Computing Confidence Intervals for Point Process Models

Citation: Sarma, Sridevi V. et al. "Computing Confidence Intervals for Point Process Models." Neural Computation 23.11 (2011): 2731-2745. Web. 20 Jan. 2012. © 2011 Massachusetts Institute of Technology

As Published: http://dx.doi.org/10.1162/NECO_a_00198

Publisher: MIT Press

Persistent URL: http://hdl.handle.net/1721.1/68624

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

VIEW **Communicated by Valerie Ventura**

Computing Confidence Intervals for Point Process Models

Sridevi V. Sarma
sree@jhu.edu
Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218, U.S.A.

David P. Nguyen
dpnguyen@alum.mit.edu
Neuroscience Statistics Research Laboratory, Department of Anesthesia and Critical Care, Massachusetts General Hospital, Boston, MA 02114, and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Gabriela Czanner
gabriela.czanner@warwick.ac.uk
Warwick Manufacturing Group and Warwick Medical School, University of Warwick, Coventry CV4 7AL, U.K.

Sylvia Wirth
swirth@isc.cnrs.fr
Institut des Sciences Cognitives, CNRS, 69675 Bron, France

Matthew A. Wilson
mwilson@mit.edu
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Wendy Suzuki
wendy@cns.nyu.edu
Center for Neural Science, New York University, New York, NY 10003, U.S.A.

Emery N. Brown
enb@neurostat.mit.edu
Neuroscience Statistics Research Laboratory, Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, MA 02114; Harvard–MIT, Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139; and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Sridevi Sarma and David Nguyen contributed equally to this view.

Characterizing neural spiking activity as a function of intrinsic and extrinsic factors is important in neuroscience. Point process models are valuable for capturing such information; however, the process of fully applying these models is not always obvious. A complete model application has four broad steps: specification of the model, estimation of model parameters given observed data, verification of the model using goodness of fit, and characterization of the model using confidence bounds. Of these steps, only the first three have been applied widely in the literature, suggesting the need to dedicate a discussion to how the time-rescaling theorem, in combination with parametric bootstrap sampling, can be generally used to compute confidence bounds of point process models. In our first example, we use a generalized linear model of spiking propensity to demonstrate that confidence bounds derived from bootstrap simulations are consistent with those computed from closed-form analytic solutions. In our second example, we consider an adaptive point process model of hippocampal place field plasticity for which no analytical confidence bounds can be derived. We demonstrate how to simulate bootstrap samples from adaptive point process models, how to use these samples to generate confidence bounds, and how to statistically test the hypothesis that neural representations at two time points are significantly different. These examples have been designed as useful guides for performing scientific inference based on point process models.

1 Introduction

Receptive fields of neurons in the brain change in response to environmental stimuli as well as learning. For example, in the cat visual system, retinal lesions lead to reorganization of cortical topography (Pettet & Gilbert, 1992). Peripheral nerve sectioning can substantially alter the receptive fields of neurons in monkey somatosensory and motor cortices (Kaas, Merzenich, & Killackey, 1983; Merzenich et al., 1984). Similarly, the directional tuning of neural receptive fields in monkey motor cortex changes as the animal learns to compensate for an externally applied force field while moving a manipulandum (Gandolfo, Li, Benda, Schioppa, & Bizzi, 2000). In the rat hippocampus, the pyramidal neurons in the CA1 region have spatial receptive fields. As a rat executes a behavioral task, a given CA1 neuron fires only in a restricted region of the experimental environment, termed the cell's spatial or place receptive field (O'Keefe & Dostrovsky, 1971). Place fields change in a reliable manner as the animal executes its task (Mehta, Barnes, & McNaughton, 1997; Mehta, Quirk, & Wilson, 2000). Analysis of such neural dynamics is crucial for understanding how different brain regions update neural representations with learning and experience.

These examples highlight the need for models that integrate neurophysiological knowledge with internal and external covariates and capture

time-dependent processes of spiking activity, such as learning, adaptation, and plasticity on the millisecond timescales. The point process framework is able to meet these needs, as demonstrated over a wide spectrum of neuronal data types (Brown, Nguyen, Frank, Wilson, & Solo, 2001; Eden, Frank, Barbieri, Solo, & Brown, 2004; Ergun, Barbieri, Eden, Wilson, & Brown, 2007; Sarma et al., 2010) because it naturally extends the Poisson process and can therefore utilize a wealth of existing statistical solutions (Ogata, 1988; Brown, Barbieri, Ventura, Kass, & Frank, 2002; Daley & Vere-Jones, 2003). Here, in particular, we focus on a general solution to the problem: What is the probable range of my model parameters, and how do I perform significance testing on my data given a point process model?

We describe a general procedure for computing confidence intervals of point process models based on the parametric bootstrap, a method that first fits a parametric statistical model to observed data and then uses that model to simulate surrogate data (Davison & Hinkley, 1997). The term *bootstrapping* generally refers to the process of using surrogate data for the purposes of estimation and inference. In the following sections, we show that by using the time-rescaling theorem (Ogata, 1988; Brown et al., 2002; Daley & Vere-Jones, 2003), a fundamental property of the point process, it is possible to draw surrogate spike trains that are statistically probable given the real observed spike train and a point process model of good fit. Surrogate spike trains are then pooled and processed to construct bootstrapped conditional probability densities from which the confidence intervals for model parameters are derived (Efron & Tibshirani, 1993; Gilks, Richardson, & Spiegelhalter, 1998; Brown et al., 2002; Gentle, 2003).

We consider two examples where parametric bootstrapping is used to infer the confidence bounds of estimated model parameters that describe spike rate. The first is an analysis of a neuron recorded from the medial temporal lobe of a monkey performing a visual saccade task (Efron & Tibshirani, 1993; Gilks et al., 1998; Brown et al., 2002; Gentle, 2003). The model of the neuron relates intrinsic effects such as the neuron's own spiking history and concurrent ensemble activity, as well as the extrinsic effects of the presented stimuli (Efron & Tibshirani, 1993; Gilks et al., 1998; Brown et al., 2002; Gentle, 2003). We chose this particular example for its closed-form solution, which allowed us to compute confidence bounds using a theoretical maximum likelihood solution and a bootstrap estimation procedure, and demonstrate that the two provide consistently similar confidence bounds. The second model is motivated by the desire to improve the temporal resolution of models that quantify neuronal dynamics. Using point process modeling and steepest-descent adaptive filtering, we characterize experience-dependent changes in one rodent hippocampal place receptive field on a millisecond timescale (Brown et al., 2001; Frank, Eden, Solo, Wilson, & Brown, 2002). Although this adaptive model has many advantages, it has no closed-form solution and therefore no closed-form solution for computing confidence intervals. Using the parametric bootstrap sampler,

we demonstrate how to obtain confidence-bound estimates for this model and perform hypothesis testing to show that receptive field plasticity is statistically significant over time.

2 Methods

2.1 Point Processes. Point process models of neural spike train data are well suited for quantifying interactions among spike trains, environmental factors, and intrinsic factors (Brown et al., 2001; Eden et al., 2004; Ergun et al., 2007; Sarma et al., 2010). Here we establish the framework for bootstrapping (or simulating) structured spike trains from point process models (we refer readers to more comprehensive works for theoretical underpinnings Cox & Isham, 1980; Snyder, Miller, & Snyder, 1991; Daley & Vere-Jones, 2003).

In a point process framework, neural spike data are represented by binary random events that occur in continuous time and over the observation interval $(0, T]$. Values of 1 represent the time of one spike, and values of 0 represent no spiking activity. The spike times are parameterized by the set $u_1, u_2, \ldots, u_{N_T}$ where N_T represents the number of counted spikes in the observation interval. Given a time t in the observation interval, the instantaneous rate of spiking is given by the conditional intensity function (CIF),

$$\lambda(t \mid H(t), \theta(t)) = \lim_{\Delta \to 0} \frac{P(N(t+\Delta) - N(t) = 1 \mid H(t), \theta(t))}{\Delta}, \tag{2.1}$$

where spiking activity is a function of spike history $H(t)$ and time-varying model parameters $\theta(t)$. The spike history may be used to impose a refractory period or oscillatory dynamics, while time-varying parameters allow the modeling of phenomena such as plasticity and learning on a single-unit level. In the time interval, $(t,t+\Delta)$, the link between the CIF and the random number of spikes emitted by a neuron is the probability density Poisson $(\lambda(t \mid H(t), \theta(t))\Delta)$.

Given an experimentally observed spike train, one goal of point process modeling is to determine the functional form of the CIF and estimate parameters values that allow the CIF to match observed spiking activity. The instantaneous log likelihood,

$$l(\theta(t) \mid H(t)) = \log[\lambda(t) \mid H(t), \theta(t))]\frac{dN(t)}{dt} - \lambda(t \mid H(t), \theta(t)), \tag{2.2}$$

defines an optimization surface that is suitable for parameter estimation of stationary and adaptive and dynamic models. For the purposes of computer implementation, we approximate the instantaneous log likelihood in discrete time for $t = i\Delta$ as

$$l(\theta_i \mid H_i) \approx \log[\lambda(H_i, \theta_i)]n_i - \lambda(H_i, \theta_i), \tag{2.3}$$

where the interval Δ is on the order of milliseconds and smaller than any observed interspike interval, and n_i is the number of spikes observed in the time interval $((i-1)\Delta, i\Delta]$. For a small Δ, it can be shown that the discrete instantaneous log likelihood is equivalent to a Bernoulli process of the form $P(n_i \mid H_i) \approx (\lambda_i \Delta)^{n_i}(1-\lambda_i\Delta)^{1-n_i}$ (Truccolo, Eden, Fellows, Donoghue, & Brown, 2005). Therefore, for adequately small time steps, instead of computing the Poisson density above, we need only compute the following approximation to characterize spiking activity for time index i:

$$P(n_i = 1 \mid H_i, \theta_i) \approx \lambda(H_i, \theta_i)\Delta. \tag{2.4}$$

The time-rescaling theorem, a cornerstone of point process theory, states that if a spike train is generated from (or is consistent with) a CIF, then it is possible to transform arbitrarily distributed interspike intervals into independent and identically distributed (i.i.d.) exponential random variables, τ_k, by essentially weighting time with the CIF:

$$\left[\tau_k = \int_{u_k}^{u_{k+1}} \lambda(t \mid \cdot)\, dt\right] \sim \mathit{Exponential}(1). \tag{2.5}$$

Additionally, the interspike intervals can be transformed into i.i.d. uniformly distributed random variables,

$$[v_k = 1 - \exp(\tau_k)] \sim \mathit{Uniform}(0, 1), \tag{2.6}$$

thus making goodness-of-fit diagnostics, such as the Kilmogorov-Smirnov (K-S) test and autocorrelation analyses, a straightforward procedure (Ogata, 1988; Brown et al., 2002; Daley & Vere-Jones, 2003).

Our focus now is to demonstrate an equally important, yet less known, application of the time-rescaling theorem. That is, the theorem may be generally applied to continuously differentiable point process models to obtain confidence bounds on model parameter estimates (Ogata, 1988; Brown et al., 2002; Daley & Vere-Jones, 2003). The statistical framework embodied in equations 2.1 to 2.6 is fully compatible with Monte Carlo methods, thus allowing bootstrapped spike trains to be generated from the nonhomogeneous Poisson process that is parameterized by a time-varying CIF, $\lambda(t \mid H(t), \theta(t))$. These samples may then be used to construct confidence bounds for any test statistic that is based on the parameters of the CIF, $\theta(t)$, and history $H(t)$.

2.2 Bootstrap Sampling of Point Process Models. The parametric bootstrap sampler, a particular type of sampling method, may be used for inference (Davison & Hinkley, 1997). Here, an analytical model is estimated from the data in order to provide a basis for deriving statistics of interest analytically or generating simulated data samples with similar statistical qualities as the original data.

In the case of neuronal spike trains, a parametric bootstrap sampler allows confidence bounds to be computed around the instantaneous rate given by the CIF estimate; such confidence bounds allows us to better quantify the statistical significance of firing rate comparisons between experimental conditions. In this section, we outline the algorithm for parametric bootstrap sampling and then apply it to two real data examples.

The goal of our bootstrap sampler is to estimate the underlying probability distribution of the CIF parameters, $\theta(\cdot)$, in light of the history of spiking $H(\cdot)$. By utilizing the time rescaling theorem described in Brown et al. (2002), we can generate $b = 1, \ldots, B$ spike train samples, and for each sample, the CIF model parameters, $\hat{\theta}_i^b$, are estimated for a fixed b and $0 < i\Delta \leq T$. The K-S and independence tests are used to validate the bootstrapped model parameters (Johnson & Kotz, 1969). From the bootstrap samples, we may generate confidence intervals for some test statistic of interest, $\varphi(\theta)$. The value of B is both large and determined online by convergence criteria described below.

Let b=1 to start, increment b by 1 with each completion of step 6, and let $b = B$ denote the last sample to be generated:

1. Let k index be the spike to be generated. Initialize the algorithm by setting $k = 1$ and $\tilde{u}_0^b = 0$.
2. Draw $\tilde{v}$ from a uniform distribution on [0, 1], and set $\tilde{\tau}_k^b = -\log(\tilde{v})$.
3. Find the sampled spike time, $\tilde{u}_k^b$, as the solution to $\tilde{\tau}_k^b = \int_{\tilde{u}_{k-1}^b}^{\tilde{u}_k^b} \lambda(t \mid H(t), \hat{\theta}(t))\, dt$. It is important to note that the spiking history, $H(t)$, in the integral expression is accumulated as you sample.
4. If $\tilde{u}_k^b > T$, then discard $\tilde{u}_k^b$, let $N^b(T) = k - 1$, and go to step 5; otherwise increment k by 1 and return to step 2. Let the collection of simulated spike times from each bootstrap iteration be denoted as $\tilde{H}^b(t) = \{\tilde{u}_k^b \text{ for } k = 1, 2, \ldots, N^b(t)\}$.
5. Using the simulated spike history, $\tilde{H}^b(t)$, reestimate the parameter vector $\tilde{\theta}^b(t)$ using the same estimation procedure used to obtain $\hat{\theta}(t)$, and check for goodness of fit.
6. Let the sample statistic be defined as $\tilde{\varphi}^b(t) = \varphi(t \mid \tilde{H}^b(t) \cdot \tilde{\theta}^b(t))$ for $b \in [1, B]$.

The topic of convergence deals with how to determine a value for B that allows the simulation to fully sample from the distribution of $\theta(t)$ for all times of interest. Many solutions have been proposed, especially in the field of Markov chain Monte Carlo (MCMC) simulation (Gilks et al., 1998). While we can learn from the MCMC literature, we are not creating a Markov chain in this simulation because each sample does not depend on the previous one. Thus, we do not consider the need to define burn-in periods and problems of being trapped in state spaces of local optima.

Hence, it is possible to monitor convergence in our simulation using a simple approach. For every tenth bootstrap sample, we computed the

emprical 2.5% and 97.5% points from our test statistic set, $[\tilde{\varphi}^1(t), \ldots, \tilde{\varphi}^b(t)]$, where time t is chosen to reflect a time of interest in the experiment. When the dynamic range of each percentile is bounded by a small interval, ε, the simulation is said to have converged (Billingsley, 1999).

When the choice of ε is not clear, it is possible to use more automated methods of convergence detection. Many of these approaches require the simulation of M parallel bootstrap simulations, with the idea that at the point of convergence, the statistical properties of each simulation will be the same. One method in particular, the potential scale reduction factor (PSRF), computes the square root of the ratio of the between-simulation posterior variance estimate and the within-simulation variance to determine convergence (Gilks et al., 1998). As the simulation progresses, the PSRF should begin to decrease and eventually reach 1 at convergence.

3 Results

3.1 Comparison of Bootstrap and Analytical Confidence Bounds. Theoretically, bootstrap estimation generally can be used with any point process model with an integrable CIF (Brown et al., 2001). However, the application of this theory is not always straightforward. We first provide an example using a generalized linear model (GLM) framework (McCullagh & Nelder, 1990), which is becoming increasingly useful in neuroscience. In particular, using the GLM form for the CIF, we compare our bootstrap-derived confidence bounds with those obtained theoretically using maximum likelihood estimation.

In a GLM, the log of the CIF is modeled as a linear function of parameters that multiply functions of the covariates $r(t)$ that describe the neural activity dependencies, that is,

$$\log(\lambda(t \mid \alpha)) = \sum_{j=1}^{J} \alpha_j f_j(r(t), t), \tag{3.1}$$

where $\alpha_j \in \mathbf{R}$ for each j and f_j are known basis functions. The model parameters are α_j for $j = 1, 2, \ldots, J$.

The GLM is an extension of the multiple linear regression model in which the variable being predicted, in this case spike times, need not be gaussian (McCullagh & Nelder, 1990). The GLM also provides an efficient computational scheme for model parameter estimation and a likelihood framework for conducting statistical inferences based on the estimated model and error bounds on model parameters (Brown, Barbieri, Eden, & Frank, 2003).

Czanner et al. (2008) set out to examine the patterns of neural activity observed during the acquisition of new associative memories in monkeys (see Figure 1). They proposed a state-space GLM (described in section 2.1) of the CIF to capture both intrinsic effects, such as the neuron's own spiking

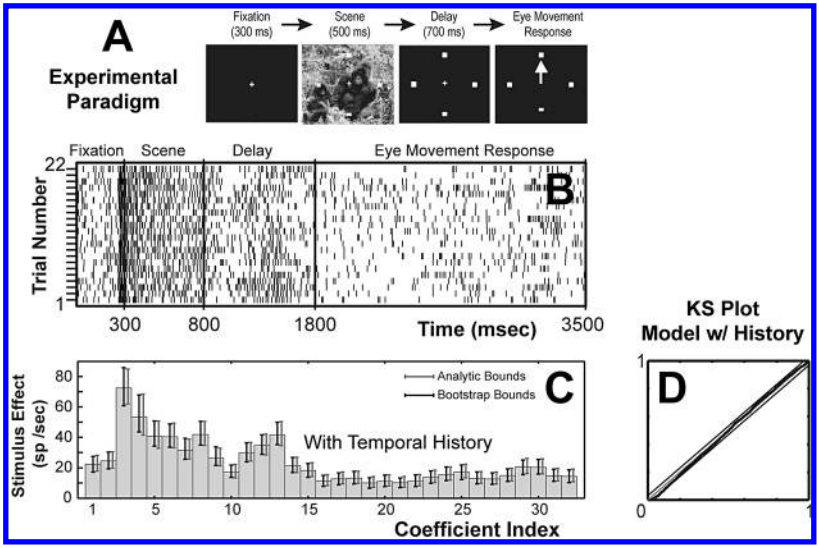

Figure 1: (A) Schematic illustration of location scene association task. (B) Raster plot from scene epoch to end of eye movement response. (C) Estimated stimulus coefficients of GLM overlaid with error bounds computed analytically (gray) and via our bootstrap (black). (D) Kilmogorov-Smirnov plot for GLM model with temporal history.

history and concurrent ensemble activity, and the extrinsic effect of the stimulus both within and across trials. Using their model applied to a single neuron spike train, we compute error bounds for a GLM point process model using simulated and analytical methods.

In Czanner et al. (2008), the discretized CIF is modeled over an interval of length T as the product of the stimulus intensity and the effect of the history dependence of the neural process as

$$\lambda(l\Delta \mid \theta^S, \theta^H, H_l) = \lambda^S(l\Delta \mid \theta^S)\lambda^H(l\Delta \mid \theta^H, H_l), \tag{3.2}$$

where $l = 1, 2, \ldots, L$ denotes the discrete time step that corresponds to continuous time $t \in ((l-1)\Delta, l\Delta]$, $\Delta = TL^{-1}$; H_l defines the spiking history up to time $l\Delta$; $\theta^S = (\theta_1^S, \theta_2^S, \ldots, \theta_L^S)$ are parameters of the GLM that characterize the stimulus or task effect on the spiking activity; and $\theta^H = (\theta_1^H, \theta_2^H, \ldots, \theta_J^H)$ is a constant vector of parameters that describe the dependence of current spiking activity on the spiking history.

The GLM model of the stimulus effect is

$$\log(\lambda^s(l\Delta \mid \theta^s)) = \sum_{r=1}^{L} \theta_r^s g_r(l\Delta), \tag{3.3}$$

where $g_r(l\Delta)$ is a sequence of unit pulse functions with equal length, that is,

$$g_r(l\Delta) = \begin{cases} 1 & \text{if } l = r \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

Since we compute only error bounds for the stimulus parameters, we omit the details of the history effects, which are described in Czanner et al., (2008).

We use data from one neuron for which a static GLM model provides a good fit. Twenty-two trials of length 3.2 seconds were used to develop the model that captured the CIF from 0.3 to 3.5 seconds— thus, the length of the stimulus parameter vector $L = 32$. The time bin length $\Delta = 1$ msec. The K-S plot for the model is shown in Figure 1D, which stays within the computed 95% confidence bounds for the degree of agreement using the distribution of the K-S statistic.

3.1.1 Analytic Approach. Since $\hat{\theta}$ is a maximum likelihood (ML) estimate of θ, it can be shown to asymptotically converge to a gaussian distribution with mean $\hat{\theta}$ (the asymptotic estimate that converges to the true θ) and variance $\sigma_i^2 = I(\hat{\theta})^{-1}$, where $I(\hat{\theta})$ is the expected Fisher information evaluated at $\hat{\theta}$ (Brown et al., 2003). Also, since unit pulse functions (g_r) are used, it is straightforward to show that the $\hat{\theta}_i^S$'s are mutually independent (Czanner et al., 2008). Hence, 95% confidence intervals can be computed individually as $\hat{\theta}_i \pm z_i$ such that $z_i = \sigma_i \Phi^{-1}(0.975)$.

3.1.2 Bootstrap Approach. Apply the bootstrap algorithm in section 2.2 to generate K samples of spike trains that comprise one bootstrap sample:

1. We first fix $\theta^H$and compute $\hat{\theta}^S$ from the bootstrap sample.
2. Repeat the process to generate $B = 1000$ bootstrapped ML estimates of θ^S.
3. Check the convergence of the bootstrap sample distribution. If needed, generate more bootstrap samples such that $B = B + 1000$.
4. Compute the 95% confidence interval from all the bootstrap samples via the 5th and 95th percentiles of the samples.
   a. Given a time step l from our model above, for each bootstrap sample indexed by s, compute the CIF conditioned on the sample: $\hat{\lambda}_l(\hat{\theta}^S) \equiv \lambda(l\Delta \mid \hat{\theta}^S, \theta^H, H_l)$. The result will be B samples of $\hat{\lambda}_l(\hat{\theta}^S)$.
   b. Given these B samples, we compute the confidence interval for each l using percentiles of 2.5 and 97.5.

Figures 1B and 1C illustrate the raw data in the form of raster plots for each neuron and corresponding stimulus coefficients overlaid with 95% confidence bounds computed both analytically and via bootstrap simulation. As seen in Figure 1C, the confidence interval obtained by using

the asymptotic normality of a maximum likelihood estimate matches the parametric bootstrap confidence interval for that estimate (which is expected because both procedures rely on the same model). This result helps to confirm that our parametric bootstrap design will provide the theoretically comparable confidence bounds.

3.2 Spline-Based Adaptive Point Process Modeling of Place Cells. Previously we developed and applied adaptive spline-based point process filters to model the receptive field plasticity of hippocampal place cells (Brown et al., 2001; Frank et al., 2002). The spline representation is flexible, requires few assumptions, and lends itself to an adaptive estimation framework. These applications demonstrated that adaptive estimation of CIF parameters may yield consistent data model agreement; however, we did not demonstrate how to obtain error bounds for our CIF parameter estimates. Here, we briefly reiterate the nonparametric, spline-based CIF model and steepest-descent estimator, and then apply the bootstrap method for obtaining error bounds.

3.2.1 Theory. Let time be discretized such that $t = i\Delta$ and let the subscript on the function denote the time index for all parameters, for example, $F_i(r) = F(r_i) = F(r(i\Delta))$. The proposed CIF for hippocampal place cells is a function that is decomposed into a spatial and temporal component,

$$\lambda_i(x, \tau \mid H_i, \theta_i^S, \theta_i^T) = \lambda_i^S(x \mid \theta_i^S)\lambda_i^T(\tau/\theta_i^T). \tag{3.5}$$

The spatial dependency, $\lambda_i^S(\cdot)$, depends on x, which is the location of the rat on a linear track (see section 3.3.2 for the experimental description). The temporal dependency, $\lambda_i^T(\cdot)$, depends on $\tau_i = i\Delta - t_{lastspike}$ which incorporates the history dependence and is defined as the time since the last spike. For convenience, we state the time dependency of the parameters in equation 3.5 using the index i.

The two components of our CIF have the same functional form, a Catmull-Rom or cardinal spline, which we first define generally as $\lambda(\rho/\theta)$ (Bartels, Beatty, & Barsky, 1987), where ρ will later be replaced by x or τ. A priori, we define the range supported by the spline, $(\rho_1, \rho_J]$, and $J + 2$ uniformly spaced, time-varying control points, $\theta_i = \{\theta_{i,0}, \theta_{i,1}, \ldots, \theta_{i,J+1}\}$, located at the fixed locations $\{\rho_0, \rho_1, \ldots, \rho_{J+1}\}$. The spline model is

$$\lambda(\rho \mid \theta) = [\beta(\rho)^3 \ \beta(\rho)^2 \ \beta(\rho) \ 1] \begin{bmatrix} -0.5 & 1.5 & -1.5 & 0.5 \\ 1 & -2.5 & 2 & -0.5 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_{i,j-1} \\ \theta_{i,j} \\ \theta_{i,j+1} \\ \theta_{i,j+2} \end{bmatrix}, \tag{3.6}$$

for $\rho \in (\rho_j, \rho_{j+1}]$ and where $\beta(\rho) = (\rho - \rho_j)/(\rho_{j+1} - \rho_j)$. At any point in the continuous range of $(\rho_1, \rho_J]$, the spline is fully determined by the four nearest control points.

The goal is to quantify the plasticity of the place-receptive field by sequentially estimating the vector θ_i for each incremental time step $i\Delta = \Delta, 2\Delta, \ldots$ to T (Brown et al., 2001). The steepest-descent solution attempts to move $\hat{\theta}_i$ toward the solution that will maximize the instantaneous log likelihood in equation 2.3 at each time step:

$$\hat{\theta}_i = \hat{\theta}_{i-1} - \varepsilon \left. \frac{\partial l_t(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}_{i-1}}. \tag{3.7}$$

The parameter ε is an adaptive gain parameter that determines the speed of adaptation. The full derivation of this adaptive filter can be found in Frank et al. (2002).

3.2.2 Application. In the following example, a rat runs back and forth on a 1D track that is 330 cm in length. The activity of a pyramidal neuron from the CA1 region of the dorsal hippocampus is observed during behavior. In Figure 2A, the trajectory of the rat is shown in space-time by the gray line, and the action potentials are shown by the black dots. We focus this example on the condition where the rat is running toward the 0 location.

The spatial spline contains 33 controls points that span from -11 cm to 342 cm. The temporal spline is defined on the $\log_{10}$ scale with 35 control points ranging from -4 to 3 (equivalently 0.1 msec to 1000 sec). We initialized the algorithm as $\Delta = 16.7$ ms, $\varepsilon^S = \varepsilon^T = 60$, $\theta_0^S = 0$, and $\theta_0^T = 0$. We run the adaptive filter forward and backward in time like a temporal smoother until the values of $\hat{\theta}_i^S$ and $\hat{\theta}_i^T$ do not change significantly over successive passes. We then validate the model using the K-S statistic (see Figure 2B). We illustrate the utility of the point process sampler by computing the 95% confidence bounds for $\hat{\lambda}_i^S(\cdot)$ and $\hat{\lambda}_i^T(\cdot)$ at time points $t_1 = 9840$ sec and $t_2 = 10{,}714$ sec by obtaining $B = 500$ bootstrap samples of $\tilde{\theta}^S$ and $\tilde{\theta}^T$ (see Figures 2C and 2D). Convergence criteria of the variance of the 5th and 95th percentiles were met for $B = 500$.

In order to state if the change in the neuron's receptive field is significant, we chose our test statistics to be $\varphi^S = \int_{x=11}^{x=342} [\lambda_{t_2}^S(x \mid \tilde{\theta}^S) - \lambda_{t_1}^S(x \mid \tilde{\theta}^S)]dx$ and $\varphi^T = \int_{\rho_\tau=-4}^{\rho_\tau=3} [\lambda_{t_2}^T(\rho_\tau \mid \tilde{\theta}^T) - \lambda_{t_1}^T(\rho_\tau \mid \tilde{\theta}^T)]d\rho_\tau$, where $\tilde{\theta}^S$ or $\tilde{\theta}^T$ denotes a bootstrap sample and $\rho_\tau = \log_{10}(\tau)$. Under a null hypothesis where the receptive has not changed between t_1 and t_2, the distributions for φ^S and φ^T should be symmetric and centered at 0. In Figures 2E and 2F, we find that the null hypotheses are not supported, with both p-values equaling 0.

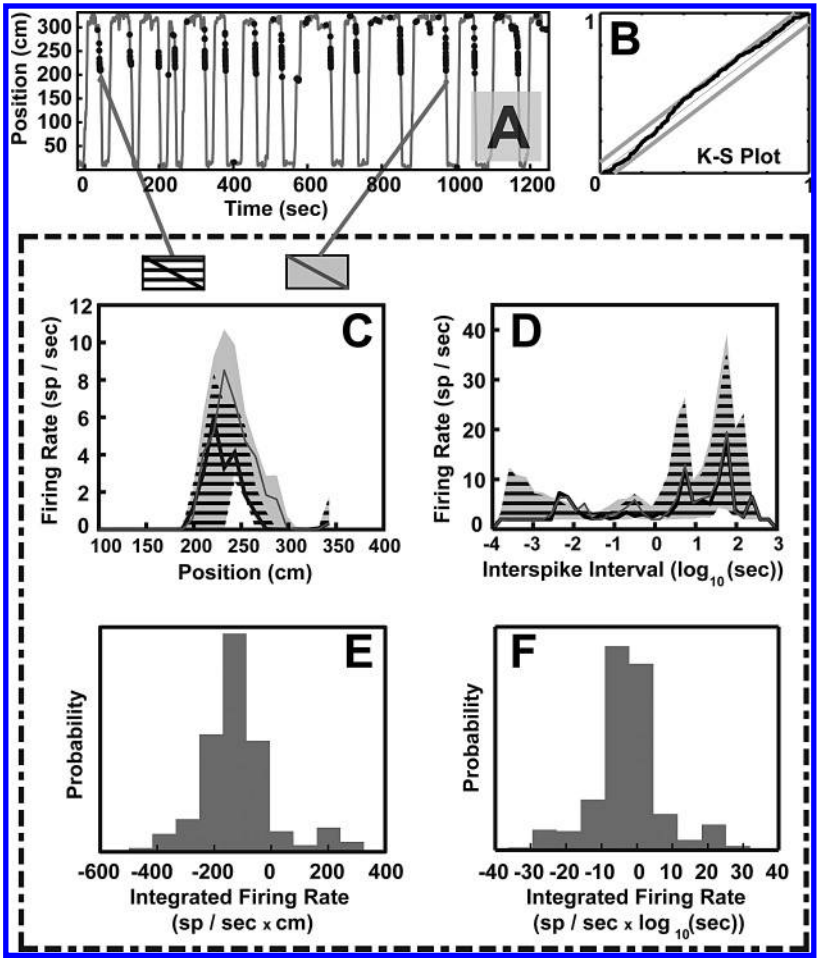

Figure 2: (A) Rat position overlaid with spike events. Each dot marks an action potential of the neuron. The two lines indicate points in time where the error bounds of the conditional intensity function were produced. (B) The Kilmogorov-Smirnov test demonstrating the adaptive filter model fit is appropriate. (C) Conditional spatial intensity function for each of the two time stamps. The shaded regions correspond to the 95% confidence bounds. (D) Conditional temporal intensity function for each of the two times tamps. The shaded regions correspond to the 95% confidence bounds. (E) Probability distribution for the difference in the test statistic, φ^s, between two times of interest shown in *C*. (F) Probability distribution for the difference in the test statistic, φ^T, between two times of interest shown in *D*.

4 Conclusions and Future Work

We have discussed a general parametric bootstrap method for computing confidence intervals of point process filters that may or may not have closed-form expressions. Hypothesis tests, which are crucial in the process of scientific inference, greatly rely on the accuracy of confidence intervals. Thus, we demonstrated in our first example that confidence intervals computed by bootstrapping are consistent with those that are theoretically derived. In the second example, we showed that the parametric bootstrap also allowed us to compute confidence bounds for an iterative filter and determine if significant changes occurred in neuronal coding.

The time-rescaling theorem, an elegant property from point process theory, is the basis for useful diagnostic measures such as the Q-Q test and Kolmogorov-Smirnov test, as well as for the inversion sampler used here to derive bootstrap samples from original data. It allows us to take any cross-validated point process model that is integrable and simulate random spike trains that have statistically similar features to the observed data. It is important to note that while we have proposed to bootstrap samples using inversion sampling, other methods of random sampling, such as rejection sampling (Efron & Tibshirani, 1993; Gilks et al., 1998; Brown et al., 2002; Gentle, 2003), may be used with possibly easier implementation. In addition, it is worth noting that the accuracy of the confidence intervals derived by parametric bootstrapping is predicated on the proper validation of the point process model by goodness-of-fit measures such as the K-S test and tests of random independence.

The method presented here increases the value of an existing point process model by allowing comparisons to be made within a data set without having to alter the functional form of the CIF. In addition, depending on the extent of the data set and model, it is entirely possible to perform multiple hypothesis tests under the same parametric bootstrap simulation. Thus, with a well-designed analysis approach using point process modeling and parametric bootstrap, it may be possible to obtain additional scientific inferences that are statistically well supported.

Acknowledgments

We thank Uri Eden for valuable discussions. Support was provided by Burrough's Wellcome Fund and L'Oreal for Women in Science fellowship to S.V.S, NIH MH59733 and NIH Pioneer Award DP1 OD003646-01 to E.N.B, NIH DA015644 to E.N.B. and W.A.S., NIH MH58847 to W.A.S., a McKnight Foundation grant to W.A.S., and a John Merck Scholars Award to W.A.S.

References

Bartels, R. H., Beatty, J. C., & Barsky, B. A. (1987). *An introduction to splines for use in computer graphics and geometric modeling*. San Francisco: Morgan Kaufmann.

Billingsley, P. (1999). *Convergence of probability measures*. Hoboken, NJ: Wiley.

Brown, E. N., Barbieri, R., Eden, U. T., & Frank, L. M. (2003). Likelihood methods for neural data analysis. In J. Feng (Ed.), *Computational neuroscience: A comprehensive approach* (pp. 253–286). Boca Raton, FL: CRC.

Brown, E. N., Barbieri, R., Ventura, V., Kass, R. E., & Frank, L. M. (2002). The time-rescaling theorem and its application to neural spike train data analysis. *Neural Computation*, *14*(2), 325–346.

Brown, E. N., Nguyen, D. P., Frank, L. M., Wilson, M. A., & Solo, V. (2001). An analysis of neural receptive field plasticity by point process adaptive filtering. *Proc. Natl. Acad. Sci. U.S.A.*, *98*(21), 12261–12266.

Cox, D. R., & Isham, V. (1980). *Point processes*. London: Chapman and Hall.

Czanner, G., Eden, U. T., Wirth, S., Yanike, M., Suzuki, W. A., & Brown, E. N. (2008). Analysis of between-trial and within-trial neural spiking dynamics. *J. Neurophysiol.*, *99*(5), 2672–2693.

Daley, D. J., & Vere-Jones, D. (2003). *An introduction to the theory of point processes*. New York: Springer.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge: Cambridge University Press.

Eden, U. T., Frank, L. M., Barbieri, R., Solo, V., & Brown, E. N. (2004). Dynamic analysis of neural encoding by point process adaptive filtering. *Neural Computation*, *16*(5), 971–998.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

Ergun, A., Barbieri, R., Eden, U. T., Wilson, M. A., & Brown, E. N. (2007). Construction of point process adaptive filter algorithms for neural systems using sequential Monte Carlo methods. *IEEE Trans. Biomed. Eng.*, *54*(3), 419–428.

Frank, L. M., Eden, U. T., Solo, V., Wilson, M. A., & Brown, E. N. (2002). Contrasting patterns of receptive field plasticity in the hippocampus and the entorhinal cortex: An adaptive filtering approach. *J. Neurosci.*, *22*(9), 3817–3830.

Gandolfo, F., Li, C., Benda, B. J., Schioppa, C. P., & Bizzi, E. (2000). Cortical correlates of learning in monkeys adapting to a new dynamical environment. *Proc. Natl. Acad. Sci. U.S.A.*, *97*(5), 2259–2263.

Gentle, J. E. (2003). *Random number generation and Monte Carlo methods*. New York: Springer-Verlag.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1998). *Markov chain Monte Carlo in practice*. Boca Raton, FL: Chapman & Hall.

Johnson, N. L., & Kotz, S. (1969). *Distributions in statistics*. Hoboken, NJ: Wiley.

Kaas, J. H., Merzenich, M. M., & Killackey, H. P. (1983). The reorganization of somatosensory cortex following peripheral-nerve damage in adult and developing mammals. *Annual Review of Neuroscience*, *6*, 325–356.

McCullagh, P., & Nelder, J. A. (1990). *Generalized linear models* (2nd ed.). Boca Raton, FL: CRC.

Mehta, M. R., Barnes, C. A., & McNaughton, B. L. (1997). Experience-dependent, asymmetric expansion of hippocampal place fields. *Proc. Natl. Acad. Sci. U.S.A.*, *94*(16), 8918–8921.

Mehta, M. R., Quirk, M. C., & Wilson, M. A. (2000). Experience-dependent asymmetric shape of hippocampal receptive fields. *Neuron*, *25*(3), 707–715.

Merzenich, M. M., Nelson, R. J., Stryker, M. P., Cynader, M. S., Schoppmann, A., & Zook, J. M. (1984). Somatosensory cortical map changes following digit amputation in adult monkeys. *J. Comp. Neurol.*, *224*(4), 591–605.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.*, *34*(1), 171–175.

Ogata, Y. (1988). Statistical-models for earthquake occurrences and residual analysis for point-processes. *Journal of the American Statistical Association*, *83*(401), 9–27.

Pettet, M. W., & Gilbert, C. D. (1992). Dynamic changes in receptive-field size in cat primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.*, *89*(17), 8366–8370.

Sarma, S. V., Eden, U. T., Cheng, M. L., Williams, Z. M., Hu, R., Eskandar, E., et al. (2010). Using point process models to compare neural spiking activity in the subthalamic nucleus of Parkinson's patients and a healthy primate. *IEEE Trans. Biomed. Eng.*, *57*(6), 1297–1305.

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

Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., & Brown, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J. Neurophysiol.*, *93*(2), 1074–1089.

Received November 10, 2009; accepted April 26, 2011.

This article has been cited by: