INFERENCE FOR THE ORNSTEIN-UHLENBECK MODEL FOR NEURAL ACTIVITY

Satish Iyengar
Statistics Department
University of Pittsburgh
Pittsburgh, PA 15208

Paul Mullowney
Tech-X Corporation
5621 Arapahoe Avenue
Boulder, CO 80303

Abstract

The Ornstein-Uhlenbeck process arises from a leaky stochastic integrate-and-fire model of the membrane potential of a neuron, in which its firing corresponds to the first time the process hits a barrier. We address the problem of estimating the parameters of the underlying process when the available data are the neuron's successive spikes, or spike train. The first-hitting time density is not tractable, so we use its Laplace transform to determine the identifiable parameters of the model, show that their maximum likelihood estimates are consistent and asymptotically normal, and describe computational methods to obtain the estimates and their standard errors.

Keywords: First-passage time, Laplace transform, maximum likelihood, parabolic cylinder function, renewal process, integrate-and-fire model.

AMS Subject Classifications: 60G15, 62F12, 92C20.

1 Introduction

It has long been recognized that many neurons exhibit stochastic activity [29]. This activity is due to complicated biophysical processes such as the neuron's membrane potential and its synaptic inputs. Since these substrates are hard to measure, most common neurophysiological experiments record the action potentials, or spikes, which are the all-or-none outputs from one or more neurons. The analysis of spike trains is done at many levels. The simplest use summaries of interspike intervals (ISIs) such as the shape of their histograms, low-order moments (or functions of them, such as the coefficient of variation or Fano factor), and autocorrelation [17].

Such summaries, however, are somewhat crude for they do not relate the data to any known or hypothesized electrophysiology or anatomy of the neuron. When such information has been gathered from experiments, more realistic mathematical models for neural activity have been proposed. The most prominent model of action potential generation is due to Hodgkin and Huxley [12], which consists of a system of four coupled nonlinear differential equations, one for the membrane potential $V = V_t$ and the others for ionic conductances. This model is quite difficult to work with, so various simplifications have been proposed and studied: for example, there are the dynamical models of Fitzhugh and Nagumo and Morris and Lecar [17]. Our work deals with an alternative approach, namely integrate-and-fire (I-F) models which were pioneered by Lapicque [17], and more recently studied by Stein [28] and others. A simple example of this technique considers the neuron as an RC circuit with single capacitance C and resistance R, so that the membrane potential is governed by the differential equation

$$C\frac{dV}{dt} + \frac{V}{R} = I(t), \tag{1}$$

with input $I(t)$ and initial condition V_0. The resistance leads to leakage of the membrane potential in the absence of inputs. In this case, that decay is exponential with time constant $\tau = RC$. This model neuron fires when V_t reaches its firing threshold V_f, resets to V_0, and the process continues. When the inputs form a stationary stochastic process, and V_t resets to the same value upon firing, the resulting spike train is well approximated by a renewal process [25,29].

The earliest diffusion model for V_t is due to Gerstein and Mandelbrot [11], who studied the special case of $\tau = \infty$ so that there is no leakage. In that case, V_t follows a Brownian motion with drift, and the time to firing has an inverse Gaussian density [4] for which estimation and testing procedures are well understood. They showed that the inverse Gaussian provided a good fit for data from many neurons in a cat's auditory cortex. Moreover, computer simulation of a modification of the Brownian motion process gave the same qualitative results as actual stimulated activity. Soon thereafter, Stein proposed a leaky I-F model, a stochastic version of Lapicque's equation:

$$dV_t = \left(\mu - \frac{V_t}{\tau}\right) dt + \sigma dW_t, \tag{2}$$

with inital condition V_0, and W_t a standard Brownian motion. Thus, V_t follows an Ornstein-Uhlenbeck (OU) process between spikes [30]. Here, μ is the net input per unit time, a difference between excitatory and inhibitory stimuli, and σ is the noise standard deviation.

Since then, many other variations and elaborations of I-F diffusion models with $\mu = \mu(V_t)$ and $\sigma = \sigma(V_t)$ have been proposed. For example, Capocelli and Ricciardi [3] suggested one that included reversal potentials, in which the effect of an input depends upon the membrane potential at that time. As V_t gets closer to the excitatory (inhibitory) reversal potential, the effect of an excitatory (inhibitory) input decreases, and in fact its effect is reversed beyond the reversal potentials. These diffusions have μ linear in V_t and σ^2 linear or quadratic in V_t. Iyengar and Liao [14] showed that the generalized inverse Gaussian family [15] provides a good fit to spike trains from goldfish retinal ganglion neurons; they also pointed to the relevance of an earlier result of Barndorff-Nielsen, et al. [2], who showed that the generalized inverse Gaussian family can arise as the first-passage time distributions for such diffusions. Others, such as the Feller process [13] are approximations that ignore one of the reversal potentials.

The main difficulty with the study of first passage times of many realistic diffusion processes is that they are not analytically tractable. Only the generalized inverse Gaussian distribution has a known probability density function (pdf) that allows standard statistical inference, such as maximum likelihood (ML) or Bayes estimation and hypothesis testing. Thus, it is hard to estimate parameters from available data to fit models to data, to judge the goodness of fit, or to compare various models.

In this paper, we study the OU process as a model for the activity of a neuron, and discuss the problem of parameter estimation when the observed data are the ISIs. We assume that the firing threshold is a constant, and that each time the neuron fires, it is instantly reset to its resting potential; that is, we ignore the refractory period, during which a neuron cannot fire. The spike train then forms a renewal process. The Laplace transform of the ISI density is known: in fact, Darling and Siegert, [7], showed that for a temporally homogeneous diffusion satisfying the Chapman-Kolmogorov equation, the Laplace transform of the first-passage time to a constant boundary can be written in terms of the Laplace transform of the diffusion's transition density. For the OU process, the first-passage time Laplace transform is sufficiently complicated, so inference for it has been difficult. Our work is similar in spirit to that of DuMouchel [10], who studied the problem of estimating the parameters of stable laws, for which the characteristic function is available, but the pdf is intractable.

Much of the earlier work on the OU first passage time dealt with inversion of the Laplace transform and the computation of moments [27]. Notable is the work of Ricciardi and Sato [17], who provided useful approximations for certain ranges of parameter values. More recently, there have been several attempts at fitting the model to data. These include the work of Inoue, et al. [13] and Ditlevsen and Ditlevsen [9] and Madec and Japhet [20], each of whom used moment methods. Paninski, et al. [24] applied the

ML technique to a closely related model using a path integral approach rather than using the Laplace transform. They sketched an argument that the likelihood is log concave. However, to date, there has been no rigorous study of identifiability and ML estimation for this problem.

In Section 2, we sketch the neurophysiological background that leads to the OU model. In Section 3, we determine the identifiable functions of the model parameters, and state our main result concerning the asymptotic properties of the ML estimates. In Section 4, we briefly describe our numerical methods for inverting the Laplace transform, computing the ML estimates, and their asymptotic covariance matrix; we also give an illustration with simulated data. In Section 5 we conclude with a discussion of this work.

2 Physiological Background

The main features of I-F models are the following. There is the neuron's membrane potential, V_t at time t relative to its resting potential, $V_{rest} = 0$ millivolts (mV). In the absence of inputs, V_t decays towards 0 due to leakage. When an excitatory (inhibitory) input arrives through a synapse, V_t moves towards (away from) the firing threshold V_f; and there are the two reversal potentials as mentioned above.

Stein's model ignores reversal potentials, and it leads to a discontinuous trajectory for V_t. One formulation of it is described by the stochastic difference equation

$$\Delta V_t = -\frac{1}{\tau} V_t \, \Delta t + \sum_{i=1}^{n_E} a_i^E dN_i^E + \sum_{j=1}^{n_I} a_j^I dN_j^I. \tag{3}$$

In (3), τ is the time constant that describes exponential decay towards zero; n_E (n_I) is the number of excitatory (inhibitory) synaptic inputs; $a_i^E > 0$ ($a_j^I < 0$) is the magnitude of that input from the corresponding synapse. Finally, the synaptic inputs follow Poisson processes N_i^E and N_j^I, with corresponding rates λ_i^E and λ_j^I. The OU approximation to this process is appropriate for neurons with many inputs (large n_E and n_I) that have small input magnitudes a_i^E and a_j^I. For example, when λ_i^E and λ_j^I both tend to infinity and the magnitudes $a_i^E > 0$ and $a_j^I < 0$ both tend to zero such that

$$\sum_{i=1}^{n_E} a_i^E \lambda_i^E + \sum_{j=1}^{n_I} a_j^I \lambda_j^I \to \mu \quad \text{and} \quad \sum_{i=1}^{n_E} \left(a_i^E\right)^2 \lambda_i^E + \sum_{j=1}^{n_I} \left(a_j^I\right)^2 \lambda_j^I \to \sigma^2,$$

and the higher order moments tend to zero, then (3) is well approximated by the OU stochastic differential equation (2). See Kallianpur [16], Ricciardi and Sacerdote [26], Ricciardi [25], and Tuckwell [29] for more careful derivations under various assumptions.

Given an initial value V_0 the time to firing is the first passage time

$$T = \inf\{t > 0 : V_t = V_f\}. \tag{4}$$

Upon firing, actual neurons exhibit a short refractory period, during which they cannot fire; it is usually regarded as a sum of a deterministic and a random time interval. We ignore this feature here, but recognize that modifications to include refractory periods by adding small lags are possible. Thus, our mathematical model assumes that whenever the neuron fires, its membrane potential instantly resets to V_0, whereupon the process starts afresh, so the successive spikes form a renewal process with ISI distribution being that of T.

The parameters of this model are

$$\eta = (V_0, V_f, \mu, \sigma, \tau).$$

The available data are the spike train, or ISIs $(T_1, \ldots T_n)$. Because the underlying OU process is not observed, not all parameters are identifiable. The problems we address are to determine the identifiable functions of these five parameters; to prove that the ML estimates are asymptotically efficient; and to show that computation of the estimates and the information matrix are feasible.

3 Inference based on the first passage time

For our purposes, a convenient expression for the solution to (2) with initial condition V_0 is

$$V_t = V_0 e^{-t/\tau} + \mu t(1 - e^{-t/\tau}) + \sigma\sqrt{\frac{\tau}{2}} e^{-t/\tau} W(e^{2t/\tau} - 1). \tag{5}$$

For $\eta_0 = (V_0, V_f, 0, 1, 1)$, let $G(t|V_0, V_f)$ be the first passage time cumulative distribution function (cdf). For general η, a simple manipulation of (5) shows that the first passage time cdf is

$$P(T \le t) = G\left(\frac{t}{\tau} | \frac{V_0 - \mu\tau}{\sigma\sqrt{\tau}}, \frac{V_f - \mu\tau}{\sigma\sqrt{\tau}}\right).$$

Thus, the identifiable functions of η appear to be given by

$$\theta = (\theta_1, \theta_2, \theta_3) = \left(\frac{V_0 - \mu\tau}{\sigma\sqrt{\tau}}, \frac{V_f - \mu\tau}{\sigma\sqrt{\tau}}, \tau\right).$$

To show that no further reduction is possible so that θ is indeed identifiable, we use the Laplace transform $E_\theta(e^{-\nu T})$ because the cdf and pdf of T are not tractable. Darling and Siegert [7] have shown that this Laplace transform can be expressed in terms of Hermite (or parabolic cylinder) functions. Among the many expressions for these functions, we use

$$M(z, \nu) = \frac{1}{2\Gamma(\nu)} \sum_{n=0}^{\infty} \Gamma\left(\frac{\nu + n}{2}\right) \frac{(2z)^n}{n!}. \tag{6}$$

Not surprisingly, the Hermite functions are closely related to other Gaussian quantities: when ν is a negative integer, $M(z, \nu)$ is a Hermite polynomial [1], and when ν is a non-negative integer, it is a derivative of Mills' ratio [1].

For the problem at hand, the Laplace transform of T is

$$\hat{g}(\nu|\theta) = E_\theta(e^{-\nu T}) = \frac{M(\theta_1, \nu\theta_3)}{M(\theta_2, \nu\theta_3)}. \tag{7}$$

Because $M(z,0) \equiv 1$ we have $\hat{g}(0|\theta) = 1$, and $P_\theta(T < \infty) = 1$ for all values of θ. Thus, even when the net input per unit time is inhibitory ($\mu < 0$) the mean reverting aspect of the OU model leads to a finite time to firing with probability one. This is in contrast with the inverse Gaussian model for which the firing time is finite with probability one only if $\mu \geq 0$.

Lemma 1: Distinct values of θ have distinct Laplace transforms; thus θ is identifiable.

Proof: From [1], we have the following asymptotic expansion for $M(z,\nu)$: for fixed z and $\nu \to \infty$,

$$M(z,\nu) \sim \frac{\sqrt{\pi}[1 + \mathcal{O}((2\nu)^{-3/2})]}{2^\nu \Gamma\left(\frac{\nu+1}{2}\right)} \exp\left[\frac{z^2}{2} + z\sqrt{2\nu} + \frac{z^3 - 3z}{6\sqrt{2\nu}} - \frac{z^2}{4(2\nu)}\right], \tag{8}$$

so that

$$\begin{aligned} \hat{g}(\nu|\theta) \;\; &\sim \;\; [1 + \mathcal{O}(\nu^{-3/2})] \exp\left[\frac{\theta_1^2 - \theta_2^2}{2}\right] \times \\ &\exp\left[-\sqrt{2\nu}(\theta_2 - \theta_1)\sqrt{\theta_3} + \frac{(\theta_1^3 - 3\theta_1) - (\theta_2^3 - 3\theta_2)}{6\sqrt{\theta_3}\sqrt{2\nu}} - \frac{\theta_1^2 - \theta_2^2}{4\theta_3(2\nu)}\right]. \end{aligned} \tag{9}$$

The three coefficients of the powers of $\sqrt{2\nu}$ in (9) are distinct if and only if the corresponding parameter vectors are distinct. Thus, $\theta = \theta'$ if and only if $\hat{g}(\nu|\theta) = \hat{g}(\nu|\theta')$, and the proof is complete.

Thus, the parameter space is the open set $\Theta = \{\theta \in R^3 : \theta_1 < \theta_2, \theta_3 > 0\}$. Next, we need certain properties of Hermite functions and the Laplace transform $\hat{g}(\nu|\theta)$. First, extending both the index ν and argument z to the complex plane, $M(z,\nu)$ is an entire function of each, it does not vanish when $\Re(\nu) \geq 0$, and the expansion (8) is valid in that half plane as $|\nu| \to \infty$. Thus, the Laplace transform decays exponentially as $\sqrt{|2\nu|} \to \infty$ because $(\theta_2 - \theta_1) > 0$ (see [21]). The inversion formula (or Bromwich integral, see [5]) is then valid, and it yields a continuous pdf for T given by

$$g(t|\theta) = \frac{1}{2\pi i} \int_{0-i\infty}^{0+i\infty} e^{t\nu} \hat{g}(\nu|\theta) d\nu. \tag{10}$$

Since $\nu^k \hat{g}(\nu|\theta)$ also decays exponentially similarly for every k , the pdf has derivatives with respect to t of all orders. Ricciardi and his colleagues [27] have used these results to propose various numerical inversion methods. They also showed that $g(t|\theta) > 0$ for all $t > 0$; that as $t \to \infty$, $g(t,\theta) \sim C(\theta)e^{-\lambda_\theta t}$; and that as $t \to 0$, $g(t,\theta) \sim D(\theta)t^{-3/2}e^{-\alpha_\theta/t}$. Below, we will need the fact that the constants here — $C(\theta), D(\theta), \lambda_\theta, \alpha_\theta$ — are all continuous in θ.

Next, we show that derivatives of the log-likelihood with respect to the parameters also exist and are appropriately bounded. First, some notation:

$$M_1 = \frac{\partial M}{\partial z}, \quad M_2 = \frac{\partial M}{\partial \nu}, \quad M_{12} = \frac{\partial^2 M}{\partial z \partial \nu}, \ldots; \hat{g}_i = \frac{\partial \hat{g}}{\partial \theta_i}, \quad \hat{g}_{ij} = \frac{\partial^2 \hat{g}}{\partial \theta_i \partial \theta_j}, \ldots,$$

and similarly for g.

Lemma 2: In the region $\Re(\nu) \geq 0$ all partial derivatives of $\hat{g}$ with respect to θ decay exponentially as $\sqrt{|2\nu|} \to \infty$.

Proof: Writing $\log \hat{g}(\nu|\theta) = \log M(\theta_1, \nu\theta_3) - \log M(\theta_2, \nu\theta_3)$, the partial derivatives with respect to θ are easy: for example,

$$\begin{aligned} \hat{g}_1(\nu|\theta) &= \hat{g}(\nu|\theta)\frac{M_1(\theta_1, \nu\theta_3)}{M(\theta_1, \nu\theta_3)}, \quad \hat{g}_3(\nu|\theta) = \theta_3 \hat{g}(\nu|\theta)\left[\frac{M_2(\theta_1, \nu\theta_3)}{M(\theta_1, \nu\theta_3)} - \frac{M_2(\theta_2, \nu\theta_3)}{M(\theta_2, \nu\theta_3)}\right], \\ \hat{g}_{13}(\nu|\theta) &= \theta_3 \hat{g}(\nu|\theta)\left[\frac{M_{12}(\theta_1, \nu\theta_3)}{M(\theta_1, \nu\theta_3)} - \frac{M_1(\theta_1, \nu\theta_3)M_2(\theta_2, \nu\theta_3)}{M(\theta_1, \nu\theta_3)M(\theta_2, \nu\theta_3)}\right], \end{aligned}$$

and similarly for the others. Thus, it suffices to show that each of the ratios M_i/M, M_{ij}/M, and M_{ijk}/M increase at most polynomially in $|\nu|$. First, for derivatives with respect to z, the recursion $M_1(z, \nu) = 2\nu M(z, \nu+1)$ and expansion (8) gives $|M_1/M| = \mathcal{O}(|\nu|^{1/2})$, $|M_{11}/M| = \mathcal{O}(|\nu|)$, and $|M_{111}/M| = \mathcal{O}(|\nu|^{3/2})$. The derivatives with respect to ν are a bit more involved. Consider the following integral representation [18]

$$M(z, \nu) = \frac{1}{\Gamma(\nu)} \int_0^\infty e^{-t^2+2tz} t^{\nu-1}\, dt,$$

which is valid for $\Re(\nu) > 0$. Then

$$M_2(z, \nu) = \frac{1}{\Gamma(\nu)} \int_0^\infty e^{-t^2+2tz} t^{\nu-1} (\log t)\, dt - \psi(\nu) M(z, \nu),$$

where $\psi(\nu) = \Gamma'(\nu)/\Gamma(\nu)$ is the digamma function, which is asymptotically equivalent to $\log(\nu)$ as $|\nu| \to \infty$. Next, using Laplace's method when there are logarithmic singularities, the integral above is $\mathcal{O}(|\log(\nu) M(z, \nu)|)$ (see [23]); thus, $M_2(z, \nu)/M(z, \nu)$ is of order $\mathcal{O}(|\log(\nu)|)$. The same method yields the orders of magnitude for higher derivatives with respect to ν:

$$M_{22}(z, \nu)/M(z, \nu) = \mathcal{O}(|\log(\nu)|^2), \quad M_{222}(z, \nu)/M(z, \nu) = \mathcal{O}(|\log(\nu)|^3),$$

and the proof is complete.

We now state and prove our main result about ML estimation of θ.

Theorem: Suppose that the true value of θ^0 is in the interior of Θ. With probability tending to 1 as $n \to \infty$, there exist solutions $\hat{\theta}_n$ based on $T_1, \ldots, T_n$ such that (i) $\hat{\theta}_{jn}$ is consistent for θ_j^0, $j = 1, 2, 3$, (ii) $\sqrt{n}(\hat{\theta}_n - \theta)$ is asymptotically normal with mean 0 and covariance matrix $I(\theta)^{-1}$, and (iii) $\hat{\theta}_{jn}$ is asymptotically efficient.

Proof: We verify the conditions set forth in Lehmann, [19]; we have modified the statements of the conditions to suit our notation. Above, we have dealt with the first three conditions concerning identifiability, common support, and the existence of a pdf; below, we treat the remaining four conditions.

(i) DIFFERENTIABILITY WITH RESPECT TO θ. Consider an open subset ω in Θ containing the true parameter θ^0, and with compact closure $\bar{\omega}$. From Lemma 2, we know that for each $\theta \in \omega$ there are positive constants A_θ and B_θ that are both continuous in θ such that

$$|\hat{g}_{ijk}(\nu|\theta)| \leq A_\theta e^{-B_\theta \sqrt{|2\nu|}} \leq A e^{-B\sqrt{|2\nu|}},$$

where A and B are the suprema of A_θ and B_θ, respectively, over $\bar{\omega}$. Thus, by [5] that $g_{ijk}(t|\theta)$ is the inverse of the Laplace transform $\hat{g}_{ijk}(\nu|\theta)$.

(ii) MEANS OF FIRST TWO LOG DERIVATIVES OF g. Just as in the part (i), there are constants A_i and B_i such that

$$A_i \hat{g}(\nu|\theta) \leq \hat{g}_i(\nu|\theta) \leq B_i \hat{g}(\nu|\theta),$$

so that g_i is integrable, and it follows that

$$\int_0^\infty \frac{\partial}{\partial \theta_i} g(t|\theta)\, dt = \frac{\partial}{\partial \theta_i} \int_0^\infty g(t|\theta)\, dt = 0.$$

Using the fact that the expected score is zero, a similar argument shows that

$$\begin{aligned} I_{jk}(\theta) &= E_\theta \left[\frac{\partial}{\partial \theta_j} \log g(T|\theta) \frac{\partial}{\partial \theta_k} \log g(T|\theta) \right] \\ &= E_\theta \left[-\frac{\partial^2}{\partial \theta_j \theta_k} \log g(T|\theta) \right]. \end{aligned}$$

(iii) THE INFORMATION MATRIX IS NONSINGULAR. We prove the equivalent statement that the statistics

$$\frac{\partial}{\partial \theta_i} \log g(T|\theta) \text{ for } i = 1, 2, 3$$

are affinely independent with probability 1 (for all θ). We must show that for fixed θ there do not exist constants a, b, c, d such that

$$a \frac{\partial}{\partial \theta_1} g(T|\theta) + b \frac{\partial}{\partial \theta_2} g(T|\theta) + c \frac{\partial}{\partial \theta_3} g(T|\theta) = d$$

with probability 1. Expressing these partial derivatives using the inversion formula, we have

$$\frac{1}{2\pi i} \int_{0-i\infty}^{0+i\infty} e^{T\nu} \left[a\hat{g}_1(\nu|\theta) + b\hat{g}_2(\nu|\theta) + c\hat{g}_3(\nu|\theta) \right] d\nu = d$$

However, by Lemma 2, the Laplace transform in this integral decreases exponentially as $|\sqrt{2\nu}| \to \infty$, so its inverse transform cannot be a constant with probability 1.

(iv) BOUNDS ON THIRD DERIVATIVES. We must show that there exist functions h_{ijk} such that

$$\left|\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\theta_k}\log g(t|\theta)\right| \leq h_{ijk}(t) \quad \text{for all } \theta \in \omega$$

where $E_{\theta^0}[h_{jkl}(T)] < \infty$. This follows as in part (i): for $\theta \in \omega$, there is a constant A_θ such that $|\hat{g}_{ijk}(\nu|\theta)| \leq A_\theta \hat{g}(\nu|\theta)$ so that $|g_{ijk}(t|\theta)| \leq A_\theta g(t|\theta)$. Now consider the bounding function h constructed thus: first use $\sup_{\theta\in\bar{\omega}} A_\theta$ in place of A_θ; next, recall that $g(t|\theta)$ has an exponential tail with parameter λ_θ, that is continuous in θ. Hence, $\lambda_{\min} = \inf_{\theta\in\omega}\lambda_\theta$ yields an integrable upper bound. The proof is now complete.

4 Computation

Maximum likelihood would be of limited interest if the computation of the estimates, their standard errors, and likelihood plots were not tractable. In this section, we show that these quantities can indeed be computed efficiently using a simulated spike train. We will only sketch our approach, and refer the reader to our companion paper [22] for details.

Figure 1 shows a histogram of the 1000 ISIs from the OU process, for which the identifiable parameters are $\theta = (-\sqrt{10}, -\sqrt{10}/4, 10) = (-3.16, -0.79, 10)$. Computation of the

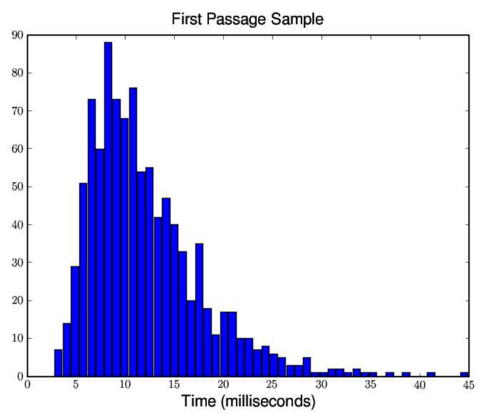

Figure 1: ISI histogram for OU parameters $(V_0, V_f, \mu, \sigma, \tau) = (0, 15, 2, 2, 10)$.

various quantities needed to proceed with inference presented several choices. We now give an account of the steps we took. Although there are series expansions for the pdf of T (see [27]), they are rather inefficient for routine calculation; in fact the coefficients of these series were infinite series themselves. We therefore dealt more directly with the Laplace transform. There are several methods for the numerical inversion of the density and its derivatives with respect to the parameters. We chose the technique of DeHoog [8] because of its efficiency for multiple time evaluations. We also needed different expansions for Hermite functions for different regions of the parameter space to avoid catastrophic cancellations that lead to errors of magnitude $\mathcal{O}(1)$. We used Newton's method to search for the ML estimate; because of the log-concavity of the log-likelihood, few iterations were needed. We used moment-based estimates [27] for θ_1 and θ_2; no such estimates for θ_3, so we searched in the biologically meaningful 5 to 20 msec range with various starting points.

We performed the inference in two ways: first assuming that the time constant $\theta_3 = \tau$ is known and then estimating it along with the other two parameters. Previous efforts ([9, 13]) at inference assumed that τ is known. This assumption is not justified without auxiliary information when the available data are the ISI only. Furthermore, our numerical work shows that the standard errors are quite sensitive to this assumption. The parameter estimates and their standard errors (SE) for the two cases are

$$(\hat{\theta}_1, \hat{\theta}_2) = (-3.01, -0.70) \quad \text{and} \quad SE = (0.097, 0.043),$$

$$(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3) = (-3.04, -0.46, 8.27) \quad \text{and} \quad SE = (0.177, .565, 3.83),$$

with corresponding asymptotic covariance matrices

$$C_2 = \begin{pmatrix} 0.0094 & 0.0038 \\ 0.0038 & 0.0019 \end{pmatrix} \quad \text{and} \quad C_3 = \begin{pmatrix} 0.0313 & -0.0784 & 0.5567 \\ -0.0784 & 0.3198 & -2.1633 \\ 0.5567 & -2.1633 & 14.7064 \end{pmatrix}.$$

Thus, when the time constant is not known, considerably greater sample sizes are needed to estimate all of the parameters; further, there is much less information about the time constant than the other parameters. These differences are not as pronounced in the pdfs that are estimated from the parameters, as seen in Figures 2 and 3.

A B

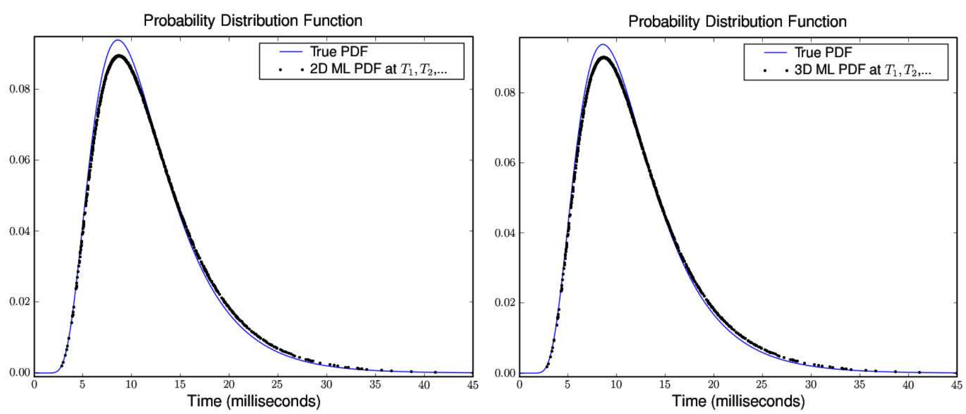

Figure 2: True and estimated pdfs: (A) 2-parameter case and (B) 3-parameter case.

Finally, for the two-parameter case we present the contours of the likelihood surface along with the approximate 95% and 99% confidence ellipses and rectangles (both of which contain the true value) in Figure 3. The elliptical contours indicate that a normal approximation is adequate.

5 Discussion

The results of our work have several implications. First, the determination of identifiable functions (θ) point to the necessary experiments needed to determine all model parameters

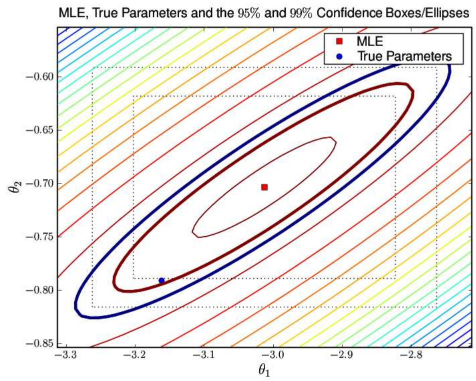

Figure 3: Contours of the log-likelihood and 95% and 99% confidence ellipses.

(η). Of course, in order to apply this model to actual data sets, we will need modifications such as the inclusion of refractory periods. Next, inference using properties of the Laplace transform (in this case through the Hermite functions) is a feasible approach; we are now extending this approach to other diffusions (such as the Feller process) that arise in the study of I-F neuron models. The feasibility of computations by inverting the Laplace transform provides a promising alternative to quadrature in high dimensions [24]. Also, the maximum likelihood approach also leads to model comparison procedures via the Kullback-Leibler divergence between two candidate model families, as we did in [14].

6 References

1. Abramowitz, M., Stegun, I.A. (1965) *Handbook of Mathematical Functions.* Dover.

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

3. Capocelli, R.M., Ricciardi, L.M. (1973) A continuous Markovian model for neuronal activity. *J. Theoretical Biology,* 40: 369–387.

4. Chhikara, R., Folks, J.L. (1988) *The Inverse Gaussian Distribution: Theory, Methodology, and Applications.* Marcel-Dekker.

5. Churchill, R.V. (1981) *Operational Mathematics.* McGraw-Hill.

6. D'Amore, L., Laccetti, G., Murli, A. (1999) An implementation of a Fourier series method for numerical inversion of the Laplace transform. *ACM Transactions on Mathematical Software,* 25: 279–305.

7. Darling, D., Siegert, A. (1953) The first passage problem for a continuous Markov process. *Annals of Mathematical Statistics,* 24: 624–639.

8. DeHoog, F.R., Knight, J.H., Stokes, A.N. (1982) An improved method for the numerical inversion of Laplace transforms. *SIAM J. of Scientific and Statistical Computing,* 3: 357–366.

9. Ditlevsen, S., Ditlevsen, O. (2006) parameter estimation from observations of first

passage times of the Ornstein-Uhlenbeck process and the Feller process. Unpublished manuscript.

10. DuMouchel, W. (1973) On the asymptotic normality of the maximum likelihood estimate when sampling from a stable distribution. *Annals of Statistics,* 1: 948–957.

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

12. Hodgkin, A.L., Huxley, A.F. (1952) A quantitative description of membrane current and its application to conduction and excitation in nerves. *J. Physiology*, 117: 500–544.

13. Inoue, J., Sato, S., Ricciardi, L.M. (1995) On the parameter estimation for diffusion models of single neuron's activities. *Biological Cybernetics,* 73: 209–221.

14. Iyengar, S., Liao, Q. (1997) Modeling neural activity using the generalized inverse Gaussian distribution. *Biological Cybernetics,* 77: 289–295.

15. Jorgensen, B. (1981) *The Generalized Inverse Gaussian Distribution.* Springer.

16. Kallianpur, G. (1983) On a diffusion approximation to a discontinuous model for a single neuron. In *Contributions to Statistics.* North Holland.

17. Koch,C. (1999) *Biophysics of Computation: Information Processing in Single Neurons.* Oxford.

18. Lebedev, N.N. (1972) *Special Functions and their Applications.* Dover.

19. Lehmann, E.L. (1983) *Theory of Point Estimation.* Wiley.

20. Madec, Y., Japhet, C. (2004) First passage time problem for a drifted Ornstein-Uhlenbeck process. *Mathematical Biosciences,* 189: 131–140.

21. Magnus, W., Oberhettinger, F., Soni, R.P. (1966) *Formulas and Theorems for the Special Functions of Mathematical Physics.* 3rd edition. Springer.

22. Mullowney, P., Iyengar, S. (2006) Maximum Likelihood Estimation of an Integrate-and-Fire Neuronal Model. Submitted to *J. Neural Computation.*

23. Olver, F.W.J. (1974) *Asymptotics and Special Functions.* Academic Press.

24. Paninski, L., Pillow, J.W., Simoncelli, E.P. (2004) Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model. *Neural Computation,* 16: 2533–2561.

25. Ricciardi, L. (1977) *Diffusion Processes and Related Topics in Biology.* Springer.

26. Ricciardi, L., Sacerdote, L. (1977) The Ornstein-Uhlenbeck process as a model for neuronal activity. *Biological Cybernetics,* 35: 1–9.

27. Ricciardi, L., Sato, S. (1988) First passage time density and moments of the Ornstein-Uhlenbeck process. *J. Applied Probability,* 25: 43–57.

28. Stein, R.B. (1965) A theoretical analysis of neuronal variability. *Biophysical J.,* 5: 173–194.

29. Tuckwell, H. (1988) *Introduction to Theoretical Neurobiology. Volume 2: Nonlinear and Stochastic Theories.* Cambridge.

30. Uhlenbeck, G., Ornstein, L. (1930) On the theory of Brownian motion. Reprinted in *Selected papers in Noise and Stochastic Processes.* (1954) N. Wax, ed. Dover.