

**INFERENCE FOR STOCHASTIC NEURONAL MODELS**

**by**

M.K. Habib and A. Thavaneswaran

The University of North Carolina at Chapel Hill and  
Temple University

Institute of Statistics Mimeo Series No. 1846

March 1988

## INFERENCE FOR STOCHASTIC NEURONAL MODELS

M.K. Habib<sup>1</sup> and A. Thavaneswaran<sup>2</sup>

<sup>1</sup>The University of North Carolina at Chapel Hill

and

<sup>2</sup>Temple University

### 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 parameter estimation of physiologically meaningful parameters of the developed models. Both ordinary and partial stochastic differential equation models are treated.

AMS 1980 Subject Classification: 60G55, 60K99, 62F99

Key Words: Asymptotic Normality, Consistency, Parameter Estimation, Neuron, Neuronal Models, Ordinary Stochastic Differential Equations, Semimartingales, Stochastic Partial Differential Equations.

Research supported by the Office of Naval Research under Contract Number N00014-03-0387.

## 1. INTRODUCTION

This paper is concerned with the development and application of stochastic models and methods of statistical inference 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 spatio-temporal 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 solution of stochastic differential equations. These models contain neurophysiologically meaningful parameters such as the effective somal membrane time constant, amplitudes and frequency of occurrence of post-synaptic 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 Section 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 methods of optimal

estimating functions is applied in this case (Godambe, 1985; Thavaneswaran and Thompson, 1986). 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 spatio-temporal 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 stochastic partial differential equations. 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 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 functions (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 terminal 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 pre-synaptic terminal of one cell (the pre-synaptic cell) and the receptive surface of the other cell (the post-synaptic cell) (Kandel and Schwartz, 1981.) Axonal terminals which 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 ( $\text{Na}^+$ ), potassium ( $\text{K}^+$ ), and chloride ( $\text{Cl}^-$ ). 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 m.V., 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 and combines with the receptor sites of the post-synaptic membrane and produces a change in potential. This potential change is called post-synaptic potential (PSP). A post-synaptic potential that polarizes the post-synaptic membrane is called an excitatory postsynaptic potential (EPSP), and one that depolarizes the neuronal membrane is called an inhibitory post-synaptic potential (IPSP). The PSPs perturb the membrane potential and if a certain potential level (called the neuron's threshold: -35 to -45 m.V.) is reached, the membrane potential goes through a rapid

stereotypical change during which it reaches about +30 m.V. and declines rapidly back to values less than -70 m.V. within 2 to 3 millisecond. 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. The time interval between the moments at which the membrane potential is at a resting level until it reaches threshold is obviously random in nature. More precisely, it is a first passage time of the membrane potential of the neuronal threshold. See Kandel and Schwartz (1981) 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  $(\Omega, F, P)$ . It is subject to instantaneous changes due to the occurrence of excitatory post-synaptic potentials (EPSPs) which are assumed to occur according to mutually independent Poisson processes  $P(\lambda_k^e; t)$  with rates  $\lambda_k^e$  ( $k=1, 2, \dots, n_1$ ), each accompanied by an instantaneous displacement of  $V(t)$  by a constant amount  $\alpha_k^e > 0$  ( $k=1, 2, \dots, n_1$ ), and inhibitory post-synaptic potentials (IPSP) which occur according to independent Poisson processes  $P(\lambda_k^i; t)$  with amplitudes  $\alpha_k^i > 0$  ( $k=1, 2, \dots, n_2$ ). Between PSPs,  $V(t)$  decays exponentially to a resting potential with time constant  $\tau$ . 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. Following the action potential,  $V(t)$  is reset to a resting potential with a time constant  $\tau$ . 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

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

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 (Stein, 1965) and is a special case of the well known Poisson driven Markov process models (see Snyder, 1975). This model has been treated in the literature by many authors among them Johannesma (1968) and Tuckwell (1979).

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. Ricciardi, 1976; 1979; 1982; Ricciardi and Sacerdote, 1979; Ricciardi, Sacerdote and Sato 1983; Sato, 1982; Lansky and Lansky, 1987). These types of approximation are justified on the grounds that for many types of neurons in the central nervous system, synapses are densely packed along the dendritic tree. For example, there exists 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 post-synaptic potentials are very large, then the approximation of the Poisson driven Markov model by a diffusion model is appropriate and is accomplished by allowing the amplitudes  $\alpha^e$ ,  $\alpha^i$  to tend to zero and the frequencies  $\lambda^e$ ,  $\lambda^i$  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 (1980). Kallianpur (1983) established a diffusion approximation for model (2.1) using the functional central limit theorem for semimartingales of Liptser and Shiriyayev (1980, 1981). Under some regularity conditions it was shown that model (2.1) can be approximated by the diffusion model

$$(2.2) \quad dV(t) = (-\rho V(t) + \mu) dt + \sigma dW(t), \quad 0 \leq t \leq T,$$

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

Next we consider a stochastic neuronal model which takes into account the influence of extensive low amplitude synaptic input as well as PSPs with large amplitudes, which may be reflecting the influence of a number of dominating synapses. These synapses may be electronically close to the initial segment. 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 Ikeda and Watanabe, 1981). 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  $\alpha_k^e$ ,  $k=1,2,\dots,n_1$ . 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  $\alpha_k^i$ ,  $k=1,2,\dots,n_2$ . This set up leads to the following extended mixed model to describe the membrane potential of a stimulus driven neuron:

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

The problem of parameter estimation of the mixed model 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 (1983, 1984) 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  $T \rightarrow \infty$ . 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  $0 \leq t \leq T$  as  $T \rightarrow \infty$ . (See e.g. Brown and Hewitt, 1975; Feigen, 1978; Baswa and Prakasa Rao, 1980). For an extensive review of such problems see Bellach (1980; 1983). Given  $n$  independent trajectories  $\{V_k(t), \tau_{k-1} \leq t \leq \tau_k\}$   $k=1,2,\dots,n$ , where,  $\tau_1, \tau_2, \dots, \tau_n$  are independent random variables with  $P(\tau_k < \infty) = 1$ ,  $k=1,2,\dots,n$ , Habib (1985) derived maximum likelihood estimators of the parameters  $\rho$  and  $\mu$  and established their large sample properties such as strong consistency and asymptotic normality assuming  $\sigma$  is known.



Now recall the diffusion neuronal model (2.2). From Sørensen (1983), the log-likelihood function is given by

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

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

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

and

$$(2.6) \quad \hat{\mu}_n = \frac{\left[ \sum_{k=1}^n \int_{\tau_{k-1}}^{\tau_k} V_k^2(t) dt \right] \left[ \sum_{k=1}^n \int_{\tau_{k-1}}^{\tau_k} dV_k(t) \right] - \left[ \sum_{k=1}^n \int_{\tau_{k-1}}^{\tau_k} V_k(t) dV_k(t) \right] \left[ \sum_{k=1}^n \int_{\tau_{k-1}}^{\tau_k} V_k(t) dt \right]}{\left[ \sum_{k=1}^n (\tau_k - \tau_{k-1}) \right] \left[ \sum_{k=1}^n \int_{\tau_{k-1}}^{\tau_k} V_k^2(t) dt \right] - \left[ \sum_{k=1}^n \int_{\tau_{k-1}}^{\tau_k} V_k(t) dt \right]^2}$$

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

$$(2.7) \quad \hat{\sigma}^2(k) = \frac{1}{(\tau_k - \tau_{k-1})} \lim_{m \rightarrow \infty} \sum_{j=1}^{2^m} \left[ V_k(\tau_{k-1} + j d_k^{2^{-m}}) - V_k(\tau_{k-1} + (j-1) d_k^{2^{-m}}) \right]^2.$$

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 Wiener process). A natural estimate of  $\sigma^2$  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  $n \rightarrow \infty$ ) have been established in Habib (1985).

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 Ferster (1987) where extracellular 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  $\rho$  and  $\mu$  of 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 (c.f. (2.5) and (2.6)). But, in practice one has to approximate these integrals with appropriate finite sums which depend on the digitization scheme or the partition mesh  $\{t_0, t_1, \dots, t_n\} \subset [0, T]$ .

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

$$(2.8) \quad V(t_{k+1}) = (-\rho V(t_k) + \mu)h + \sigma(W(t_{k+1}) - W(t_k))$$

where  $h = \frac{T}{N}$ ,  $t_k = kh$ ,  $k=1, 2, \dots, N$ . It is well known that the solution of (2.8) converges to  $V(t)$ . For instance, if we set  $V_N(t) = V(t_k)$  for  $t \in [t_k, t_{k+1})$ , then

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

as  $N \rightarrow \infty$  (see Gihman and Skorokhod, 1979). This and other kinds of discretization, specially Rung-Kutta schemes, have been extensively studied (see e.g. Janssen, 1984; Paradox and Taly, 1985; Haworth and Pope, 1986; Magshoodi and Harris, 1987).

If  $\hat{\theta}_{n,N,T}$  ( $\tilde{\theta}_{n,N,T}$ ), where  $\theta = (\rho, \mu)$ , denotes the resulting estimate (respectively, optimal or ML estimates) based on observations at  $t_0, t_1, \dots, t_N$ . LeBreton (1976) showed that when  $\max(t_{i+1} - t_i) = \delta_N \rightarrow 0$ ,

$$P - \lim_{n \rightarrow \infty} \hat{\theta}_{n,N,T} = P - \lim_{n \rightarrow \infty} \tilde{\theta}_{n,N,T} = \hat{\theta}_{n,T},$$

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 the following tables.

Table 2.1: Parameter estimates using a simulated diffusion process observed n-times over a fixed period [0,T] and sampled every  $\sigma$  units:

$$T = 10 \text{ m.s.}, \delta = 0.10.$$

Parameters	True Value	Estimated Value n=1	Estimated Value n=10	Estimated Value n=50
$\rho = \tau^{-1}$	0.33333	0.30336	0.33000	0.33427
$\mu$	5.00000	4.63803	4.84648	4.88702
$\sigma$	0.31623	0.67566	0.67364	0.67583

Table 2.2: Parameter estimates using a simulated diffusion process observed n-times over a fixed period [0,T] and sampled every  $\delta$  units:

$$T = 20 \text{ m.s.}, \delta = 0.10.$$

Parameters	True Value	Estimated Value n=1	Estimated Value n=10	Estimated Value n=50
$\rho = \tau^{-1}$	0.33333	0.30369	0.32705	0.32399
$\mu$	5.00000	4.86121	4.77822	4.71001
$\sigma$	0.31623	0.33012	0.51796	0.33537

It is clear from Table 2.1 that for processes which are observed over a period  $[0, T]$ , with  $T=10$  m.s. the estimates of all parameters are very close and that they improve as the number of observed trajectories,  $n$ , increases. From Table 2.2, 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 2.2 the period of observation  $[0, T]$  was longer,  $T = 15$  m.s. Therefore, one may conclude that for action potentials with long durations, one does not gain much by recording a large number of spikes, but for action potentials with relatively short durations, one can expect that the parameter estimators will improve as the number of observed action potentials increases.

Tables (2.3) and (2.4) show similar behavior even when the sampling interval of the simulated trajectories reduces to  $\delta = 0.05$ . However, Table (2.5) shows that the estimates deteriorate when the diffusion coefficient  $\sigma$  increases to  $\sigma = 3.0$ .

Table 2.3: Parameter estimates using a simulated diffusion process observed  $n$ -times over a fixed period  $[0, T]$  and sampled every  $\delta$  units:

$T=10$  m.s.,  $\delta = 0.05$ .

Parameters	True Value	Estimated Value $n=1$	Estimated Value $n=10$	Estimated Value $n=50$
$\rho = \tau^{-1}$	0.33333	0.30627	0.32898	0.33620
$\mu$	5.00000	4.87798	4.75184	4.83044
$\sigma$	0.31623	0.53005	0.51272	0.51828

Table 2.4: Parameter estimates using a simulated diffusion process observed n-times over a fixed period [0,T] and sampled every  $\delta$  units:

$$T = 15 \text{ m.s.}, \delta = 0.05$$

Parameters	True Value	Estimated Value n=1	Estimated Value n=10	Estimated Value n=50
$\rho = \tau^{-1}$	0.33333	0.34615	0.32371	0.33153
$\mu$	5.00000	4.93817	4.68444	4.77649
$\sigma$	0.31623	0.46003	0.45486	0.45778

Table 2.5: Parameter estimates using a simulated diffusion process observed n-times over a fixed period [0,T] and sampled every  $\delta$  units:

$$T = 15 \text{ m.s.}, \delta = 0.05.$$

Parameters	True Value	Estimated Value n=1	Estimated Value n=10	Estimated Value n=50
$\rho = \tau^{-1}$	0.33333	0.339721	0.24716	0.24202
$\mu$	5.00000	4.34994	2.63800	2.58367
$\sigma$	*3.00000	2.89161	2.91949	2.91685

#### 2.4 A Neuronal Model with Time Dependent Parameters

In this section, 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 membrane potential clearly reveal that the drift parameter  $\mu$  in (2.2) is a function of time rather than a constant. Furthermore, replacing the membrane's rate of  $\rho$  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 neuronal membrane potential is considered.

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

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

$$\{X_k(t), \tau_{k-1} \leq t \leq \tau_k\}, \quad k=1, 2, \dots, n.$$

The log-likelihood function is given by

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

It should be noted here that the technique for estimating finite dimensional parameters usually fails in the finite dimensional case, and we are forced to consider the method of sieves (see Grenander, 1981). 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 Geman and Hwang (1982).

Following Nguyen and Pham (1982), 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  $U_n \leq U_{n+1}$ ,  $V_n \leq V_{n+1}$ , and  $\bigcup_{n \geq 1} U_n$  and  $\bigcup_{n \geq 1} V_n$  are dense in  $L^2([0, T], dt)$  such that  $(\theta_1, \dots, \theta_{d_n})$  and  $(\psi_1, \dots, \psi_{d'_n})$  form the basis of  $U_n$  and  $V_n$  respectively, for all  $n \geq 1$ . For  $\theta \in U_n$  and  $\mu \in V_n$  with

$$\theta(\cdot) = \sum_{i=1}^{d_n} \theta_i \phi_i(\cdot),$$

and

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

The likelihood function in this case takes the form

$$(2.10) \quad L_n(\theta, \mu) = \sum_{k=1}^n \left\{ \int_{\tau_{k-1}}^{\tau_k} \left[ \left( \sum_{i=1}^{d_n} \theta_i \phi_i(t) \right) X_k(t) + \sum_{j=1}^{d'_n} \mu_j \psi_j(t) \right] dX_k(t) - 1/2 \int_{\tau_{k-1}}^{\tau_k} \left[ \left( \sum_{j=1}^{d_n} \theta_j \phi_j(t) \right) X_k(t) + \sum_{j=1}^{d'_n} \mu_j \psi_j(t) \right]^2 dt \right\}.$$

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 McKeague (1986a,b).

## 2.5 A Semimartingale Neuronal Model

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

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

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

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

$\theta$  is the parameter of interest, and  $b$  is a non-negative predictable process. (For the relevant definitions see Habib and Thavaneswaran 1988.)  
Furthermore, assume that  $f$  is a continuously differentiable function.

Notice that no distributional assumptions have been imposed on the driving martingale  $M$  in model (2.11). Therefore, in order to estimate the model's parameters  $\theta$ , we must employ a likelihood free estimation method. We therefore, apply the method of optimal estimating function which was considered by Godambe (1960, 1985). This method was applied to continuous time stochastic process by Hutton and Nelson (1986), Thavaneswaran and Thompson (1986) and Heyde (1988).

Godambe's optimality criterion: Following Godambe (1985), consider parameter  $\theta$  to be a function of  $P \in \mathcal{P}$  (a family of probability measures). Let  $G(V, \theta) = (G_t(V, \theta), \mathcal{F}_t; t \geq 0)$  represent a family of processes indexed by  $\theta$  such that  $E_P G_t(V, \theta) = 0$  for each  $t \geq 0$ , for each  $P$ , and  $\theta = \theta(P)$ . This corresponds to



the unbiasedness property of Godambe (1960). Godambe's optimality criteria, adapted to this situation, reads:  $G^\circ$  is optimal in  $L$ , the class of unbiased estimating functions, if  $Q - A_h - A_h^\circ$  is non-negative definite for all  $G \in L$  and for all  $P \in P$ , where

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

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

and  $A$  is the variance-covariance matrix for  $h$  under  $\theta_0$  (the true value of  $\theta$ ). The following sufficient condition for optimality to hold is due to M.E. Thompson.

Lemma 3.1  $G^\circ$  is optimal in  $L$  if

$$E \left[ \frac{\partial G}{\partial \theta} \right] = K E[GG^{\circ'}].$$

Proof (c.f. Thavaneswaran, 1985, p.57)

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

$$(2.12) \quad G_{t,\theta} = \int_0^t a_{s,\theta} dM_{s,\theta}$$

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

$$E[G^{\circ'} G] = E \int_0^t a_{s,\theta} d\langle M \rangle_{s,\theta} a_{s,\theta}^{\circ'}$$

Where  $A'$  denotes the transpose of  $a$ , and

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

Hence

$$E[G_0 G_{\theta}^{\circ'}] = E \int_0^t a_{s,\theta} b_{s,\theta} \alpha'_{s,\theta} (dA_s)$$

Moreover

$$\begin{aligned} E \left[ \frac{\partial G_{\theta}}{\partial \theta} \right] &= E \int_0^t a_{s,\theta} \frac{\partial}{\partial \theta} (dM_{s,\theta}) \\ &= E \int_0^t a_{s,\theta} f'_{s,\theta} dA_s, \end{aligned}$$

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

$$(2.13) \quad G_{\theta}^{\circ} = \int_0^t f'_{s,\theta} b_{s,\theta}^+ dM_{s,\theta}$$

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

Example 2.1: In this example, model (2.1) is extended by replacing the Wiener process  $(W_t)$  in (2.2) 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  $\theta = \begin{bmatrix} -\rho \\ \mu \end{bmatrix}$ ,  $f(V_t) = \begin{bmatrix} V_t \\ 1 \end{bmatrix}$ ,  $\langle M \rangle_t = t$ , and  $\sigma > 0$  (assumed to be known). The optimal estimating function is given by

$$(2.14) \quad G_{n, \theta}^o = \sum_{k=1}^n \int_0^{\tau_k} \begin{bmatrix} V^{(k)} \\ 1 \end{bmatrix} dM_s^{(k)} \text{ a.s.}$$

Note:

(i) Equation (2.13) is the same as the m.l.e. equation in Habib (1985) 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 Visual 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 spatio-temporal aspects of synaptic input to a single neuron. Therefore, we develop spatio-temporal stochastic models of the subthreshold behavior of the somal membrane potential. These models should serve in evaluating the role of the geometry of the dendrites 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 spatio-temporal models.

An important neuronal characteristic is the dependence of both the magnitude and time course of the post-synaptic potential, evoked by 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 (triggering 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 the dendrites. Never the less, 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 directly and strongly neuronal firing behavior. A major exception of this view was suggested by Rall (1978), 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 spatio-temporal 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 model (2.2) in order to incorporate the important feature of spatial synaptic distribution. This extension is based on Rall's model neuron (Rall, 1978). In Rall's model neuron the cable properties of a system of branched dendrites are reduced to a one-dimensional equivalent dendrite, with synapses made 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 post-synaptic 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. Wan and Tuckwell, 1979; Walsh, 1981; Kallianpur and Wolpert, 1984). 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. This type of input 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

$s \leq t$ . The probability that exactly  $k$  such pulses arrive during the period  $(0, t]$  is  $e^{-\lambda} \lambda^k / k!$  with intensity  $\lambda = t \cdot \mu(A \times B)$ . That is,  $N$  is a Poisson measure on  $R \times X \times R^+$  with intensity measure  $dt \mu(da, dx)$ .

Let  $W(t, x)$  be a two parameter Wiener process defined on  $R_+ \times X$ . 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 parameters diffusion process:

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

(See Nualart, 1983, for a discussion of two parameters diffusion processes.) 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  $(V(t, x), 0 \leq x < L)$  of the stochastic differential equation:

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

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

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

Now, consider the differential operator

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

Then (3.3) takes the form

$$(3.4) \quad dV(t,x) = TV(t,x) dt + dM(t,x)$$

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 the infinite dimensional stochastic differential equations (SDE) of the type

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

where  $\theta$  belongs to a compact parameter set  $\Theta$  contained in the positive reals,  $T$  is the infinitesimal generator of a strongly continuous semigroup  $(T_t)$ ,  $t > 0$  acting on a real separable Hilbert space  $H$  with scalar product  $\{.,.\}$ , and the norm  $\|.\|$ .  $W_t$  is an  $H$ -valued Wiener process, i.e.  $W_t$  is a stochastic process defined on a complete probability space  $(\Omega, F, P)$  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 suppressed the dependence of the processes on the optical variable  $x$ .

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

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

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

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

where  $\langle \cdot, \cdot \rangle$  denotes 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 Godambe (1985).

Assume that  $I_{i,s,\theta} = \frac{\partial \beta_{i,s,\theta}}{\partial \theta}$  exists for all  $s$  (a.s.) and all  $i$ . Let  $\mathcal{U}$

be the class of  $G$  in (3.5) such that

- (a) for each  $\theta$  the process  $a_{s,\theta}$  is differentiable with respect to  $\theta$ ,
- (b) 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,

- (c) for all  $t$ ,  $\frac{\partial G_{t,\theta}}{\partial \theta}$  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

- (d) for all  $t$ ,  $E G_{t,\theta}^2$  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

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

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

$$(3.7) \quad E[G_{t,\theta} G_{t,\theta}^*] / E\left[\frac{\partial G_{t,\theta}}{\partial \theta}\right]$$

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

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

## REFERENCES

- Basawa, I.V. and Prakasa Rao, B.L.S. (1980) Statistical Inference for Stochastic Process, Academic Press, New York.
- Bellach, B. (1980) Consistency, asymptotic normality and efficiency of the maximum likelihood estimator in linear stochastic differential equations. Math. Operationsforsch. Statist. Ser. Statistics, 11, 227-266.
- Bellach, B. (1983) Parameter estimators in linear stochastic differential equations and their asymptotic properties. Math Operationforsch. Statist. Ser. Statistics, 14, 141-191.
- Brown, B.M. and Hewitt, J.I. (1975) Asymptotic likelihood theory for diffusion processes. J. Appl. Prob. 12, 228-238.
- Capocelli, R.M. and Ricciardi, L.M. (1971) Diffusion approximation and first passage time problem for a model neuron. Kybernetik 8, 214-223.
- Feigen, P.D. (1978) Maximum likelihood estimation for stochastic processes. Adv. Appl. Prob. 8, 712-736.
- Ferster, D. (1987) Origin of orientation selective EPSPs in neurons of cat visual cortex. J. Neurosci. 7, 1780-1791.
- Geman, S. and Hwang, C. (1982) Nonparametric maximum likelihood estimation by the method of sieves. Ann. Statist. 10, 401-414.
- Gihman, I.I. and Skorokhod, A.V. (1979) Stochastic Differential Equations, Springer-Verlag, New York.
- Godambe, V.P. (1960) An optimum property of regular maximum likelihood equation. Ann. Math. Statist. 31, 1208-11.
- Godambe, V.P. (1985) The foundations of finite sample estimation in stochastic processes. Biometrika 12, 419-428.
- Godambe, V.P. and Heyde, C.C. (1986) Quasilikelihood and optimal estimation. Intern. Statist. Rev. 55: 231-244.
- Grenander, V. (1981) Abstract Inference. John Wiley and Sons. New York.
- Habib, M.K. (1985) Parameter estimation for randomly stopped processes and neuronal modeling. UNC Institute of Statistics, Mimeo Series No. 1492.
- Habib, M.K. and Sen, P.K. (1985) Nonstationary stochastic point process models in neurophysiology with applications to learning. Biostatistics: Statistical, Biomedical, Public Health and Environmental Sciences. North Holland Amsterdam. 481-509.

Habib, M.K. and Thavaneswaran, A. (1988) Optimal Estimation for Semimartingale Neuronal Models. Proceeding of a conference on "Stochastic Methods for Biological Intelligence" Plenum Publishing Corp: Editors M.K. Habib and J. Davis.

Haworth, D.C. and Pope, S.B. (1986) A second-order Monte-Carlo method for the solution of the Itô stochastic differential equation. Stochastic Analysis Appl. 40, 151-186.

Heyde, C.C. (1988) New developments in inference for temporal stochastic processes. Proceeding of a conference on "Stochastic Methods for Biological Intelligence" Plenum Publishing Corp: Editors M.K. Habib and J. Davis.

Hutton, J.E. and Nelson, P.I. (1986) Quasilikelihood estimation for semimartingales. Stoch. Proc. Appl. 22, 245-257.

Ikeda, N. and Watanabe, S. (1981) Stochastic Differential Equations and Diffusion Processes. North Holland, Kodansha. New York.

Janssen, R. (1984). Discretization of the Wiener-process in difference methods for stochastic differential equations. Stochast. Proc. Appl. 18, 361-369.

Johanssema, P.I.M. (1968) Diffusion models for the stochastic activity of neurons. In Caianello, E.R., (ed.) Neuronal Networks. New York, Springer Verlag.

Kallianpur, G. (1983) On the diffusion approximation to a discontinuous model for a single neuron. In: Sen P.K. (ed) Contributions to Statistics. North-Holland, Amsterdam.

Kallianpur, G. and Wolpert, R. (1984) Infinite dimensional stochastic differential equation models for spatially distributed neurons. Appl. Math. Optim. 12, 125-172.

Kandel, E.R. and Schwartz, J.H. (1981) Principles of Neural Science. Elsevier North Holland. New York.

Lansky, P. (1983) Inference for the diffusion models of neuronal activity. Math. Biosci. 67, 247-260.

Lansky, P. (1984) On approximations of Stein's neuronal model. J. Theor. Biol. 107, 631-647.

Lansky, P. and Lanska, V. (1987) Diffusion approximation of the neuronal model with synaptic reverse potentials. Biol. Cybern. 56, 19-26.

Liptser and Shiriyayev (1980) A functional central limit theorem for semimartingales. Theor. Prob. Appl. XXV, 4, 683-703.

Liptser and Shiriyayev (1981) On necessary and sufficient conditions on a central limit theorem for semi-martingales. Theor. Prob. Appl. XXVI, 1, 132-137.

Magshoodi, Y. and Harris, C.J. (1987) On probability approximation and simulation of non-linear jump-diffusion stochastic differential equations. IMA J. Math. Cont. Inform. 1, 1-28.

McKeague, I.W. (1986a) Estimation for a semimartingale model using the method of sieves. Ann. Statist. 13, 579-589.

McKeague, I.W. (1986b) Asymptotic theory for seive estimator and semimartingale regression models. Tech Report #M722, Florida State University.

Nguyen, H.T. and Pham, T.D. (1982) Identification of non-stationary diffusion models by the method of sieves. SIAM J. Control Optim. 20, 603-611.

Paradoux, E. and Talay, D. (1985) Discretization and simulation of stochastic differential equations. Acta. Applic. Math. 3, 23-47.

Rall, W. (1978) Core conductor theory and cable properties of neurons. Handbook of Physiology - The Nervous System I. Vol. 1. American Physiological Society, Bethesda, Maryland.

Ricciardi, L.M. (1976) Diffusion approximation for a multi-input model neuron. Biol. Cybernet. 24, 237-240.

Ricciardi, L.M. (1979) Diffusion processes and related topics in biology. Lecture Notes in Biomathematics 14. Springer, Berlin Heidelberg New York.

Ricciardi, L.M. (1982) Diffusion approximations and computational problems for single neurons activity. In: Amari S, Arbib MA (eds) Competition and cooperation in neural networks. Lecture Notes in Biomathematics 45, 143-154.

Ricciardi, L.M. and Sacerdote, L. (1979) The Ornstein-Uhlenbeck process as a model for neuronal activity. Biol. Cybernet. 35, 1-9.

Ricciardi, L.M., Sacerdote, L., and Sato, S. (1983) Diffusion approximation and first passage time problem for a model neuron. Math Biosci. 64, 29-44.

Sato, S. (1982) Diffusion approximation for the stochastic activity of a neuron and moments of the interspike interval distribution. In: Pichler R.F., Trapp R. (eds) Progress in Cybernetics and System Research, Vol VI. Hemisphere, Washington.

Shiryayev, A.N. (1981) Martingales: Recent developments, results and applications. Int. Statist. Rev. 199-233.

Snyder, D.L. (1975) Random Point Processes. J. Wiley and Sons. New York.

Sørensen, L.M. (1983) On maximum likelihood estimation in randomly stopped diffusion-type processes, Intern. Statist. Rev. 51, 93-110.

Stein, R.B. (1965) Some models of neuronal variability. Biophysical J. 5, 173-195.

Tuckwell, H.C. (1979) Synaptic transmission in a model for stochastic neural activity. J. Theor. Biol. 77, 65-81.

Tuckwell, H.C. and Cope, D.K. (1980) Accuracy of neuronal interspike times calculated from a diffusion approximation. J. Theor. Biol. 83, 377-387.

Thavaneswaran, A. (1985) Estimation for semimartingales. Ph.d. Thesis. University of Waterloo, Waterloo, Ontario, Canada.

Thavaneswaran, A. and Thompson, M.E. (1986) Optimal estimation for semimartingales. J. Appl. Prob. 23, 409-417.

Walsh, J.B. (1981) A stochastic model of neuronal response. Adv. Appl. Prob., 231-281.

Wan, F.Y.M. and Tuckwell, H.C. (1979) The response of a spatially distributed neuron to white noise current injection. Biol. Cybern 33, 39-55.