Modeling neural activity using the generalized inverse Gaussian distribution

Satish Iyengar[1], **Qiming Liao**[2]

[1] Department of Statistics, University of Pittsburgh, Pittsburgh, PA 15260, USA
[2] Department of Biostatistics, Harvard University, Cambridge, Massachusetts, USA

Abstract. Spike trains from neurons are often used to make inferences about the underlying processes that generate the spikes. Random walks or diffusions are commonly used to model these processes; in such models, a spike corresponds to the first passage of the diffusion to a boundary, or firing threshold. An important first step in such a study is to fit families of densities to the trains' interspike interval histograms; the estimated parameters, and the families' goodness of fit can then provide information about the process leading to the spikes. In this paper, we propose the generalized inverse Gaussian family because its members arise as first passage time distributions of certain diffusions to a constant boundary. We provide some theoretical support for the use of these diffusions in neural firing models. We compare this family with the lognormal family, using spike trains from retinal ganglion cells of goldfish, and simulations from an integrate-and-fire and a dynamical model for generating spikes. We show that the generalized inverse Gaussian family is closer to the true model in all these cases.

1 Introduction

The analysis of neural spike trains has a long history (Tuckwell 1988). The study of spike trains has several purposes: for example, to draw inferences about the nature of the inputs to the neuron, and the processes within the neuron that convert those inputs into impulses. There are two main approaches for such analyses. The first is to formulate a stochastic model for the neuron's activity and inputs, derive the distribution of the interspike interval (ISI), and use it to fit the model to data (Gerstein and Mandelbrot 1964; Inoue et al. 1995). The stochastic process is often modeled as a random walk or diffusion, and the spike corresponds to its first passage time to a boundary, or firing threshold. When such a stochastic model is not known, a second approach is to fit standard families of distributions, such as the lognormal or gamma, using maximum likelihood or some ad hoc technique. In this case, inferences about the underlying

process are more difficult, and at times not attempted at all (Tuckwell 1988; Levine 1991).

In this paper, we study the adequacy of the generalized inverse Gaussian family for ISI data. We propose this family for several reasons. First, for certain values of the parameters of this family, its members are the first passage time distributions of certain diffusion processes to a constant boundary. Also, it is an exponential family which contains commonly used subfamilies such as the gamma, inverse Gaussian, and their reciprocals. Using spike trains from retinal ganglion cells of goldfish, we compare this family with the lognormal family, which is commonly used to fit ISI histograms. For problems of this sort, it is typically not known that the true model (the one that represents the process which generates the data) belongs to one of the families under consideration. We therefore use a quasi-maximum likelihood approach to identify the family that is closer to the true model, using the Kullback-Leibler information criterion (Nishii 1988). We also apply this technique to spike trains simulated from an integrate-and-fire model, and from a dynamical model.

In Sect. 2, we present the spike train data, give the details of the two families, and describe the neurophysiological significance of the diffusion processes underlying the generalized inverse Gaussian family. We then describe the method of model comparison, and derive the test statistics needed. In Sect. 3, we give the results of the tests, and show that the generalized inverse Gaussian family is the closer model for all cases. We then conclude with a discussion of our findings in Sect. 4.

2 Materials and methods

2.1 Spike train data

The simulated data are from two models of spike generation. The first is Stein's integrate-and-fire model (Tuckwell 1988). We use noise uniformly distributed on the interval [0, 1] to perturb a mean bias signal to create the integrated variable. The summary statistics of the noise, i.e. [mean, standard deviation, skewness, kurtosis], for the two runs are [0.501, 0.288, 0.001, −1.21] and [0.501, 0.289, 0.001, −1.18], respectively; these values are close to their expected values. The integration proceeds at 1-ms time steps until a selected firing threshold is reached. This model uses a leaky integrator with time constant $\tau = 5$ ms, and approximately 30 000 noise inputs

Correspondence to: S. Iyengar (Tel: +1 412 624 8341,
Fax: +1 412 624 8397, e-mail: si@stat.pitt.edu)

Table 1. Summary statistics for integrate-and-fire simulations

Run:	1	2
Spikes:	1159	1102
Mean:	25.9	26.4
Median:	22.0	23.0
SD:	15.9	11.9
CV:	0.61	0.45

Table 2. Summary statistics for dynamical model simulations

Run:	1	2
Spikes:	6383	6364
Mean:	3.13	3.14
Median:	2.89	2.91
SD:	1.06	1.02
CV:	0.34	0.32

to produce more than 1100 simulated spikes. Table 1 contains summary statistics for these two simulated spike trains. The ISI means are around 26 ms, the ISI histograms are skewed to the right, and their coefficients of variation (CVs) are 0.45 and 0.61. Figure 1 shows a smoothed histogram of the latter, using a 1-ms binwidth.

The second set of simulated data is from two runs of a dynamical model which is a variant of a model due to Morris and Lecar 1981; for background and a description of the simulation methods. see Ermentrout 1996; Ermentrout and Rinzel 1989; Gutkin and Ermentrout 1997. This dynamical approach uses a phase variable ϕ to model the location along the spike trajectory in the phase plane. The phase evolves according to the stochastic differential equation

$$d\phi = [\beta(1 + \cos\phi) + (1 - \cos\phi)]\,dt + \sigma(1 + \cos\phi)\,dW_t$$

for $0 \leq \phi \leq 2\pi$. Here, dW_t designates white noise, σ is the noise standard deviation, and β is the bias parameter which models the strength of the input current; a spike occurs every time ϕ passes through π. In the simulations we used $\beta = \sigma = 1$. Table 2 contains summary statistics for these two simulated spike trains. In each spike train, there are about 6375 spikes, the ISI means are around 3.1, and the ISI histograms are slightly skewed to the right. These two data sets have low CVs, both around 0.33. Figure 2 shows a smoothed histogram for one of the data sets, using a 0.06-ms binwidth.

The physiological data are maintained discharges of retinal ganglion cells of goldfish. These data were provided by Dr. Michael Levine of the University of Illinois; he has detailed his laboratory methods elsewhere (Levine 1991; Levine et al. 1988), so we give a brief summary here. The retinae were removed from the goldfish, and maintained in a flow of moist oxygen. Recordings of the retinal ganglion cells were made with an extracellular microelectrode. A level detector signaled the nerve impulses, whose time of occurrence was recorded to an accuracy of within 1 ms. Continuous blocks of data were recorded for at least 20 s in the absence of any change in the illumination of the retina. When the illumination was changed, enough time was allowed so that the behavior appeared stationary. For this paper, we have used data from ten different cells, labeled 68, 72, 75, 78, 80, 82, 83, 87, 90, and 98; Table 3 contains summary statistics for each cell. The recording time for each cell was 30 s. The number of spikes ranges from 703 to 1028. The ISI means range from 29.1 to 42.7 ms. The ISI histograms are all skewed to the right, and the CVs all exceed 1. Figures 3 and 4 show representative smoothed histograms for the data from cells 75 and 78, respectively.

2.2 Models for the ISI distribution

We compare the following two models for the ISI histograms: the lognormal and generalized inverse Gaussian. For $y \geq 0$, the lognormal has density

$$f(y|\alpha) = \frac{1}{y\sqrt{2\pi\alpha_2}} \exp\left[-\frac{(\log y - \alpha_1)^2}{2\alpha_2}\right] \tag{1}$$

where $-\infty < \alpha_1 < \infty$ and $\alpha_2 > 0$. The generalized inverse Gaussian has density

$$h(y|\psi, \chi, \lambda) = \frac{(\psi/\chi)^{\lambda/2}}{2K_\lambda(\sqrt{\psi\chi})}\, y^{\lambda-1} \exp\left[-\frac{1}{2}(\psi y + \frac{\chi}{y})\right], \tag{2}$$

where K_λ is the modified Bessel function of the third kind with index λ (Abramowitz and Stegun 1965),

$$K_\lambda(u) = \frac{1}{2}\int_0^\infty x^{\lambda-1} \exp\left[-\frac{u}{2}(x + \frac{1}{x})\right]dx$$

Let $\theta = (\psi, \chi, \lambda)$, and let Θ be the parameter space. Then for $-\infty < \lambda < \infty$, $\Theta = \bigcup_\lambda \Theta_\lambda$, where

$$\Theta_\lambda = \begin{cases} \{(\psi, \chi) : \psi > 0, \chi \geq 0\} \text{ if } \lambda > 0 \\ \{(\psi, \chi) : \psi > 0, \chi > 0\} \text{ if } \lambda = 0 \\ \{(\psi, \chi) : \psi \geq 0, \chi > 0\} \text{ if } \lambda < 0. \end{cases}$$

It is convenient to reparametrize this model thus: $\omega = \sqrt{\psi\chi}$ and $\eta = \sqrt{\chi/\psi}$; η is a scale parameter and ω is a concentration parameter when $\omega > 0$. When $\omega = 0$, these interpretations are no longer valid. In that case, it is more appropriate to regard (2) as a gamma or reciprocal gamma density; the norming constants are obtained from the asymptotic formula for $\lambda > 0$ and $\omega \downarrow 0$:

$$K_\lambda(\omega) \sim \frac{\Gamma(\lambda)2^{\lambda-1}}{\omega^\lambda}$$

and the relation $K_\lambda(\omega) = K_{-\lambda}(\omega)$. Below, we express the test statistics in terms of the estimates of $\beta = (\omega, \eta, \lambda)$ rather than $\theta = (\psi, \chi, \lambda)$.

The lognormal model has been used before to fit spike train data from the cat cerebral cortex and other preparations (Tuckwell 1988). Also, Levine (1991) compared the lognormal with the hyperbolic normal, hyperbolic gamma, and inverse Gaussian [which has $\lambda = -\frac{1}{2}$ in (2) above] using ad hoc curve fitting methods. He argued that the lognormal was the preferred model among them. On the other hand, the generalized inverse Gaussian family has only sporadically been applied in the neuroscience context. Jorgensen (1982) used it to study a single spike train, and Lánský et al. (1990) suggested its use, but a systematic study of this family in this context has not been done.

We propose the generalized inverse Gaussian model for several reasons. First, diffusions are commonly used to model the underlying processes of neurons which lead to their firing. For $\lambda \leq 0$, Barndorff-Nielsen et al. (1978) have shown that the generalized inverse Gaussian distribution arises as the first passage time of a diffusion (with certain infinitesimal drift and variance parameters) to a constant boundary. Thus, the estimates of the parameters will inform us about possible mechanisms leading to the spike train. Next, it contains as special cases models that have been proposed for ISI distributions, such as the exponential, gamma, reciprocal gamma, and inverse Gaussian; thus, it allows for a unified approach to the analysis of ISI histograms. Finally, the generalized inverse Gaussian has the following statistical properties (Jorgensen 1982): It is an exponential family, so it is an analytically convenient and flexible family for model selection. All its members are unimodal with a positive mode unless $\chi = 0$ and $0 < \lambda \leq 1$, in which case the density is infinite at zero. The range of CVs for this family spans all positive numbers; the correspondence is not one-to-one, as many values of (ψ, χ, λ) give rise to the same CV. Based on extensive numerical work, it is conjectured, but not yet proven, that all members of this family have positive skewness and kurtosis.

2.3 The underlying diffusions

In this section, we describe the diffusions that give rise to the generalized inverse Gaussian family, and then discuss their potential neurophysiological significance. Suppose that $\theta = (\psi, \chi, \lambda) \in \Theta$, and $\lambda \leq 0$ (so that $\chi > 0$). For $x > 0$, let $\sigma^2(x)$ be a strictly positive differentiable function, and suppose that

$$\gamma(x) = \int_0^x \frac{du}{\sigma(u)}$$

is finite for all $x > 0$, and that $\gamma(x) \to \infty$ as $x \to \infty$. Also, let

Table 3. Summary statistics for the ten goldfish retinal ganglion cells

Cell:	68	72	75	78	80	82	83	87	90	98
Spikes:	814	811	823	975	984	1019	1028	1020	998	703
Mean:	36.9	36.9	36.4	30.8	30.4	29.4	29.1	29.4	30.0	42.7
Median:	15.0	18.0	16.0	10.0	10.0	12.0	12.0	13.0	13.0	20.0
SD:	39.6	37.6	40.7	38.8	37.1	33.1	33.5	31.3	32.9	44.0
CV:	1.1	1.0	1.1	1.3	1.2	1.1	1.2	1.1	1.1	1.0

$$\mu(x) = \begin{cases} -\sigma(x)\left[\dfrac{2\lambda-1}{2\gamma(x)} + \sqrt{\psi}\,\dfrac{K_{\lambda-1}(\gamma(x)\sqrt{\psi})}{K_{\lambda}(\gamma(x)\sqrt{\psi})}\right] + \dfrac{1}{4}\dfrac{d}{dx}\sigma^2(x) \\ \quad \text{for } \psi > 0, \lambda \leq 0, \\ \sigma(x)\dfrac{2\lambda+1}{2\gamma(x)} + \dfrac{1}{4}\dfrac{d}{dx}\sigma^2(x) \\ \quad \text{for } \psi = 0, \lambda < 0 \end{cases}$$

The result of Barndorff-Nielsen et al. (1978) is the following:

Theorem. Let $\{X(t) : t \geq 0\}$ be a time-homogeneous diffusion process with state space $[0, \infty)$, infinitesimal variance $\sigma^2(x)$, and infinitesimal mean $\mu(x)$ given above. Let $x_0 > 0$ satisfy $\gamma^2(x_0) = \chi$. If $\lambda \leq 0$, then the first passage time $T = \inf\{t > 0 : X(t) = 0\}$ has the generalized inverse Gaussian distribution with parameter (ψ, χ, λ).

The ordinary inverse Gaussian was first used by Gerstein and Mandelbrot (1964); its parameters $(\psi, \chi, -\frac{1}{2})$ correspond to $\sigma^2(x) \equiv \sigma^2$, $\mu(x) \equiv -\sigma\sqrt{\psi}$, and $x_0 = -\sigma\sqrt{\chi}$. This elementary example emphasizes the fact that many choices of $\sigma^2(x)$ yield the same first passage time distribution. This non-uniqueness allows for a wide range of diffusions to be covered by this theorem. For instance, Barndorff-Nielsen et al. (1978) showed that the reciprocal gamma arises from two diffusions with distinct infinitesimal parameters: one with constant $\sigma^2(x)$ and $\mu(x)$ proportional to x^{-1}, and another with $\sigma^2(x)$ a power of x and $\mu(x) \equiv 0$.

Given any diffusion $\{X(t) : t > 0\}$ with infinitesimal parameters $\mu(x)$ and $\sigma^2(x)$, we now describe a general method of constructing other diffusions leading to same first passage time distribution. For any function $g(x)$ that is strictly monotone and differentiable for $x > 0$, let $\tilde{X}(t) = g(X(t))$ be a transformed process. Its state space is the range of g, and its infinitesimal mean and variance are

$$\tilde{\mu}(y) = \frac{1}{2}\sigma^2(x)g''(x) + \mu(x)g'(x)$$

and

$$\tilde{\sigma}^2(y) = \sigma^2(x)[g'(x)]^2$$

respectively, where $y = g(x)$ (Karlin and Taylor 1981). Since g is strictly monotone,

$$T = \inf\{t > 0 : X(t) = 0\} = \inf\{t > 0 : \tilde{X}(t) = g(0)\}$$

when $X(0) = x_0$ and $\tilde{X}(0) = g(x_0)$, the same generalized inverse Gaussian distribution arises from two distinct infinitesimal mean and variance functions.

For example, the values $\lambda = -\frac{1}{2}$ and $\psi = 0$ can lead to $\sigma^2(x) \equiv 1$ and $\mu(x) = 0$ for $x > 0$. Applying the transformation $g(x) = e^{-x}$, we have $\tilde{\sigma}^2(y) = y^2$ and $\tilde{\mu}(y) = y/2$ for the transformed state space $0 < y \leq 1$. Thus, the ordinary inverse Gaussian distribution is applicable to this diffusion with more complicated infinitesimal parameters. Several models of neural firing that include reversal potentials or shunting inhibition lead to a linear infinitesimal mean, and quadratic infinitesimal variance on a bounded state space (see, for example, Hanson and Tuckwell 1983; Lánský et al. 1990). Thus, a subclass of the generalized inverse Gaussian family arises out of diffusions that have been independently derived from models for neural firing. Of course, some elementary transformations are needed to depict more realistically the state space, which for the Hanson and Tuckwell model is $\{y : V_I < y \leq V_E\}$, where V_I is the inhibitory reversal potential, and V_E is the excitatory reversal potential (which exceeds the firing threshold).

2.4 Model comparison method

In many cases, there is not enough information to assume that the candidate families contain the true model that is generating the data. This is certainly the case for the problem at hand, so we leave the true model for the ISI histograms unspecified; using the data, we then seek to identify the model which is closer to the unknown true model. The approach here is based on the results of Foutz and Srivastava (1977), Nishii (1988) and White (1982). We now describe the approach in general, and then specialize to the case of the lognormal versus generalized inverse Gaussian families at the end of this subsection.

To assess the closeness of two densities f and p on the real line, we use the Kullback-Leibler information criterion (or divergence)

$$I[p : f] = \int_0^\infty p(x) \log \frac{p(x)}{f(x)}\, dx$$

It is known that $I[p : f] \geq 0$, with equality only when $p = f$ (Kullback 1968; White 1994). For example, if $p(x) = \mu e^{-\mu x}$ and $f(x) = \lambda e^{-\lambda x}$ for $x > 0$ and zero otherwise, then $I[p : f] = \lambda/\mu - 1 - \log(\lambda/\mu)$, which is 0.09 when the ratio of the means of these exponential densities is $\lambda/\mu = 1.5$. Next, if p and f are normal densities with means μ and ν, respectively, and both variances equal to 1, then $I[p : f] = (\mu - \nu)^2/2$, which is 0.125 when $|\mu - \nu| = 0.5$.

The Kullback-Leibler information criterion is not a metric, but it does have an advantage over metrics for the problem of comparing two approximations f_1 and f_2 to an unknown density p. Specifically,

$$I[p : f_1] - I[p : f_2] = \int_0^\infty p(x) \log \frac{f_2(x)}{f_1(x)}$$
$$= E_p\left[\log \frac{f_2(Y)}{f_1(Y)}\right] \tag{3}$$

where E_p refers to expectation with respect to the density p, which generates the random variable Y. In practice, given independent and identically distributed data $(Y_1, \ldots, Y_n)$ from p, an estimate of this quantity is

$$\frac{1}{n}\sum_1^n \log \frac{f_2(Y_i)}{f_1(Y_i)}$$

If this estimate is positive, we conclude that f_2 is closer to p than f_1.

Next, let

$$\mathscr{F} = \{f(y|\alpha): \alpha \in \mathscr{A}\}$$

and

$$\mathscr{G} = \{g(y|\beta): \beta \in \mathscr{B}\}$$

be two families of densities under consideration, and let $p(y)$ denote the true unknown density that is generating the data. Let α_* be the quasi-true parameter in $\mathscr{A}$ associated with the true density $p(y)$: $f(y|\alpha_*)$ is the member of $\mathscr{F}$ that minimizes $I[p : f(\cdot|\alpha)]$, over $\alpha \in \mathscr{A}$, and define $\beta_* \in \mathscr{G}$ similarly. Thus,

$$I_{\mathscr{F}} = I_{\mathscr{F}}[p : \alpha_*] = \int_0^\infty p(y) \log \frac{p(y)}{f(y|\alpha_*)}\, dy$$

characterizes the proximity of $\mathscr{F}$ to p, and $I_{\mathscr{G}} = I_{\mathscr{G}}[p : \beta_*]$ (defined similarly) characterizes the proximity of $\mathscr{G}$ to p. Thus, the difference

$$I_{\mathscr{G}} - I_{\mathscr{F}} = \int_0^\infty p(y) \log \frac{f(y|\alpha_*)}{g(y|\beta_*)}\, dy$$

allows for a comparison of $\mathscr{F}$ and $\mathscr{G}$, using estimates of α_* and β_*, which we describe next.

Given a random sample $Y = (Y_1, \ldots, Y_n)$ generated from $p(y)$, let the quasi-log-likelihood function of α under $\mathscr{F}$ be

$$L_n(\alpha|Y) = \sum_{i=1}^{n} \log f(Y_i|\alpha)$$

and a quasi-maximum likelihood estimator (QMLE) $\hat{\alpha}_n$ be a value which maximizes $L_n(\alpha|Y)$ over $\alpha \in \mathscr{A}$; for $\mathscr{G}$, define $L_n(\beta|Y)$ and $\hat{\beta}_n$ similarly.

A natural estimate of $I_{\mathscr{G}} - I_{\mathscr{F}}$ is

$$T_n = \frac{1}{n} \sum_{i=1}^{n} \log \frac{f(Y_i|\hat{\alpha}_n)}{g(Y_i|\hat{\beta}_n)}$$

which is asymptotically normal:

$$\sqrt{n}[T_n - (I_{\mathscr{G}} - I_{\mathscr{F}})] \xrightarrow{d} N(0, \sigma^2) \tag{4}$$

where the variance $\sigma^2 = E_p \left[\log f(Y|\alpha_*)/g(Y|\beta_*) \right]^2$ is estimated by

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^{n} \left[\log \frac{f(Y_i|\hat{\alpha}_n)}{g(Y_i|\hat{\beta}_n)} \right]^2 \tag{5}$$

The regularity conditions that guarantee the existence of the quasi-true parameters α_* and β_*, the consistency of the QMLEs, the asymptotic normality of T_n, and the consistency of $\hat{\sigma}_n^2$ are described in the papers by Foutz and Srivastava (1977), White (1982) and Nishii (1988). These conditions are satisfied for the comparison below; we omit the details of that verification.

The decision about which of the families, $\mathscr{F}$ or $\mathscr{G}$, is closer to the unknown true model $p(y)$, now proceeds in the following way. We first pose the null hypothesis $H_0: I_{\mathscr{G}} = I_{\mathscr{F}}$ against the alternative hypothesis $H_1: I_{\mathscr{G}} \neq I_{\mathscr{F}}$. We then compute the statistic $\sqrt{n}T_n$ and its estimated standard deviation $\hat{\sigma}_n$. Given a confidence coefficient l, we construct the approximate confidence interval for $(I_{\mathscr{G}} - I_{\mathscr{F}})$,

$$I_{n,l} = \left(T_n - c_l \frac{\hat{\sigma}_n}{\sqrt{n}}, T_n + c_l \frac{\hat{\sigma}_n}{\sqrt{n}} \right)$$

where c_l is the upper $100(1 - l/2)$-percent point of the standard normal distribution (if $l = 0.95$, $c_l = 1.96$). If $0 \in I_{n,l}$, then we cannot reject the null hypothesis that the two models are equally close to the true model; otherwise we reject the null hypothesis. Furthermore, if $I_{n,l} \subseteq (0, \infty)$, we conclude that $\mathscr{F}$ is closer to the true model; if $I_{n,l} \subseteq (-\infty, 0)$, we conclude that $\mathscr{G}$ is closer to the true model.

We now turn to the specifics for the lognormal and generalized inverse Gaussian families. Let $\mathscr{F}$ and $\mathscr{G}$ refer to the lognormal and generalized inverse Gaussian family, respectively. The QMLEs $\hat{\alpha}_n = (\hat{\alpha}_{1n}, \hat{\alpha}_{2n})$ of the lognormal parameters are

$$\hat{\alpha}_{1n} = \frac{1}{n} \sum_{1}^{n} \log Y_i$$

and

$$\hat{\alpha}_{2n} = \frac{1}{n} \sum_{1}^{n} [\log Y_i - \hat{\alpha}_{1n}]^2$$

respectively. The QMLEs $\hat{\beta}_n = (\hat{\eta}_n, \hat{\omega}_n, \hat{\lambda}_n)$ of the generalized inverse Gaussian parameters are not analytically tractable, so we used IMSL (Visual Numerics 1994) to evaluate them numerically, by solving the likelihood equations described in Jorgensen (1982). Let

$$\bar{Y} = \frac{1}{n} \sum_{1}^{n} Y_i \quad \text{and} \quad \tilde{Y} = \frac{1}{n} \sum_{1}^{n} \frac{1}{Y_i}$$

Then,

Table 4. Comparison of the two runs of the integrate-and-fire model and QMLEs for the generalized inverse Gaussian model

Run:	1	2
Comparison of the two families		
$\sqrt{n}T_n$	−0.33	−0.45
$\hat{\sigma}_n$	0.14	0.11
QMLEs for the generalized inverse Gaussian model		
$\hat{\eta}_n$	46.66	51.18
$\hat{\omega}_n$	2.44	2.78
$\hat{\lambda}_n$	−2.26	−2.87

$$T_n = \hat{\lambda}_n \log \hat{\eta}_n + \log[2K_{\hat{\lambda}_n}(\hat{\omega}_n)] + \frac{\hat{\omega}_n}{2}[\hat{\eta}_n^{-1}\bar{Y} + \hat{\eta}_n\tilde{Y}]$$
$$-\hat{\lambda}_n\hat{\alpha}_{1n} - \log\left[\sqrt{2\pi\hat{\alpha}_{2n}}\right] - \frac{1}{2} \tag{6}$$

and

$$\hat{\sigma}_n^2 = \left[\hat{\lambda}_n \log \hat{\eta}_n + \log(2K_{\hat{\lambda}_n}(\hat{\omega}_n)) - \log\sqrt{2\pi\hat{\alpha}_{2n}} \right]^2 \tag{7}$$
$$+ \left[\hat{\lambda}_n \log \hat{\eta}_n + \log(2K_{\hat{\lambda}_n}(\hat{\omega}_n)) - \log\sqrt{2\pi\hat{\alpha}_{2n}} \right]$$
$$\times \left[\hat{\omega}_n(\hat{\eta}_n^{-1}\bar{Y} + \hat{\eta}_n\tilde{Y}) - 2\hat{\lambda}_n\hat{\alpha}_{1n} - 1 \right]$$
$$+ \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\hat{\omega}_n}{2}(\hat{\eta}_n^{-1}Y_i + \hat{\eta}_n Y_i^{-1}) - \frac{(\log Y_i - \hat{\alpha}_{1n})^2}{2\hat{\alpha}_{2n}} - \hat{\lambda}_n \log Y_i \right]^2$$

3 Results

Table 4 contains the results of the comparisons for the two runs of the integrate-and-fire model. For both cases the estimates of $(I_{\mathscr{G}} - I_{\mathscr{F}})$ are approximately 0.01, and the 95% confidence intervals are in the interval $(-\infty, 0)$; hence, we conclude that for both runs the generalized inverse Gaussian family is closer than the lognormal family to the true model. Table 4 also contains the QMLEs for the generalized inverse Gaussian model. For both runs, $\hat{\lambda}_n < 0$, so the Theorem above suggests that diffusion processes are useful approximations for the integrate-and-fire model's subthreshold behavior. For both runs, the estimated value of λ is less than -2, so that the ordinary inverse Gaussian is not a reasonable fit for either run. Figure 1 depicts a smoothed histogram, along with the fitted densities for the first run. While both families fit rather well in the right-hand tail (where ISI > 50), the generalized inverse Gaussian is better near the mode. The small value of the estimated difference in the Kullback-Leibler criterion is reflected in the rather good fits of both models to the data. The picture for the second run is similar, so we omit it.

Similarly, Table 5 contains the results of the comparisons for the two runs of a variant of the Morris-Lecar dynamical model. The estimates of $(I_{\mathscr{G}} - I_{\mathscr{F}})$ are approximately 0.02; once again, their 95% confidence intervals lead us to conclude that for both runs the generalized inverse Gaussian family is closer than the lognormal family to the true model. Table 5 also contains the QMLEs for the generalized inverse Gaussian model. For both runs, $\hat{\lambda}_n < 0$, so the Theorem above applies once again. However, this time the QMLE of χ is zero, so the special case of the reciprocal gamma is the best fit here; the estimated values of λ are less than -12. Figure 2 depicts a smoothed histogram and

Table 5. Comparison of the two runs of the dynamical model and QMLEs for the generalized inverse Gaussian model (reciprocal gamma)

Run:	1	2
Comparison of the two families		
$\sqrt{n}T_n$	−1.73	−1.47
$\hat{\sigma}_n$	0.14	0.10
QMLEs for the generalized inverse Gaussian model (reciprocal gamma)		
$\hat{\chi}_n$	67.87	71.91
$\hat{\lambda}_n$	−11.86	−12.46
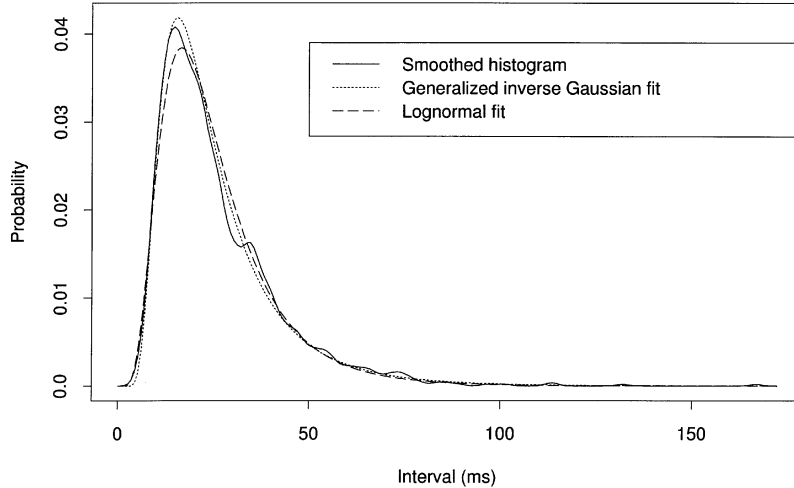
Fig. 1. Simulated integrate-and-fire spike train

the fitted densities for the first run of the dynamical model simulation. While neither family reproduces the smoothed histogram, the generalized inverse Gaussian is a better fit at the mode. The extreme tail (ISI > 8) is well fit by both, but the intermediate tail (5 < ISI < 8) is not well fit by either, due to several lesser modes there. The picture for the second run is similar, so we omit it.

Table 6 contains the results of the comparisons for the ten goldfish retinal ganglion cells. Just as for the simulated spike trains, we conclude that for each cell the generalized inverse Gaussian family is closer than the lognormal family to the true model. The estimates of $(I_{\mathscr{G}} - I_{\mathscr{F}})$ are all between 0.05 and 0.08; once again, their 95% confidence intervals lead us to conclude that for both runs the generalized inverse Gaussian family is closer than the lognormal family to the true model. Table 6 also contains the QMLEs for the generalized inverse Gaussian model. For all cells $\hat{\lambda}_n < 0$. In fact, except for cells 72, 78, and 98 the estimated value of λ is near $-\frac{1}{2}$, indicating that the ordinary inverse Gaussian is a reasonable fit for seven of the ten cells. Since the results for the ten cells are otherwise similar, in Figs. 3 and 4 we illustrate cells 75 and 78, respectively, as representatives. The smoothed histograms are considerably noisier than those of the simulations, as evidenced by the many small modes in the 50–150 ms range. The generalized inverse Gaussian is a better fit near the mode, but both models fail to reproduce the tail behavior, since the two families of densities are unimodal. It is unclear at present whether the smaller modes in the tail are the result of noise or indicate some oscillatory activity.

4 Discussion

The use of diffusion processes to model a neuron's membrane potential has a long history. Interest in such models is still quite strong, as evidenced by recent work that attempts to understand the mechanisms that lead to highly irregular firing patterns (Softky and Koch 1993). The most thoroughly studied diffusions in this context are the Wiener and Ornstein-Uhlenbeck processes; the latter arises as the continuous limit of Stein's integrate-and-fire model. The Wiener process yields the inverse Gaussian distribution for the ISI, which is analytically tractable, so that standard inferential procedures (parameter estimation, confidence intervals, goodness of fit tests) have been well developed. That is not yet the case for the Ornstein-Uhlenbeck or other models. This paucity of statistical methods for biologically more realistic models has led to the use of rather ad hoc procedures, such as the use of moments (plots of coefficient of variation versus the mean), or curve fitting using generic models for positive random variables, such as the lognormal. Also, Levine (1991) has suggested that it is preferable to use such procedures because it is hard to estimate separately the various parameters, such as the threshold, leakage time constant, and input rates, of biologically more realistic models. Recent work (Inoue et al. 1995; Iyengar 1997; Iyengar and Liao 1997) has alleviated some of these difficulties, but much still remains to be done.

In this paper, we have applied a formal inferential procedure for comparing models to assess the adequacy of the fit of the generalized inverse Gaussian family to ISI histograms. For spike trains from simulations and from retinal ganglion cells of goldfish, it is closer, in terms of the Kullback-Leibler information criterion, to the true ISI density than is the log-

Table 6. Comparison of the ten goldfish retinal ganglion cells and QMLEs for the generalized inverse Gaussian model

Cell	68	72	75	78	80	82	83	87	90	98
Comparison of the two families										
$\sqrt{n}T_n$	−1.61	−1.45	−1.63	−2.57	−2.26	−1.90	−1.96	−1.65	−1.89	−1.46
$\hat{\sigma}_n$	0.11	0.09	0.12	0.18	0.12	0.14	0.14	0.13	0.12	0.09
Cell	68	72	75	78	80	82	83	87	90	98
QMLEs for the generalized inverse Gaussian model										
$\hat{\eta}_n$	31.00	25.08	31.61	46.07	27.13	32.11	30.68	25.85	27.86	28.10
$\hat{\omega}_n$	0.47	0.51	0.45	0.33	0.42	0.47	0.44	0.51	0.46	0.47
$\hat{\lambda}_n$	−0.36	−0.17	−0.39	−0.76	−0.40	−0.57	−0.54	−0.39	−0.44	−0.16
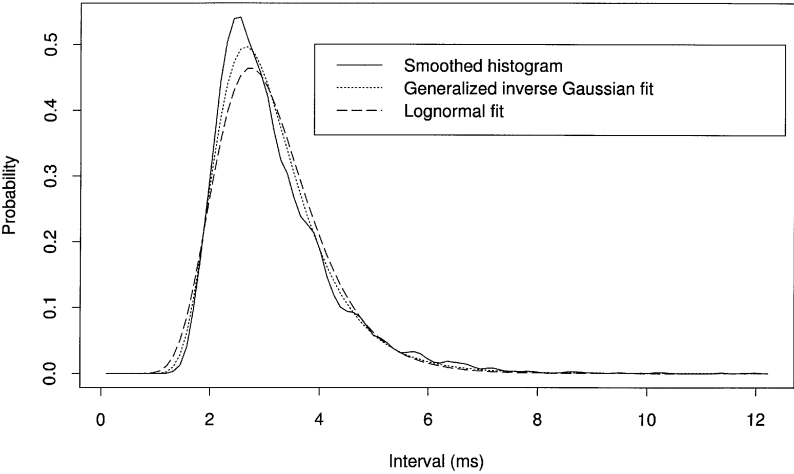
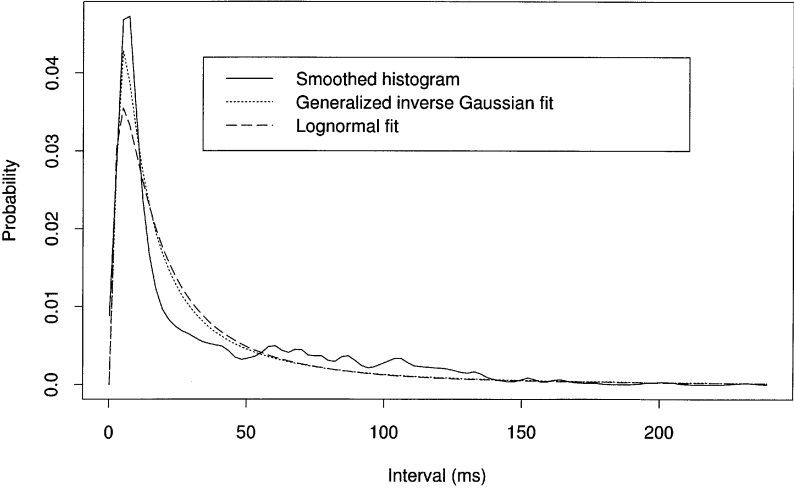**Fig. 2.** Simulated spike train from the dynamical model

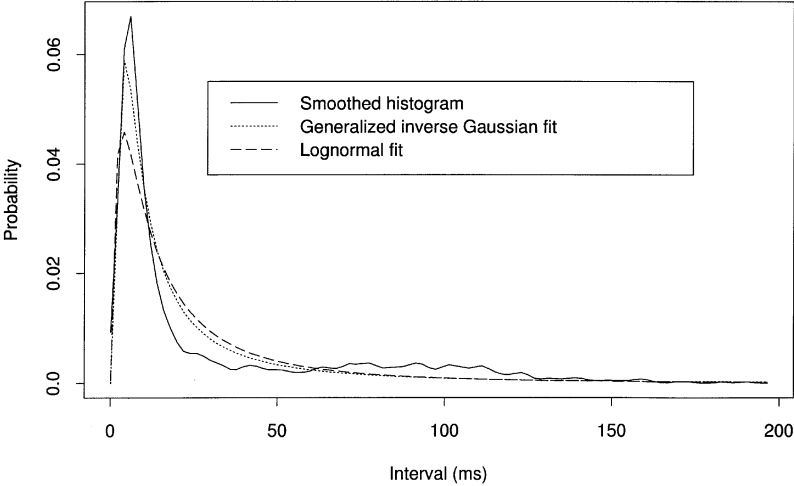**Fig. 3.** Spike train for goldfish retinal ganglion cell 75

Fig. 4. Spike train for goldfish retinal ganglion cell 78

normal. The fit of this model to the simulated data is better than that for the actual spike trains, largely because there are several modes in the right-hand tail of the latters' ISI histograms. The source of these modes is not clear. They could be the result of further noise that is not accounted for, or of some oscillatory (periodic) activity. The theoretical justification given in Sect. 2.3 points to one possible resolution of this difficulty, namely, the use of mixtures of these densities to account for the other modes. In particular, given the infinitesimal variance $\sigma^2(x)$ the parameter χ, determines the initial point, which might be the resting potential, or some other intrinsic feature of the neuron; the other two paramters, λ and ψ determine the infinitesimal mean, which largely reflects the inputs to the neuron. Thus, a mixture over χ could model variations in the reset voltage. We are presently developing fitting procedures for such mixtures.

The theoretical justification above also motivates a more detailed study of the underlying diffusions themselves, with an eye towards more detailed modeling of physiological quantities. For instance, we are studying the problem of fitting parameters of the transformed diffusion above to the Hanson and Tuckwell (1983) model. Also, while the case $\lambda = -\frac{1}{2}$ fits many of the ganglion cells, other cases yield estimates of λ that are much less than that. Approximations to these diffusions for the limiting case $\lambda \to -\infty$ may yield simpler subfamilies. Finally, there have been numerous reports that the exponential and gamma densities provide good fits for other ISI histograms, and that (for certain limiting parameter values) these simple densities approximate the Ornstein-Uhlenbeck first passage time density. Thus, members of the generalized inverse Gaussian family with $\lambda > 0$ may provide good approximations to other diffusions not covered by the result of Barndorff-Nielsen et al. (1978).

References

Abramowitz M, Stegun IA (1965) Handbook of mathematical functions. New York, Dover

Barndorff-Nielsen O, Blaesild P, Halgreen C (1978) First hitting time models for the generalized inverse Gaussian distribution. Stoch Processes Appl 7:49–54

Ermentrout GB (1996) Type I membranes, phase resetting curves, and synchrony. Neural Comput 8: 979–1001

Ermentrout GB, Rinzel J (1989) Analysis of neural excitability and oscillations. In: Koch C, Segev I (eds) Methods in neuronal modeling. MIT Press, Cambridge, Mass, pp 135–169

Foutz RV, Srivastava RC (1977) The performance of the likelihood ratio test when the model is incorrect. Ann Statist 5: 1183–1194

Gerstein G, Mandelbrot B (1964) Random walk models for the spike activity of a single Neuron. Biophys J 4: 41–68

Gutkin B, Ermentrout GB (1997) Type I membrane excitability as a mechanism for cortical spike train statistics. Technical report, Mathematics Department, University of Pittsburgh

Hanson FB, Tuckwell HC (1983) Diffusion approximations to neuronal activity using synaptic reversal potentials. J Theor Neurobiol 2: 127–153

Inoue J, Sato S, Ricciardi L (1995) On the parameter estimation for diffusion models of single neurons. Biol Cybern 73: 209–221

Iyengar S (1997) Maximum likelihood estimation for the Ornstein-Uhlenbeck process. Technical report, Statistics Department, University of Pittsburgh

Iyengar S, Liao Q (1997) Model selection for neural spike train data. Technical report, Statistics Department, University of Pittsburgh

Jorgensen B (1982) Statistical properties of the generalized inverse Gaussian distribution. Springer, Berlin Heidelberg New York

Karlin S, Taylor HM (1981) A second course in stochastic processes. Academic Press, New York

Kullback S (1968) Information theory and statistics. Wiley, New York

Lánský P, Smith CE, Ricciardi L (1990) One-dimensional stochastic diffusion models of neuronal activity and related first passage time problems. Trends Biol Cybern 1: 153–162

Levine MW (1991) The distribution of intervals between neural impulses in the maintained discharges of retinal ganglion cells. Biol Cybern 65: 459–467

Levine MW, Saleh EJ, Yarnold P (1988) Statistical properties of the maintained discharge of chemically isolated ganglion cells in goldfish retina. Vis Neurosci 1: 31–46

Morris C, Lecar H (1981) Voltage oscillations in the barnacle giant muscle fiber. Biophys J 35: 193–213

Nishii R (1988) Maximum likelihood principle and model selection when the true model is unspecified. In: Rao CR, Rao MM (eds) Multivariate statistics and probability. Academic Press, New York, pp 392–403

Softky WR, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. J Neurosci 13(1): 334–349

Visual Numerics (1994) IMSL Math/Library. Visual Numerics, Houston, Texas

Tuckwell H (1988) Introduction to theoretical neurobiology, vols 1 and 2. Cambridge University Press, New York

White H (1982) Maximum likelihood estimation of misspecified models. Econometrica 50: 1–26

White H (1994) Estimation, inference, and specification analysis. Cambridge University Press, New York