) \right] \frac{v_i(t)}{p_i(t)}, \tag{5.22}$$

and the dynamics of $\boldsymbol{p}_i(t)$ and $\boldsymbol{p}_{ij}(t \mid t')$ follows equations 5.3 and 5.5. As is the case without the Markov refractoriness, calculating this finite size effect is computationally expensive for time-dependent input, although in principle, this equation should work under that case as well. We plot the mean firing intensities of two neurons and cross-correlations in Figure 8. The finite size effect well captures the cross-correlation functions. It also successfully captures a peak in the cross-correlation function if two neurons are not directly connected but receive a common input from a third neuron (see Figure 8). Again, if the values of the above equations are initially set to the solution in the limit of $N_c \to \infty$ given by equations 5.16 to 5.18, only a couple of iterations of the above equations are sufficient for the evaluation of cross-correlation functions.

6 Discussion

We have introduced mean-field methods for analyzing the dynamics of a coupled population of neurons whose activity may be well approximated as the output of a generalized linear point process model. Our approximations for the mean time-varying firing rate and correlations in the population are exact in the mean-field limit ($N_c \to \infty$ under $J \sim 1/N_c$ scaling), though we have found numerically that the finite size correction to the mean-field equations is useful even for physiologically relevant connectivity strengths. The approximations may be computed by solving a set of coupled ordinary differential equations. This approach is much more computationally tractable than direct Monte Carlo sampling and may lead to greater analytical insight into the behavior of the network. In addition, we have introduced a new model, the generalized linear point process model with Markovian refractoriness, that captures strong refractoriness, retains all of the easy-fitting properties of the standard generalized linear model, and whose firing rate dynamics are much more amenable to mean-field analysis than the standard modells. Note that most of our illustrations involved the simulation of only a couple of mutually connected neurons; this small-N_c case can be considered a kind of worst-case analysis, since with more asynchronous neurons (larger N_c), the central limit theorem becomes progressively applicable, the distribution of recurrent input approaches the assumed gaussian distribution, and our approximations become more accurate.

This kind of mean-field analysis—or more generally, approaches for reducing the complexity of the dynamics of networks of noisy, nonlinear neurons—has a long history in computational neuroscience, as reviewed (Hertz et al., 1991; Gerstner & Kistler, 2002; Renart et al., 2003). While the point-process models we have discussed here are somewhat distinct from the noisy integrate-and-fire-type models that have been analyzed most extensively in this literature (e.g., Mattia & Del Giudice, 2002; Shriki, Hansel, & Sompolinsky, 2003; Moreno-Bote & Parga, 2004, 2006; Fourcaud-Trocme, Hansel, van Vreeswijk, & Brunel, 2003; Chizhov & Graham, 2007; Doiron,

Lindner, Longtin, Maler, & Bastian, 2004; Lindner, Doiron, & Longtin, 2005), it is more important to stress the differences in the motivation of this work and previous work. Our main goal here was to provide analytical tools that an experimentalist who has fit a point process model to his or her data (as in, e.g., Paninski, 2004; Truccolo et al., 2005; Okatan et al., 2005; Pillow et al., 2008) may use to understand the behavior of the network or single-cell models that have just been inferred. In particular, we were interested in predicting, for example, the mean firing rate of a specific GLM network to a novel arbitrary dynamic input stimulus. This contrasts with the literature cited above, which has for the most part focused on the mean firing rates, first-order response, and fixed-point stability properties behavior of idealized, infinitely large populations (Mattia & Del Giudice, 2002, is an exception here) with homogeneous membrane and connectivity properties. In addition, most analytical studies of the input-output property of model neurons with a dynamical input start from a Fokker-Planck formalism and rely on a direct numerical simulation of the partial differential equation (Nykamp & Tranchina, 2000; Knight, Omurtag, & Sirovich, 2000; Chizhov & Graham, 2007; this direct approach is feasible only in the case that the system dynamics may be reduced to a state space of dimension at most two or so, and therefore does not apply in the networks studied here), or on linear response theory that assumes a small time-dependent component relative to the baseline component (Shriki et al., 2003; Fourcaud-Trocme et al., 2003; Doiron et al., 2004; Lindner et al., 2005), or on some quasi-stationary assumption (Knight et al., 2000; Mattia & Del Giudice, 2002; Fourcaud-Trocme & Brunel, 2005; Moreno-Bote & Parga, 2004, 2006) restricting the applicable input stimulus to slowly changing one or two simple step stimuli with sufficiently long gaps between steps that the network may reach equilibrium. Thus, it is difficult to apply the insights developed in these previous analyses directly to obtain a quantitative prediction of the dynamic firing rates of a specific nonhomogeneous, nonsymmetric network.

We derived, in this article, the finite size correction to the mean-field equation that captures cross-correlation functions between neurons. This finite size effect originates from the $J_{ij} \sim 1/N_c$ scaling of synaptic strengths that guarantees $O(1/\sqrt{N_c})$ fluctuation of inputs under asynchronous states. Hence, the finite size effect described in this article is the contribution of small fluctuation in the input around the mean. Another kind of finite size effect has also been discussed in the literature. In a sparsely connected network, correlations between two neurons disappear if N_c is sufficiently smaller than the total number of neurons, N (Derrida, Gardner, & Zippelius, 1987). The finite N_c effect has been evaluated using stochastic Fokker-Planck equations (Brunel & Hakim, 1999; Mattia & Del Giudice, 2002). One should note, however, that the evaluation of this finite size effect for an experimentally estimated network structure is generally not straightforward because the evaluation of a Fokker-Planck equation with many state-space variables is computationally hard. Finally, $J_{ij} \sim 1/N_c$ is not the unique way to

scale synapses. Under the balanced input assumption, $J_{ij} \sim 1/\sqrt{N_c}$ yields order one fluctuation of input even in the limit of $N_c \to \infty$ (Sompolinksy, Crisanti, & Sommers, 1988; van Vreeswijk & Sompolinsky, 1996, 1998). However, nontrivial cross-correlation functions cannot be captured under the standard asynchronous assumption of mean-field analysis in the $N_c \to \infty$ limit. The calculation of the finite size effect of a network with $J_{ij} \sim 1/\sqrt{N_c}$ is not the scope of this article.

Two works that bear a stronger mathematical resemblance to ours are Ginzburg and Sompolinsky (1994) and Meyer and van Vreeswijk (2002). These authors applied mean-field approximations to coupled model neurons to study the mean firing rates and autocorrelation and cross-correlation functions in the network. However, the two-state neuron model discussed in Ginzburg and Sompolinsky (1994) (where the neurons flip between "active" and "inactive" firing states according to a Markov process with some finite rate constant) is somewhat distinct from the point-process models we have treated here (where the neuron remains in the "active" state for a negligible time—i.e., the neuron spikes instantaneously), and we have not been able to translate these earlier results to the problems of interest in this article. Renewal point spiking neuorns are analyzed in Meyer and van Vreeswijk (2002), and their results are closer to ours. However, according to their refractory model, one has to consider infinitely many refractory states in the continuous time limit, whereas our spiking neuron model has M refractory state; M can be as small as 2 while keeping the order 1 strong refractory effect. Accordingly, their cross-correlations should be evaluated by integral equations that are computationally (and conceptually) somewhat more involved, while the methods developed here require us to evaluate just a couple of differential equations.

Finally, the effect of common input on the cross-correlation function was previously studied in a related model (Nykamp, 2007), using an expansion of the output $f(.)$ nonlinearity arround $J = 0$. One major difference between the analysis we have presented here and this previous work is that a simple expansion around $J = 0$ does not lead to a good approximation of the firing rate, even in the mean-field limit of $N_c \to \infty$, because the recurrent input changes the baseline firing rate (cf. Figure 5); thus, it is much more accurate to expand around the zeroth-order firing rate given by the roots of equation 4.13, as discussed in section 4.3. (On the other hand, our mean-field approach does require a gaussian approximation that is known to be inaccurate in the case of large J terms.) It would be interesting to explore whether the methods developed here could help lead to more accurate inference of the common-input effects discussed in Nykamp (2007).

There are many possible applications of this mean-field method for problems that require fast evaluation of mean firing rates and cross-correlation functions. One example is to evaluate the information coded by spiking of a recurrently connected network about the input stimulus. This kind of

information calculation usually requires averaging over recent spike history (Toyoizumi, Aihara, & Amari, 2006), but the gaussian approximation of the input described here could greatly ease the computational complexity to evaluate the information of stimulus coded by the network. We hope to pursue this direction in future work.

Appendix: Mathematical Details ⸺⸺⸺⸺⸺⸺⸺⸺

In this appendix, we collect the details of the expansion analyses summarized in the main text.

A.1 Calculation of β_i^k, β_{ij}^k, and B_i^{kl}. In this appendix, we use $D_i^k s = e^{-(t-s)/\tau_i^k} ds$ to simplify some expressions. Direct evaluation of equation 5.14 yields

$$
\begin{aligned}
\left[\beta_i^k(t)\right]_m &= E\left[h_i^k(t) \mid x_i(t) = m\right] \cdot \left[p_i(t)\right]_m \\
&= E\left[h_i^k(t) I(x_i(t) = m)\right] \\
&= \int_{-\infty}^t D_i^k s\, E\left[S_k(s) I(x_i(t) = m)\right] \\
&= \int_{-\infty}^t D_i^k s\, \left[p_{ik}(t \mid s)\right]_m v_k(s), \\
\left[\beta_{ij}^k(t \mid t')\right]_m &= E[h_i^k(t) \mid x_i(t) = m, S_j(t')] \cdot \left[p_{ij}(t \mid t')\right]_m \qquad\text{(A.1)} \\
&= E[h_i^k(t) I(x_i(t) = m) \mid S_j(t')], \\
&= \int_{-\infty}^t D_i^k s\, E[S_k(s) I(x_i(t) = m) \mid S_j(t')], \\
&= \int_{-\infty}^t D_i^k s\, \left[p_{ikj}(t \mid s, t')\right]_m v_{kj}(s \mid t'),
\end{aligned}
$$

where $\left[p_{ik}(t \mid s)\right]_m = E[I(x_i(t) = m) \mid S_k(s)]$ and $\left[p_{ikj}(t \mid s, t')\right]_m = E[I(x_i(t) = m) \mid S_k(s), S_j(t')]$. Similarly, we find

$$
\begin{aligned}
\left[B_i^{kl}(t)\right]_m &= \mathrm{Cov}\left[h_i^k(t), h_i^l(t) \mid x_i(t) = m\right] \cdot \left[p_i(t)\right]_m \\
&= \left\{ E\left[h_i^k(t) h_i^l(t) \mid x_i(t) = m\right] - E\left[h_i^k(t) \mid x_i(t) = m\right] \right. \\
&\quad \left. \times E\left[h_i^l(t) \mid x_i(t) = m\right]\right\} \cdot \left[p_i(t)\right]_m \\
&= E\left[h_i^k(t) h_i^l(t) I(x_i(t) = m)\right] - \frac{\left[\beta_i^k(t)\right]_m \left[\beta_i^l(t')\right]_m}{\left[p_i(t)\right]_m}
\end{aligned}
$$

$$= \int_{-\infty}^{t} D_i^k s \int_{-\infty}^{t} D_i^l s' \; \{E[S_k(s)S_l(s')I(x_i(t) = m)]$$

$$- \frac{[p_{ik}(t \mid s)]_m [p_{il}(t \mid s')]_m}{[p_i(t)]_m} v_k(s) v_l(s') \}$$

$$= \int_{-\infty}^{t} D_i^k s \int_{-\infty}^{t} D_i^l s' \; \{[p_{ikl}(t \mid s, s')]_m \phi_{kl}(s, s')$$

$$- \frac{[p_{ik}(t \mid s)]_m [p_{il}(t \mid s')]_m}{[p_i(t)]_m} v_k(s) v_l(s') \}. \tag{A.2}$$

A.2 Approximation of $\beta_i^k(t)$, $\beta_{ij}^k(t \mid t')$ and $B_i^{kl}(t)$. In order to evaluate the finite size effect—order $1/N_c$ terms of the mean firing intensity and cross-covariance functions—we need to evaluate $\mu_i(t) p_i(t) = \sum_k J_i^k \beta_i^k(t)$, $\mu_{ij}(t \mid t') p_{ij}(t \mid t') = \sum_k J_i^k \beta_{ij}^k(t \mid t')$, and $\sigma_i^2(t) p_i(t) = \sum_{k,l} J_i^k J_i^l B_i^{kl}(t)$ to the first order of $1/N_c$. Assuming that the synaptic strengths scale as $J \sim 1/N_c$, we need to evaluate

$$\beta_i^k(t) = \int_{-\infty}^{t} D_i^k s \; p_{ik}(t \mid s) v_k(s)$$

to the first order of $1/N_c$,

$$\beta_{ij}^k(t \mid t') = \int_{-\infty}^{t} D_i^k s \; p_{ikj}(t \mid s, t') v_{kj}(s \mid t')$$

to the zeroth order of $1/N_c$ if $k = i$ or $k = j$ because $J \sim 1/N_c$ but to the first order of $1/N_c$ otherwise (in the following, we divide these into five cases and consider each case separately), and

$$[B_i^{kl}(t)]_m = \int_{-\infty}^{t} D_i^k s \int_{-\infty}^{t} D_i^l s' \; \{[p_{ikl}(t \mid s, s')]_m \phi_{kl}(s, s')$$

$$- \frac{[p_{ik}(t \mid s)]_m [p_{il}(t \mid s')]_m}{[p_i(t)]_m} v_k(s) v_l(s') \}$$

to the first order of $1/N_c$ if i, j, k are all different, and to the zeroth order of $1/N_c$ if $k = l$, $k = i$, or $l = i$.

First, direct differentiation of $\beta_i^k(t)$ yields

$$\frac{d}{dt} \beta_i^k(t) = -\frac{\beta_i^k(t)}{\tau_i^k} + p_{ik}(t \mid t^-) v_k(t) + \int_{-\infty}^{t} D_i^k s \; \overline{W}_{ik}(t \mid s) p_{ik}(t \mid s) v_k(s)$$

$$= \left(-\frac{1}{\tau_i^k} + \overline{W}_i(t) \right) \beta_i^k(t) + p_{ik}(t \mid t^-) v_k(t)$$

$$+ \int_{-\infty}^{t} D_i^k s \, \Delta \overline{W}_{ik}(t \mid s) \, p_{ik}(t \mid s) v_k(s)$$

$$= \left(-\frac{1}{\tau_i^k} + \overline{W}_i(t) \right) \boldsymbol{\beta}_i^k(t) + p_{ik}(t \mid t^-) v_k(t)$$

$$+ m \int_{-\infty}^{t} D_i^k s \left(v_{ik}(t \mid s) - \frac{v_i(t)}{p_i(t)} p_{ik}(t \mid s) \right) v_k(s)$$

$$= \left(-\frac{1}{\tau_i^k} + \overline{W}_i(t) \right) \boldsymbol{\beta}_i^k(t) + p_{ik}(t \mid t^-) v_k(t)$$

$$+ m \int_{-\infty}^{t} D_i^k s \left(v_{ki}(s \mid t) - \frac{v_k(s)}{p_i(t)} p_{ik}(t \mid s) \right) v_i(t)$$

$$= \left(-\frac{1}{\tau_i^k} + \overline{W}_i(t) \right) \boldsymbol{\beta}_i^k(t) + p_{ik}(t \mid t^-) v_k(t)$$

$$+ m \left(\alpha_{ii}^k(t^- \mid t) - \frac{\beta_i^k(t)}{p_i(t)} \right) v_i(t), \tag{A.3}$$

where $\boldsymbol{m} = (1, 0, \dots, 0, -1)^T$, and we used in the above calculation, from equations 5.4 and 5.6,

$$\Delta \overline{W}_{ik}(t \mid s) \, p_{ik}(t \mid s) = [\overline{W}_{ik}(t \mid s) - \overline{W}_i(t)] p_{ik}(t \mid s)$$

$$= m \left(\frac{v_{ik}(t \mid s)}{p_{ik}(t \mid s)} - \frac{v_i(t)}{p_i(t)} \right) p_{ik}(t \mid s). \tag{A.4}$$

Next, to evaluate $\boldsymbol{\beta}_{ij}^k(t \mid t')$, we use $\Delta p_{ikl}(t \mid s, s') = [p_{ikl}(t \mid s, s') - p_i(t) - \Delta p_{ik}(t \mid s) - \Delta p_{il}(t \mid s')] \sim 1/N_c^2$ if i, k, l are all different; $\Delta p_{ik}(t \mid s) = p_{ik}(t \mid s) - p_i(t) \sim 1/N_c$ for $i \neq k$. As we discussed at the beginning of this section, in order to evaluate $\mu_{ij}(t \mid t')$ to the first order of $1/N_c$, we have to consider the number of combinations of indexes. For example, there are N_c terms of $\boldsymbol{\beta}_{ii}^i$ and N_c^2 terms of $\boldsymbol{\beta}_{ij}^i$. Hence, we consider the following five possibilities for combining indexes:

- Case 1 $(i \neq j, i \neq k)$; evaluated up to the first order of $1/N_c$:

$$\boldsymbol{\beta}_{ij}^k(t \mid t') \approx \int_{-\infty}^{t} D_i^k s \, [p_i(t) + \Delta p_{ik}(t \mid s) + \Delta p_{ij}(t \mid t')] v_{kj}(s \mid t')$$

$$= \int_{-\infty}^{t} D_i^k s \, [p_i(t) v_{kj}(s \mid t') + \Delta p_{ik}(t \mid s)(v_k(s) + \Delta v_{kj}(s \mid t'))$$

$$+ \Delta p_{ij}(t \mid t')(v_k(s) + \Delta v_{kj}(s \mid t'))]$$

$$\approx \int_{-\infty}^{t} D_i^k s \, [p_i(t) v_{kj}(s \mid t') + \Delta p_{ik}(t \mid s) v_k(s)$$

$$+\Delta p_{ij}(t \mid t')v_k(s)]$$

$$= \int_{-\infty}^{t} D_i^k s \, [\, p_i(t)v_{kj}(s \mid t') + (p_{ik}(t \mid s) - p_i(t))v_k(s)$$

$$+\Delta p_{ij}(t \mid t')v_k(s)]$$

$$= p_i(t)\alpha_{ij}^k(t \mid t') + (\beta_i^k(t) - p_i(t)\alpha_i^k(t)) + \Delta p_{ij}(t \mid t')\alpha_i^k(s)$$

$$= \beta_i^k(t) + p_i(t)\Delta\alpha_{ij}^k(t \mid t') + \Delta p_{ij}(t \mid t')\alpha_i^k(s)$$

$$\approx \beta_i^k(t) + p_{ij}(t \mid t')\alpha_{ij}^k(t \mid t') - p_i(t)\alpha_i^k(s). \tag{A.5}$$

Note that in the second line, we used $\Delta p_{ik} \sim 1/N_c$ and $\Delta v_{ij} \sim 1/N_c$, for example, and neglected higher-order terms such as $\Delta p_{ik}\Delta v_{kj}$.

- Case 2 ($i = j, k \neq i$); evaluated up to the first order of $1/N_c$:

$$\beta_{ii}^k(t \mid t') \approx \int_{-\infty}^{t} D_i^k s \, [\, p_{ii}(t \mid t') + \Delta p_{ik}(t \mid s)]v_{ki}(s \mid t')$$

$$= \int_{-\infty}^{t} D_i^k s \, [\, p_{ii}(t \mid t')v_{ki}(s \mid t') + (p_{ik}(t \mid s) - p_i(t))v_k(s)]$$

$$= p_{ii}(t \mid t')\alpha_{ii}^k(t \mid t') + (\beta_i^k(t) - p_i(t)\alpha_i^k(t))$$

$$= \beta_i^k(t) + p_{ii}(t \mid t')\alpha_{ii}^k(t \mid t') - p_i(t)\alpha_i^k(t). \tag{A.6}$$

- Case 3 ($i = k, i \neq j$); evaluated up to the zeroth order of $1/N_c$:

$$\beta_{ij}^i(t \mid t') \approx \int_{-\infty}^{t} D_i^i s \, p_{ii}(t \mid s)v_i(s)$$

$$= \beta_i^i(t)$$

$$\approx \beta_i^i(t) + p_{ij}(t \mid t')\alpha_{ij}^i(t \mid t') - p_i(t)\alpha_i^i(s). \tag{A.7}$$

- Case 4 ($j = k, i \neq j$); evaluated up to the zeroth order of $1/N_c$:

$$\beta_{ij}^j(t \mid t') \approx \int_{-\infty}^{t} D_i^j s \, [\, p_i(t) + \Delta p_{ij}(t \mid s) + \Delta p_{ij}(t \mid t')]v_{jj}(s \mid t')$$

$$\approx \int_{-\infty}^{t} D_i^j s \, p_i(t)v_{jj}(s \mid t')$$

$$= p_i(t)\alpha_{ij}^j(t \mid t')$$

$$\approx \beta_i^j(t) + p_i(t)\alpha_{ij}^j(t \mid t') - p_i(t)\alpha_i^j(t). \tag{A.8}$$

- Case 5 ($i = j = k$); evaluated up to the first order of $1/N_c$:

$$\beta_{ii}^i(t \mid t') = \int_{-\infty}^{t} D_i^i s \, p_{iii}(t \mid s, t')v_{ii}(s \mid t'). \tag{A.9}$$

This is a difficult situation to analyze precisely. However, unless the neuron i receives extremely strong input, the conditional firing intensity $v_{ii}(s \mid t')$ is almost zero if s and t' are as close as the refractory time constant of the neuron $(M-1)\tau_r$, so the contribution of this interval on the integral is small. If s and t' are separated more than the refractory time constant, the interaction between the spike at s and t' becomes small, that is, $p_{iii}(t \mid s, t') \approx p_i(t) + \Delta p_{ii}(t \mid s) + \Delta p_{ii}(t \mid t')$, and $\Delta v_{ii}(s \mid t') = v_{ii}(s \mid t') - v_i(s)$ becomes small. Hence, to a good approximation,

$$
\begin{aligned}
\boldsymbol{\beta}_{ii}^i(t \mid t') &\approx \int_{-\infty}^t D_i^i s\,[\,\boldsymbol{p}_{ii}(t \mid t') + \Delta \boldsymbol{p}_{ii}(t \mid s)]v_{ii}(s \mid t') \\
&= \int_{-\infty}^t D_i^i s\,[\,\boldsymbol{p}_{ii}(t \mid t')v_{ii}(s \mid t') + \Delta \boldsymbol{p}_{ii}(t \mid s)v_i(s) \\
&\quad + \Delta \boldsymbol{p}_{ii}(t \mid s)\Delta v_{ii}(s \mid t')] \\
&\approx \int_{-\infty}^t D_i^i s\,[\,\boldsymbol{p}_{ii}(t \mid t')v_{ii}(s \mid t') + \Delta \boldsymbol{p}_{ii}(t \mid s)v_i(s)] \\
&= \boldsymbol{\beta}_i^i(t) + \boldsymbol{p}_{ii}(t \mid t')\alpha_{ii}^i(t \mid t') - \boldsymbol{p}_i(t)\alpha_i^i(t).
\end{aligned} \tag{A.10}
$$

The validity of this intuitive approximation should be examined by numerical simulations. Note that $\boldsymbol{\beta}_{ii}^i(t \mid t')$ contributes to the order $1/N_c$ term of the autocovariance function $\Delta\phi_{ii}(t, t')$ but contributes only to order $1/N_c^2$ terms of the mean firing intensity $v_i(t)$ and cross-covariance functions $\Delta\phi_{ij}(t, t')$ with $i \neq j$.

Finally, we can rewrite $\boldsymbol{B}_i^{kl}$ as

$$
\begin{aligned}
[\boldsymbol{B}_i^{kl}(t)]_m &= \int_{-\infty}^t D_i^k s \int_{-\infty}^t D_i^l s'\,\Big\{ [\,\boldsymbol{p}_{ikl}(t \mid s, s')]_m \phi_{kl}(s, s') \\
&\quad - \frac{[\,\boldsymbol{p}_{ik}(t \mid s)]_m [\,\boldsymbol{p}_{il}(t \mid s')]_m}{[\,\boldsymbol{p}_i(t)]_m}\, v_k(s)v_l(s') \Big\} \\
&= \int_{-\infty}^t D_i^k s \int_{-\infty}^t D_i^l s'\,\Big\{ [\,\boldsymbol{p}_{ikl}(t \mid s, s')]_m \Delta\phi_{kl}(s, s') \\
&\quad + \Big([\,\boldsymbol{p}_{ikl}(t \mid s, s')]_m - \frac{[\,\boldsymbol{p}_{ik}(t \mid s)]_m [\,\boldsymbol{p}_{il}(t \mid s')]_m}{[\,\boldsymbol{p}_i(t)]_m} \Big) v_k(s)v_l(s') \Big\} \\
&= \int_{-\infty}^t D_i^k s \int_{-\infty}^t D_i^l s'\,\Big\{ [\,\boldsymbol{p}_{ikl}(t \mid s, s')]_m \Delta\phi_{kl}(s, s') \\
&\quad + \Big([\Delta \boldsymbol{p}_{ikl}(t \mid s, s')]_m - \frac{[\Delta \boldsymbol{p}_{ik}(t \mid s)]_m [\Delta \boldsymbol{p}_{il}(t \mid s')]_m}{[\,\boldsymbol{p}_i(t)]_m} \Big) v_k(s)v_l(s') \Big\},
\end{aligned} \tag{A.11}
$$

where we used $\Delta p_{ik}(t \mid s) = p_{ik}(t \mid s) - p_i(t)$ and $\Delta p_{ikl}(t \mid s, s') = p_{ikl}(t \mid s, s') - p_i(t) - \Delta p_{ik}(t \mid s) - \Delta p_{il}(t \mid s')$. We need to evaluate B_i^{kl} up to the first order of $1/N_c$ if i, k, l are all different and up to the zeroth order of $1/N_c$ if $k = l$, $k = i$, or $l = i$. Up to these orders, B_i^{kl} can be approximated as

$$B_i^{kl}(t) \approx \int_{-\infty}^t D_i^k s \int_{-\infty}^t D_i^l s' \, p_i(t) \Delta\phi_{kl}(s, s')$$

$$= p_i(t) A_i^{kl}(t). \tag{A.12}$$

Note that $\Delta p_{ik} \sim 1/N_c$ for $i \neq k$, $\Delta p_{ikl}(t \mid s, s') \sim 1/N_c^2$ if i, k, l are different, and $\Delta p_{iik}(t \mid s, s') \sim 1/N_c$ for $i \neq k$.

A.3 Calculation of the Linear Filter in the Generalized Linear Model with Markov Refractoriness. We assume that the input is given by

$$I_i(t) = \int K_i(t - s)\xi_i(s)\,ds \tag{A.13}$$

with stimulus $\xi_i = \xi_i^{(0)} + \delta\xi_i(t)$, and calculate the linearized input-output filter for small stimulus fluctuation $\delta\xi_i(t)$ about a constiant baseline $\xi_i^{(0)}$.

In the limit of $N_c \to \infty$, the mean-field equation is given by

$$v_i(t) = f(I_i(t) + \mu_i(t))p_i(t),$$

$$\mu_i(t) = \sum_k J_i^k \frac{\beta_i^k(t)}{p_i(t)},$$

$$\frac{d}{dt}p_i(t) = \overline{W}_i(t)p_i(t),$$

$$\frac{d}{dt}\beta_i^k(t) = \left(-\frac{1}{\tau_i^k} + \overline{W}_i(t)\right)\beta_i^k(t) + p_i(t)v_k(t). \tag{A.14}$$

First, for a constant input $I_i^{(0)}$, we obtain

$$v_i^{(0)} = f_i^{(0)} p_i^{(0)}, \tag{A.15}$$

$$\mu_i^{(0)} = \sum_k J_i^k \frac{\beta_i^{k(0)}}{p_i^{(0)}},$$

$$0 = \overline{W}_i^{(0)} p_i^{(0)},$$

$$\beta_i^{k(0)} = \left[\frac{1}{\tau_i^k} - \overline{W}_i^{(0)}\right]^{-1} p_i^{(0)} v_k^{(0)},$$

where $f_i^{(0)} = f(I_i^{(0)} + \mu_i^{(0)})$ and the second to last equation is easily solvable with respect to $p_i^{(0)}$; we find

$$[p_i^{(0)}]_m = \begin{cases} \dfrac{\tau_r f_i^{(0)}}{1 + (M-1)\tau_r f_i^{(0)}} & (m = 1, \ldots, M-1) \\[4mm] \dfrac{1}{1 + (M-1)\tau_r f_i^{(0)}} & (m = M). \end{cases} \tag{A.16}$$

Next, the linear response to a small perturbation $\delta I_i(t) = \int K_i(t - s)\delta\xi_i(s)\,ds$ is given by

$$\delta v_i(t) = f_i'^{(0)}(\delta I_i(t) + \delta\mu_i(t))p_i^{(0)} + f_i^{(0)}\delta p_i(t),$$

$$\delta\mu_i(t) = \sum_k J_i^k \left[\frac{\delta\beta_i^k(t)}{p_i^{(0)}} - \frac{\beta_i^{k(0)}\delta p_i(t)}{(p_i^{(0)})^2} \right],$$

$$\frac{d}{dt}\delta p_i(t) = \delta\overline{W}_i(t)p_i^{(0)} + \overline{W}_i^{(0)}\delta p_i(t),$$

$$\frac{d}{dt}\delta\beta_i^k(t) = \left(-\frac{1}{\tau_i^k} + \overline{W}_i^{(0)} \right)\delta\beta_i^k(t) + \delta\overline{W}_i(t)\beta_i^{k(0)}$$

$$+\delta p_i(t)v_k^{(0)} + p_i^{(0)}\delta v_k(t), \tag{A.17}$$

where $f_i'^{(0)} = f'(I_i^0 + \mu_i^{(0)})$ and $\delta\overline{W}_i(t) = m e_M^T[\delta v_i(t)/p_i^{(0)} - v_i^{(0)}\delta p_i(t)/(p_i^{(0)})^2]$. Remember that $m = (1, 0, \ldots, 0, -1)^T$ and $e_M = (0, \ldots, 0, 1)^T$. Fourier transformation of the above equations gives

$$\delta\hat{v}_i(\omega) = f_i'^{(0)}(\delta\hat{I}_i(\omega) + \delta\hat{\mu}_i(\omega))p_i^{(0)} + f_i^{(0)}\delta\hat{p}_i(\omega),$$

$$\delta\hat{\mu}_i(\omega) = \sum_k J_i^k \left[\frac{\delta\hat{\beta}_i^k(\omega)}{p_i^{(0)}} - \frac{\beta_i^{k(0)}\delta\hat{p}_i(\omega)}{(p_i^{(0)})^2} \right],$$

$$\delta\hat{p}_i(\omega) = \hat{L}_i(\omega)m\delta\hat{v}_i(\omega),$$

$$\delta\hat{\beta}_i^k(\omega) = \hat{q}_i^k(\omega)\delta\hat{v}_i(\omega) + \hat{R}_i^k(\omega)\delta\hat{v}_k(\omega), \tag{A.18}$$

where $\hat{L}_i(\omega) = [i\omega + m e_M^T v_i^{(0)}/p_i^{(0)} - \overline{W}_i^{(0)}]^{-1}$, $\hat{q}_i^k(\omega) = [i\omega + 1/\tau_i^k - \overline{W}_i^{(0)}]^{-1}$ $[m\beta_i^{k(0)}/p_i^{(0)} - m e_M^T\hat{L}_i(\omega)m\beta_i^{k(0)}v_i^{(0)}/(p_i^{(0)})^2 + \hat{L}_i(\omega)mv_k^{(0)}]$, and $\hat{R}_i^k(\omega) = [i\omega + 1/\tau_i^k - \overline{W}_i^{(0)}]^{-1}p_i^{(0)}$. Note that i is the imaginary unit.

The above equation is still very complicated. So it is worth thinking about the special case $J_i^k = 0$. In this case, the linear response is described

by

$$\delta \hat{v}_i(\omega) = \frac{f_i'^{(0)} p_i^{(0)}}{1 - f_i^{(0)} e_M^T \hat{L}_i(\omega) m} \delta \hat{I}_i(\omega). \tag{A.19}$$

The calculation of $\hat{L}_i(\omega)$ is straightforward thanks to the bidiagonal nature of $[i\omega + m e_M^T v_i^{(0)}/p_i^{(0)} - \overline{W}_i^{(0)}]$. After direct matrix inversion, we find

$$\hat{L}_i(\omega) = \tau_r \rho \begin{pmatrix} 1 & 0 & 0 & \cdots & & & 0 \\ \rho & 1 & 0 & & & & \vdots \\ \rho^2 & \rho & 1 & & & & \\ \vdots & & & \ddots & & & \\ \rho^{M-2} & \cdots & & & \rho & 1 & 0 \\ \rho^{M-1}\eta & \cdots & & & \rho^2\eta & \rho\eta & \eta \end{pmatrix}, \tag{A.20}$$

with $\rho = 1/(1 + i\omega\tau_r)$ and $\eta = 1 + 1/(i\omega\tau_r)$. Hence, the linear response is simplified as

$$\delta \hat{v}_i(\omega) = \frac{f_i'^{(0)} p_i^{(0)}}{1 - f_i^{(0)} \tau_r \rho \eta (\rho^{M-1} - 1)} \delta \hat{I}_i(\omega). \tag{A.21}$$

In particular, if $M = 2$, we find that the Fourier inverse transformation of the gain function $\hat{G}(\omega) = \delta \hat{v}_i(\omega)/\delta \hat{I}_i(\omega) = \frac{f_i'^{(0)} p_i^{(0)}}{1 + f_i^{(0)} \tau_r/(1+i\omega\tau_r)}$ has an analytical form, that is, using $p_i^{(0)} = 1/(1 + \tau_r f_i^{(0)})$, we find

$$G(t) = \frac{f_i'^{(0)}}{1 + \tau_r f_i^{(0)}} \left[\delta(t) - f_i^{(0)} e^{-(f_i^{(0)} + 1/\tau_r)t} \Theta(t) \right]. \tag{A.22}$$

Note that in the limit of $\tau_r \to 0$, the trivial gain function is $G(t) \to f_i'^{(0)} \delta(t)$. Hence, we can see that the suppressive kernel is added to this instantaneous gain function due to the Markov refractoriness. This linear response for $J_i^k = 0$ case is a special case of a more general linear response for a renewal neuron (Gerstner & Kistler, 2002). Note that in equation A.18, we discussed a more general case and considered the interaction between the refractory effect and spike interaction effect of a network of recurrently connected neurons.

Acknowledgments

References

Amit, D. J., & Tsodyks, M. V. (1991). Quantitative study of attractor neural networks retrieving at low spike rates. I: Substrate—spikes, rates, and neuronal gain. *Network*, 2, 259–273.

Berry, M., & Meister, M. (1998). Refractoriness and neural precision. *J. Neurosci.*, *18*, 2200–2211.

Brillinger, D. (1988). Maximum likelihood analysis of spike trains of interacting nerve cells. *Biological Cyberkinetics*, *59*, 189–200.

Brown, E., Kass, R., & Mitra, P. (2004). Multiple neural spike train data analysis: State-of-the-art and future challenges. *Nature Neuroscience*, 7, 456–461.

Brunel, N., & Hakim, V. (1999). Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Comput.*, *11*, 1621–1671.

Bryant, H., & Segundo, J. (1976). Spike initiation by transmembrane current: A white-noise analysis. *Journal of Physiology*, *260*, 279–314.

Chizhov, A. V., & Graham, L. J. (2007). Population model of hippocampal pyramidal neurons, linking a refractory density approach to conductance-based neurons. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, *75*, 011924.

Chornoboy, E., Schramm, L., & Karr, A. (1988). Maximum likelihood identification of neural point process systems. *Biological Cybernetics*, *59*, 265–275.

Derrida, B., Gardner, E., & Zippelius, A. (1987). An exactly solvable asymmetric neural network model. *Europhys. Lett.*, *4*, 167–173.

Doiron, B., Lindner, B., Longtin, A., Maler, L., & Bastian, J. (2004). Oscillatory activity in electrosensory neurons increases with the spatial correlation of the stochastic input stimulus. *Phys. Rev. Lett.*, *93*(4), 048101.

Escola, S., & Paninski, L. (2007). *Hidden Markov models applied toward the inference of neural states and the improved estimation of linear receptive fields*. Manuscript submitted for publication.

Fourcaud-Trocme, N., & Brunel, N. (2005). Dynamics of the instantaneous firing rate in response to changes in input statistics. *J. Comput. Neurosci.*, *18*, 311–321.

Fourcaud-Trocme, N., Hansel, D., van Vreeswijk, C., & Brunel, N. (2003). How spike generation mechanisms determine the neuronal response to fluctuating inputs. *J. Neurosci.*, *23*, 11628–11640.

Gerstner, W., & Kistler, W. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge: Cambridge University Press.

Ginzburg, I., & Sompolinsky, H. (1994). Theory of correlations in stochastic neural networks. *Phys. Rev. E*, *50*(4), 3171–3191.

Harris, K., Csicsvari, J., Hirase, H., Dragoi, G., & Buzsaki, G. (2003). Organization of cell assemblies in the hippocampus. *Nature*, *424*, 552–556.

Hatsopoulos, N., Ojakangas, C., Paninski, L., & Donoghue, J. (1998). Information about movement direction obtained by synchronous activity of motor cortical neurons. *PNAS*, *95*, 15706–15711.

Hertz, J., Krogh, A., & Palmer, R. G. (1991). *Introduction to the theory of neural computation*. Redwood City, CA: Addison-Wesley.

Kappen, H. J., & Spanjers, J. J. (2000). Mean field theory for asymmetric neural networks. *Phys. Rev. E, 61*(5 Pt B), 5658–5663.

Kass, R., & Ventura, V. (2001). A spike-train probability model. *Neural Comput., 13*, 1713–1720.

Keat, J., Reinagel, P., Reid, R., & Meister, M. (2001). Predicting every spike: A model for the responses of visual neurons. *Neuron, 30*, 803–817.

Knight, B., Omurtag, A., & Sirovich, L. (2000). The approach of a neuron population firing rate to a new equilibrium: An exact theoretical result. *Neural Comput., 12*, 1045–1055.

Kohn, A., & Smith, M. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *Journal of Neuroscience, 25*, 3661–3673.

Kulkarni, J., & Paninski, L. (2007). Common-input models for multiple neural spike-train data. *Network: Computation in Neural Systems, 18*, 375–407.

Lindner, B., Doiron, B., & Longtin, A. (2005). Theory of oscillatory firing induced by spatially correlated noise and delayed inhibitory feedback. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys., 72*(6 Pt 1), 061919.

Marmarelis, P., & Marmarelis, V. (1978). *Analysis of physiological systems: The white-noise approach*. New York: Plenum Press.

Martignon, L., Deco, G., Laskey, K., Diamond, M., Freiwald, W., & Vaadia, E. (2000). Neural coding: Higher-order temporal patterns in the neurostatistics of cell assemblies. *Neural Comput., 12*, 2621–2653.

Mattia, M., & Del Giudice, P. (2002). Population dynamics of interacting spiking neurons. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys., 66*, 051917.

Meyer, C., & van Vreeswijk, C. (2002). Temporal correlations in stochastic networks of spiking neurons. *Neural Comput., 14*, 369–404.

Moreno-Bote, R., & Parga, N. (2004). Role of synaptic filtering on the firing response of simple model neurons. *Phys. Rev. Lett., 92*, 028102.

Moreno-Bote, R., & Parga, N. (2006). Auto- and cross-correlograms for the spike response of leaky integrate-and-fire neurons with slow synapses. *Phys. Rev. Lett., 96*, 028101.

Nirenberg, S., Carcieri, S., Jacobs, A., & Latham, P. (2002). Retinal ganglion cells act largely as independent encoders. *Nature, 411*, 698–701.

Nykamp, D. (2007). A mathematical framework for inferring connectivity in probabilistic neuronal networks. *Math. Biosci., 205*, 204–251.

Nykamp, D., & Tranchina, D. (2000). A population density approach that facilitates large-scale modeling of neural networks: Analysis and an application to orientation tuning. *Journal of Computational Neuroscience, 8*, 19–50.

Okatan, M., Wilson, M., & Brown, E. (2005). Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. *Neural Comput., 17*, 1927–1961.

Oksendal, B. (2002). *Stochastic differential equations*. Berlin: Springer.

Oram, M., Hatsopoulos, N., Richmond, B., & Donoghue, J. (2001). Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. *Journal of Neurophysiology, 86*, 1700–1716.

Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Computation in Neural Systems*, *15*, 243–262.

Paninski, L., Fellows, M., Shoham, S., Hatsopoulos, N., & Donoghue, J. (2004). Superlinear population encoding of dynamic hand trajectory in primary motor cortex. *J. Neurosci.*, *24*, 8551–8561.

Paninski, L., Pillow, J., & Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design. In P. Cisek, T. Drew, & J. Kalaska (Eds.), *Computational neuroscience: Progress in brain research*. Dordrecht: Elsevier.

Pillow, J., Paninski, L., Uzzell, V., Simoncelli, E., & Chichilnisky, E. (2005). Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *Journal of Neuroscience*, *25*, 11003–11013.

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., & Chichilnisky, et al. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature, 454*, 995–999.

Renart, A., Brunel, N., & Wang, X.-J. (2003). *Computational neuroscience: A comprehensive approach*. Boca Raton, FL: CRC Press.

Riehle, A., Grün, S., Diesmann, M., & Aertsen, A. (1997). Spike synchronization and rate modulation differentially involved in motor cortical function. *Science*, *278*, 1950–1953.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Rigat, F., de Gunst, M., & van Pelt, J. (2006). Bayesian modelling and analysis of spatio-temporal neuronal networks. *Bayesian Analysis*, *1*, 733–764.

Schnitzer, M., & Meister, M. (2003). Multineuronal firing patterns in the signal from eye to brain. *Neuron*, *37*, 499–511.

Shriki, O., Hansel, D., & Sompolinsky, H. (2003). Rate models for conductance-based cortical neuronal networks. *Neural Comput.*, *15*, 1809–1841.

Snyder, D., & Miller, M. (1991). *Random point processes in time and space*. Berlin: Springer-Verlag.

Sompolinksy, H., Crisanti, A., & Sommers, H. (1988). Chaos in random neural networks. *Physical Review Letters*, *61*, 259–262.

Toyoizumi, T., Aihara, K., & Amari, S. (2006). Fisher information for spike-based population decoding. *Phys. Rev. Lett.*, *97*, 098102.

Truccolo, W., Eden, U., Fellows, M., Donoghue, J., & Brown, E. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble and extrinsic covariate effects. *Journal of Neurophysiology*, *93*, 1074–1089.

van Vreeswijk, C., & Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, *274*, 1724–1726.

van Vreeswijk, C., & Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Comput.*, *10*, 1321–1371.

Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, *12*(1), 1–24.