Statistical Assessment of Time-Varying Dependence Between Two Neurons

Can Cai, Robert E. Kass, and Valérie Ventura

June 24, 2004

Department of Statistics

and

Center for the Neural Basis of Cognition
Carnegie Mellon University

Running Head: Smoothed time-varying synchrony

ABSTRACT

The joint peristimulus time histogram (JPSTH) provides a visual representation of the dynamics of correlated activity for a pair of neurons. There are many ways to adjust the JPSTH for the time-varying firing-rate modulation of each neuron, and then to define a suitable measure of time-varying correlated activity. Our approach is to introduce a statistical model for the time-varying joint spiking activity so that the joint firing rate can be estimated more efficiently. We have applied an adaptive smoothing method, which has been shown to be effective in capturing sudden changes in firing rate, to the ratio of joint firing probability to the probability of firing predicted by independence. A Bootstrap procedure, applicable to both Poisson and non-Poisson data, was used to define a statistical significance test of whether a large ratio could be due to chance alone. A numerical simulation showed that the Bootstrap-based significance test has very nearly the correct rejection probability, and can have markedly better power to detect departures from independence than does an approach based on testing contiguous bins in the JPSTH. In a companion paper (Cai *et al.* 2004b) we show how this formulation can accommodate latency and time-varying excitability effects, which can confound spike timing effects.

Keywords: Bootstrap, Correlation, Cross-correlogram, Excursion significance test, Firing rate, Gamma process, IMI process, Non-Poisson Spiking, Peri-Stimulus Time Histogram, Point Process, Smoothing, Statistical model, Spike Train Analysis

1 Introduction

Time-varying coincident spiking activity for a pair of neurons may be displayed with the joint peri-stimulus time histogram (JPSTH). In order to take account of firing rate variation in the neurons, Aersten *et al.* (1989) proposed a normalized version of the JPSTH and showed how statistical significance of joint spike counts could be assessed under the hypothesis that the two neurons fired independently. Such examination of time trends in neuronal activity necessarily involves estimation of instantaneous firing rates, and the statistical efficiency of this estimation may be greatly increased by smoothing (Kass *et al.* 2003). Furthermore, a statistical test for excessive (or deficient) synchronous spiking, above that predicted by independence, can take advantage of the available smoothed firing-rate functions. In this paper we use a smoothed version of the JPSTH to define an alternative significance test, and report results that show the new test to be substantially more powerful than the earlier test. The method developed here is based on the smoothed ratio of joint spike rates to the product of smoothed individual spike rates. This ratio would be 1, apart from random variation, under independence. The statistical test of excess synchronous activity is based on the magnitude of the excursion, across time, of this smoothed ratio outside certain bounds. When a ratio is either high above its expected value of 1 for a brief period of time, or moderately above for a substantial period of time, the magnitude of the excursion becomes large. We have used the Bootstrap (Davison and Hinkley 1997; Efron and Tibshirani 1993) both to define the excursion boundaries and to compute statistical significance. See Ventura (2004) for an overview of Bootstrap testing and model selection in the analysis of spike train data.

The excursion test based on efficient smoothing procedures was developed with the goal of being useful for small or moderate numbers of trials. In related work, Pipa and Grün (2003) show how to use both a permutation test and a Bootstrap test to assess the significance of synchrony using the unitary event coincidence count as the test statistic. Their approach, however, considers only the total coincidence count across the trial interval and

does not attempt to assess time-varying excess firing. The method here provides temporal information in the spirit of the normalized JPSTH, and may be extended to adjust for excess trial-to-trial variability (Cai *at al.* 2004b). To smooth the joint spike rates and the individual spike rates, we have applied a recently-developed method, called BARS, that has been shown to perform very well in similar smoothing applications (DiMatteo *et al.* 2001; Kass and Wallstrom 2002; Zhang *et al.* 2003).

Suppose discharges from two neurons are recorded simultaneously over some time interval across R repeated trials, and the interval is decomposed into small subintervals of length Δt (such as 1 ms) at times $t = 1, 2, \ldots, t_{max}$. Letting $Y^i(u), i = 1, 2$ be the number of trials during which neuron i fires at time u, $Y^i(u)/R$ estimates $P^i(u)$, the probabilities that neuron i fires at time u. Similarly, letting $Y^{12}(u, v)$ be the number of trials during which neuron 1 fires at time u and neuron 2 fires at time v, for all combinations of u and v, $Y^{12}(u, v)/R$ is an estimate of $P^{12}(u, v)$, the probability that both neuron 1 fires at time u and neuron 2 fires at time v. (More precisely, $P^{12}(u, v)$ is the probability that neurons 1 and 2 fire in the intervals $(u - \frac{\Delta u}{2}, u + \frac{\Delta u}{2})$ and $(v - \frac{\Delta v}{2}, v + \frac{\Delta v}{2})$.) The JPSTH is the two-dimensional display of values $Y^{12}(u, v)/R$. The quantities $Y^1(u)/R$ and $Y^2(v)/R$ are typically used to adjust the JPSTH for variations in firing rate.

The goal is to understand the time-varying relatedness of the firing of the two neurons. In particular, if the two neurons were to fire independently we would have

$$P^{12}(u, v) = P^1(u) \cdot P^2(v). \tag{1}$$

The problems are (i) to measure departure from independence using probability, (ii) to define a corresponding statistical assessment in terms of the data $Y^{12}(u, v)$ rather than the probabilities in (1), and (iii) to calibrate the assessment and understand how well it works.

In Section 5.2 will consider a general point processes formulation (Barbieri *et al.* 2001; Daley and Vere-Jones 2003; Johnson 1996) of our methodology but intially we will assume

that the two individual neuron spike trains and the joint spike trains follow inhomogeneous Poisson processes. These processes are determined by their firing-rate intensity functions, which we write as $\lambda^1(t)$, $\lambda^2(t)$, and $\lambda^{12}(u,v)$, the latter defining the joint firing rate for neuron 1 at time u and and time v for neuron 2. We thus have $P^1(u) = \lambda^1(u)\Delta u$, $P^2(v) = \lambda^2(v)\Delta v$ and, similarly, $P^{12}(u,v) = \lambda^{12}(u,v)\Delta u \Delta v$. To capture potential departures from independence in (1) we define $\zeta(u,v)$ by

$$\lambda^{12}(u,v) = \lambda^1(u) \cdot \lambda^2(v) \cdot \zeta(u,v). \tag{2}$$

The quantity $\zeta(u,v) - 1$ may be interpreted as the excess proportion, above what is predicted by independence, in the probability that both neuron 1 and neuron 2 will fire at respective times u and v. Ito and Tsuji (2000) observed that there is no uniquely compelling normalization for the JPSTH, and that each normalization effectively models the occurrence of excess joint spiking activity, beyond what would be predicted under independence. We use the probability model (2) as a statistically convenient specification of excess spiking activity.

In estimating $\zeta(u,v)$ we focus on the diagonals $\zeta(u, u+\delta)$, each of which represents the time-varying tendency for neuron 2 to fire more frequently than predicted by independence at a time lag of δ following the firing of neuron 1. Let us write the diagonal at lag δ as a function ζ_δ, i.e., $\zeta_\delta(t) = \zeta(t, t+\delta)$. Rewriting Equation (2) as a ratio we have

$$\zeta_\delta(t) = \frac{\lambda^{12}(t, t+\delta)}{\lambda^1(t)\lambda^2(t+\delta)}. \tag{3}$$

This is the way we solve problem (i) outlined above: the function $\zeta_\delta(t)$ measures time-varying departure from independence. Such departures may vary rapidly, with onset of excess joint firing ramping up over several milliseconds, but it is reasonable to assume that $\zeta_\delta(t)$ is a continuous function of t. This assumption motivates two features of the statistical methodology presented here. First, smooth estimates $\hat{\zeta}_\delta(t)$ of $\zeta_\delta(t)$ are used. Second, statistical significance is judged from the magnitude of contiguous excursions of $\hat{\zeta}_\delta(t)$ above (or below) $\zeta_\delta(t) = 1$.

We next consider problem (ii). As we explain in detail in Sections 2–4, we begin by smoothing across time the histograms for estimating each of f_δ, λ^1, and λ^2, where $f_\delta(t) = \lambda^{12}(t, t+\delta)$, and apply (3) to obtain a smooth estimate $\hat{\zeta}_\delta$ of ζ_δ. We then must assess whether $\hat{\zeta}_\delta(t)$ deviates from $\zeta_\delta(t) = 1$ more than would be found by chance, under independence. To do this we compute 95% probability boundaries for $\hat{\zeta}_\delta$ based on the assumption of independence. This computation uses a set of Bootstrap samples. We evaluate the magnitude of the excursion of $\hat{\zeta}_\delta$ outside these boundaries and compute a p-value for the excursion, thereby testing the hypothesis of independence, by re-using the same Bootstrap samples. This provides our solution to problem (ii). We discuss in Section 2 the smoothing method, and in Section 3 the bootstrap calculations.

Problem (iii) involves evaluation of the statistical procedure. We present in Section 4 results from simulation studies under the assumption of Poisson spiking. In Section 5 we show how the procedure may be extended to non-Poisson spiking, which is important in many settings, and we report additional simulation studies of the non-Poisson case.

2 The Statistical Efficiency of Smoothing

The function $\zeta_\delta(t)$ was introduced in Equation (3) as a measure of time-varying departure from independence. Figure 1 displays the result of a simulation of two correlated Poisson neurons in which the correlated activity occurs at lag $\delta = 0$, i.e., synchronously. The algorithm used to generate the data is described in the Appendix. Figure 1 also shows the estimate $\hat{\zeta}_0(t)$ of $\zeta_0(t)$ obtained by smoothing the sequences $Y^1(t)/R$, $Y^2(t)/R$ and $Y_0^{12}(t)/R$, where, for example, $Y_0^{12}(t)$ is the mean number of times neuron 1 fires at time t and neuron 2 fires at time $t = t + \delta$, with $\delta = 0$, i.e.,

$$Y_0^{12}(t) = \sum_r X_r^{12}(t, t)$$

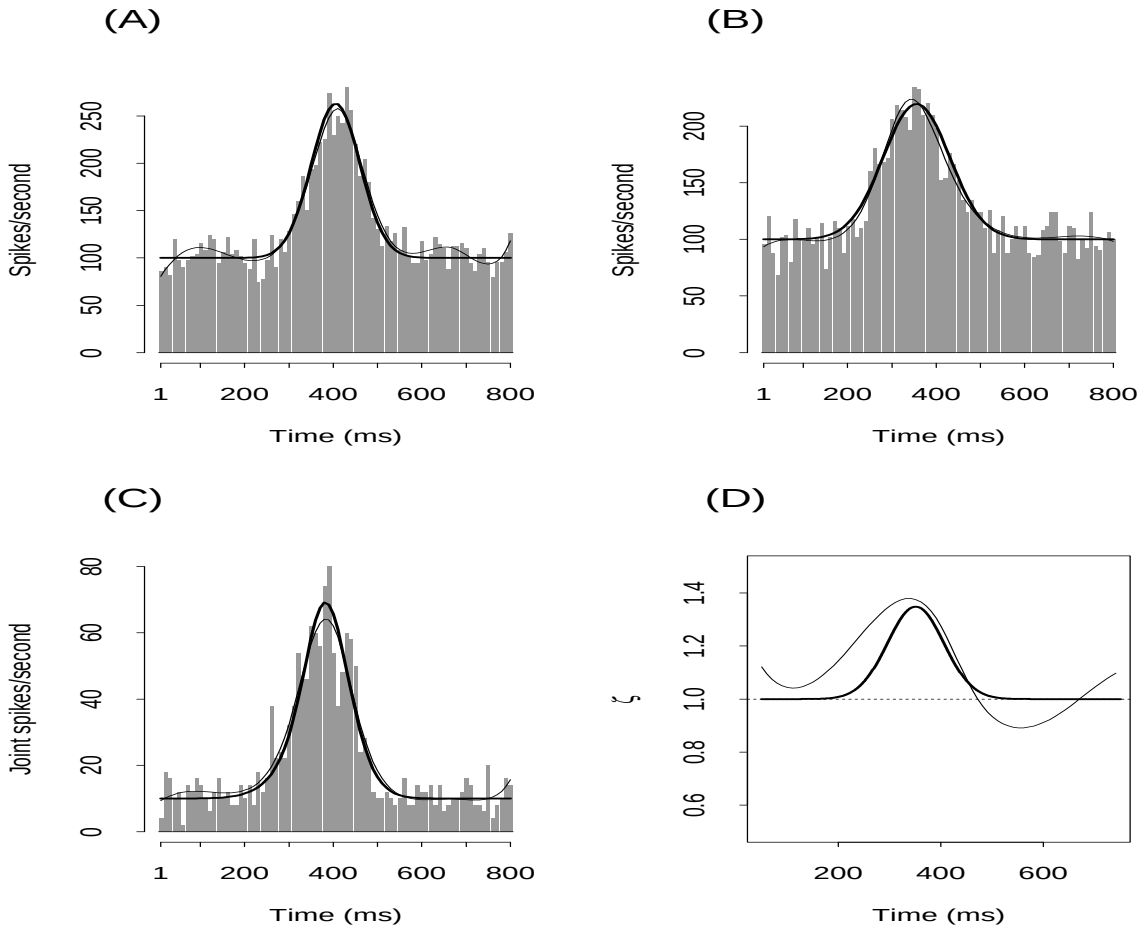

Figure 1: *Simulated data. (A) and (B): PSTH based on 200 trials, together with true firing rates* $\lambda^1(t)$ *and* $\lambda^2(t)$ *overlaid in bold. (C): Main diagonal of the JPSTH, with the true joint firing rate taken from the right-hand side of Equation* (2) *with* $u = v$ *overlaid in bold. (D): The true function* $\zeta_0(t)$ *defined in Equation* (3)*. In each panel, the thick line is the truth and the thinner line is its corresponding estimate.*

where $X_r^{12}(t,t)$ is 1 if both neurons fire at time t on trial r and is 0 otherwise. To do the smoothing we prefer a spline-based method called BARS (DiMatteo *at al.* 2001) because it produces relatively good statistical estimates in many contexts, and works particularly well when a firing-rate function varies rapidly in some part of the time domain. However, this particular smoothing method is not essential to the methodology presented here: any other smoothing method could be used, such as Gaussian filtering (see Kass *at al.* 2003).

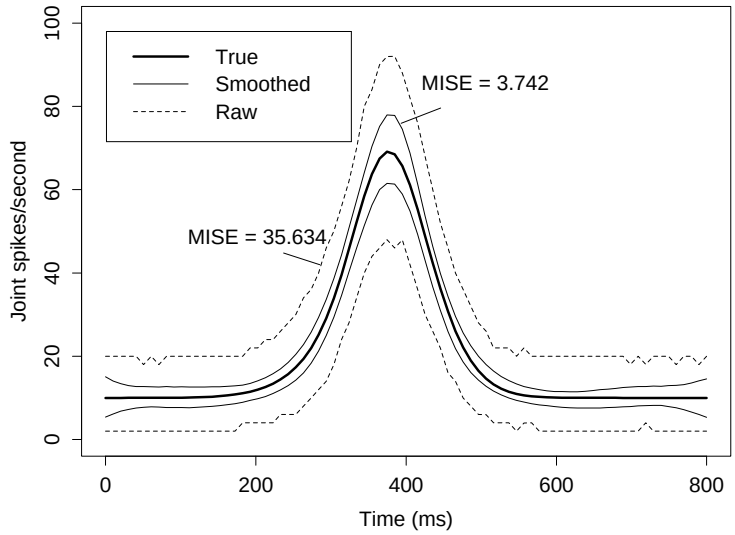

Figure 2: *The 95% simulation bands for diagonal with and without smoothing, based on 1000 simulated neuron pairs, each having 50 trials. Bands are shown for the raw JPSTH diagonals (without smoothing) and for the smoothed version, using BARS. The bands for BARS are much closer to the true diagonal, indicating much higher accuracy after smoothing.*

The advantage of introducing $\zeta_\delta(t)$ is that we can estimate it relatively efficiently. That is, the mean integrated squared error (MISE) of the estimator $\hat{\zeta}_\delta(t)$,

$$MISE = \int (\hat{\zeta}_\delta(t) - \zeta_\delta(t))^2 dt$$

will be smaller than the MISE of the corresponding diagonal of the JPSTH, which in turn will provide increased power to detect synchrony. The gain in efficiency obtained

from smoothing is illustrated in Figure 2, which shows the true diagonal of the JPSTH we used in Figure 1, along with 95% simulation bands obtained from 1000 simulations from our model for the raw diagonal of the JPSTH and for $\hat{\zeta}_0(t)$. The MISE values are also indicated on the figure, and suggest that for the diagonal of the JPSTH to have the same accuracy as $\hat{\zeta}_0(t)$, one would need to collect approximately 10 times as much data. See also Kass *at al.* (2003) for additional examples. The increase in power of subsequent statistical procedures is studied in Section 4.

3 Bootstrap Significance Test

Under the null hypothesis of independence $\hat{\zeta}_\delta(t)$ will, because of random fluctuations, differ from $\zeta_\delta(t) = 1$. Real departures of $\zeta_\delta(t)$ from 1, the kind we wish to detect, will be statistically substantial and will be sustained over some interval of time. We therefore measure the deviation of $\hat{\zeta}_\delta(t)$ from 1 by assessing the magnitude of its excursion beyond 95% probability boundaries for $\hat{\zeta}_\delta$ based on the assumption of independence. Note that some excursions beyond the boundaries remain likely to occur due to chance alone, under the null hypothesis. We will evaluate the probability of large excursions.

The computation begins with a set of Bootstrap samples, which are first used to compute the upper and lower 95% probability boundaries and are subsequently re-used to compute the probability of large excursions. The upper boundary $h_U(t)$, at time t, is defined to be the value such that the probability of $\hat{\zeta}_\delta(t) > h_U(t)$, under independence, is .025; similarly, the lower boundary $h_L(t)$, at time t, is defined to be the value such that the probability of $h_L(t) < \hat{\zeta}_\delta(t)$, under independence, is .025. We refer to these boundaries as pointwise null bands. This is explained in Section 3.1. Section 3.2 defines a test statistic based on the null bands, and then specifies how the Bootstrap sample drawn to compute the bands is re-used to compute the p-value of the test statistic. Because this procedure is computationally intensive, in Section 3.3 we also discuss the use of Normal approximations

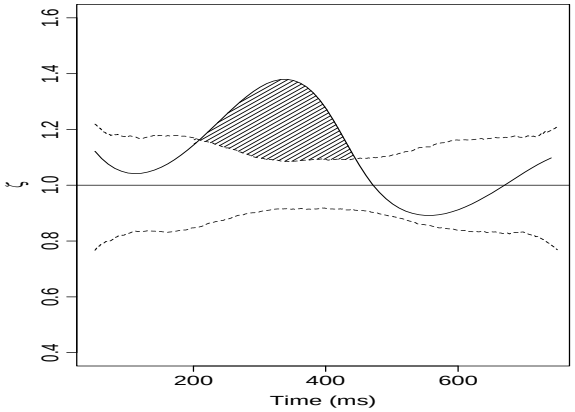

Figure 3: *Estimated $\hat{\zeta}_\delta(t)$ (dark curve) with 95% bootstrap confidence bands. The bands $h_{L=.025}$ and $h_{U=.975}$ are shown. The excursion $\hat{G}_{obs}$ of $\hat{\zeta}(t)$ beyond the bands is the area of the shaded region.*

in creating the null bands, based on a smaller Bootstrap sample.

3.1 Pointwise Null Bands for $\hat{\zeta}_\delta(t)$

Assuming the data set consists of R trials in which spiking events for a pair of neurons are observed, the following steps may be used to obtain 95% pointwise null bands for $\zeta_\delta(t)$:

1. Obtain smooth estimated firing rate functions $\hat{\lambda}^1(t)$ and $\hat{\lambda}^2(t)$;

2. Simulate R trials of spike trains for the two neurons independently using firing rate functions $\hat{\lambda}^1(t)$ and $\hat{\lambda}^2(t)$;

3. Obtain a smooth estimate $\hat{\zeta}_\delta(t)$ for each δ of interest;

4. Repeat steps 2-3 N times to get N estimates $\hat{\zeta}_\delta(t)$;

5. For each time t, define $h_L(t)$ and $h_U(t)$ to be the .025 and .975 quantiles of the N values $\hat{\zeta}_\delta(t)$ (so that, for example, 2.5% of the sampled $\hat{\zeta}_\delta(t)$ values lie below $h_L(t)$).

Note that Step 2 is an implementation of a *parametric* Bootstrap. An alternative nonparametric Bootstrap procedure is also straightforward: form R joint spike trains by sampling at random with replacement the observed spike trains of neuron 1, and separately of neuron 2. This nonparametric option is appropriate if there are a large number of trials and there is no excess trial-to-trial variability (discussed in Cai *et al.* 2004b).

To determine the bands accurately we used Bootstrap sample sizes of $N = 1000$. Figure 3 shows the 95% Bootstrap pointwise null bands for the estimated $\zeta_0(t)$ of the simulated neuron pairs in Figure 1. Here, in carrying out step 2, we assumed Poisson variation. Although this is a simplification, as we discuss in Section 5, it is not an essential assumption of the method. Of course, $(1-Q)\%$ bands could be used instead of 95% bands by taking $h_L(t)$ and $h_U(t)$ to be the $\frac{Q}{2}$ and $1-\frac{Q}{2}$ quantiles.

3.2 Significance Test for Assessing Time-Varying Synchrony

Focusing on a particular diagonal specified by δ, the null hypothesis is that $\zeta_\delta(t) = 1$ for all t. (The complete null hypothesis of independence is that $\zeta_\delta(t) = 1$ for all t and all δ). A statistical test statistic is a formal ordering of the deviations from the null hypothesis, that is, a way of saying which deviations are greater than others. Based on the assumption that deviations of interest will be those that affect many contiguous values of time, we define the test statistic to be the magnitude of the largest excursion either above the null band $h_U(t)$ or below the null band $h_L(t)$, evaluated as the area between the fitted curve and the null band. This is pictured as the shaded area in Figure 3 for our simulated example. More precisely, we define the observed value of the test statistic

$$G_{obs} = max\left(max_{t_s,t_e}\int_{t_s}^{t_e}\hat{\zeta}_\delta(t) - h_U(t)dt,\ max_{t_s,t_e}\int_{t_s}^{t_e} h_L(t) - \hat{\zeta}_\delta(t)dt\right) \qquad (4)$$

where t_s and t_e are the starting and ending times of the periods during which $\hat{\zeta}_\delta(t)$ is outside the null bands, and the maximum is taken across all t_s and t_e. Then G_{obs} is the largest area of $\hat{\zeta}(t)$ exceeding the band. To calculate the p-value, let $\hat{\zeta}_\delta^{(n)}(t), n = 1, \ldots, N$ stand for the estimate of $\zeta_\delta(t)$ obtained from the nth bootstrap sample. For each Bootstrap sample we compute $G_{boot}^{(n)}$, the value of the statistic defined in (4) with $\hat{\zeta}_\delta^{(n)}(t)$ replacing $\hat{\zeta}_\delta(t)$. We then calculate the p-value for G_{obs}

$$p = \frac{\text{number of Bootstrap samples for which } \ G_{boot}^{(n)} > G_{obs}}{N+1},$$

as is standard for Bootstrap tests. As an illustration, we obtained $p < .0001$ for the shaded area in Figure 3, so we would conclude correctly that there exists synchrony between this pair of neurons at time lag $\delta = 0$.

3.3 Normal Approximations

As a rough guideline we would recommend $N = 1000$ Bootstrap samples to construct the 95% null bands. To reduce computing time, if desirable, a Normal approximation may be used: for large R (large number of trials) one may assume $\hat{\zeta}_\delta(t) \sim \text{Normal}(0, \sigma_z^2(t))$ where $\sigma_z^2(t)$ is the variance of $\hat{\zeta}_\delta(t)$, which can be estimated by the sample variance of the $\hat{\zeta}_\delta(t)$ computed from a small Bootstrap sample (e.g., $N = 50$). In practice, to make Normal approximations more accurate a change of variable (or "transformation") is usually helpful. Our simulations suggest that the square root transformation improves Normality, that is, we may use

$$\sqrt{\hat{\zeta}_\delta(t)} \sim \text{Normal}(0, \sigma^2(t))$$

where, again, $\sigma^2(t)$ is estimated from a small Bootstrap sample. The adequacy of the Normal approximation depends on neuronal firing rates, but we have found it to be quite accurate for this purpose with $R \approx 50$.

4 Properties of the Excursion Test

In this section we evaluate the operating characteristics of the Bootstrap significance test, and compare it to a simpler test based on the normalized JPSTH of Aertsen *et al.* (1989). We consider first, in Section 4.1, whether the Bootstrap test produces nearly the correct p-value, i.e., if $\hat{\alpha}$ is the proportion of newly simulated data samples for which $p < \alpha$, we see whether $\hat{\alpha} \approx \alpha$. Then, in Section 4.2 we consider the power of the test (the probability of rejecting the null when it is truly false) in comparison to that based on the normalized JPSTH.

4.1 Significance Level – Probability of a False Negative

We examine only the case $\delta = 0$, corresponding to synchrony, because we expect performance for other values of δ to be similar: the procedures and data are essentially identical (except that for large $|\delta|$ the data become sparse, as we would be examining data that go into the upper left and lower right corners of the JPSTH, and there the statistical power drops off sharply). We simulated 1000 pairs of independent neurons, i.e. $\zeta_0(t) = 1$ for all t. The firing rates of the neuron pairs are as in the left panel of Figure 4. For each simulated neuron pair, we calculated the p-value specified in Section 3.2. If we reject the null hypothesis based on $\alpha = 0.05$, then the percentage of the p-values less than 0.05 is the empirical type I error $\hat{\alpha}$ of the test. The right panel of Figure 4 displays a plot of $\hat{\alpha}$ versus α, for values of α ranging up to .1. We find, in this case, that the empirical type I error matches the actual α level quite closely.

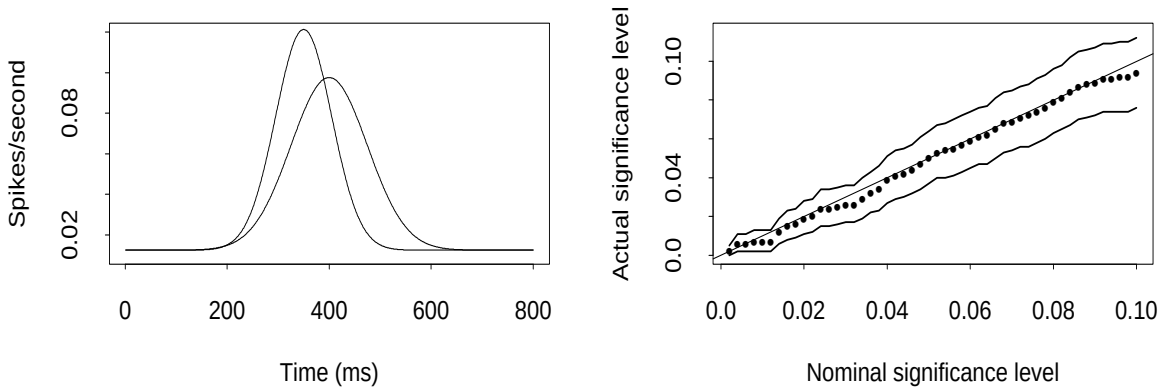

Figure 4: *Type I error versus its nominal value for the excursion test. Having set the nominal type I error to α (shown on the x-axis of the right-hand plot) we compute the proportion among 1000 simulated data samples for which the p-value satisfies $p < \alpha$. The plot on the left panel shows the firing rate functions used for the two simulated neurons. The plot on the right displays extremely good agreement between $\hat{\alpha}$ and the nominal type I error α. The solid line corresponds to perfect agreement, and the irregular bands are 95% confidence intervals for the simulated p values.*

4.2 Power – Probability of a False Positive

To examine the power of significance tests for time-varying dependence we must define time-varying alternatives to the null independence model. In our framework (and again confining attention to $\delta = 0$) this means we must define some set of functions $\zeta_0(t)$ that are not identically equal to 1. We continue to use two neurons with the firing rates displayed in Figure 4A, and set $\zeta(t) = 1 + 4\beta f(t)$, where $f(t)$ is a bell shaped function we took to be the Normal density function with mean 350 msec and standard deviation 55, and β is a gain coefficient taking increasing values 1, 2, 3, 4, 6, 8, and 12, which correspond to a maximum percentage of excess joint firing rate of 2.9, 5.8, 8.7, 11.6, 17.4, 23.2, 34.8 respectively. The $\zeta_0(t)$ functions we used here are proportional to the $\zeta_0(t)$ plotted in

Figures 1(D) and 2.

As a comparison, we used a significance test based on the normalized JPSTH of Aertsen *et al.* (1989), the (u, v) pixel of which is defined to be

$$\rho(u,v) = \frac{P^{12}(u,v) - P^1(u)P^2(v)}{\sqrt{P^1(u)(1-P^1(u))P^2(v)(1-P^2(v))}}. \tag{5}$$

Let us define $\rho_\delta(t) = \rho(t, t+\delta)$ and then take $\hat{\rho}_\delta(t)$ to be the value of $\rho_\delta(t)$ when $P^1(t)$, $P^2(t)$ and $P^{12}(t, t+\delta)$ are replaced with their observed-data counterparts $Y^1(t)/R$, $Y^2(t+\delta)/R$, and $Y^{12}(t, t+\delta)/R$. We consider the case $\delta = 0$. Because large magnitudes $|\hat{\rho}_0(t)|$ for isolated values of t would likely be interpreted as chance fluctuations, we define the test to reject the null hypothesis whenever $|\hat{\rho}_0(t)| > \rho^{**}$ for two contiguous values of t, where ρ^{**} is a threshold value so that the type I error of this procedure is $\alpha = .05$. Note that in order for two alternative statistical tests, here the Bootstrap-based excursion test versus the contiguous-bin JPSTH test, to be comparable with respect to power they must be defined to have the same type I error. The value ρ^{**} is not the same as the .05-level threshold for a single t: it is computed by comparing many alternatives thresholds under the null hypothesis and choosing the value that rejects (falsely) 5% of the time.

Figure 5 displays the results of the power calculation. The Bootstrap test is able to achieve very good power when the maximum percentage of excess joint spiking activity is around 20%. The JPSTH-based method using (5) is very much less powerful. In fact, when the maximum excess firing rate is 17.5%, four times as many trials would be needed using the JPSTH-based method than using the Bootstrap significance test.

5 Non-Poisson Variability

The foregoing material assumed Poisson variation either implicitly, by ignoring spiking history (timing of sequences of spikes) within trials, or explicitly, by generating from

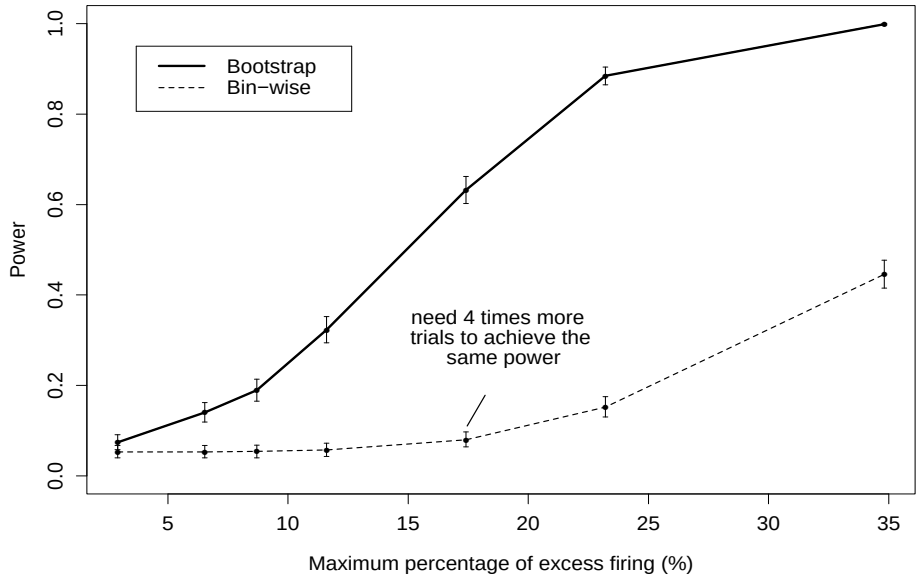

Figure 5: *Power of the Bootstrap-based excursion test compared with the test using two contiguous bins in the JPSTH. Power (probability of rejecting H_0) is plotted against maximum percentage excess firing (max $\zeta_0(t) - 1$). Both tests have the same probability of rejecting H_0 when H_0 holds (type I error) $\alpha = .05$, indicated by the coincidence of the two power graphs when the percentage excess firing is zero.*

Poisson distributions, both within the Bootstrap procedure and in our simulation studies. Two remaining issues are (1) the extent to which the Poisson-based Bootstrap p-value becomes inaccurate in the presence of non-Poisson variability, and (2) the manner in which the excursion test be modified for non-Poisson variability, and its resulting statistical properties. We consider these issues in this section.

5.1 Simulation Study

To examine the effect of departures from Poisson spiking on the excursion test for synchrony we performed further simulation studies. These consisted of applying the test to non-Poisson data and recording the observed significance level $\hat{\alpha}$.

We considered two types of non-Poisson data. The first type were Poisson spike trains with constant rate ranging from 25 to 125 Hz, that were pruned back to enforce a hard refractory period of either 1, 2, 5, or 10 msec. The second type of non-Poisson data were gamma spike trains with parameter q ranging from 0.05 (much more variable than Poisson processes) to 16 (less variable). An easy way to understand a gamma process with q an integer is to simulate one. A Poisson spike train is simulated as follows: divide the time in small intervals of length dt, centered on times t_i, and for all i, generate a spike in the interval centered on t_i with probability $\lambda(t_i)dt$, where $\lambda(t_i)$ denotes the spiking rate at time t_i. A gamma process of order q is defined as the waiting time until the qth event of a Poisson process; therefore to generate such a process with rate $\lambda(t)$, we generate a Poisson process with rate $q \cdot \lambda(t)$, and create the gamma spike train by retaining only every qth spikes. For more general q's (integer or not, above or below 1), we generate gamma processes using the time rescaling theorem, as described in Brown *et al.* (2002).

To further illustrate the properties of gamma processes, note that for a Poisson process with rate λ, the inter-spike intervals (ISI) have an exponential distribution with mean λ^{-1} and variance λ^{-2}. For a gamma(q) process with same rate λ, the distribution distribution

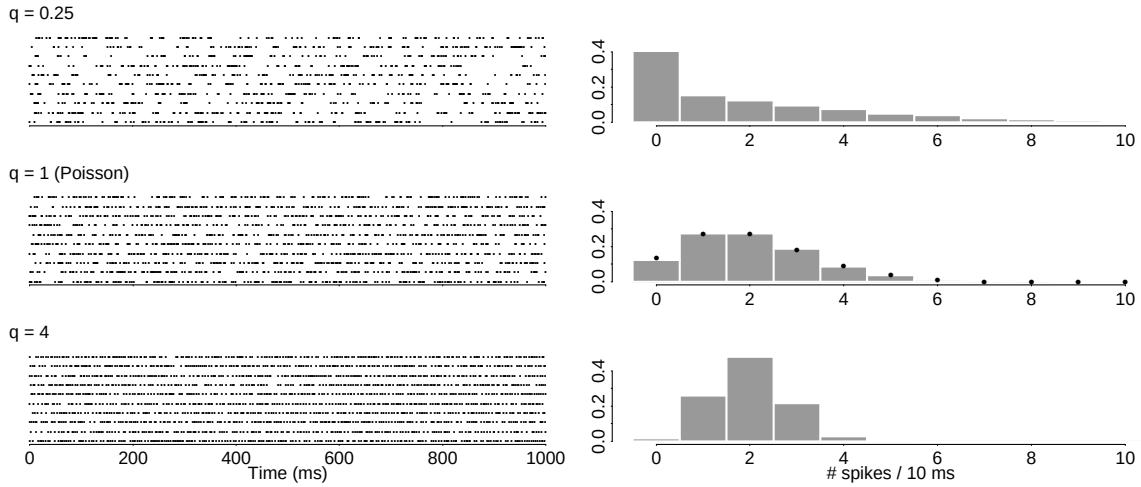

Figure 6: *Properties of gamma processes. Left: raster plots of gamma(q) spike trains with* $q = 0.25, 1$ *and 4, and firing rate 200 Hz. Right: simulated distributions of the number of spikes in 10 msec windows. The black dots in the middle panel are the theoretical distribution of the counts for a Poisson process; simulated and true distributions match.*

of the ISI is $1/q$ times a gamma distribution with mean $q \cdot \lambda^{-1}$ and variance $q \cdot \lambda^{-2}$; hence the ISI have mean λ^{-1}, as do the ISI of the Poisson process, but variance λ^{-2}/q that is smaller/larger if q is larger/small than 1. Therefore the spikes of a gamma process with $q > 1$ ($q < 1$) occur with more (less) regularity than the spikes of a Poisson process with the same firing rate.

Figure 6 illustrates this. It shows raster plots, each based on 10 spike trains generated from gamma processes with the same rate $\lambda = 200$ Hz, and $q = 0.25$, 1 (Poisson), and 4 respectively. The average ISI for all spike trains is 5 msec, so that on average, we would expect to observe 2 spikes in a 10 msec window of time. The right panels of Figure 6 show the distributions of the number of spikes in a 10 msec window; each distribution has a mean of 2, as expected. The larger q becomes, the smaller the ISI variability about the mean 2. On the other hand, when q is small, the variability increases, and the probability of observing 10 spikes in a 10 msec window also increases.

Figure 7 shows the observed significance level $\hat{\alpha}$ from the following simulations. In (A), both neurons were Poisson with the same hard refractory period equal to 0, 1, 5, or 10 msec, and equal firing rates ranging from 25 to 125Hz. For each combination of the firing rate and refractory period we simulated 1000 independent neuron pairs. We applied the joint spiking model to those simulated neuron data, under the Poisson assumption, and calculated the p-value of $\zeta_\delta(t)$ for $\delta = 0$. Note that when the hard refractory period is equal to 0ms, the two spike trains follow Poisson processes and $\hat{\alpha}$ should not and indeed does not differ significantly from α.

From (A) we can see that for firing rates up to 50Hz the type I error remains correct with the Poisson model for non-Poisson spike trains having a hard refractory period as long as 10ms. With a hard refractory period of 5 ms the type I error does not go above .10 until the firing rate exceeds 100Hz. We would conclude that for firing rates below 50Hz the Poisson version of the test will perform well despite the presence of a refractory period, while for higher firing rates the type I error begins to drift upward. For high firing rates (above 100Hz) the test should not be trusted. Note that in Figure 7(A), we took both neurons to have the same refractory period.

Figure 7(B) also shows $\hat{\alpha}$ for Poisson spike trains with hard refractory periods with the following characteristics. Neurons 1 has a constant firing rate of 200 Hz and refractory period of 10 msec, while neuron 2 has rate either 130 or 200 msec, and refractory period ranging from 0 to 10 msec in steps on 1 msec. Figure 7(B), which uses different refractory periods, suggests that the departure of $\hat{\alpha}$ from α is only acute under the conditions just described, when the 2 neurons have similar refractory periods.

Figure 7(C) shows $\hat{\alpha}$ for a variety of gamma point processes. Neuron 1 is gamma(q_1) with rate 200Hz, while neuron 2 is gamma(q_2) with rate 130Hz. We observe that the Poisson procedure applied to gamma spike trains produces approximately the correct observed significance level $\hat{\alpha}$, unless q_1 and q_2 are both either very large, or very small.

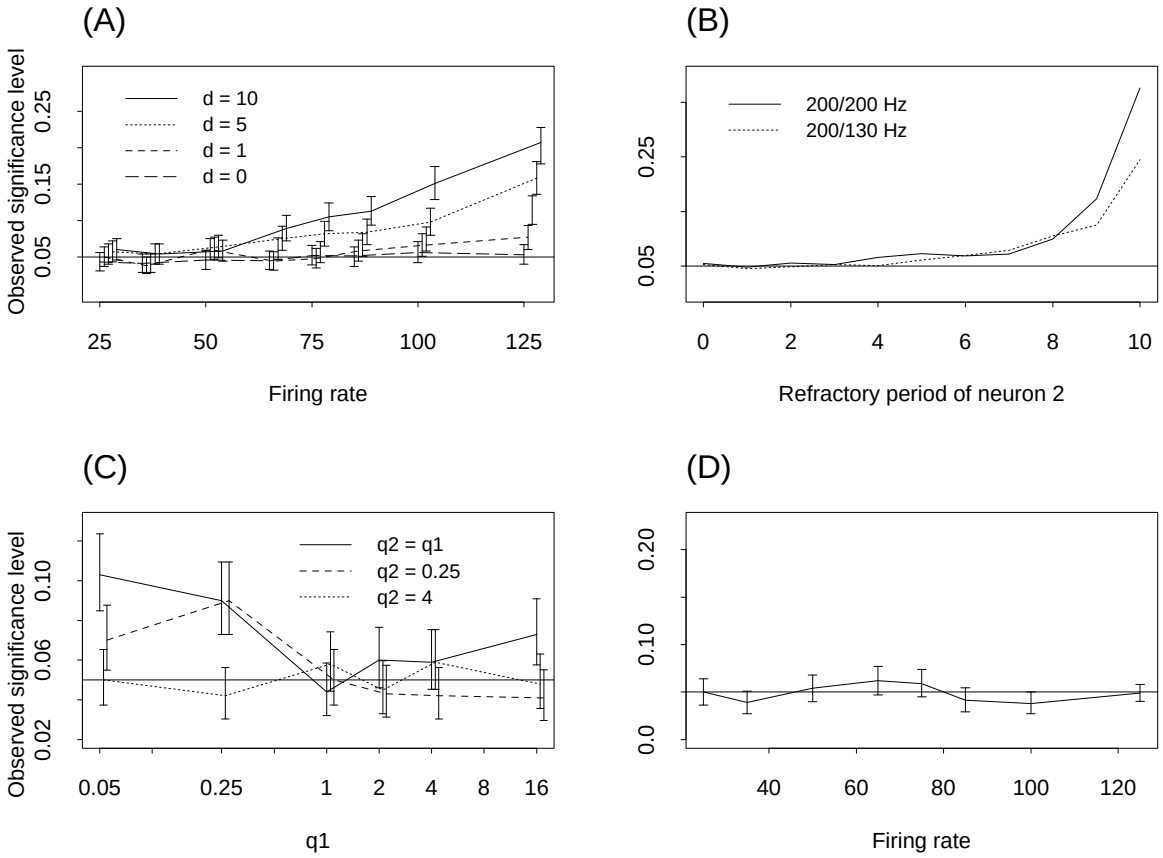

Figure 7: *Type I error of excursion test under the Poisson assumption when (A,B) the data are non-Poisson due to a refractory period, and when (C) the data are gamma. (A) Both neurons have the same firing rate (x-axis) and the same refractory period d. The horizontal line at the bottom is at* $\alpha = .05$*, the type I error the test is supposed to achieve. Vertical bars show the simulation error in the form of 95% confidence intervals for each plotted* $\hat{\alpha}$*. (The vertical bars have been separated horizontally to make them more distinct.) (B) Neuron 1 has firing rate 200Hz and refractory period 10 msec; neuron 2 has firing rate 200Hz or 130Hz, and refractory period on the x-axis. (C) Neurons 1 is gamma*(q_1) *with rate 200Hz; neurons 2 is gamma*(q_2) *with rate 130Hz. The x-axis is on the log scale for clarity. (D) Type I error of excursion test for the Poisson spike trains with hard refractory period of 10 ms in (A), when model (7) is used in the bootstrap simulation. The dotted line is at* $\alpha = .05$*, the type I error the test is supposed to achieve. Vertical bars show the simulation error in the form of 95% confidence intervals for each plotted* $\hat{\alpha}$*.*

5.2 Non-Poisson Models

For a general (not-necessarily Poisson) point process the firing rate is governed by the conditional intensity function $\lambda(t|\mathcal{H})$, where $\mathcal{H}$ is the spiking history (for a given trial) up to time t. Various non-Poisson alternatives may be used, including the inhomogeneous Markov interval (IMI) models discussed by Kass and Ventura (2001) or Gamma process models or other inhomogeneous versions of renewal processes (Barbieri, *et al.*, 2001; Brown *et al.*, 2002). The joint spiking model under the alternative non-Poisson assumption may be fitted just as under the Poisson assumption except the form for the marginal firing rate of each neuron must be changed. Equation (2) is replaced by

$$\lambda^{12}(u,v) = \lambda^1(u|\mathcal{H}^1) \cdot \lambda^2(v|\mathcal{H}^2) \cdot \zeta(u,v) \tag{6}$$

where $\mathcal{H}^i$ for $i = 1, 2$ are the respective spiking histories for the two neurons up to time t. Equation (6) assumes that while the overall joint spiking depends on the spiking history for each neuron through their individual intensity functions, their excess joint spiking above that predicted by independence $\zeta(u,v)$ is not itself mediated by recent spiking activity. The smoothed estimate of $\zeta_\delta(t)$ is obtained, as before, by smoothing the joint spike counts to obtain a smoothed version of the numerator of (3) and then dividing by the similarly estimated product $\lambda^1(t|\mathcal{H}^1)\lambda^2(t+\delta|\mathcal{H}^2)$. In our work we have again used spline-based likelihood methods to estimate the conditional intensity functions.

In the special case of IMI models each neuron's conditional intensity function may be written in the form

$$\lambda(t \mid \mathcal{H}) = \lambda(t, s_*(t))$$

where $s_*(t)$ is the most recent spike time prior to t so that Equation (6) becomes

$$\lambda^{12}(u,v) = \lambda^1(u, s^1_*(u)) \cdot \lambda^2(v, s^2_*(v)) \cdot \zeta(u,v). \tag{7}$$

We repeated the simulation study of the previous subsection, with a hard refractory period of 10 ms, but we instead used (7) to define the Bootstrap excursion test. Figure 7 displays

the results. We can see from the figure that the type I errors are now indistinguishably close to the desired level $\alpha = .05$, even for high firing rates.

6 Discussion

The function $\zeta_\delta(t)$ uses probability to characterize theoretically the evolving dependence in the firing of two neurons. A smooth estimate $\hat{\zeta}_\delta(t)$, like a smoothed diagonal of the JPSTH, describes evolving dependence in the data. Because the data are noisy, a statistical procedure must be used to judge whether an observed deviations of $\hat{\zeta}_\delta(t)$ away from the value of 1, which would hold under independence, is due to chance fluctuations. The magnitude of the excursion of $\hat{\zeta}_\delta(t)$ beyond the 95% null bounds is intended to match the intuition that increases in joint spiking activity *over continuous intervals of time* above that expected under independence should provide evidence of dependence. Creation of an approximately correct p-value for any excursion test of this type could be difficult. With the Bootstrap, however, it is quite straightforward. A substantial literature on the Bootstrap indicates that it has good statistical properties (e.g., Davison and Hinkley, 1997, Sections 2.6 and 5.4, and references therein), and the results presented here indicate that the Bootstrap excursion test performs well in the sense of yielding accurate p-values and having good power for moderate sample sizes.

The particular choice of $\zeta_\delta(t)$ is not essential to the Bootstrap approach applied here, and alternative measures of departure from independence could be used instead. As Ito and Tsuji (2000) observed, there is no uniquely compelling normalization of the JPSTH and each normalization effectively models the occurrence of excess joint spiking activity beyond what would be predicted by independence. Thus, when dependence evolves over time different measures could lead to distinct pictures of the phenomenon. For example, in the presence of time-varying marginal firing rates, excess joint spiking activity that is constant in time when measured in the ratio form of $\zeta_\delta(t)$ would appear non-constant

when measured in an additive form.

Non-Poisson neurons can produce a greater or smaller number of joint spikes than that predicted under Poisson firing. We discussed and reported on applications of the Bootstrap excursion test to non-Poisson spike trains. For mild departures from non-Poisson spiking, and low firing rates, the effects on performance of the Poisson-based test are not large. However, good data-analytical practice would involve checking for non-Poisson behavior by fitting non-Poisson models and, if indicated, applying the non-Poisson versions of the Bootstrap excursion test discussed in Section 5.

The benefit of the statistical approach adopted here is that it used the information in the data efficiently. In addition, the framework allows us to extend the significance test to include effects such as excess trial-to-trial variation, which is discussed in the companion paper (Cai *at al.* 2004b).

Appendix

We here describe how we simulated spike trains for two correlated neurons with marginal rates $\lambda^1(t)$, $\lambda^2(t)$, and joint rate $\lambda^{12}(t) = \lambda^1(t)\lambda^2(t)\zeta_0(t)$. This extends immediately to correlations at other lags. It does not extend to correlations that spread across several lags.

1. Simulate a spike train for neuron A with firing rate $\lambda^1(t)$.

2. Simulate the spike train of neuron B conditional on the spike train of neuron A.

   That is, if at time t, neuron A had a spike, generate a spike for neuron B at time t from a Bernoulli distribution with probability

$$P^2(t \mid \text{neuron A spiked at t}) \;=\; \lambda^2(t \mid \text{neuron A spiked at t})\Delta t$$

$$\begin{aligned} &= \frac{\lambda^{12}(t,t)\Delta t\Delta t}{\lambda^1(t)\Delta t} \\ &= \frac{\lambda^1(t)\Delta t\lambda^2(t)\Delta t\zeta_0(t)}{\lambda^1(t)\Delta t} \\ &= \lambda^2(t)\zeta_0(t)\Delta t. \end{aligned}$$

If, on the other hand, neuron A did not fire at time t, generate a spike for neuron B at time t from a Bernoulli distribution with probability

$$P^2(t \mid \text{neuron A did not spike at t}) = \frac{(\lambda^2(t) - \lambda^2(t)\lambda^1(t)\zeta_0(t))\Delta t}{1 - \lambda^1(t)\Delta t}.$$

Grants

Cai and Kass: MH64537

Ventura: MH64537 and N01-NS-2-2346

References

Aersten, A, Gerstein, G, Habib, M, and Palm, G. Dynamics of neuronal firing correlation–Modulation in effective connectivity. *Journal Of Neurophysiology* 61:900–917, 1989.

Barbieri, R, Quirk, MC, Frank, LM, Wilson, MA, and Brown, EN. Non-Poisson stimulus-response models of neural spike train activity. *J. Neurosci. Methods* 105:25–37, 2001.

Brown, EN, Barbieri, R, Ventura, V, Kass, RE, Frank, LM. A note on the time-rescaling theorem and its implications for neural data analysis. *Neural Computation* 14:325–46, 2002.

Cai, C, Kass, RE, and Ventura, V. Trial-to-trial variability and its effect on time-varying dependence between two neurons. 2004b. *Submitted*

Daley, DJ and Vere-Jones, D. *An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods, Second Edition.* New York: Springer, 2003.

Davison, AC and Hinkley, DV. *Bootstrap Methods and Their Applications.* Cambridge University Press, 1997.

DiMatteo, I, Genovese, CR, and Kass, RE. Bayesian curve fitting with free–knot splines. *Biometrika*, 88:1055-1073, 2001.

Efron, B and Tibshirani, R. *An Introduction to the Bootstrap*, New York: Chapman and Hall, 1993.

Ito, H and Tsuji, S. Model dependence in quantification of spike interdependence by joint peri-stimulus time histogram. *Neural Computation* 12:195–217, 2000.

Johnson, D. Point process models of single-neuron discharges. *J. Comput. Neurosci.* 3:275–299, 1996.

Kass, RE and Ventura, V. A spike-train probability model. *Neural Comput.* 13:1713–1720, 2001.

Kass, RE, Ventura, V, and Cai, C. Statistical smoothing of neuronal data. *Network: Comput. Neural Syst.* 14:5-15, 2003.

Pipa, G and Grün, S. Non-parametric significance estimation of joint-spike events by shuffling and resampling. *Neurocomputing*, 52–54:31–37, 2003.

Ventura, V. Testing for, and estimating latency effects for Poisson and non-Poisson spike trains. *Neural Comput.*, 2004, to appear.

List of Figures