NOTE Communicated by Jonathan Victor

The Time-Rescaling Theorem and Its Application to Neural Spike Train Data Analysis

Emery N. Brown
brown@srlb.mgh.harvard.edu
Neuroscience Statistics Research Laboratory, Department of Anesthesia and Critical Care, Massachusetts General Hospital, Boston, MA 02114, U.S.A., and Division of Health Sciences and Technology, Harvard Medical School/Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Riccardo Barbieri
barbieri@srlb.mgh.harvard.edu
Neuroscience Statistics Research Laboratory, Department of Anesthesia and Critical Care, Massachusetts General Hospital, Boston, MA 02114, U.S.A.

Valérie Ventura
vventura@stat.cmv.edu
Robert E. Kass
kass@stat.cmu.edu
Department of Statistics, Carnegie Mellon University, Center for the Neural Basis of Cognition, Pittsburg, PA 15213, U.S.A.

Loren M. Frank
loren@neurostat.mgh.harvard.edu
Neuroscience Statistics Research Laboratory, Department of Anesthesia and Critical Care, Massachusetts General Hospital, Boston, MA 02114, U.S.A.

Measuring agreement between a statistical model and a spike train data series, that is, evaluating goodness of fit, is crucial for establishing the model's validity prior to using it to make inferences about a particular neural system. Assessing goodness-of-fit is a challenging problem for point process neural spike train models, especially for histogram-based models such as perstimulus time histograms (PSTH) and rate functions estimated by spike train smoothing. The time-rescaling theorem is a well-known result in probability theory, which states that any point process with an integrable conditional intensity function may be transformed into a Poisson process with unit rate. We describe how the theorem may be used to develop goodness-of-fit tests for both parametric and histogram-based point process models of neural spike trains. We apply these tests in two examples: a comparison of PSTH, inhomogeneous Poisson, and inhomogeneous Markov interval models of neural spike trains from the sup-

plementary eye field of a macque monkey and a comparison of temporal and spatial smoothers, inhomogeneous Poisson, inhomogeneous gamma, and inhomogeneous inverse gaussian models of rat hippocampal place cell spiking activity. To help make the logic behind the time-rescaling theorem more accessible to researchers in neuroscience, we present a proof using only elementary probability theory arguments. We also show how the theorem may be used to simulate a general point process model of a spike train. Our paradigm makes it possible to compare parametric and histogram-based neural spike train models directly. These results suggest that the time-rescaling theorem can be a valuable tool for neural spike train data analysis.

1 Introduction

The development of statistical models that accurately describe the stochastic structure of neural spike trains is a growing area of quantitative research in neuroscience. Evaluating model goodness of fit—that is, measuring quantitatively the agreement between a proposed model and a spike train data series—is crucial for establishing the model's validity prior to using it to make inferences about the neural system being studied. Assessing goodness of fit for point neural spike train models is a more challenging problem than for models of continuous-valued processes. This is because typical distance discrepancy measures applied in continuous data analyses, such as the average sum of squared deviations between recorded data values and estimated values from the model, cannot be directly computed for point process data. Goodness-of-fit assessments are even more challenging for histogram-based models, such as peristimulus time histograms (PSTH) and rate functions estimated by spike train smoothing, because the probability assumptions needed to evaluate model properties are often implicit. Berman (1983) and Ogata (1988) developed transformations that, under a given model, convert point processes like spike trains into continuous measures in order to assess model goodness of fit. One of the theoretical results used to construct these transformations is the time-rescaling theorem.

A form of the time-rescaling theorem is well known in elementary probability theory. It states that any inhomogeneous Poisson process may be rescaled or transformed into a homogeneous Poisson process with a unit rate (Taylor & Karlin, 1994). The inverse transformation is a standard method for simulating an inhomogeneous Poisson process from a constant rate (homogeneous) Poisson process. Meyer (1969) and Papangelou (1972) established the general time-rescaling theorem, which states that any point process with an integrable rate function may be rescaled into a Poisson process with a unit rate. Berman and Ogata derived their transformations by applying the general form of the theorem. While the more general time-rescaling theorem is well known among researchers in point process theory (Brémaud, 1981; Jacobsen, 1982; Daley & Vere-Jones, 1988; Ogata, 1988; Karr, 1991), the

theorem is less familiar to neuroscience researchers. The technical nature of the proof, which relies on the martingale representation of a point process, may have prevented its significance from being more broadly appreciated.

The time-rescaling theorem has important theoretical and practical implications for application of point process models in neural spike train data analysis. To help make this result more accessible to researchers in neuroscience, we present a proof that uses only elementary probability theory arguments. We describe how the theorem may be used to develop goodness-of-fit tests for both parametric and histogram-based point process models of neural spike trains. We apply these tests in two examples: a comparison of PSTH, inhomogeneous Poisson, and inhomogeneous Markov interval models of neural spike trains from the supplementary eye field of a macque monkey and a comparison of temporal and spatial smoothers, inhomogeneous Poisson, inhomogeneous gamma, and inhomogeneous inverse gaussian models of rat hippocampal place cell spiking activity. We also demonstrate how the time-rescaling theorem may be used to simulate a general point process.

2 Theory

2.1 The Conditional Intensity Function and the Joint Probability Density of the Spike Train. Define an interval $(0, T]$, and let $0 < u_1 < u_2 < , \ldots, < u_{n-1} < u_n \leq T$ be a set of event (spike) times from a point process. For $t \in (0, T]$, let $N(t)$ be the sample path of the associated counting process. The sample path is a right continuous function that jumps 1 at the event times and is constant otherwise (Snyder & Miller, 1991). In this way, $N(t)$ counts the number and location of spikes in the interval $(0, t]$. Therefore, it contains all the information in the sequence of events or spike times. For $t \in (0, T]$, we define the conditional or stochastic intensity function as

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

where $H_t = \{0 < u_1 < u_2, \ldots, u_{N(t)} < t\}$ is the history of the process up to time t and $u_{N(t)}$ is the time of the last spike prior to t. If the point process is an inhomogeneous Poisson process, then $\lambda(t \mid H_t) = \lambda(t)$ is simply the Poisson rate function. Otherwise, it depends on the history of the process. Hence, the conditional intensity generalizes the definition of the Poisson rate. It is well known that $\lambda(t \mid H_t)$ can be defined in terms of the event (spike) time probability density, $f(t \mid H_t)$, as

$$\lambda(t \mid H_t) = \frac{f(t \mid H_t)}{1 - \int_{u_{N(t)}}^{t} f(u \mid H_t)\, du}, \tag{2.2}$$

for $t > u_{N(t)}$ (Daley & Vere-Jones, 1988; Barbieri, Quirk, Frank, Wilson, & Brown, 2001). We may gain insight into equation 2.2 and the meaning of the

conditional intensity function by computing explicitly the probability that given H_t, a spike u_k occurs in $[t, t + \Delta t)$ where $k = N(t) + 1$. To do this, we note that the events $\{N(t + \Delta t) - N(t) = 1 \mid H_t\}$ and $\{u_k \in [t, t + \Delta t) \mid H_t\}$ are equivalent and that therefore,

$$\begin{aligned}
\Pr(N(t + \Delta t) - N(t) = 1 \mid H_t) &= \Pr(u_k \in [t, t + \Delta t) \mid u_k > t, H_t) \\
&= \frac{\Pr(u_k \in [t, t + \Delta t) \mid H_t)}{\Pr(u_k > t \mid H_t)} \\
&= \frac{\int_t^{t+\Delta t} f(u \mid H_t)\, du}{1 - \int_{u_{N(t)}}^t f(u \mid H_t)\, du} \\
&\approx \frac{f(t \mid H_t)\Delta t}{1 - \int_{u_{N(t)}}^t f(u \mid H_t)\, du} \\
&= \lambda(t \mid H_t)\Delta t. \qquad (2.3)
\end{aligned}$$

Dividing by Δt and taking the limit gives

$$\begin{aligned}
\lim_{\Delta t \to 0} \frac{\Pr(u_k \in [t, t + \Delta t) \mid H_t)}{\Delta t} &= \frac{f(t \mid H_t)}{1 - \int_{u_{N(t)}}^t f(u \mid H_t)\, du} \\
&= \lambda(t \mid H_t), \qquad (2.4)
\end{aligned}$$

which is equation 2.1. Therefore, $\lambda(t \mid H_t)\Delta t$ is the probability of a spike in $[t, t + \Delta t)$ when there is history dependence in the spike train. In survival analysis, the conditional intensity is termed the *hazard function* because in this case, $\lambda(t \mid H_t)\Delta t$ measures the probability of a failure or death in $[t, t+\Delta t)$ given that the process has survived up to time t (Kalbfleisch & Prentice, 1980).

Because we would like to apply the time-rescaling theorem to spike train data series, we require the joint probability density of exactly n event times in $(0, T]$. This joint probability density is (Daley & Vere-Jones, 1988; Barbieri, Quirk, et al., 2001)

$$\begin{aligned}
&f(u_1, u_2, \ldots, u_n \cap N(T) = n) \\
&\quad = f(u_1, u_2, \ldots, u_n \cap u_{n+1} > T) \\
&\quad = f(u_1, u_2, \ldots, u_n \cap N(u_n) = n) \Pr(u_{n+1} > T \mid u_1, u_2, \ldots, u_n) \\
&\quad = \prod_{k=1}^{n} \lambda(u_k \mid H_{u_k}) \exp\left\{ - \int_{u_{k-1}}^{u_k} \lambda(u \mid H_u)\, du \right\} \\
&\qquad \cdot \exp\left\{ - \int_{u_n}^{T} \lambda(u \mid H_u)\, du \right\}, \qquad (2.5)
\end{aligned}$$

where

$$f(u_1, u_2, \ldots, u_n \cap N(u_n) = n) = \prod_{k=1}^{n} \lambda(u_k \mid H_{u_k}) \exp\left\{-\int_{u_{k-1}}^{u_k} \lambda(u \mid H_u)\, du\right\} \tag{2.6}$$

$$\Pr(u_{n+1} > T \mid u_1, u_2, \ldots, u_n) = \exp\left\{-\int_{u_n}^{T} \lambda(u \mid H_u)\, du\right\}, \tag{2.7}$$

and $u_0 = 0$. Equation 2.6 is the joint probability density of exactly n events in $(0, u_n]$, whereas equation 2.7 is the probability that the $n + 1st$ event occurs after T. The conditional intensity function provides a succinct way to represent the joint probability density of the spike times. We can now state and prove the time-rescaling theorem.

Time-Rescaling Theorem. *Let $0 < u_1 < u_2 <, \ldots, < u_n < T$ be a realization from a point process with a conditional intensity function $\lambda(t \mid H_t)$ satisfying $0 < \lambda(t \mid H_t)$ for all $t \in (0, T]$. Define the transformation*

$$\Lambda(u_k) = \int_0^{u_k} \lambda(u \mid H_u)\, du, \tag{2.8}$$

for $k = 1, \ldots, n$, and assume $\Lambda(t) < \infty$ with probability one for all $t \in (0, T]$. Then the $\Lambda(u_k)$'s are a Poisson process with unit rate.

Proof. Let $\tau_k = \Lambda(u_k) - \Lambda(u_{k-1})$ for $k = 1, \ldots, n$ and set $\tau_T = \int_{u_n}^{T} \lambda(u \mid H_u)\, du$. To establish the result, it suffices to show that the τ_ks are independent and identically distributed exponential random variables with mean one. Because the τ_k transformation is one-to-one and $\tau_{n+1} > \tau_T$ if and only if $u_{n+1} > T$, the joint probability density of the τ_k's is

$$f(\tau_1, \tau_2, \ldots, \tau_n \cap \tau_{n+1} > \tau_T) = f(\tau_1, \ldots, \tau_n) \Pr(\tau_{n+1} > \tau_T \mid \tau_1, \ldots, \tau_n). \tag{2.9}$$

We evaluate each of the two terms on the right side of equation 2.9. The following two events are equivalent:

$$\{\tau_{n+1} > \tau_T \mid \tau_1, \ldots, \tau_n\} = \{u_{n+1} > T \mid u_1, u_2, \ldots, u_n\}. \tag{2.10}$$

Hence

$$\begin{aligned} \Pr(\tau_{n+1} > \tau_T \mid \tau_1, \tau_2, \ldots, \tau_n) &= \Pr(u_{n+1} > T \mid u_1, u_2, \ldots, u_n) \\ &= \exp\left\{-\int_{u_n}^{T} \lambda(u \mid H_{u_n})\, du\right\} \\ &= \exp\{-\tau_T\}, \end{aligned} \tag{2.11}$$

where the last equality follows from the definition of τ_T. By the multivariate change-of-variable formula (Port, 1994),

$$f(\tau_1, \tau_2, \ldots, \tau_n) = |J| f(u_1, u_2, \ldots, u_n \cap N(u_n) = n), \tag{2.12}$$

where J is the Jacobian of the transformation between u_j, $j = 1, \ldots, n$ and τ_k, $k = 1, \ldots, n$. Because τ_k is a function of $u_1, \ldots, u_k$, J is a lower triangular matrix, and its determinant is the product of its diagonal elements defined as $|J| = |\prod_{k=1}^{n} J_{kk}|$. By assumption $0 < \lambda(t \mid H_t)$ and by equation 2.8 and the definition of τ_k, the mapping of u into τ is one-to-one. Therefore, by the inverse differentiation theorem (Protter & Morrey, 1991), the diagonal elements of J are

$$J_{kk} = \frac{\partial u_k}{\partial \tau_k} = \lambda(u_k \mid H_{u_k})^{-1}. \tag{2.13}$$

Substituting $|J|$ and equation 2.6 into equation 2.12 yields

$$\begin{aligned} f(\tau_1, \tau_2, \ldots, \tau_n) &= \prod_{k=1}^{n} \lambda(u_k \mid H_{u_k})^{-1} \prod_{k=1}^{n} \lambda(u_k \mid H_{u_k}) \\ &\quad \cdot \exp\left\{ - \int_{u_{k-1}}^{u_k} \lambda(u \mid H_u)\, du \right\} \\ &= \prod_{k=1}^{n} \exp\{-[\Lambda(u_k) - \Lambda(u_{k-1})]\} \\ &= \prod_{k=1}^{n} \exp\{-\tau_k\}. \end{aligned} \tag{2.14}$$

Substituting equations 2.11 and 2.14 into 2.9 yields

$$\begin{aligned} f(\tau_1, \tau_2, \ldots, \tau_n \cap \tau_{n+1} > \tau_T) &= f(\tau_1, \ldots, \tau_n) \Pr(\tau_{n+1} > \tau_T \mid \tau_1, \ldots, \tau_n) \\ &= \left(\prod_{k=1}^{n} \exp\{-\tau_k\} \right) \exp\{-\tau_T\}, \end{aligned} \tag{2.15}$$

which establishes the result.

The time-rescaling theorem generates a history-dependent rescaling of the time axis that converts a point process into a Poisson process with a unit rate.

2.2 Assessing Model Goodness of Fit. We may use the time-rescaling theorem to construct goodness-of-fit tests for a spike data model. Once a

model has been fit to a spike train data series, we can compute from its estimated conditional intensity the rescaled times

$$\tau_k = \Lambda(u_k) - \Lambda(u_{k-1}). \tag{2.16}$$

If the model is correct, then, according to the theorem, the $\tau_k s$ are independent exponential random variables with mean 1. If we make the further transformation

$$z_k = 1 - \exp(-\tau_k), \tag{2.17}$$

then $z_k s$ are independent uniform random variables on the interval $(0, 1)$. Because the transformations in equations 2.16 and 2.17 are both one-to-one, any statistical assessment that measures agreement between the $z_k s$ and a uniform distribution directly evaluates how well the original model agrees with the spike train data. Here we present two methods: Kolmogorov-Smirnov tests and quantile-quantile plots.

To construct the Kolmogorov-Smirnov test, we first order the $z_k s$ from smallest to largest, denoting the ordered values as $z_{(k)} s$. We then plot the values of the cumulative distribution function of the uniform density defined as $b_k = \frac{k-\frac{1}{2}}{n}$ for $k = 1, \ldots, n$ against the $z_{(k)} s$. If the model is correct, then the points should lie on a 45-degree line (Johnson & Kotz, 1970). Confidence bounds for the degree of agreement between the models and the data may be constructed using the distribution of the Kolmogorov-Smirnov statistic. For moderate to large sample sizes the 95% (99%) confidence bounds are well approximated as $b_k \pm 1.36/n^{1/2}$ ($b_k \pm 1.63/n^{1/2}$) (Johnson & Kotz, 1970). We term such a plot a Kolmogorov-Smirnov (KS) plot.

Another approach to measuring agreement between the uniform probability density and the $z_k s$ is to construct a quantile-quantile (Q-Q) plot (Ventura, Carta, Kass, Gettner, & Olson, 2001; Barbieri, Quirk, et al., 2001; Hogg & Tanis, 2001). In this display, we plot the quantiles of the uniform distribution, denoted here also as the $b_k s$, against the $z_{(k)} s$. As in the case of the KS plots, exact agreement occurs between the point process model and the experimental data if the points lie on a 45-degree line. Pointwise confidence bands can be constructed to measure the magnitude of the departure of the plot from the 45-degree line relative to chance. To construct pointwise bands, we note that if the $\tau_k s$ are independent exponential random variables with mean 1 and the $z_k s$ are thus uniform on the interval $(0, 1)$, then each $z_{(k)}$ has a beta probability density with parameters k and $n - k + 1$ defined as

$$f(z \mid k, n-k+1) = \frac{n!}{(n-k)!\,(k-1)!} z^{k-1}(1-z)^{n-k}, \tag{2.18}$$

for $0 < z < 1$ (Johnson & Kotz, 1970). We set the 95% confidence bounds by finding the 2.5th and 97.5th quantiles of the cumulative distribution

associated with equation 2.18 for $k = 1, \ldots, n$. These exact quantiles are readily available in many statistical software packages. For moderate to large spike train data series, a reasonable approximation to the 95% (99%) confidence bounds is given by the gaussian approximation to the binomial probability distribution as $z_{(k)} \pm 1.96[z_{(k)}(1 - z_{(k)})/n]^{1/2}$ ($z_{(k)} \pm 2.575[z_{(k)}(1 - z_{(k)})/n]^{1/2}$). To our knowledge, these local confidence bounds for the Q-Q plots based on the beta distribution and the gaussian approximation are new.

In general, the KS confidence intervals will be wider than the corresponding Q-Q plot intervals. To see this, it suffices to compare the widths of the two intervals using their approximate formulas for large n. From the gaussian approximation to the binomial, the maximum width of the 95% confidence interval for the Q-Q plots occurs at the median: $z_{(k)} = 0.50$ and is $2[1.96/(4n)^{1/2}] = 1.96n^{-1/2}$. For n large, the width of the 95% confidence intervals for the KS plots is $2.72n^{-1/2}$ at all quantiles. The KS confidence bounds consider the maximum discrepancy from the 45-degree line along all quantiles; the 95% bands show the discrepancy that would be exceeded 5% of the time by chance if the plotted data were truly uniformly distributed. The Q-Q plot confidence bounds consider the maximum discrepancy from the 45-degree line for each quantile separately. These pointwise 95% confidence bounds mark the amount by which each value $z_{(k)}$ would deviate from the true quantile 5% of the time purely by chance. The KS bounds are broad because they are based on the joint distribution of all n deviations, and they consider the distribution of the largest of these deviations. The Q-Q plot bounds are narrower because they measure the deviation at each quantile separately. Used together, the two plots help approximate upper and lower limits on the discrepancy between a proposed model and a spike train data series.

3 Applications

3.1 An Analysis of Supplementary Eye Field Recordings. For the first application of the time-rescaling theorem to a goodness-of-fit analysis, we analyze a spike train recorded from the supplementary eye field (SEF) of a macaque monkey. Neurons in the SEF play a role in oculomotor processes (Olson, Gettner, Ventura, Carta, & Kass, 2000). A standard paradigm for studying the spiking properties of these neurons is a delayed eye movement task. In this task, the monkey fixates, is shown locations of potential target sites, and is then cued to the specific target to which it must saccade. Next, a preparatory cue is given, followed a random time later by a go signal. Upon receiving the go signal, the animal must saccade to the specific target and hold fixation for a defined amount of time in order to receive a reward. Beginning from the point of the specific target cue, neural activity is recorded for a fixed interval of time beyond the presentation of the go signal. After a brief rest period, the trial is repeated. Multiple trials from an experiment

such as this are jointly analyzed using a PSTH to estimate firing rate for a finite interval following a fixed initiation point. That is, the trials are time aligned with respect to a fixed initial point, such as the target cue. The data across trials are binned in time intervals of a fixed length, and the rate in each bin is estimated as the average number of spikes in the fixed time interval.

Kass and Ventura (2001) recently presented inhomogeneous Markov interval (IMI) models as an alternative to the PSTH for analyzing multiple-trial neural spike train data. These models use a Markov representation for the conditional intensity function. One form of the IMI conditional intensity function they considered is

$$\lambda(t \mid H_t) = \lambda(t \mid u_{N(t)}, \theta) = \lambda_1(t \mid \theta)\lambda_2(t - u_{N(t)} \mid \theta), \tag{3.1}$$

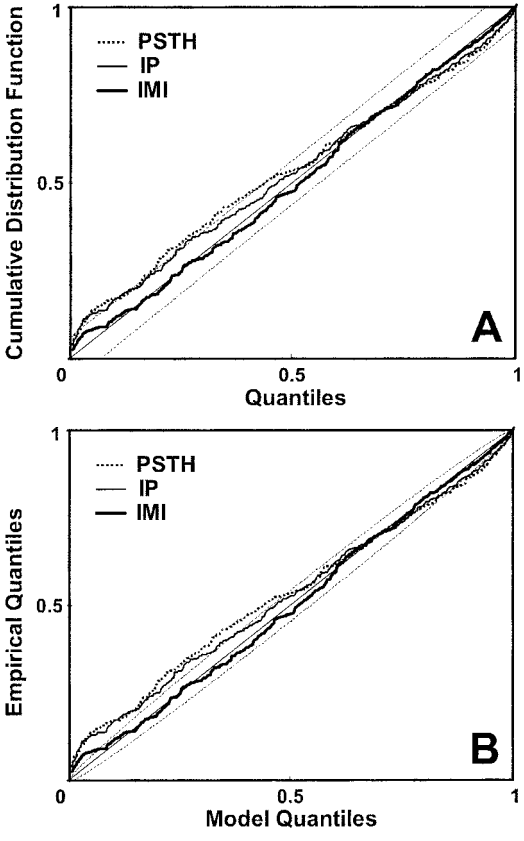
where $u_{N(t)}$ is the time of the last spike prior to t, $\lambda_1(t \mid \theta)$ modulates firing as a function of the experimental clock time, $\lambda_2(t - u_{N(t)} \mid \theta)$ represents Markov dependence in the spike train, and θ is a vector of model parameters to be estimated. Kass and Ventura modeled $\log \lambda_1(t \mid \theta)$ and $\log \lambda_2(t - u_{N(t)} \mid \theta)$ as separate piecewise cubic splines in their respective arguments t and $t - u_{N(t)}$. The cubic pieces were joined at knots so that the resulting functions were twice continuously differentiable. The number and positions of the knots were chosen in a preliminary data analysis. In the special case $\lambda(t \mid u_{N(t)}, \theta) = \lambda_1(t \mid \theta)$, the conditional intensity function in equation 3.1 corresponds to an inhomogeneous Poisson (IP) model because this assumes no temporal dependence among the spike times.

Kass and Ventura used their models to analyze SEF data that consisted of 486 spikes recorded during the 400 msec following the target cue signal in 15 trials of a delayed eye movement task (neuron PK166a from Olson et al., 2000, using the pattern condition). The IP and IMI models were fit by maximum likelihood, and statistical significance tests on the spline coefficients were used to compare goodness of fit. Included in the analysis were spline models with higher-order dependence among the spike times than the first-order Markov dependence in equation 3.1. They found that the IMI model gave a statistically significant improvement in the fit relative to the IP and that adding higher-order dependence to the model gave no further improvements. The fit of the IMI model was not improved by including terms to model between-trial differences in spike rate. The authors concluded that there was strong first-order Markov dependence in the firing pattern of this SEF neuron. Kass and Ventura did not provide an overall assessment of model goodness of fit or evaluate how much the IMI and the IP models improved over the histogram-based rate model estimated by the PSTH.

Using the KS and Q-Q plots derived from the time-rescaling theorem, it is possible to compare directly the fits of the IMI, IP, and PSTH models and to determine which gives the most accurate description of the SEF spike train structure. The equations for the IP rate function, $\lambda_{IP}(t \mid \theta_{IP}) = \lambda_1(t \mid \theta_{IP})$,

and for the IMI conditional intensity function, $\lambda_{IMI}(t \mid u_{N(t)}, \theta_{IMI}) = \lambda_1(t \mid \theta_{IMI})\lambda_2(t - u_{N(t)}, \theta_{IMI})$, are given in the appendix, along with a discussion of the maximum likelihood procedure used to estimate the coefficients of the spline basis elements. The estimated conditional intensity functions for the IP and IMI models are, respectively, $\lambda_{IP}(t \mid, \hat{\theta}_{IP})$ and $\lambda_{IMI}(t \mid u_{N(t)}, \hat{\theta}_{IMI})$, where $\hat{\theta}$ is the maximum likelihood estimate of the specific spline coefficients. For the PSTH model, the conditional intensity estimate is the PSTH computed by averaging the number of spikes in each of 40 10 msec bins (the 400 msec following the target cue signal) across the 15 trials. The PSTH is the fit of another inhomogeneous Poisson model because it assumes no dependence among the spike times.

The results of the IMI, IP, and PSTH model fits compared by KS and Q-Q plots are shown in Figure 1. For the IP model, there is lack of fit at lower quantiles (below 0.25) because in that range, its KS plot lies just outside the 95% confidence bounds (see Figure 1A). From quantile 0.25 and beyond, the IP model is within the 95% confidence bounds, although be-



yond the quantile 0.75, it slightly underpredicts the true probability model of the data. The KS plot of the PSTH model is similar to that of the IP except that it lies entirely outside the 95% confidence bands below quantile 0.50. Beyond this quantile, it is within the 95% confidence bounds. The KS plot of the PSTH underpredicts the probability model of the data to a greater degree in the upper range than the IP model. The IMI model is completely within the 95% confidence bounds and lies almost exactly on the 45-degree line of complete agreement between the model and the data. The Q-Q plot analyses (see Figure 1B) agree with KS plot analyses with a few exceptions. The Q-Q plot analyses show that the lack of fit of the IP and PSTH models is greater at the lower quantiles (0–0.50) than suggested by the KS plots. The Q-Q plots for the IP and PSTH models also show that the deviations of these two models near quantiles 0.80 to 0.90 are statistically significant. With the exception of a small deviation below quantile 0.10, the Q-Q plot of the IMI lies almost exactly on the 45-degree line.

Figure 1: *Facing page.* (A) Kolmogorov-Smirnov (K-S) plots of the inhomogeneous Markov interval (IMI) model, inhomogeneous Poisson (IP), and perstimulus time histogram (PSTH) model fits to the SEF spike train data. The solid 45-degree line represents exact agreement between the model and the data. The dashed 45-degree lines are the 95% confidence bounds for exact agreement between the model and experimental data based on the distribution of the Kolmogorov-Smirnov statistic. The 95% confidence bounds are $b_k \pm 1.36n^{-\frac{1}{2}}$, where $b_k = (k - \frac{1}{2})/n$ for $k = 1, \ldots, n$ and n is the total number of spikes. The IMI model (thick, solid line) is completely within the 95% confidence bounds and lies almost exactly on the 45-degree line. The IP model (thin, solid line) has lack of fit at lower quantiles (< 0.25). From the quantile 0.25 and beyond, the IP model is within the 95% confidence bounds. The KS plot of the PSTH model (dotted line) is similar to that of the IP model except that it lies outside the 95% confidence bands below quantile 0.50. Beyond this quantile, it is within the 95% confidence bounds, yet it underpredicts the probability model of the data to a greater degree in this range than the IP model does. The IMI model agrees more closely with the spike train data than either the IP or the PSTH models. (B) Quantile-quantile (Q-Q) plots of the IMI (thick, solid line), IP (thin, solid line), and PSTH (dotted line) models. The dashed lines are the local 95% confidence bounds for the individual quantiles computed from the beta probability density defined in equation 2.18. The solid 45-degree line represents exact agreement between the model and the data. The Q-Q plots suggests that the lack of fit of the IP and PSTH models is greater at the lower quantiles (0–0.50) than suggested by the KS plots. The Q-Q plots for the IP and PSTH models also show that the deviations of these two models near quantiles 0.80 to 0.90 are statistically significant. With the exception of a small deviation below quantile 0.10, the Q-Q plot of the IMI lies almost exactly on the 45-degree line.

We conclude that the IMI model gives the best description of these SEF spike trains. In agreement with the report of Kass and Ventura (2001), this analysis supports a first-order Markov dependence among the spike times and not a Poisson structure, as would be suggested by either the IP or the PSTH models. This analysis extends the findings of Kass and Ventura by showing that of the three models, the PSTH gives the poorest description of the SEF spike train. The IP model gives a better fit to the SEF data than the PSTH model because the maximum likelihood analysis of the parametric IP model is more statistically efficient than the histogram (method of moments) estimate obtained from the PSTH (Casella & Berger, 1990). That is, the IP model fit by maximum likelihood uses all the data to estimate the conditional intensity function at all time points, whereas the PSTH analysis uses only spikes in a specified time bin to estimate the firing rate in that bin. The additional improvement of the IMI model over the IP is due to the fact that the former represents temporal dependence in the spike train.

3.2 An Analysis of Hippocampal Place Cell Recordings. As a second example of using the time-rescaling theorem to develop goodness-of-fit tests, we analyze the spiking activity of a pyramidal cell in the CA1 region of the rat hippocampus recorded from an animal running back and forth on a linear track. Hippocampal pyramidal neurons have place-specific firing (O'Keefe & Dostrovsky, 1971); a given neuron fires only when the animal is in a certain subregion of the environment termed the neuron's place field. On a linear track, these fields approximately resemble one-dimensional gaussian surfaces. The neuron's spiking activity correlates most closely with the animal's position on the track (Wilson & McNaughton, 1993). The data series we analyze consists of 691 spikes from a place cell in the CA1 region of the hippocampus recorded from a rat running back and forth for 20 minutes on a 300 cm U-shaped track. The track was linearized for the purposes of this analysis (Frank, Brown, & Wilson, 2000).

There are two approaches to estimating the place-specific firing maps of a hippocampal neuron. One approach is to use maximum likelihood to fit a specific parametric model of the spike times to the place cell data as in Brown, Frank, Tang, Quirk, and Wilson (1998) and Barbieri, Quirk, et al. (2001). If $x(t)$ is the animal's position at time t, we define the spatial function for the one-dimensional place field model as the gaussian surface

$$s(t) = \exp\left\{\alpha - \frac{\beta(x(t) - \mu)^2}{2}\right\}, \tag{3.2}$$

where μ is the center of place field, β is a scale factor, and $\exp\{\alpha\}$ is the maximum height of the place field at its center. We represent the spike time probability density of the neuron as either an inhomogeneous gamma (IG)

model, defined as

$$f(u_k \mid u_{k-1}, \theta) = \frac{\psi s(u_k)}{\Gamma(\psi)} \left[\int_{u_{k-1}}^{u_k} \psi s(u)\, du \right]^{\psi-1} \exp\left\{ - \int_{u_{k-1}}^{u_k} \psi s(u)\, du \right\}, \tag{3.3}$$

or as an inhomogeneous inverse gaussian (IIG) model, defined as

$$f(u_k \mid u_{k-1}, \theta) = \frac{s(u_k)}{\left[2\pi \left[\int_{u_{k-1}}^{u_k} s(u)\, du \right]^3 \right]^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} \frac{\left(\int_{u_{k-1}}^{u_k} s(u)\, du - \psi \right)^2}{\psi^2 \int_{u_{k-1}}^{u_k} s(u)\, du} \right\}, \tag{3.4}$$

where $\psi > 0$ is a location parameter for both models and $\theta = (\mu, \alpha, \beta, \psi)$ is the set of model parameters to be estimated from the spike train. If we set $\psi = 1$ in equation 3.3, we obtain the IP model as a special case of the IG model. The parameters for all three models—the IP, IG, and the IIG—can be estimated from the spike train data by maximum likelihood (Barbieri, Quirk, et al., 2001). The models in equations 3.3 and 3.4 are Markov so that the current value of either the spike time probability density or the conditional intensity (rate) function depends on only the time of the previous spike. Because of equation 2.2, specifying the spike time probability density is equivalent to specifying the conditional intensity function. If we let $\hat{\theta}$ denote the maximum likelihood estimate of θ, then the maximum likelihood estimate of the conditional intensity function for each model can be computed from equation 2.2 as

$$\lambda(t \mid H_t, \hat{\theta}) = \frac{f(t \mid u_{N(t)}, \hat{\theta})}{1 - \int_{u_{N(t)}}^{t} f(u \mid u_{N(t)}, \hat{\theta})\, du}. \tag{3.5}$$

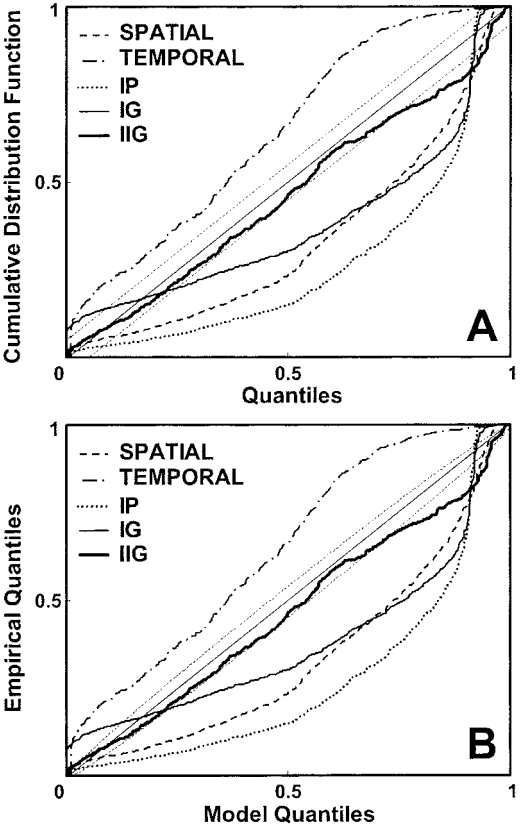
for $t > u_{N(t)}$. The estimated conditional intensity from each model may be used in the time-rescaling theorem to assess model goodness of fit as described in Section 2.1.

The second approach is to compute a histogram-based estimate of the conditional intensity function by using either spatial smoothing (Muller & Kubie, 1987; Frank et al., 2000) or temporal smoothing (Wood, Dudchenko, & Eichenbaum, 1999) of the spike train. To compute the spatial smoothing estimate of the conditional intensity function, we followed Frank et al. (2000) and divided the 300 cm track into 4.2 cm bins, counted the number of spikes per bin, and divided the count by the amount of time the animal spends in the bin. We smooth the binned firing rate with a six-point gaussian window with a standard deviation of one bin to reduce the effect of

running velocity. The spatial conditional intensity estimate is the smoothed spatial rate function. To compute the temporal rate function, we followed Wood et al. (1999) and divided the experiment into time bins of 200 msec and computed the rate as the number of spikes per 200 msec.

These two smoothing procedures produce histogram-based estimates of $\lambda(t)$. Both are histogram-based estimates of Poisson rate functions because neither the estimated spatial nor the temporal rate functions make any history-dependence assumption about the spike train. As in the analysis of the SEF spike trains, we again use the KS and Q-Q plots to compare directly goodness of fit of the five spike train models for the hippocampal place cells. The IP, IG, and IIG models were fit to the spike train data by maximum likelihood. The spatial and temporal rate models were computed as described. The KS and Q-Q plot goodness-of-fit comparisons are in Figure 2.

The IG model overestimates at lower quantiles, underestimates at intermediate quantiles, and overestimates at the upper quantiles (see Figure 2A).



The IP model underestimates the lower and intermediate quantiles and overestimates the upper quantiles. The KS plot of the spatial rate model is similar to that of the IP model yet closer to the 45-degree line. The temporal rate model overestimates the quantiles of the true probability model of the data. This analysis suggests that the IG, IP, and spatial rate models are most likely oversmoothing this spike train, whereas the temporal rate model undersmooths it. Of the five models, the one that is closest to the 45-degree line and lies almost entirely within the confidence bounds is the IIG. This model disagrees only with the experimental data near quantile 0.80. Because all of the models with the exception of the IIG have appreciable lack of fit in terms of the KS plots, the findings in the Q-Q plot analyses are almost identical (see Figure 2B). As in the KS plot, the Q-Q plot for the IIG model is close to the 45-degree line and within the 95% confidence bounds with the exception of quantiles near 0.80. These analyses suggest that IIG rate model gives the best agreement with the spiking activity of this pyramidal neuron. The spatial and temporal rate function models and the IP model have the greatest lack of fit.

3.3 Simulating a General Point Process by Time Rescaling. As a second application of the time-rescaling theorem, we describe how the theorem may be used to simulate a general point process. We stated in the Introduction that the time-rescaling theorem provides a standard approach for simulating an inhomogeneous Poisson process from a simple Poisson

Figure 2: *Facing page.* (A) KS plots of the parametric and histogram-based model fits to the hippocampal place cell spike train activity. The parametric models are the inhomogeneous Poisson (IP) (dotted line), the inhomogeneous gamma (IG) (thin, solid line), and the inhomogeneous inverse gaussian (IIG) (thick, solid line). The histogram-based models are the spatial rate model (dashed line) based on 4.2 cm spatial bin width and the temporal rate model (dashed and dotted line) based a 200 msec time bin width. The 45-degree solid line represents exact agreement between the model and the experimental data, and the 45-degree thin, solid lines are the 95% confidence bounds based on the KS statistic, as in Figure 1. The IIG model KS plot lies almost entirely within the confidence bounds, whereas all the other models show significant lack of fit. This suggests that the IIG model gives the best description of this place cell data series, and the histogram-based models agree least with the data. (B) Q-Q plots of the IP (dotted line), IG (thin, solid line), and IIG (thick, solid line) spatial (dashed line) temporal (dotted and dashed line) models. The 95% local confidence bounds (thin dashed lines) are computed as described in Figure 1. The findings in the Q-Q plot analyses are almost identical to those in the KS plots. The Q-Q plot for the IIG model is close to the 45-degree line and within the 95% confidence bounds, with the exception of quantiles near 0.80. This analysis also suggests that IIG rate model gives the best agreement with the spiking activity of this pyramidal neuron. The spatial and temporal rate function models and the IP model have the greatest lack of fit.

process. The general form of the time-rescaling theorem suggests that any point process with an integrable conditional intensity function may be simulated from a Poisson process with unit rate by rescaling time with respect to the conditional intensity (rate) function. Given an interval $(0, T]$, the simulation algorithm proceeds as follows:

1. Set $u_0 = 0$; Set $k = 1$.
2. Draw τ_k an exponential random variable with mean 1.
3. Find u_k as the solution to $\tau_k = \int_{u_{k-1}}^{u_k} \lambda(u \mid u_0, u_1, \ldots, u_{k-1})\, du$.
4. If $u_k > T$, then stop.
5. $k = k + 1$
6. Go to 2.

By using equation 2.3, a discrete version of the algorithm can be constructed as follows. Choose J large, and divide the interval $(0, T]$ into J bins each of width $\Delta = T/J$. For $k = 1, \ldots, J$ draw a Bernoulli random variable u_k^* with probability $\lambda(k\Delta \mid u_1^*, \ldots, u_{k-1}^*)\Delta$, and assign a spike to bin k if $u_k^* = 1$, and no spike if $u_k^* = 0$.

While in many instances there will be faster, more computationally efficient algorithms for simulating a point process, such as model-based methods for specific renewal processes (Ripley, 1987) and thinning algorithms (Lewis & Shedler, 1978; Ogata, 1981; Ross, 1993), the algorithm above is simple to implement given a specification of the conditional intensity function. For an example of where this algorithm is crucial for point process simulation, we consider the IG model in equation 3.2. Its conditional intensity function is infinite immediately following a spike if $\psi < 1$. If in addition, ψ is time varying ($\psi = \psi(t) < 1$ for all t), then neither thinning nor standard algorithms for making draws from a gamma probability distribution may be used to simulate data from this model. The thinning algorithm fails because the conditional intensity function is not bounded, and the standard algorithms for simulating a gamma model cannot be applied because ψ is time varying. In this case, the time-rescaling simulation algorithm may be applied as long as the conditional intensity function remains integrable as ψ varies temporally.

4 Discussion

Measuring how well a point process model describes a neural spike train data series is imperative prior to using the model for making inferences. The time-rescaling theorem states that any point process with an integrable conditional intensity function may be transformed into a Poisson process with a unit rate. Berman (1983) and Ogata (1988) showed that this theorem may be used to develop goodness-of-fit tests for point process models of seismologic data. Goodness-of-fit methods for neural spike train models based on

time-rescaling transformations but not the time-rescaling theorem have also been reported (Reich, Victor, & Knight, 1998; Barbieri, Frank, Quirk, Wilson, & Brown, 2001). Here, we have described how the time-rescaling theorem may be used to develop goodness-of-fit tests for point process models of neural spike trains.

To illustrate our approach, we analyzed two types of commonly recorded spike train data. The SEF data are a set of multiple short (400 msec) series of spike times, each measured under identical experimental conditions. These data are typically analyzed by a PSTH. The hippocampus data are a long (20 minutes) series of spike time recordings that are typically analyzed with either spatial or temporal histogram models. To each type of data we fit both parametric and histogram-based models. Histogram-based models are popular neural data analysis tools because of the ease with which they can be computed and interpreted. These apparent advantages do not override the need to evaluate the goodness of fit of these models. We previously used the time-rescaling theorem to assess goodness of fit for parametric spike train models (Olson et al., 2000; Barbieri, Quirk, et al., 2001; Ventura et al., 2001). Our main result in this article is that the time-rescaling theorem can be used to evaluate goodness of fit of parametric and histogram-based models and to compare directly the accuracy of models from the two classes. We recommend that before making an inference based on either type of model, a goodness-of-fit analysis should be performed to establish how well the model describes the spike train data. If the model and data agree closely, then the inference is more credible than when there is significant lack of fit. The KS and Q-Q plots provide assessments of overall goodness of fit. For the models fit by maximum likelihood, these assessments can be applied along with methods that measure the marginal value and marginal costs of using more complex models, such as Akaikie's information criterion (AIC) and the Bayesian information criterion (BIC), in order to gain a more complete evaluation of model agreement with experimental data (Barbieri, Quirk, et al., 2001).

We assessed goodness of fit by using the time-rescaling theorem to construct KS and Q-Q plots having, respectively, liberal and conservative confidence bounds. Together, the two sets of confidence bounds help characterize the range of agreement between the model and the data. For example, a model whose KS plot lies consistently outside the 95% KS confidence bounds (the IP model for the hippocampal data) agrees poorly with the data. On the other hand, a model that is within all the 95% confidence bounds of the Q-Q plots (the IMI model for the SEF data) agrees closely with the data. A model such as the IP model for the SEF data, that is, within nearly all the KS bounds, may lie outside the Q-Q plot intervals. In this case, if the lack of fit with respect to the Q-Q plot intervals is systematic (i.e., is over a set of contiguous quantiles), this suggests that the model does not fit the data well.

As a second application of the time-rescaling theorem, we presented an algorithm for simulating spike trains from a point process given its condi-

tional intensity (rate) function. This algorithm generalizes the well-known technique of simulating an inhomogeneous Poisson process by rescaling a Poisson process with a constant rate.

Finally, to make the reasoning behind the time-rescaling theorem more accessible to the neuroscience researchers, we proved its general form using elementary probability arguments. While this elementary proof is most certainly apparent to experts in probability and point process theory (D. Brillinger, personal communication; Guttorp, 1995) its details, to our knowledge, have not been previously presented. The original proofs of this theorem use measure theory and are based on the martingale representation of point processes (Meyer, 1969; Papangelou, 1972; Brémaud, 1981; Jacobsen, 1982). The conditional intensity function (see equation 2.1) is defined in terms of the martingale representation. Our proof uses elementary arguments because it is based on the fact that the joint probability density of a set of point process observations (spike train) has a canonical representation in terms of the conditional intensity function. When the joint probability density is represented in this way, the Jacobian in the change of variables between the original spike times and the rescaled interspike intervals simplifies to a product of the reciprocals of the conditional intensity functions evaluated at the spike times.

The proof also highlights the significance of the conditional intensity function in spike train modeling; its specification completely defines the stochastic structure of the point process. This is because in a small time interval, the product of the conditional intensity function and the time interval defines the probability of a spike in that interval given the history of the spike train up to that time (see equation 2.3). When there is no history dependence, the conditional intensity function is simply the Poisson rate function. An important consequence of this simplification is that unless history dependence is specifically included, then histogram-based models, such as the PSTH, and the spatial and temporal smoothers are implicit Poisson models.

In both the SEF and hippocampus examples, the histogram-based models gave poor fits to the spike train. These poor fits arose because these models used few data points to estimate many parameters and because they do not model history dependence in the spike train. Our parametric models used fewer parameters and represented temporal dependence explicitly as Markov. Point process models with higher-order temporal dependence have been studied by Ogata (1981, 1988), Brillinger (1988), and Kass and Ventura (2001) and will be considered further in our future work. Parametric conditional intensity functions may be estimated from neural spike train data in any experiments where there are enough data to estimate reliably a histogram-based model. This is because if there are enough data to estimate many parameters using an inefficient procedure (histogram/method of moments), then there should be enough data to estimate a smaller number of parameters using an efficient one (maximum likelihood).

Using the K-S and Q-Q plots derived from the time-rescaling theorem, it is possible to devise a systematic approach to the use of histogram-based

models. That is, it is possible to determine when a histogram-based model accurately describes a spike train and when a different model class, temporal dependence, or the effects of covariates (e.g., the theta rhythm and the rat running velocity in the case of the place cells) should be considered and the degree of improvement the alternative models provide.

In summary, we have illustrated how the time-rescaling theorem may be used to compare directly goodness of fit of parametric and histogram-based point process models of neural spiking activity and to simulate spike train data. These results suggest that the time-rescaling theorem can be a valuable tool for neural spike train data analysis.

Appendix

A.1 Maximum Likelihood Estimation of the IMI Model. To fit the IMI model, we choose J large, and divide the interval $(0, T]$ into J bins of width $\Delta = T/J$. We choose J so that there is at most one spike in any bin. In this way, we convert the spike times $0 < u_1 < u_2 <, \ldots, < u_{n-1} < u_n \leq T$ into a binary sequence u_j^*, where $u_j^* = 1$ if there is a spike in bin j and 0 otherwise for $j = 1, \ldots, J$. By definition of the conditional intensity function for the IMI model in equation 3.1, it follows that each u_j^* is a Bernoulli random variable with the probability of a spike at $j\Delta$ defined as $\lambda_1(j\Delta \mid \theta)\lambda_2(j\Delta - u^*_{N(j\Delta)} \mid \theta)\Delta$. We note that this discretization is identical to the one used to construct the discretized version of the simulation algorithm in Section 3.3. The log probability of a spike is thus

$$\log(\lambda_1(j\Delta \mid \theta)) + \log(\lambda_2(j\Delta - u^*_{N(j\Delta)} \mid \theta)\Delta), \tag{A.1}$$

and the cubic spline models for $\log(\lambda_1(j\Delta \mid \theta))$ and $\log(\lambda_2(j\Delta - u^*_{N(j\Delta)} \mid \theta))$ are, respectively,

$$\begin{aligned}\log(\lambda_1(j\Delta \mid \theta)) = &\sum_{l=1}^{3} \theta_l (j\Delta - \xi_1)_+^l \\ &+ \theta_4 (j\Delta - \xi_2)_+^3 + \theta_5 (j\Delta - \xi_3)_+^3\end{aligned} \tag{A.2}$$

$$\begin{aligned}\log(\lambda_2(j\Delta - u^*_{N(j\Delta)} \mid \theta)) = &\sum_{l=1}^{3} \theta_{l+5} (j\Delta - u^*_{N(j\Delta)} - \gamma_1)_+^l \\ &+ \theta_9 (j\Delta - u^*_{N(j\Delta)} - \gamma_2)_+^3,\end{aligned} \tag{A.3}$$

where $\theta = (\theta_1, \ldots, \theta_9)$ and the knots are defined as $\xi_l = lT/4$ for $l = 1, 2, 3$ and γ_l is the observed $100l/3$ quantile of the interspike interval distribution of the trial spike trains for $l = 1, 2$. We have for Section 3.1 that $\theta_{IP} = (\theta_1, \ldots, \theta_5)$

and $\theta_{IMI} = (\theta_1, \ldots, \theta_9)$ are the coefficients of the spline basis elements for the IP and IMI models, respectively.

The parameters are estimated by maximum likelihood with $\Delta = 1$ msec using the **gam** function in S-PLUS (MathSoft, Seattle) as described in Chapter 11 of Venables and Ripley (1999). The inputs to **gam** are the u_j^*s, the time arguments $j\Delta$ and $j\Delta - u^*_{N(j\Delta)}$, the symbolic specification of the spline models, and the knots. Further discussion on spline-based regression methods for analyzing neural data may be found in Olson et al. (2000) and Ventura et al. (2001).

Acknowledgments

We are grateful to Carl Olson for allowing us to use the supplementary eye field data and Matthew Wilson for allowing us to use the hippocampal place cell data. We thank David Brillinger and Victor Solo for helpful discussions and the two anonymous referees whose suggestions helped improve the exposition. This research was supported in part by NIH grants CA54652, MH59733, and MH61637 and NSF grants DMS 9803433 and IBN 0081548.

References

Barbieri, R., Frank, L. M., Quirk, M. C., Wilson, M. A., & Brown, E. N. (2001). Diagnostic methods for statistical models of place cell spiking activity. *Neurocomputing, 38–40,* 1087–1093.

Barbieri, R., Quirk, M. C., Frank, L. M., Wilson, M. A., & Brown, E. N. (2001). Construction and analysis of non-Poisson stimulus-response models of neural spike train activity. *J. Neurosci. Meth., 105,* 25–37.

Berman, M. (1983). Comment on "Likelihood analysis of point processes and its applications to seismological data" by Ogata. *Bulletin Internatl. Stat. Instit., 50,* 412–418.

Brémaud, P. (1981). *Point processes and queues: Martingale dynamics.* Berlin: Springer-Verlag.

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

Brown, E. N., Frank, L. M., Tang, D., Quirk, M. C., & Wilson, M. A. (1998). A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *Journal of Neuroscience, 18,* 7411–7425.

Casella, G., & Berger, R. L. (1990). *Statistical inference.* Belmont, CA: Duxbury.

Daley, D., & Vere-Jones, D. (1988). *An introduction to the theory of point process.* New York: Springer-Verlag.

Frank, L. M., Brown, E. N., & Wilson, M. A. (2000). Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron, 27,* 169–178.

Guttorp, P. (1995). *Stochastic modeling of scientific data.* London: Chapman & Hall.

Hogg, R. V., & Tanis, E. A. (2001). *Probability and statistical inference* (6th ed.). Englewood Cliffs, NJ: Prentice Hall.

Jacobsen, M. (1982). *Statistical analysis of counting processes*. New York: Springer-Verlag.

Johnson, A., & Kotz, S. (1970). *Distributions in statistics: Continuous univariate distributions—2*. New York: Wiley.

Kalbfleisch, J., & Prentice, R. (1980). *The statistical analysis of failure time data*. New York: Wiley.

Karr, A. (1991). *Point processes and their statistical inference* (2nd ed.). New York: Marcel Dekker.

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

Lewis, P. A. W., & Shedler, G. S. (1978). Simulation of non-homogeneous Poisson processes by thinning. *Naval Res. Logistics Quart., 26*, 403–413.

Meyer, P. (1969). Démonstration simplifiée d'un théorème de Knight. In *Séminaire probabilité V* (pp. 191–195). New York: Springer-Verlag.

Muller, R. U., & Kubie, J. L. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *Journal of Neuroscience, 7*, 1951–1968.

Ogata, Y. (1981). On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory, IT-27*, 23–31.

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

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*, 171–175.

Olson, C. R., Gettner, S. N., Ventura, V., Carta, R., & Kass, R. E. (2000). Neuronal activity in macaque supplementary eye field during planning of saccades in response to pattern and spatial cues. *Journal of Neurophysiology, 84*, 1369–1384.

Papangelou, F. (1972). Integrability of expected increments of point processes and a related random change of scale. *Trans. Amer. Math. Soc., 165*, 483–506.

Port, S. (1994). *Theoretical probability for applications*. New York: Wiley.

Protter, M. H., & Morrey, C. B. (1991). *A first course in real analysis* (2nd ed.). New York: Springer-Verlag.

Reich, D., Victor, J., & Knight, B. (1998). The power ratio and the interval map: Spiking models and extracellular recordings. *Journal of Neuroscience, 18*, 10090–10104.

Ripley, B. (1987). *Stochastic simulation*. New York: Wiley.

Ross, S. (1993). *Introduction to probability models* (5th ed.). San Diego, CA: Academic Press.

Snyder, D., & Miller, M. (1991). *Random point processes in time and space* (2nd ed.). New York: Springer-Verlag.

Taylor, H. M., & Karlin, S. (1994). *An introduction to stochastic modeling* (rev. ed.). San Diego, CA: Academic Press.

Venables, W., & Ripley, B. (1999). *Modern applied statistics with S-PLUS* (3rd ed.). New York: Springer-Verlag.

Ventura, V., Carta, R., Kass, R., Gettner, S., & Olson, C. (2001). Statistical analysis of temporal evolution in single-neuron firing rate. *Biostatistics*. In press.

Wilson, M. A., & McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science, 261*, 1055–1058.

Wood, E. R., Dudchenko, P. A., & Eichenbaum, H. (1999). The global record of memory in hippocampal neuronal activity. *Nature, 397*, 613–616.

Received October 27, 2000; accepted May 2, 2001.