Inference for Stochastic Neuronal Models*

M. K. Habib
Center for Computational Statistics & Probability
George Mason University
Fairfax, Virginia 22030-4444

and

A. Thavaneswaran

University of Manitoba
Winnepeg, Manitoba, Canada

ABSTRACT

Stochastic models of some aspects of the electrical activity in the nervous system at the cellular level are developed. In particular, models of the subthreshold behavior of the membrane potential of neurons are considered along with the problem of estimation of physiologically meaningful parameters of the developed models. Both ordinary and partial stochastic differential equation models are treated.

1. INTRODUCTION

This paper is concerned with the development of stochastic models and methods of statistical inference and their application to studies of the electrical activity in the nervous system at the cellular level. These methods provide neuroscientists with quantitative means to estimate physiologically meaningful parameters of appropriate temporal and spatiotemporal stochastic models that describe certain aspects of the electrical activity of nerve cells (or neurons) using experimentally generated data.

In Section 2, the temporal subthreshold behavior of the somal membrane potential of a single neuron is modeled as a solution of stochastic differential equations. These models contain neurophysiologically meaningful parameters

*Research supported by the Office of Naval Research under Contracts N00014-83-K-0387 and N00014-89-J-1502.

655 Avenue of the Americas, New York, NY 10010 0096-3003/90/$03.50

such as the effective somal membrane time constant, the amplitudes and frequency of occurrence of postsynaptic potentials, and measures of variability of synaptic input. Also developed in this section are methods of estimation of these parameters using experimentally generated intracellular recordings of the somal membrane potential of single neurons. In Section 2.1, a brief description of some basic features of neurons is given. In this subsection the membrane potential is modeled as a solution of temporal Itô-type stochastic differential equations (SDEs) driven by various stochastic processes such as point process, Wiener process, and a mixture of Wiener and point processes. In Section 2.2, the maximum-likelihood (ML) estimates of the parameters of a diffusion neuronal model, observed over random intervals, are treated. The results of a simulation study to evaluate the performance of the ML estimators are presented in Section 2.3. A diffusion neuronal model with time-dependent parameters is considered in Section 2.4, and the method of sieves for estimating these parameters is briefly discussed. In Section 2.5, we propose a semimartingale neuronal model which extends all previously discussed models. No distributional assumptions are imposed on this model, and a distribution-free estimation method called the method of optimal estimating functions is applied in this case [11, 51]. These stochastic methods are appropriate for neurophysiological studies in which no quantitative data concerning the spatial distribution of synaptic input are available.

In Section 3, the spatiotemporal aspects of synaptic neuronal input are considered. In Section 3.1, the subthreshold behavior of the somal membrane potential is modeled as a solution of a stochastic partial differential equation. The method of optimal estimating functions is applied to this model in Section 3.2.

2. TEMPORAL STOCHASTIC NEURONAL MODELS

2.1. Temporal Neuronal Models

A basic functional unit for transmitting and receiving information in the nervous system is the nerve cell or neuron. Morphologically, a neuron consists of three essential regions: the cell body (or soma), the dendrites, and the axon. The dendrites form a series of highly branched outgrowths from the soma. The dendrites and the soma are the sites of most specialized junctions (or synapses) where input is received from other neurons. The axon is an extension of the soma which branches near its end into numerous fine branches, each of which has a specialized ending called the presynaptic terminal. The terminals contacts the receptive surface of other cells and transmits, by chemical or electrical means, information about the activity of the neuron to other neurons or to effector cells. The point of contact is known as the synapse. It is formed by the presynaptic terminal of one cell (the

presynaptic cell) and the receptive surface of the other cell (the postsynaptic cell) [27]. Axonal terminals are responsible for transmitting electrical signals generated by the neuron to other neurons. The entire surface of the neuron is bounded by an electrochemically sensitive membrane which is selectively permeable to ions. Across the neuronal membrane there exists a difference in potential due to the presence of organic and inorganic electrically charged ions with different concentrations inside and outside the cell. Among the important inorganic ions are sodium (Na^+), potassium (K^+), and chloride (Cl^-). In the absence of input to the neuron, the membrane potential is kept at a certain level called the resting potential, which is usually between -60 and -70 mV, with the inside of the cell negative with respect to the outside.

When a (chemical) synapse is activated, a chemical substance called neural transmitter is released into the synaptic cleft. The transmitter then crosses the synaptic cleft, combines with the receptor sites of the postsynaptic membrane, and produces a change in potential. This potential change is called postsynaptic potential (PSP). A postsynaptic potential that depolarizes the postsynaptic membrane is called an excitatory postsynaptic potential (EPSP), and one that hyperpolarizes the neuronal membrane is called an inhibitory postsynaptic potential (IPSP). The PSPs perturb the membrane potential, and if a certain potential level (called the neuron's threshold: -35 to -45 mV) is reached, the membrane potential goes through a rapid stereotypical change during which it reaches about $+30$ mV and declines rapidly back to values less than -70 mV within 2 to 3 msec. After this rapid change in amplitude, the membrane potential reaches its resting value and the process starts all over again. This event is called an action potential. In our formulation, a single observation period is the time interval between the moment at which the membrane potential is at a resting level until it reaches threshold. This interval is obviously random in nature. More precisely, it is a first passage time of the membrane potential through the neuronal threshold. See [27] for more details.

Now assume that the state of the neuron is characterized by the difference in potential across its membrane near a spatially restricted area of the soma called the initial segment (or spike initiation region.) The membrane potential is modeled by a stochastic process $V(t)$ defined on a probability space (Ω, F, P). It is subject to instantaneous changes due to the occurrence of (1) excitatory postsynaptic potentials (EPSPs), which are assumed to occur according to mutually independent Poisson processes $P(\lambda_k^e; t)$ with rates λ_k^e $(k = 1,2,\ldots,n_1)$, each accompanied by an instantaneous displacement of $V(t)$ by a constant amount $\alpha_k^e > 0$ $(k = 1,2,\ldots,n_1)$, and (2) inhibitory postsynaptic potentials (IPSPs), which occur according to independent Poisson processes $P(\lambda_k^i; t)$ with amplitudes $\alpha_k^i > 0$ $(k = 1,2,\ldots,n_2)$. Between PSPs, $V(t)$ decays exponentially to a resting potential with time constant τ. As a

first approximation the PSPs are assumed to sum linearly at the trigger zone, and when $V(t)$ reaches the neuron's threshold, an action potential takes place. Based on this simplified model neuron and considering n_1 excitatory synapses and n_2 inhibitory ones, the membrane potential $V(t)$ is modeled as a solution of the stochastic differential equation

$$dV(t) = -\rho V(t)\,dt + \sum_{k=1}^{n_1} \alpha_k^e\, dP(\lambda_k^e; t) - \sum_{k=1}^{n_2} \alpha_k^i\, dP(\lambda_k^i; t), \quad (2.1)$$

where $V(0) = V_0$ and $\rho = \tau^{-1}$. Under certain conditions the solution of (2.1) is a homogeneous Markov process with discontinuous sample paths. This model is known as Stein's model [47] and is a special case of the well-known Poisson-driven Markov-process models (see [45]). This model has been treated in the literature by many authors, among them Johannesma [23] and Tuckwell [48].

Diffusion models in which the discontinuities of $V(t)$ are smoothed out have been sought as approximations to the discontinuous model (2.1) (see e.g. [38–43, 29]). These types of approximations are justified on the grounds that for many types of neurons in the central nervous system, synapses are densely packed along the dentritic tree. For example, there exist on the order of 20,000 synapses on the surface of a typical motor neuron. If the jumps of $V(t)$ are small and the rates of occurrence of the postsynaptic potentials are high, then the approximation of the Poisson-driven Markov model by a diffusion model is appropriate and is accomplished by allowing the amplitudes α^e, α^i to tend to zero and the frequencies λ^e, λ^i to become large in a certain manner. The accuracy of the diffusion approximation (and its use in studies of interspike interval calculations) is discussed by Tuckwell and Cope [49]. Kallianpur [25] established a diffusion approximation for the model (2.1) using the functional central limit theorem for semimartingales of Liptser and Shiryayev [30, 31]. Under some regularity conditions it was shown that the model (2.1) can be approximated by the diffusion model

$$dV(t) = [-\rho V(t) + \mu]\,dt + \sigma\,dW(t), \quad (2.2)$$

$V(0) = V_0$, where W is the standard Wiener process (or Brownian motion).

Next we consider a stochastic neuronal model which takes into account the influence of rapidly occurring low-amplitude synaptic input as well as PSPs with large amplitudes, which may be reflecting the influence of a number of dominating synapses responding to a certain stimulus. The activity of these synapses will be modeled by a linear combination of independent point processes. That is, the model is driven by diffusion as well as point

processes. This mixed model is a special case of a well-known class of stochastic processes called Itô-Markov processes (see [21]). Now assume that in addition to the extensive synaptic input leading to the diffusion model (2.2), there are n_1 EPSPs arriving according to independent point processes $N(\lambda_k^e(t),t)$ with random intensities $\lambda_k^e(t)$ and EPSP amplitudes α_k^e, $k = 1,2,\ldots,n_1$. In addition, IPSPs are arriving according to the independent processes $N(\lambda_k^i(t),t)$, with the corresponding parameters $\lambda_k^i(t)$ and α_k^i, $k = 1,2,\ldots,n_2$. This setup leads to the following extended mixed model to describe the membrane potential of a stimulus-driven neuron:

$$dV(t) = [-\rho V(t) + \mu]\,dt + \sigma\,dW(t) + \sum_{k=1}^{n_1} \alpha_k^e\,dN(\lambda_k^e(t),t) - \sum_{k=1}^{n_2} \alpha_k^i\,dN(\lambda_k^i(t),t). \tag{2.3}$$

The problem of parameter estimation of the mixed model (2.3) has not been sufficiently addressed in the literature. In the next section we treat the simpler problem of parameter estimation of the diffusion model (2.2).

2.2. *Parameter Estimation of a Diffusion Neuronal Model*

Lansky [28] considered the problem of parameter estimation for diffusion neuronal models observed over a fixed interval $[0,T]$ and discussed the asymptotic properties of the estimators as $T \to \infty$. It should be noted that large-sample properties of maximum-likelihood estimators of drift parameters of diffusion processes have been extensively discussed in the literature for the case of observing one trajectory over time interval $0 \leqslant t \leqslant T$ as $T \to \infty$ (see e.g. [4, 6, 1]). For an extensive review of such problems see [2, 3].

Given n independent trajectories $\{V_k(t),\ \tau_{k-1} \leqslant t \leqslant \tau_k\}$, $k = 1,2,\ldots,n$, where, $\tau_1,\tau_2,\ldots,\tau_n$ are independent random variables with $P(\tau_k < \infty) = 1$, $k = 1,2,\ldots,n$, Habib [14] derived maximum-likelihood estimators of the parameters ρ and μ and established their large-sample properties, such as strong consistency and asymptotic normality, assuming σ is known. Now recall the diffusion neuronal model (2.2). From [46], the log-likelihood function is given by

$$L_n(\rho,\mu) = \sum_{k=1}^{n} \left\{ \int_{\tau_{k-1}}^{\tau_k} [-\rho V_k(t) + \mu]\,dV_k(t) - \frac{1}{2}\int_{\tau_{k-1}}^{\tau_k} [-\rho V_k(t) + \mu]^2\,dt \right\}. \tag{2.4}$$

The maximum-likelihood estimators (MLE) of $\hat{\rho}_n$ and $\hat{\mu}_n$ of ρ and μ respectively are simply those values of ρ and μ which maximize (2.4) and are given by

$$\hat{\rho}_n = \frac{\left[\sum_{k=1}^{n}(\tau_k - \tau_{k-1})\right]\left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} V_k(t)\,dV_k(t)\right] - \left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} V_k(t)\,dt\right]\left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} dV_k(t)\right]}{\left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} V_k(t)\,dt\right]^2 - \left[\sum_{k=1}^{n}(\tau_k - \tau_{k-1})\right]\left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} V_k^2(t)\,dt\right]} \tag{2.5a}$$

and

$$\hat{\mu}_n = \frac{\left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} V_k^2(t)\,dt\right]\left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{k} dV_k(t)\right] - \left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} V_k(t)\,dV_k(t)\right]\left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} V_k(t)\,dt\right]}{\left[\sum_{k=1}^{n}(\tau_k - \tau_{k-1})\right]\left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} V_k^2(t)\,dt\right] - \left[\sum_{k=1}^{n}\int_{\tau_{k-1}}^{\tau_k} V_k(t)\,dt\right]^2}. \tag{2.5b}$$

Using the fact that the membrane potential $V(t)$ is observed continuously over random intervals, the diffusion coefficient σ^2 may be obtained from an observed trajectory V_k $(k = 1,2,\ldots,n)$ by the formula

$$\hat{\sigma}^2(k) = \frac{1}{d_k}\lim_{m\to\infty}\sum_{j=1}^{2^{m_k}}\left[V_k(\tau_{k-1} + jd_k\,2^{-m_k}) - V_k(\tau_{k-1} + (j-1)d_k\,2^{-m_k})\right]^2, \tag{2.6}$$

where $d_k = \tau_k - \tau_{k-1}$, and by partitioning the kth observation period $(\tau_{k-1}, \tau_k]$ into 2^{m_k} subintervals. This result may be proved using the corresponding result of Lévy for Brownian motion by transforming V_k via time substitutions into Brownian motion (or a Wiener process). A natural estimate

of σ^2 which employs all the observed trajectories is given by

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{k=1}^{n} \hat{\sigma}^2(k).$$

The consistency and asymptotic normality of $\hat{\rho}_n$ and $\hat{\mu}_n$ (as $n \to \infty$) have been established in [14].

The methods of parameter estimation described here may be employed in neurophysiological studies in which intracellular recordings of the somal membrane potential of single neurons are obtained. An excellent example of these experimental studies is given in [7], where intracellular recordings were obtained to study the origin of orientation selectivity in the cat visual cortex.

2.3. *Simulation Studies*

In this section we briefly discuss the results of a simulation study to evaluate the performance and efficiency of estimates of the parameters ρ and μ of the model (2.2). This study provides general guidelines for the choice of the number of observed trajectories and the length of the observation period of every trajectory. For simplicity, we consider the diffusion model (2.2). Assume for the moment that the period of observation is fixed, $[0, T]$ say. In this case the estimators $\hat{\rho}_{n,T}$ and $\hat{\mu}_{n,T}$ are defined in terms of stochastic and ordinary integrals [cf. (2.5a) and (2.5b)]. But in practice one has to approximate these integrals with appropriate finite sums which depend on a digitization scheme or the partition mesh $\{t_0, t_1, \ldots, t_n\} \subset [0, T]$.

In order to evaluate the performance of the estimates $\hat{\rho}_{n,T}$ and $\hat{\mu}_{n,T}$ we simulated the solution of the model (2.2) using the difference equation

$$V(t_{k+1}) = V(t_k) + [-\rho V(t_k) + \mu]h + \sigma[W(t_{k+1}) - W(t_k)], \quad (2.7)$$

where $h = T/N$, $t_k = kh$, $k = 1, 2, \ldots, N$. It is well known that the solution of (2.7) converges to $V(t)$. For instance, if we set $V_N(t) = V(t_k)$ for $t \in [t_k, t_{k+1})$, then

$$E\left(\sup_{0 \leqslant t \leqslant T} |V(t) - V_N(t)|^2 \right) \to 0$$

as $N \to \infty$ (see [9]). This and other kinds of discretization, especially Runge-Kutta schemes, have been extensively studied (see e.g. [22, 36, 18, 32]).

TABLE 1

PARAMETER ESTIMATES USING A SIMULATED DIFFUSION PROCESS OBSERVED n TIMES OVER A FIXED PERIOD $[0,T]$ AND SAMPLED EVERY δ UNITS:

$T = 10$ msec, $\delta = 0.10$

Parameter	True value	Estimated value		
		$n = 1$	$n = 10$	$n = 50$
$\rho = \tau^{-1}$	0.33333	0.30336	0.33000	0.33427
μ	5.00000	4.63803	4.84648	4.88702
σ	0.31623	0.67566	0.67364	0.67583

If $\hat{\theta}_{n,N,T}$ ($\tilde{\theta}_{n,N,T}$), where $\theta = (\rho, \mu)$, denotes the resulting estimate (respectively, optimal or ML estimate) based on observations at $t_0, t_1, \ldots, t_N$, LeBreton [54] showed that when $\max(t_{i+1} - t_i) = \delta_N \to 0$,

$$\underset{n\to\infty}{\text{P-lim}}\,\hat{\theta}_{n,N,T} = \underset{n\to\infty}{\text{P-lim}}\,\tilde{\theta}_{n,N,T} = \hat{\theta}_{n,T},$$

and that $\delta_N^{-1/2}(\hat{\theta}_{n,N,T} - \hat{\theta}_{n,T})$ and $\delta_N^{-1/2}(\tilde{\theta}_{n,N,T} - \hat{\theta}_{n,T})$ are both bounded in probability. The results of our simulations are summarized in Tables 1 and 2.

It is clear from Table 1 that for processes which are observed over a period $[0,T]$, with $T = 10$ msec, the estimates of all parameters are very close and that they improve as the number of observed trajectories, n, increases. Form Table 2, one does not notice any improvement in the estimators as the number of observed trajectories n increases (in fact, they deteriorate). This apparently happens because for Table 2 the period of observation $[0,T]$ was longer, $T = 20$ msec. Therefore, one may conclude that or slowly firing

TABLE 2

PARAMETER ESTIMATES USING A SIMULATED DIFFUSION PROCESS OBSERVED n TIMES OVER A FIXED PERIOD $[0,T]$ AND SAMPLED EVERY δ UNITS:

$T = 20$ msec, $\delta = 0.10$

Parameter	True value	Estimated value		
		$n = 1$	$n = 10$	$n = 50$
$\rho = \tau^{-1}$	0.33333	0.30369	0.32705	0.32399
μ	5.00000	4.86121	4.77822	4.71001
σ	0.31623	0.33012	0.51796	0.33537

TABLE 3

PARAMETER ESTIMATES USING A SIMULATED DIFFUSION PROCESS OBSERVED n TIMES OVER A FIXED PERIOD $[0, T]$ AND SAMPLED EVERY δ UNITS:

$T = 10$ msec, $\delta = 0.05$

		Estimated value		
Parameter	True value	$n = 1$	$n = 10$	$n = 50$
$\rho = \tau^{-1}$	0.33333	0.30627	0.32898	0.33620
μ	5.00000	4.87798	4.75184	4.83044
σ	0.31623	0.53005	0.51272	0.51828

TABLE 4

PARAMETER ESTIMATES USING A SIMULATED DIFFUSION PROCESS OBSERVED n TIMES OVER A FIXED PERIOD $[0, T]$ AND SAMPLED EVERY δ UNITS:

$T = 15$ msec, $\delta = 0.05$

		Estimated value		
Parameter	True value	$n = 1$	$n = 10$	$n = 50$
$\rho = \tau^{-1}$	0.33333	0.34615	0.32371	0.33153
μ	5.00000	4.93817	4.68444	4.77649
σ	0.31623	0.46003	0.45486	0.45778

neurons, one does not gain much by recording a large number of spikes, but for rapidly firing neurons, one can expect that the parameter estimators will improve as the number of observed trajectories increases.

Tables 3 and 4 show similar behavior even when the sampling interval of the simulated trajectories reduces to $\delta = 0.05$. However, Table 5 shows that the estimates deteriorate when the diffusion coefficient σ increases to 3.0.

TABLE 5

PARAMETER ESTIMATES USING A SIMULATED DIFFUSION PROCESS OBSERVED n TIMES OVER A FIXED PERIOD $[0, T]$ AND SAMPLED EVERY δ UNITS:

$T = 15$ msec, $\delta = 0.05$

		Estimated value		
Parameter	True value	$n = 1$	$n = 10$	$n = 50$
$\rho = \tau^{-1}$	0.33333	0.339721	0.24716	0.24202
μ	5.00000	4.34994	2.63800	2.58367
σ	3.00000	2.89161	2.91949	2.91685

2.4. A Neuronal Model with Time-Dependent Parameters

In this subsection, we consider the problem of maximum-likelihood estimation of infinite-dimensional parameters in randomly stopped diffusion processes. This is a more general model of the membrane potential of a neuron than (2.2), since close inspection of records of subthreshold trajectories of the membrane potential clearly reveal that the drift parameter μ in (2.2) is a function of time rather than a constant. Furthermore, replacing the decay rate ρ in (2.2) by a function of t compensates for considering only temporal aspects of synaptic input and ignoring their spatial properties. For these reasons the following more general model of the neuronal membrane potential is considered:

$$dX(t) = [-\rho(t)X(t) + \mu(t)]\,dt + \sigma\,dW(t), \qquad 0 \leqslant t \leqslant T, \quad (2.8)$$

where $X(0) = X_0$ is a random variable which is assumed to be independent of the standard Wiener processes W. Also assume that $\rho(\cdot)$ and $\mu(\cdot)$ are members of the space $L^2([0,T], dt)$ of all square-integrable functions defined on $[0,T]$. This is a Hilbert space with the inner product

$$(f,g) = \int_0^T f(t)\,g(t)\,dt.$$

The *statistical problem* at hand then is to estimate the $L^2([0,T], dt)$-unknown functions $\rho(t), \mu(t)$, $t \in [0,T]$, from the observation of n independent trajectories

$$\{X_k(t),\ \tau_{k-1} \leqslant t \leqslant \tau_k\}, \qquad k = 1,2,\ldots,n.$$

The log-likelihood function is given by

$$L_n(\rho,\mu) = \sum_{k=1}^{n} \left\{ \int_{\tau_{k-1}}^{\tau_k} [-\rho(t)X_k(t) + \mu(t)]\,dX_k(t) - \frac{1}{2}\int_{\tau_{k-1}}^{\tau_k} [-\rho(t)X_k(t) + \mu(t)]^2\,dt \right\}. \quad (2.9)$$

It should be noted here that the techniques for estimating finite-dimensional parameters usually fail in the infinite-dimensional case, and we are forced to consider the method of sieves (see [13]). In this method, for each

sample size n (n is the number of observed trajectories) a sieve is chosen which is, roughly speaking, a suitable subset of the parameter space. The likelihood function is maximized on the sieves, yielding a sequence of estimators. For a discussion of some general results on the existence of sieves leading to estimators with interesting asymptotic properties see [8].

Following Nguyen and Pham [35], one uses as sieves increasing sequences U_n and V_n of finite-dimensional subspaces of $L^2([0,T], dt)$ with dimensions d_n and d_n' such that $U_n \subset U_{n+1}$, $V_n \subset V_{n+1}$, and $\bigcup_{n \geqslant 1} U_n$ and $\bigcup_{n \geqslant 1} V_n$ are dense in $L^2([0,T], dt)$ and such that $\{\rho_1, \ldots, \rho_{d_n}\}$ and $\{\psi_1, \ldots, \psi_{d_n'}\}$ form the bases of U_n and V_n, respectively, for all $n \geqslant 1$. For $\rho \in U_n$ and $\mu \in V_n$ we have

$$\rho(\cdot) = \sum_{i=1}^{d_n} \rho_i \phi_i(\cdot)$$

and

$$\mu(\cdot) = \sum_{i=1}^{d_n'} \mu_i \psi_i(\cdot).$$

The likelihood function in this case takes the form

$$L_n(\theta, \mu) = \sum_{k=1}^{n} \left\{ \int_{\tau_{k-1}}^{\tau_k} \left[\left(- \sum_{i=1}^{d_n} \rho_i \phi_i(t) \right) X_k(t) + \sum_{j=1}^{d_n'} \mu_i \psi_i(t) \right] dX_k(t) \right.$$

$$\left. - \frac{1}{2} \int_{\tau_{k-1}}^{\tau_k} \left[\left(- \sum_{j=1}^{d_n} \rho_i \phi_i(t) \right) X_k(t) + \sum_{j=1}^{d_n'} \mu_i \psi_i(t) \right]^2 dt \right\}. \quad (2.10)$$

The objective now is to maximize the likelihood function (2.21) on the sieves to yield a sequence of estimators. Sufficient conditions for the consistency and asymptotic normality of the sieve estimators are given in [33, 34].

2.5. *A Semimartingale Neuronal Model*

A more general stochastic model which encompasses the models (2.1)–(2.3) and (2.8) is given by

$$dV(t) = f(t, V, \theta)\, dA(t) + dM(t), \quad (2.11)$$

where A is a real, monotonic, nondecreasing, right-continuous process with $A(0) = 0$; f is a predictable process; M is a cadlag, locally square-integrable martingale with a predictable variation process

$$\langle M \rangle_{t,\theta} = \int_0^t b(t,\theta)\, dA(t);$$

θ is the parameter of interest; and b is a nonnegative predictable process. (For the relevant definitions see [17].) Furthermore, assume that f is a twice continuously differentiable function.

Notice that no distributional assumptions have been imposed on the driving martingale M in the model (2.11). Therefore, in order to estimate the model's parameters θ, we must employ a likelihood-free estimation method. We therefore apply the method of optimal estimating functions which was considered by Godambe [10, 11]. This method was applied to continuous-time stochastic processes by Hutton and Nelson [20], Thavaneswaran and Thompson [51], and Heyde [19].

Godambe's Optimality Criterion. Following Godambe [11], consider the parameter θ to be a function of $P \in \mathscr{P}$ (a family of probability measures). Let $G(V,\theta) = (G_t(V,\theta), F_t;\ t \geqslant 0)$ represent a family of processes indexed by θ such that $E_p G_t(V,\theta) = 0$ for each $t \geqslant 0$ and for each P, and $\theta = \theta(P)$. This corresponds to the unbiasedness property of Godambe [10]. Godambe's optimality criterion, adapted to this situation, reads: G° is optimal in L, the class of unbiased estimating functions, if $Q = A_h - A_h^\circ$ is nonnegative definite for all $G \in L$ and for all $P \in P$, where

$$h(V) = \left(E\frac{\partial G}{\partial \theta}\right)^{-1} G(V,\theta),$$

$$h^\circ(V) = \left[E\left(\frac{\partial G^\circ}{\partial \theta}\right)^{-1}\right] G^\circ(V,\theta),$$

and A is the variance-covariance matrix for h under θ_0 (the true value of θ). The following sufficient condition for optimality to hold is due to M. E. Thompson.

LEMMA 3.1. *G° is optimal in L if*

$$E\frac{\partial G}{\partial \theta} = KE(GG^{\circ\prime}).$$

PROOF. See [50, p. 57].

Now, for the model (2.11) we choose an estimating function which is a martingale. In particular we consider an estimating function G of the form

$$G_{t,\theta} = \int_0^t a_{s,\theta}\, dM_{s,\theta} \tag{2.12}$$

generated by a predictable process $\{a_{s,\theta}\}$. Then, using the properties of Itô integrals, we have

$$E(G^{\circ\prime}G) = E\int_0^t a_{s,\theta}\, d\langle M\rangle_{s,\theta}\, a^{\circ\prime}_{s,\theta},$$

where A' denotes the transpose of A, and

$$G^{\circ}_{\theta} = \int_0^t a^{\circ}_{s,\theta}\, dM_{s,\theta}.$$

Hence

$$E(G_0 G^{\circ\prime}_{\theta}) = E\int_0^t a_{s,\theta} b_{s,\theta} a'_{s,\theta}(dA_s).$$

Moreover

$$E\frac{\partial G_\theta}{\partial \theta} = E\int_0^\tau a_{s,\theta}\frac{\partial}{\partial\theta}(dM_{s,\theta})$$

$$= -E\int a_{s,\theta} f'_{s,\theta}\, dA_s,$$

where $f'_{s,\theta} = (\partial/\partial\theta) f_{s,\theta}$. It follows that the optimal estimating function is given by

$$G^{\circ}_{\theta} = \int_0^\tau f'_{s,\theta} b^{+}_{s,\theta}\, dM_{s,\theta}, \tag{2.13}$$

provided that $b^{+}_{s,\theta}$, the inverse of $b_{s,\theta}$, exists.

EXAMPLE 2.1. In this example, the model (2.2) is extended by replacing the Wiener process $\{W_t\}$ by a square-integrable martingale $\{M_t\}$:

$$\begin{aligned} dV_t &= (-\rho V_t + \mu)\, dt + \sigma\, dM_t \\ &= (-\rho, \mu)\begin{bmatrix} V_t \\ 1 \end{bmatrix} dt + \sigma\, dM_t \\ &= \theta' f(V_t)\, dt + \sigma\, dM_t, \end{aligned}$$

where

$$\theta = \begin{bmatrix} -\rho \\ \mu \end{bmatrix}, \qquad f(V_t) = \begin{bmatrix} V_t \\ 1 \end{bmatrix}, \qquad \langle M \rangle_t = t,$$

and $\sigma > 0$ (assumed to be known). The optimal estimating function is given by

$$G^{\circ}_{n,\theta} = \sum_{k=1}^{n} \int_0^{\tau_k} \begin{bmatrix} V^{(k)} \\ 1 \end{bmatrix} dM_s^{(k)} \qquad \text{a.s.} \tag{2.14}$$

NOTE. Equation (2.13) is the same as the m.l.e. equation in [14], where it is assumed that $M_t = W_t$ (the standard Wiener process). Therefore, it is clear that without imposing any distributional assumption on the noise (driving process M), the method of optimal estimation allows for the estimation of the unknown parameter under the assumption that $\langle M \rangle_t = t$. M_t may be a purely discontinuous square-integrable martingale with $\langle M \rangle_t = t$, e.g. $M_t = N_t - t$ (where N_t is the standard Poisson process).

3. STOCHASTIC ANALYSIS OF SPATIAL NEURAL MODELS

3.1 Overview

In this section, our goal is to extend the stochastic analysis (proposed in Section 2) of temporal neural models to take into account the spatiotemporal aspects of synaptic input to a single neuron. Therefore, we develop spatiotemporal stochastic models of the subthreshold behavior of the somal membrane potential. These models should serve in evaluating the role of the geometry of the dentrites and synaptic location on the neuronal surface in influencing the neuronal behavior in response to stimulation. We will also

discuss methods of parameter estimation for these extended spatiotemporal models.

An important neuronal characteristic is the dependence of both the magnitude and time course of the postsynaptic potential, evoked at a given synapse, on the spatial location of the active synaptic junction. In Section 2, it was assumed that the synaptic inputs to a neuron can be treated as inputs delivered to a single summing point on the neuron's surface (trigger zone). However, it is a well-established anatomical fact that a great number of the neurons in the central nervous system have extensively branched dendritic receptive surfaces, and that synaptic inputs may occur both on the somatic region and on the dendrites. Nevertheless, a common assumption is that synapses located on distal dendritic branches have little effect on the spike initiation zone of a neuron. According to this view, distally located synapses would merely set the overall excitability of the neuron and would be ineffective in generating neural discharge activity. Synapses located near the soma of a neuron, on the other hand, are widely believed to influence neuronal firing behavior directly and strongly.

A major exception to this view was suggested by Rall [37], based on calculations of passive electronic current spread through the dendrites. Rall's work showed that distal synapses can play a functionally much more interesting role than previously assumed. More specifically, if the synaptic input to the dendrite has the appropriate spatiotemporal characteristics, distal synapses can influence neuronal firing to a much greater extent than is predicted on the basis of their dendritic location. In view of Rall's demonstration and in recognition of the suggestions (based on experimental evidence) that such a mechanism plays an important role in feature extraction by single sensory neurons, it seems necessary to carry out modeling studies to evaluate the potential for different spatial distributions of synaptic inputs to influence sensory-neuron behavior.

We therefore extend the model (2.2) in order to incorporate the important feature of spatial synaptic distribution. This extension is based on Rall's model neuron [37]. In Rall's model neuron the cable properties of a system of branched dendrites are reduced to a one-dimensional equivalent dendrite, with synapses formed at specific distances along the equivalent dendrite.

In what follows, the difference in potential across the membrane surface of a neuron at time $t \in \mathbb{R}_+$ (where $\mathbb{R}_+ = [0, \infty)$) and location $x \in X$ (where X represents the neuronal surface) is denoted by $V(t, x)$ and is modeled as the solution of a stochastic partial differential equation. In this model, we consider two types of synaptic input. The first is extensive, rapidly occurring postsynaptic potentials with relatively low amplitudes arriving at random times and locations on the neuronal surface. This input may then be approximated by a Gaussian noise process (see e.g. [53, 52, 26]). This type of

input reflects the influence of the electrotonically remote synaptic input as well as synaptic input resulting from spontaneously active presynaptic neurons. The second type of synaptic input considered here is assumed to occur with relatively low rate (or intensity) and relatively large amplitudes. This input may be modeled by a spatial stochastic Poisson process. It may reflect the electrical activity of a few influential synapses which have been activated in response to an effective stimulus.

Let $P((0,t]\times A\times B)$ be the number of postsynaptic potentials of amplitudes $a\in A$ arriving at sites $x\in B$ (where B is a Borel set of X) at times $s\leqslant t$. The probability that exactly k such pulses arrive during the period $(0,t]$ is $e^{-\lambda}\lambda^k/k!$ with intensity $\lambda = t\mu(A\times B)$. That is, N is a Poisson measure on $\mathbb{R}\times X\times\mathbb{R}^+$ with intensity measure $dt\,\mu(da,dx)$.

Let $W(t,x)$ be a two-parameter Wiener process defined on $\mathbb{R}^+\times X$. Now, using Rall's model and considering the nerve cell as a line segment of finite length L (i.e. $X=[0,L]$), the subthreshold behavior of the membrane potential $V(t,x)$ may be modeled in the absence of external stimuli by the two-parameter diffusion process

$$dV(t,x)=\left(-\frac{1}{\tau}V(t,x)+\frac{\partial^2V(t,x)}{\partial x^2}+\mu\right)dt+\sigma\,dW(t,x). \quad (3.1)$$

On the other hand, in the presence of an effective stimulus we will assume that only a limited number of the presynaptic neurons are firing in response to the stimulus and that the rest of the presynaptic neurons are spontaneously active. The membrane potential, in this case, may be modeled as a solution $\{V(t,x),\ 0\leqslant x<L\}$ of the stochastic differential equation:

$$dV(t,x)=\left(-\frac{1}{\tau}V(t,x)+\frac{\partial^2V(t,x)}{\partial x^2}\right)dt + \sigma\,dW(t,x)+\int_A uP(dt,da,dx). \quad (3.2)$$

If we compensate the Poisson process P by its intensity, the model (3.2) may be written as

$$dV(t,x)=\left(-\frac{1}{\tau}V(t,x)+\frac{\partial^2V(t,x)}{\partial x^2}\right)dt+\sigma\,dW(t,x) + \int_A u[P(dt,da\,dx)-dt\,\mu(da\,dx)]. \quad (3.3)$$

Now, consider the differential operator

$$T = -\frac{1}{\tau} + \frac{\partial^2}{\partial x^2}.$$

Then (3.3) takes the form

$$dV(t,x) = TV(t,x)\,dt + dM(t,x) \tag{3.4}$$

where $M(t,x)$ is also a semimartingale, given by

$$dM(t,x) = \sigma\,dW(t,x) + \int_A u[P(dt\,da\,dx) - dt\,\mu(da\,dv)].$$

Our goal now is to develop a method to estimate the parameters of models similar to (3.1)–(3.4).

3.2. *Parameter Estimation for Infinite-Dimensional Systems*

In this section we discuss the problem of parameter estimation for infinite-dimensional stochastic differential equations (SDE) of the type

$$dX(t) = \theta TX(t)\,dt + dW_{t,\theta}, \tag{3.5}$$

where θ belongs to a compact parameter set Θ contained in the positive reals, and T is the infinitesimal generator of a strongly continuous semigroup (T_t), $t > 0$, acting on a real separable Hilbert space H with scalar product $\{\cdot,\cdot\}$ and norm $\|\cdot\|$. W_t is an H-valued Wiener process, i.e., W_t is a stochastic process defined on a complete probability space (Ω, F, P) with stationary increments such that the associated (incremental) covariance operator W is nuclear and W_t has zero mean for all $t > 0$. Notice that in (3.5) we have suppressed the dependence of the processes on the optical variable x.

Let $W_0 = 0$, and assume that all eigenvalues λ_i, $i \in N$, of W are strictly positive. Then we can write $W_t = \sum_{i=1}^{\infty} \beta_{i,t} e_i$ P-a.s., where $\{e_i : i \in N\}$ is a complete orthonormal basis for H consisting of eigenvectors (w.r.t. λ_i) of W. The temporal processes $\beta_{i,t} = \langle W_t, e_i \rangle$ (for $t \in [0,\tau]$) are mutually independent real Wiener processes, $i \in N$, having variance processes $[\beta_{i,t}] = \lambda_i t$. Then the projection of the observed process in the direction e_i satisfies

$$d(\langle X(t), e_i \rangle) = \theta \langle TX(t), e_i \rangle\,dt + d(\langle W_{t,\theta}, e_i \rangle).$$

Since the parameter of interest $\theta \in \mathbb{R}$, it is convenient to consider estimating functions of the form

$$G_{t,\theta} = \sum_{i=1}^{\infty} \int_0^t a_{s,\theta} d\langle W_{s,\theta}, e_i \rangle = \sum_{i=1}^{\infty} \int_0^t a_{s,\theta}\, d\beta_{i,s,\theta}, \tag{3.6}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in H. Notice that $G_{t,\theta}$ as defined in (3.6) is the continuous-time version of the estimating function considered in [11].

Assume that $I_{i,s,\theta} = \partial \beta_{i,s,\theta} / \partial \theta$ exists for all s (a.s.) and all i. Let U be the class of G in (3.5) such that

(1) for each θ the process $a_{s,\theta}$ is differentiable with respect to θ,
(2) for all t and i the Itô stochastic integrals

$$\int_0^t \frac{\partial a_{s,\theta}}{\partial \theta}\, d\beta_{i,s,\theta} \quad \text{and} \quad G_{t,\theta} = \int_0^t a_{s,\theta}\, d\beta_{i,s,\theta}$$

exist,

(3) for all t, $\partial G_{t,\theta} / \partial \theta$ has nonzero expectation and can be expressed as

$$\int_0^t \sum_{i=1}^{\infty} \frac{\partial a_{s,\theta}}{\partial \theta}\, d\beta_{i,s,\theta} + \sum_{i=1}^{\infty} \int_0^t a_{s,\theta}\, dI_{i,s,\theta},$$

and

(4) for all t, $EG_{t,\theta}^2$ is finite.

Using the properties of the stochastic integral with respect to real-valued Wiener processes, we have

$$E(G_{t,\theta} G_{t,\theta}^*) = \sum_{i=1}^{\infty} \int_0^t a_{s,\theta} a_{s,\theta}^* \lambda_i\, ds$$

and

$$E \frac{\partial G_{t,\theta}}{\partial \theta} = E \sum_{i=1}^{\infty} \int_0^t a_{s,\theta}\, dI_{i,s,\theta}.$$

A sufficient condition for the process $G_{t,\theta} = \sum_{i=1}^{\infty} \int_0^t a^*_{s,\theta}\, d\beta_{i,s,\theta}$ to be optimal is that $G^* \in U$ and

$$\frac{E[G_{t,\theta} G^*_{t,\theta}]}{E\,\partial G_{t,\theta}/\partial\theta} \tag{3.7}$$

is the same for all $G \in U$. Without loss of generality we may take the constant value to be -1. Using the properties of stochastic integrals, the quantity in (3.7) becomes

$$\frac{E \sum_{i=0}^{\infty} \int_0^t a^*_{s,\theta} a_{s,\theta} \lambda_i\, ds}{E \sum_{i=0}^{\infty} \int_0^t a_{s,\theta}\, dI_{i,s,\theta}}.$$

Therefore, $a^*_{s,\theta}$ is representable as

$$\frac{-\,dI_{i,s,\theta}}{\lambda_i\, ds} = \frac{-\langle TX(s), e_i\rangle}{\lambda_i}$$

and is optimal. The optimal estimating equation takes the form

$$G^*_{t,\theta} = \sum_{i=1}^{\infty} \int_0^t \frac{-\langle TX(s), e_i\rangle}{\lambda_i} d(\langle W_s, e_i\rangle) = 0,$$

and the optimal estimate can be written as

$$\theta_t^* = \frac{\sum_{i=0}^{\infty} \frac{1}{\lambda_i} \int_0^t \langle TX(s), e_i\rangle\, d\langle X(s), e_i\rangle}{\sum_{i=0}^{\infty} \frac{1}{\lambda_i} \int_0^t \langle TX(s), e_i\rangle^2\, ds}.$$

4. CONCLUSION

In this paper we have introduced the semimartingale model (2.11) to describe the subthreshold behavior of the membrane potential of a single neuron. This model extends all previously proposed neuronal models such as

the models driven by a linear combination of independent homogeneous Poisson processes (Stein's model), stochastic differential equations driven by Wiener processes (the diffusion model), and mixed models driven by Poisson and Wiener processes (the Itô-Markov model). The driving process in our semimartingale model is a martingale, and hence no parametric distributional assumptions were imposed in this case. This is a desirable feature of the model because of the difficulty of verifying the parametric distributional assumption frequently imposed in the theoretical neurobiology literature. Also, no assumptions of homogeneity or independence were imposed on the driving process of the model. This is also a desirable feature in the model because of the well-established observation that processes involved in the information processing of neuron systems are, in general, time-dependent (and hence nonstationary) and also dependent on each other (see e.g. [24, 16]).

Furthermore, because of the absence of parametric assumptions, in our model we were forced to apply a nonlikelihood method to estimate the model's parameters, namely the method of optimal estimating functions. This method has not been applied before to models in neurophysiology. Here it was also applied to the spatial stochastic model (3.1).

The result of the simulation study, that for long periods of observation the estimates of the parameters deteriorate as the number of observed trajectories increases, is counterintuitive. This phenomenon should be investigated in a more comprehensive simulation study to find out if it is intrinsic to the diffusion model used in the simulation, or due to the estimators or to the simulation algorithm.

REFERENCES

1 I. V. Basawa and B. L. S. Prakasa Rao, *Statistical Inference for Stochastic Processes*, Academic, New York, 1980.

2 B. Bellach, Consistency, asymptotic normality and efficiency of the maximum likelihood estimator in linear stochastic differential equations, *Math. Operationsforsch. Statist. Ser. Statist.* 11:227–266 (1980).

3 B. Bellach, Parameter estimators in linear stochastic differential equations and their asymptotic properties, *Math. Operationforsch. Statist. Ser. Statist.* 14:141–191 (1983).

4 B. M. Brown and J. I. Hewitt, Asymptotic likelihood theory for diffusion processes, *J. Appl. Probab.* 12:228–238 (1975).

5 R. M. Capocelli and L. M. Ricciardi, Diffusion approximation and first passage time problem for a model neuron, *Kybernetik* 8:214–223 (1971).

6 P. D. Feigen, Maximum likelihood estimation for stochastic processes, *Adv. in Appl. Probab.* 8:712–736 (1978).

7 D. Ferster, Origin of orientation selective EPSPs in neurons of cat visual cortex, *J. Neurosci.* 7:1780–1791 (1987).
8 S. Geman and C. Hwang, Nonparametric maximum likelihood estimation by the method of sieves, *Ann. Statist.* 10:401–414 (1982).
9 I. I. Gihman and A. V. Skorokhod, *Stochastic Differential Equations*, Springer-Verlag, New York, 1979.
10 V. P. Godambe, An optimum property of regular maximum likelihood equations, *Ann. Math. Statist.* 31:1208–1211 (1960).
11 V. P. Godambe The foundations of finite sample estimation in stochastic processes, *Biometrika* 12:419–428 (1985).
12 V. P. Godambe and C. C. Heyde, Quasilikelihood and optimal estimation, *Internat. Statist. Rev.* 55:231–244 (1986).
13 V. Grenander, *Abstract Inference*, Wiley, New York, 1981.
14 M. K. Habib, Parameter estimation for randomly stopped processes and neuronal modeling, Mimeo Series No. 1492, UNC Inst. of Statistics, 1985.
15 M. K. Habib and P. K. Sen, Nonstationary stochastic point process models in neurophysiology with applications to learning, in *Biostatistics: Statistical, Biomedical, Public Health and Environmental Sciences*, North Holland, Amsterdam, 1985, pp. 481–509.
16 M. K. Habib and P. K. Sen, Nonstationary stochastic point-process models in neurophysiology with application to learning and memory, in *Biostatistics: Statistics in Biomedical, Public Health and Environmental Sciences* (P. K. Sen, Ed.), Elsevier Science (North Holland), 1988, pp. 481–509.
17 M. K. Habib and A. Thavaneswaran, Optimal estimation for semimartingale neuronal models, in *Stochastic Methods for Biological Intelligence*, Special Issue of J. Stat. Plann. & Inf., (1990).
18 D. C. Haworth and S. B. Pope, A second-order Monte Carlo method for the solution of the Itô stochastic differential equation, *Stochastic Anal. Appl.* 40:151–186 (1986).
19 C. C. Heyde, New developments in inference for temporal stochastic processes, in *Stochastic Methods for Biological Intelligence*, Special Issue of J. Stat. Plann. & Inf., (1990).
20 J. E. Hutton and P. I. Nelson, Quasilikelihood estimation for semimartingales, *Stochastic Processes Appl.* 22:245–257 (1986).
21 N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North Holland, Kodansha, New York, 1981.
22 R. Janssen, Discretization of the Wiener-process in difference methods for stochastic differential equations, *Stochastic Processes Appl.* 18:361–369 (1984).
23 P. I. M. Johanssema, Diffusion models for the stochastic activity of neurons, in *Neuronal Networks* (E. R. Caianello, Ed.), Springer-Verlag, New York, 1968.
24 D. H. Johnson and A. Swami, The transmission of signals by auditory-nerve fiber discharge patterns, *J. Acoust. Soc. Amer.* 74:493–501 (1983).
25 G. Kallianpur, On the diffusion approximation to a discontinuous model for a single neuron, in *Contributions to Statistics* (P. K. Sen, Ed.), North-Holland, Amsterdam, 1983.

26 G. Kallianpur and R. Wolpert, Infinite dimensional stochastic differential equation models for spatially distributed neurons, *Appl. Math. Optim.* 12:125–172 (1984).
27 E. R. Kandel and J. H. Schwartz, *Principles of Neural Science*, Elsevier North Holland, New York, 1981.
28 P. Lansky, Inference for the diffusion models of neuronal activity, *Math. Biosci.* 67:247–260 (1983).
29 P. Lansky and V. Lanska, Diffusion approximation of the neuronal model with synaptic reverse potentials, *Biol. Cybernet.* 56:19–26 (1987).
30 Liptser and A. N. Shiryayev, A functional central limit theorem for semimartingales, *Theory Probab. Appl.* XXV(4):683–703 (1980).
31 R. S. Liptser and A. N. Shiryayev, On necessary and sufficient conditions on a central limit theorem for semi-martingales, *Theory Probab. Appl.* XXVI(1):132–137 (1981).
32 Y. Magshoodi and C. J. Harris, On probability approximation and simulation of non-linear jump-diffusion stochastic differential equations, *IMA J. Math. Control Inform.* 1:1–28 (1987).
33 I. W. McKeague, Estimation for semimartingale model using the method of sieves, *Ann. Statist.* 13:579–589 (1986).
34 I. W. McKeague, Asymptotic Theory for Sieve Estimator and Semimartingale Regression Models, Tech. Report M722, Florida State Univ., 1986.
35 H. T. Nguyen and T. D. Pham, Identification of non-stationary diffusion models by the method of sieves, *SIAM J. Control Optim.* 20:603–611 (1982).
36 E. Paradoux and D. Talay, Discretization and simulation of stochastic differential equations, *Acta Appl. Math.* 3:23–47 (1985).
37 W. Rall, Core conductor theory and cable properties of neurons, in *Handbook of Physiology—The Nervous System I*, Vol. 1, Amer. Physiol. in Soc., Bethesda, Md., 1978.
38 L. M. Ricciardi, Diffusion approximation for a multi-input model neuron, *Biol. Cybernet.* 24:237–240 (1976).
39 L. M. Ricciardi, *Diffusion Processes and Related Topics in Biology*, Lecture Notes in Biomath. 14, Springer-Verlag, New York, 1979.
40 L. M. Ricciardi, Diffusion approximations and computational problems for single neuron activity, in *Competition and Cooperation in Neural Networks* (S. Amari and M. A. Arbib, Eds.), Lecture Notes in Biomath. 45, Springer-Verlag, New York, 1982, pp. 143–154.
41 L. M. Ricciardi and L. Sacerdote, The Ornstein-Uhlenbeck process as a model for neuronal activity, *Biol. Cybernet.* 35:1–9 (1979).
42 L. M. Ricciardi, L. Sacerdote, and S. Sato, Diffusion approximation and first passage time problem for a model neuron, *Math Biosci.* 64:29–44 (1983).
43 S. Sato, Diffusion approximation for the stochastic activity of a neuron and moments of the interspike interval distribution, in *Progress in Cybernetics and System Research*, Vol. VI (R. F. Pichler and R. Trappl, Eds.), Hemisphere, Washington, 1982.
44 A. N. Shiryayev, Martingales: Recent developments, results and applications, *Internat. Statist. Rev.*, 1981, pp. 199–233.

45 D. L. Snyder, *Random Point Processes*, Wiley, New York, 1975.
46 L. M. Sørensen, On maximum likelihood estimation in randomly stopped diffusion-type processes, *Internat. Statist. Rev.* 51:93–110 (1983).
47 R. B. Stein, Some models of neuronal variability, *Biophys. J.* 5:173–195 (1965).
48 H. C. Tuckwell, Synaptic transmission in a model for stochastic neural activity, *J. Theoret. Biol.* 77:65–81 (1979).
49 H. C. Tuckwell and D. K. Cope, Accuracy of neuronal interspike times calculated from a diffusion approximation, *J. Theoret. Biol.* 83:377–387 (1980).
50 A. Thavaneswaran, Estimation for Semimartingales, Ph.D. Thesis, Univ. of Waterloo, Waterloo, Ontario, Canada, 1985.
51 A. Thavaneswaran and M. E. Thompson, Optimal estimation for semimartingales, *J. Appl. Probab.* 23:409–417 (1986).
52 J. B. Walsh, A stochastic model of neuronal response, *Adv. in Appl. Probab.* 1981, pp. 231–281.
53 F. Y. M. Wan and H. C. Tuckwell, The response of a spatially distributed neuron to white noise current injection, *Biol. Cybernet.* 33:39–55 (1979).
54 A. LeBreton, On continuous and discrete sampling for parameters estimation in diffusion processes, Math. Prog. Study 5:124–144 (1976).