Nerve Cell Spike Train Data Analysis: A Progression of Technique

DAVID R. BRILLINGER*

Collections of occurrence times of events taking place irregularly in time provide a fairly common, but not broadly discussed, data type. This article is concerned with the particular circumstance of firing times in nerve cells that interact and form networks. The article reviews a progression of statistical analysis techniques: description, association as measured by moments and correlation, regression, and finally likelihood. The data is point process, but may be seen as that of regression and of multivariate analysis in standard parlance. A simple description of data collected simultaneously for one or more cells is provided.

KEY WORDS: Binary data; Nerve cell; Network; Point process; Probit analysis; Semiparametric model.

". . . the purpose of inductive reasoning, based on empirical observations, is to improve our understanding of the systems from which these observations are drawn."
Sir R. A. Fisher (1956)

The above statement sets down the spirit of applied statistics. The related goal of this article is a better understanding of the nerve cell system and the construction of better quantitative models of the neuronal firing phenomenon. On the substantive side, the author's collaborator J. P. Segundo has remarked that "the biological goal is understanding in strictly biological terms." This may be viewed as an ultimate goal. The models will change, but the biology will remain.

R. A. Fisher was central to the development of statistics, in particular to the progression of data analysis techniques from description and simple measures of association to the tools of association and regression analysis and finally to likelihood analysis. This article aims to illustrate the same progression for a data type of some contemporary interest—point process data—and to continue on to nonparametric and semiparametric likelihood analysis.

The article is concerned with a particular biological system—small networks of neurons communicating with each other and responding to stimuli. The system studied is of basic interest on both scientific and theoretical grounds. Scientific interest follows from a concern as to how the nervous system works; theoretical interest results in part from the system's strong nonlinearity.

Data from two different living preparations are studied. First discussed are some data for the cat collected by A. E. P. Villa at Lausanne, Switzerland. In Villa's experiments, cats were subjected to sound stimuli and data for eight nerve cells recorded simultaneously (Villa 1988, 1990). Also studied are simultaneous data for networks of two and three identified nerve cells (in particular cells L2, L3, L5, and L10) of *Aplysia californica* (the sea hare) collected by J. P. Segundo at the University of California, Los Angeles (Bryant, Ruiz Marcos, and Segundo 1973; Bryant and Se-

gundo 1976). *Aplysia* is commonly studied by neurophysiologists because the nerve cells are large and accessible and a number are repeatedly identifiable.

As is the pleasant feature of most time series analyses, a broad variety of figures are presented. These figures are central to the analysis.

Important aspects of nerve cell firing not addressed in this article include spatial effects and intracellular data collection and analysis.

1. WHAT IS A NERVE CELL?

Neurons (or nerve cells) are basic building blocks of an animal's central communication system. They are input-output systems of a particular structure having important functions. It is pertinent to discuss both structure and function, because in biology often the two seem directly related. Functions include accumulating, processing, and transmitting information. A nerve cell receives messages through its *dendrites,* root-like strings susceptible to chemical stimulus. The messages propagate to the cell body, or *soma.* Out of the soma grows the axon, with many branches ending at *synapses,* the junctions of neural networks. Figure 1, taken from Cajal (1895), shows a collection of neighboring neurons. The arrows indicate the flow of information. The cell bodies are the five blobs, four of which are labeled A, B, C, and D. The axons run vertically downward from the bodies—except for E, which is an axon entering from a distance. The dendrites include the three treelike structures at the top susceptible to influence from E.

The dendrites absorb input from other neurons through chemical processes that change ionic conductances and thereby induce current flows. The input is thence converted to a *membrane potential* throughout the soma. At the *axon hillock* (or trigger zone), the membrane potential occasionally reaches a threshold and the neuron fires, that is, generates an *action potential* (or spike). This action potential propagates along the axon to synapses, at which point a chemical transmitter is released to affect other neurons. The action potentials are of near-identical size and shape; see the spikes in Figure 2, which shows measured voltage fluctuations within cell R2 of *Aplysia* (Bryant and Segundo 1976). It may be argued that, because of reduced sensitivity to noise, the firing
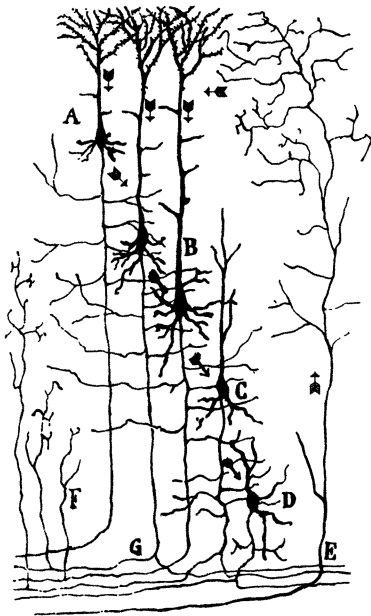
Figure 1. Drawing of Cajal (1895) Illustrating a Network of Five Cells of the Cerebral Cortex. The arrows (↧) suggest that the input arrives along the fiber E and progresses from it both directly and indirectly to the cells A, B, C, and D of different types.

times are the crucial variates in communication among neurons. Some discussion of the reduction to point processes is given in Segundo, Altshuler, Stiber, and Garfinkel (1991).

Synaptic connections may be *excitatory* or *inhibitory*; that is, depending on the type of connection, the firing of one neuron may make a second neuron either more likely or less likely to fire. Neurons also may fire spontaneously with no outside stimulus. Further is the phenomenon of *refractoriness,* wherein after a neuron has fired, the chance of it firing again is reduced (perhaps to zero) for a period.

Questions of interest include:

1. Can an analytic model incorporating the basic features of neuron behavior be developed and fit?

2. Given the firing times of a network of neurons, can one infer their causal connections?
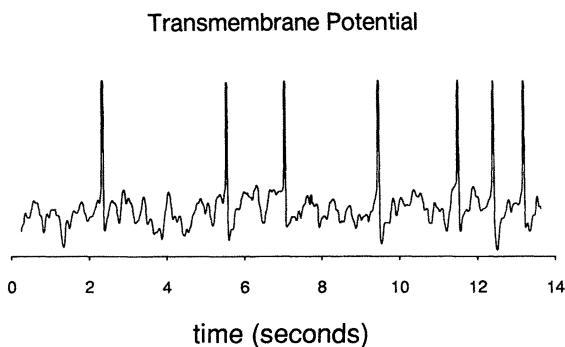
Transmembrane Potential

time (seconds)

Figure 2. Fluctuating Intracellular Voltage of the Cell R2 of Aplysia Showing the Occurrence of Point Process Data. The amplitudes of the spikes are approximately 100 millivolts. Figure adapted from Bryant and Segundo (1976).
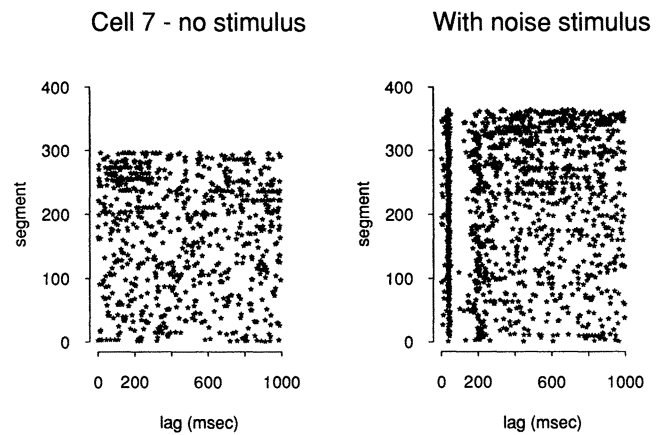

Figure 3. Rastor Plots Providing Nerve Cell 7 Firing Times. Each row of asterisks represents the events in a time interval of 1,000 msec. In the left panel, there was no experimental stimulus. In the right panel, a noise stimulus was applied at the beginning of each time interval, so each column represents the events occurring at the indicated lag after the noise stimulus.

General references for pertinent neurophysiological background information include Koch and Segev (1989), Segundo (1968, 1984, 1986), Segundo et al. (1991), and Stein (1972).

2. WHAT ARE POINT PROCESS DATA?

A stretch of point process data is a set of ordered numbers

$$\tau_1 \leq \tau_2 \leq \cdots \leq \tau_K,$$

to be thought of as the times of events that occurred in some time interval, say $(0, T)$. Usual examples are the times of telephone calls and the times of particle emission by some radioactive material. A naive descriptive statistic derived from such data is the observed rate, given here by K/T. This statistic has dimensions of counts per unit of time and is useful in elementary comparisons of point process behavior. For the data studied in this article, the rates range from about 1 spike per second to about 20 spikes per second. Figure 2 shows 7 spikes in about 14 seconds.

Descriptive statistics conducive to insight are provided by the plots in Figure 3. These plots are based on data collected in experiments studying the auditory system of the cat. Microelectrodes were inserted in a cat's brain at a location related to hearing. The plots refer to firing times for a single particular nerve cell (cell 7) which the probe happened upon. In the case of the lefthand plot, there is no applied stimulus. To describe the plot, suppose that the observation period is broken up into L segments, each 1,000 milliseconds long. Let τ_{kl} refer to the time elapsed since the start of the lth segment, of the kth spike of that segment. The points plotted are now $\{(\tau_{kl}, k), k = 1, \ldots, K_l\}$ for $l = 1, \ldots, L$. No dramatic structure is apparent in the lefthand panel. The second panel of Figure 3 refers to the same experiment but with a noise stimulus introduced into the ears of the cat every 1,000 milliseconds. The points are plotted as before, with τ referring to the time elapsed since each stimulus presentation. This picture shows that this neuron typically fires a short time after the application of the stimulus. Then there is a time period during which the neuron is unlikely to fire
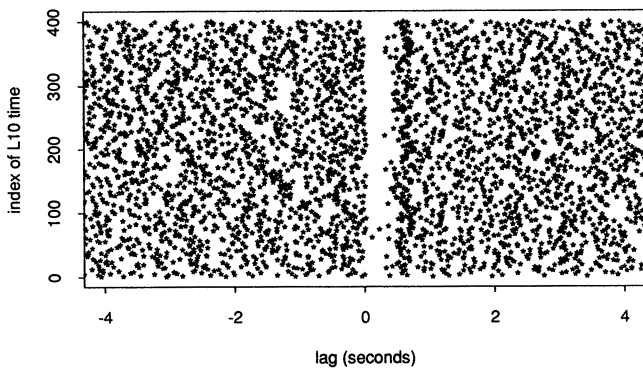
L3 spikes following L10's

Figure 4. Times of Neuron L3's Firings Relative to L10's Firings. Each row corresponds to a single firing of L10.

and perhaps then a rebound period when the cell is more likely to fire. Plots such as those in Figure 3 are known as *rastor plots*.

A second set of experimental data of some interest comes from experiments with *Aplysia,* the sea hare. Suppose that firing times are available for two related neurons—in the analysis to be presented, neurons L3 and L10 of *Aplysia.* Let $\{\sigma_j\}$ represent the firing times of L10 and $\{\tau_k\}$ represent the firing times of L3. In the case of these neurons, it "has been demonstrated almost beyond reasonable doubt" that L10 drives L3 (see Bryant, Ruiz, Marcos, and Segundo 1973 p. 205). Figure 4 plots the points $\{(\tau_k - \sigma_j, j), k = 1, 2, \ldots, K_j\}$ for $j = 1, 2, 3, \ldots$. This plot is consistent with the idea that firing of L10 tends to inhibit firing of L3. There is an indication of a brief acceleration or rebound at a lag of about .5 second. The bulk of the points appear randomly distributed.

To progress with the analysis, it is convenient to introduce some probability structure. A *stochastic point process* is a random process whose realizations are collections of points $\{\tau_k\}$, ordered by $\tau_k \leq \tau_{k+1}$, on the interval $(-\infty, \infty)$. Such a process can be described by giving the joint distributions of all the $N(I_1), \ldots, N(I_J)$, where I_j is a Borel set and $N(I_j)$ is the number of points falling in I_j for $j = 1, \ldots, J$ and $J = 1, 2, \ldots$. The process is said to be *stationary* when the joint distributions are unaffected by simple time translation, $I \rightarrow I + t$. An alternate way to describe a point process is via the joint distributions of the intervals $Y_k = \tau_{k+1} - \tau_k$ between successive points. In the stationary case, the *rate* of the process is given by $E\{N(I)\}/|I|$, where $|I|$ is the length of the interval.

It is worth remarking that there are many similarities between the concepts and techniques of time series analysis and those of point process analysis; see Brillinger (1978), as well as the classic reference for the analysis of point process data, Cox and Lewis (1966).

3. ASSOCIATION—SECOND ORDER MOMENTS

In the case of a bivariate stochastic point process (M, N) with components $M \equiv \{\sigma_j\}$ and $N \equiv \{\tau_k\}$, one can define the *cross-intensity function*

$$\lim_{h\downarrow0} \Pr\{N \text{ point in } (u + t, u + t + h] \,|\, M \text{ point at } t\}/h.$$

This will be a function of lag u alone in the stationary case. This parameter may be estimated by

$$\frac{\#\{u + \sigma_j < \tau_k \leq u + \sigma_j + h\}}{\#\{\sigma_j\}h} \qquad (3.1)$$

for small $h > 0$. Figure 5 gives the estimate for the data of Figure 4. It is essentially the histogram of the $\{\tau_k - \sigma_j\}$ and comes from counting the points in vertical strips of Figure 4. In fact, because of simpler sampling properties (including more stable variance, more symmetric distribution, and more near normal distribution), it is often more convenient to plot the square root of the estimate (Brillinger 1976; this was done here. The horizontal dashed lines provide ± 2 standard error limits set about 0. The diagram shows a period of initial inhibition after L10's firing, followed by a rebound at about .3 second. In some sense, Figure 5 does not add new information to that of Figure 4; but it does provide a specific way to interpret and assess the phenomena that occur. This cross-intensity function provides a precise measure of second-order association in the stationary case.

If two processes are associated, one can anticipate that functions of their realizations will be correlated. A particular function to study, because of its simplifying characteristics, is the empirical Fourier transform. Consider the Fourier transforms of two stretches of point process data, specifically $\sum_{0<\sigma_j\leq T} e^{-i\lambda\sigma_j}$, $\sum_{0<\tau_k\leq T} e^{-i\lambda\tau_k}$ for $0 \leq \lambda < \infty$. The quantity

$$R_{MN}(\lambda) = \lim_{T\to\infty} \mathrm{corr}\{\sum e^{-i\lambda\sigma_j}, \sum e^{-i\lambda\tau_k}\}$$

is called the *coherency* at frequency λ. Its modulus-squared, $|R_{MN}|^2$, is called the *coherence*. The coherence lies between 0 and 1 and measures the extent of linear time invariant association between two processes (Brillinger, Bryant, and Segundo 1976).

Figure 6 provides an estimate of the coherence for the L10-L3 data above. The estimate is seen to be highest for frequencies $\lambda/2\pi$ less than 1 cycle per second. The dashed line in the figure gives the (approximate) 95% upper point

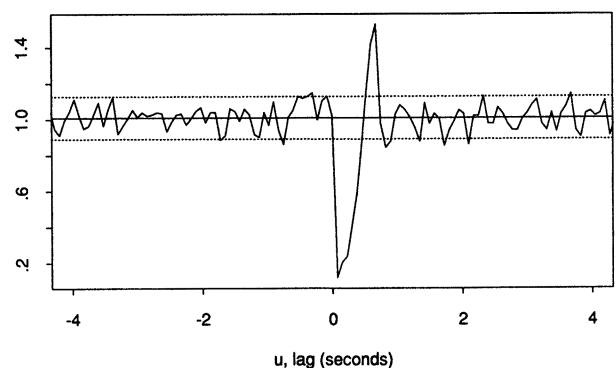## Square root crossintensity L3 given L10

Figure 5. The Square Root of the Cross-intensity Statistic (3.1). The dashed lines give upper and lower two standard error limits placed about 0 level.
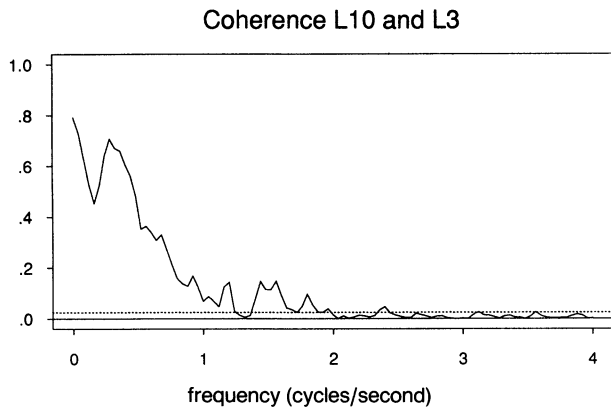
Coherence L10 and L3

Figure 6. An Estimate of the Coherence of Neurons L10 and L3 Obtained in the Fashion Described in Brillinger, Bryant, and Segundo (1976). Dashed line gives 95% null point.

of the null distribution of the estimate. Except in the case of simple translation of all the events by a common amount, mappings between realizations of point processes are inherently nonlinear. In view of this, the high magnitude here of the coherence estimate at the low frequencies is surprising.

4. REGRESSION—A LINEAR MODEL

Consider next a model

$$\lim_{h\downarrow 0} \Pr\{N \text{ point in } (t, t + h]|M\}/h = \mu + \sum_j a(t - \sigma_j).$$

$$(4.1)$$

This model is linear and time-invariant. The function $a(\cdot)$ is meant to represent the various chemical, electrical, spatial, and temporal delay processes involved in the influence of neuron M's firing on neuron N's firing. For example, if the τ's were given by $\tau_j = \sigma_j + Y_j$, with the Y's independent and of density function $a(\cdot)$, then the result (4.1) would hold with $\mu = 0$, see Brillinger (1974). The model (4.1) may be fit by cross-spectral analysis (Brillinger 1974). The resulting estimate of $a(\cdot)$ for the Aplysia data addressed in Section 3 is given in Figure 7. The estimate is seen to mimic that in
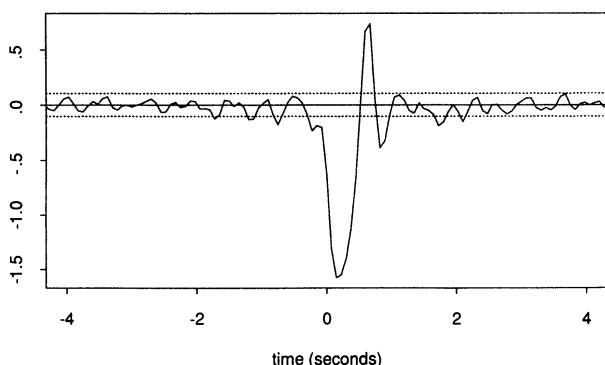
Impulse response of L3 given L10

Figure 7. An Estimate of the Function a(·) of (4.1) Obtained in the Fashion Described in Brillinger, Bryant, and Segundo (1976).

Figure 5. The distinction is that, as is the case in ordinary regression analysis, one is nearer to an object unaffected by elementary reexpressions. This analysis for this particular data set is not dramatically enlightening, but interesting examples may be found in Brillinger, Bryant, and Segundo (1976). The following section presents a more satisfying analysis of the present data in any case.

5. LIKELIHOOD—CONCEPTUAL MODELING

A model with a long history in neurophysiology involves a neuron firing when the membrane potential at its trigger zone exceeds a *threshold*. The threshold is a time-varying quantity that is reset to a high level on the neuron's firing and then is subject to slow (although not always monotonic) decay. The effect of the reset is to prevent firing from recurring immediately, and thus to incorporate the phenomenon of refractoriness. The model may be described in formal terms as follows: Let $M = \{\sigma_j\}$ refer to the times at which a first (or input) neuron fires. Given the function $a(\cdot)$, consider the following time-varying state variable

$$U(t) = \sum_{\sigma_j \leq t} a(t - \sigma_j).$$

$$(5.1)$$

The quantity $U(t)$ is meant to represent the membrane potential at time t at the trigger zone of the neuron whose firing is of interest. Here, $a(\cdot)$ is a *summation function*, meant to represent the various processes involved in the influence of M's firing on N's firing. The character of the function affects whether the firing of the neuron M increases (excites) or decreases (inhibits) the chance of the neuron N firing. The threshold decay is represented by the function $b(\cdot)$.

Figure 8 provides a layout of the situation. The bottom two panels give hypothetical $a(\cdot)$ and $b(\cdot)$ for the case of an inhibitory synapse. (Shortly, empirical estimates of $a(\cdot)$ and $b(\cdot)$ will be provided.) The vertical asterisks of the top plot are the firing times of the input neuron, M. The hook-shaped curves are the translates of the function $b(\cdot)$, with a new translate introduced with each firing of the principal neuron, N. If γ_t denotes the time elapsed since N's last firing, then the threshold curve may be represented by $\theta(t) = b(\gamma_t)$. The lower continuous curve of the figure is $U(t)$. One is concerned with the membrane potential, $U(t)$, crossing $\theta(t)$.

Consideration turns to developing a stochastic version of this model and of a corresponding likelihood function to employ in analyzing available data. Suppose first that the point processes are simplified to discrete time ($t = 0, \pm 1, \pm 2, \cdots$) and to 0-1 valued series. That is, a sampling interval of small length is selected such that only 0 or 1 points occur within each interval, and one defines $M_t = 1$ if there is a point in the unit interval starting at t and $M_t = 0$ if there is no point, for $t = 0, \pm 1, \pm 2, \ldots$ Corresponding discrete versions of N and $a(\cdot)$ are similarly defined. Now

$$U(t) = \sum a(t - \sigma_j) \approx \sum a_{t-u}M_u,$$

$$(5.2)$$

and it is convenient to represent the effect of the threshold by

$$\theta_t = \sum_{v=1}^{\gamma_t} b_v N_{t-v},$$

$$(5.3)$$

Membrane potential and threshold function

time (seconds)

Summation function

time (seconds)

Decay function

time (seconds)
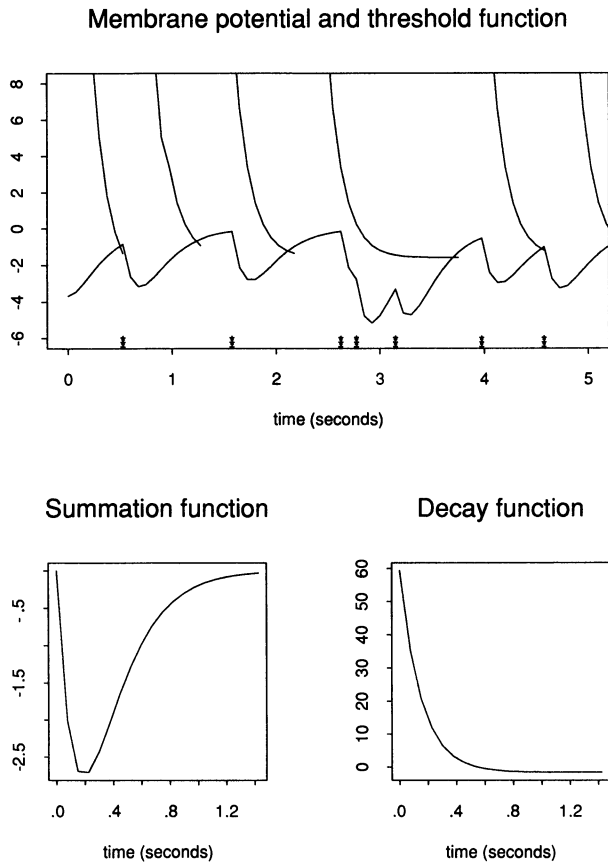
Figure 8. The Threshold Model. The lower curve of the top panel gives
U(t) of (5.1) with a(·) given by the lower left function. The hook-shaped
functions of the top panel are translates of the function of the lower right
panel initiated each time the curve U(t) is crossed. The spikes of the top
panel are the times of M firing.

with γ_t again the time elapsed since the last N firing. (That
the expression (5.3) is linear in the parameters aids in their
estimation.)

Suppose that there is noise, with c.d.f. $P(\cdot)$, superposed
on the threshold. This makes the model stochastic. The con-
ditional probability of the neuron firing, given the past, is
taken to be

$$P_t = \Pr\{N_t = 1 \mid \text{the past}\} = P(\psi_t), \qquad (5.4)$$

where

$$\psi_t = \sum a_u M_{t-u} - \theta_t.$$

The log-likelihood is

$$\sum [N_t \log P_t + (1 - N_t)\log(1 - P_t)]. \qquad (5.5)$$

Estimates of the a's and b's may now be determined by the
maximization of (5.5), employing iteratively reweighted least
squares algorithms such as those described in McCullagh
and Nelder (1989).

Figure 9 presents the results of these computations, taking
$P(\cdot)$ to be $\Phi(\cdot)$, the standard normal cumulative (as in probit
analysis) and the sampling interval to be .075 seconds. The
estimated summation function $\hat{a}_u$ is seen to swing negative
directly. This corresponds to M's (or L10's) firing inhibiting

the N's (or L3's) firing. This effect of L10 appears to last for
approximately one second. No apparent rebound effect is
present. The estimate of the decay function $\hat{b}_u$ is ∞ for the
first five coefficients, reflecting the fact that no output spikes
occurred closer than .49 second for this particular data set.
The standard errors are estimated via the usual formulas of
probit analysis. For convenience of display, in the case of $\hat{a}_u$
the errors are graphed about the horizontal axis.

The preceding analysis involved the assumption that the
perturbing noise values had a standard normal distribution.
Suppose, however, that the noise comes from an unknown
distribution and that it is desired to estimate that distribution.
It is convenient to write that distribution as

$$P(\psi) = \Phi(g(\psi)), \qquad (5.6)$$

with the consequence that $g(\cdot)$ will be linear if the noise is
in fact normal. (The function $g(\cdot)$ is not assumed monotonic
here.)

The estimation procedure employed in this case is locally
weighted maximum likelihood. The computations are carried
out recursively. To begin, set $\hat{g}(\psi) = \psi$ and $\hat{g}'(\psi) = 1$.

Step 1. Given N_t, $\hat{g}(\cdot)$, $\hat{g}'(\cdot)$ obtain estimates of the re-
maining parameters of the model, and in particular $\hat{\psi}_t$, by
ordinary maximum likelihood.

Step 2. Given N_t, $\hat{\psi}_t$ obtain $\hat{g}(\cdot)$, $\hat{g}'(\cdot)$ to maximize the
locally weighted log-likelihood

Summation function

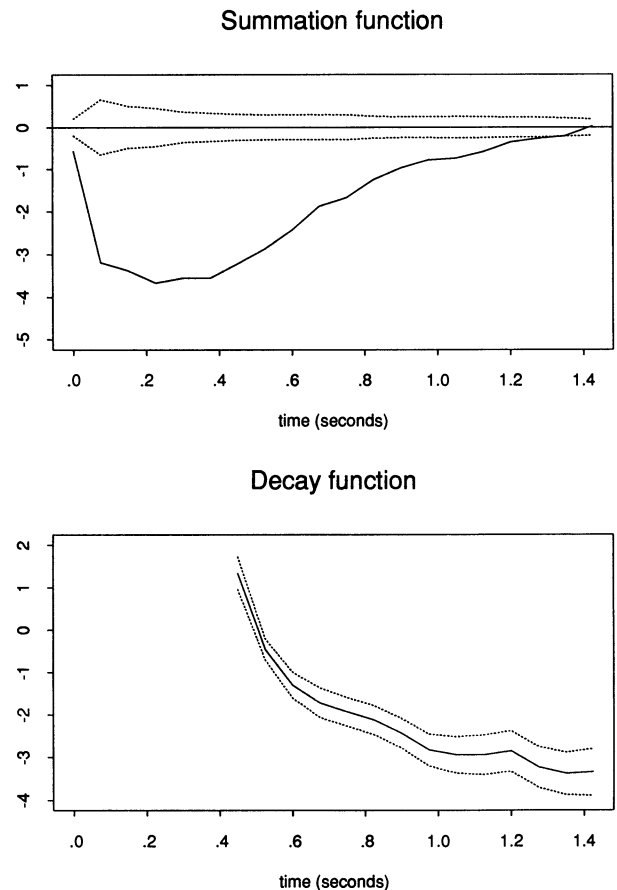time (seconds)

Decay function

time (seconds)

Figure 9. Estimates of the Functions a_u and b_u of (5.2) and (5.3). The
dashed lines provide two standard error limits.

$$\sum w(\psi - \hat{\psi}_t)[N_t \log P_t + (1 - N_t)\log(1 - P_t)], \quad (5.7)$$

with $w(\cdot)$ a weight function concentrated near 0 and with $g(\psi) = \alpha + \beta\psi$ assumed (locally) linear. (This assumption of linearity means that except for the additional weight term, the computations are usual probit ones.) The weight function focuses the local estimation towards the center of the function's support. The estimate of $g(\psi)$ is now taken to be $\hat{\alpha}_\psi + \hat{\beta}_\psi\psi$; the estimate of the derivative, $\hat{\beta}_\psi$.

Step 3. Return to Step 1 until convergence is achieved.

The function estimation procedure of Step 2 here may be found, at various stages of development, in Gilchrist (1967), Cleveland and Kleiner (1975), Brillinger (1977), Cleveland (1979), Hastie and Tibshirani (1984), Tibshirani and Hastie (1987), and Staniswalis (1989). An early version of GAIM (Almudevar and Tibshirani 1990) gave the author confidence that this procedure was feasible for the present situation. The weight function of (5.7) was taken to be the tricube, as in Cleveland and Devlin (1988).

Figures 10 and 11 present the results of these computations. The dashed lines give estimated ± 2 standard error limits. In the case of $\hat{g}'(\cdot)$, they are placed about the level 1.0. The derivative estimate $\hat{g}'(\cdot)$ is seen to not deviate much from 1.0 in the region of apparent probability mass. The computations are seen to support an assumption of linearity of $g(\cdot)$ and hence of normality. This is further reflected in
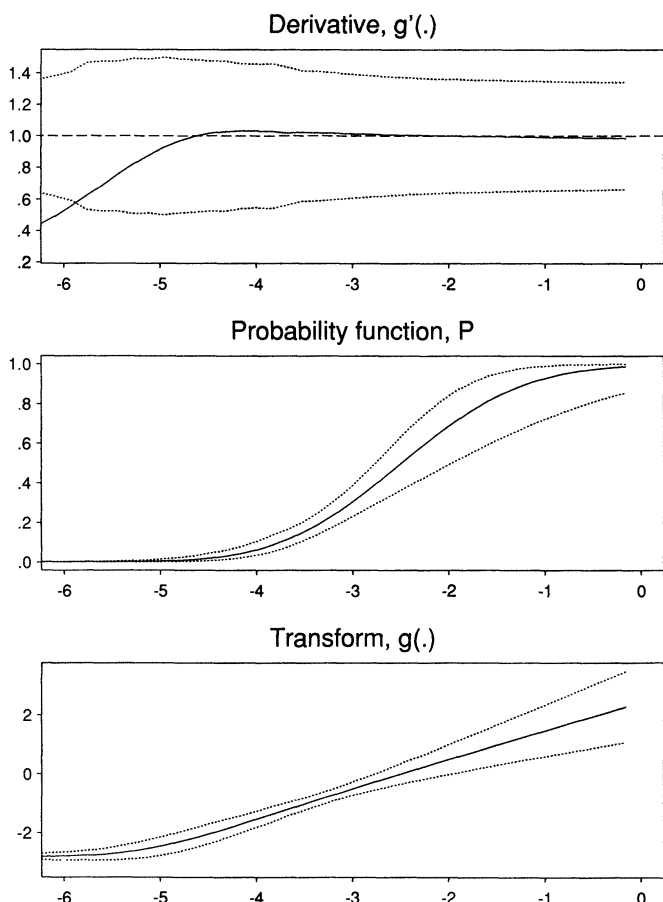
Summation function

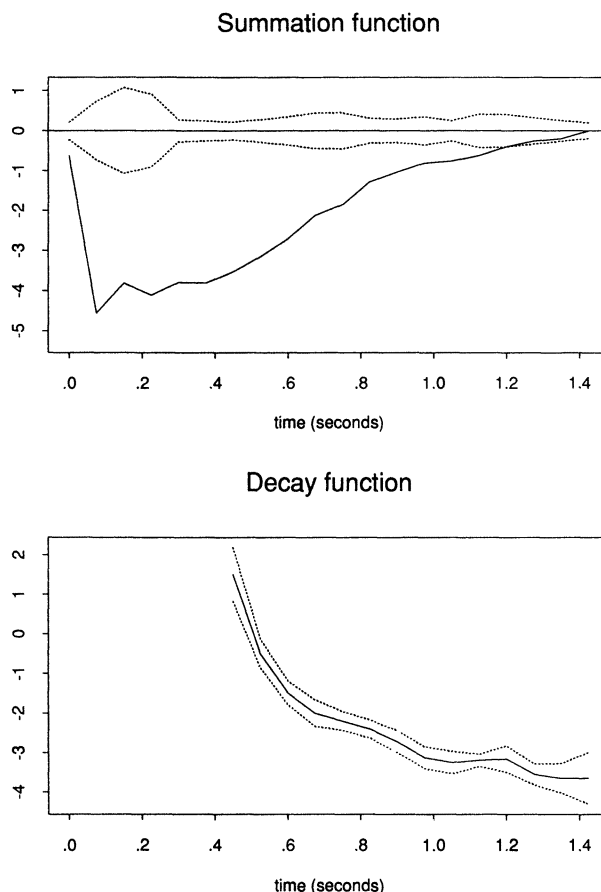time (seconds)

Decay function

time (seconds)

Figure 11. Estimates of a_u and b_u for the Case of Unknown $P(\cdot)$

the similarity of Figures 9 and 11 giving the respective estimates of a_u and b_u. The approximate standard errors were determined via the jackknife (Mosteller and Tukey 1977). In this case, replicates were based on 99% of the data, and 20 replicates were formed.

Consideration next turns to an alternate type of experiment involving *Aplysia* with a different stimulus and a correspondingly altered state variable. In the experiment, noise current is fed directly into the neuron L5 and evoked spike times are recorded. Some input and corresponding output are provided in Figure 12. Numerous neurophysiological experiments have suggested that neuronal firing depends on more than a single-state variable, such as the membrane potential's crossing a threshold. For example, the speed of the crossing, perhaps quantified via the derivatives of the functions involved, also appears to be pertinent (Segundo 1968). The preceding threshold model suggests consideration of the state variable

$$U(t) \approx \sum_{u \leq t} a_{t-u} X_u, \quad (5.8)$$

with X the input noise and ψ_t the corresponding linear predictor

$$\psi_t = U_t - d - e\gamma_t - f\gamma_t^2 - g\gamma_t^3, \quad (5.9)$$

where θ_t is here restricted to have cubic form. (In these computations it was convenient to take the threshold decay func-

Derivative, g'(.)

Probability function, P

Transform, g(.)

Figure 10. Estimates of the Functions g(·) and P(·) of (5.6) and of the Derivative of g(·)
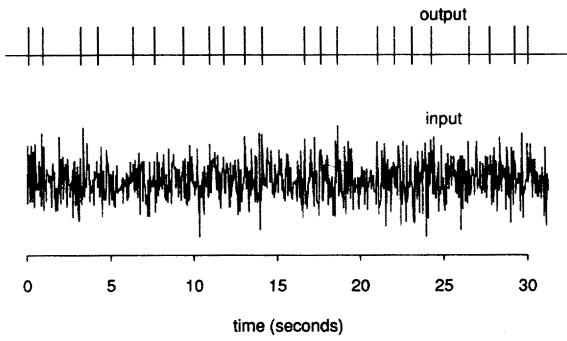
Neuron L5 - Noise Driven

Figure 12. Input and Output of a Neuron. The neuron L5 of Aplysia is stimulated directly by the Gaussian noise of the lower panel and fires as in the upper panel.
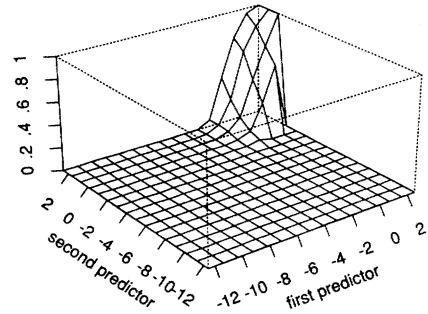
tion to be cubic in order to avoid excessive computations.) Consider also a second state variable

$$\nu_t = \sum c_u X_{t-u}. \tag{5.10}$$

Suppose further that

$$\Pr\{N_t = 1 \mid \text{the past}\} = \Phi(\psi_t)\Phi(\nu_t) \tag{5.11}$$

Empirical Firing Probability

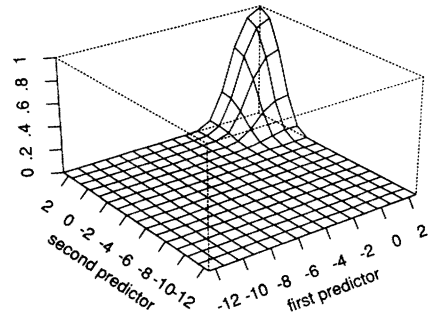

Theoretical Firing Probability

Figure 14. Firing Probability. The bottom panel gives the right side of (5.11). The top panel provides the observed proportion of times the neuron fires as a function of the first and second linear predictor values.
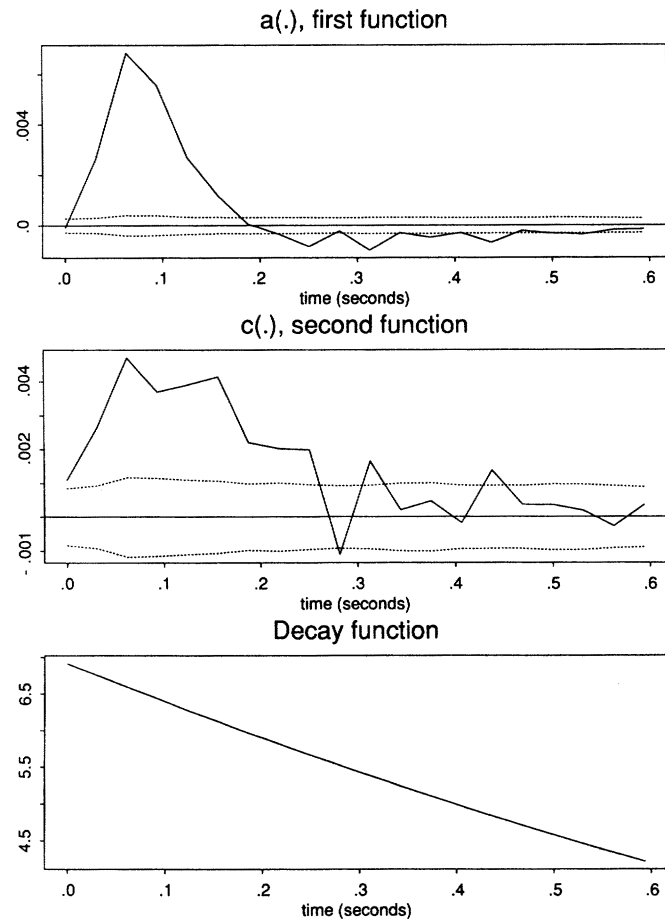
a(.), first function

c(.), second function

Decay function

Figure 13. Estimates of a_u and c_u of (5.8) and (5.10) and of the Cubic Decay Function of (5.9).

as a naive extension of (5.4). It is assumed that approximating the actions of the two state variables as independent will not lead to wildly deviating estimates. Figure 13 gives the results of fitting this model. The fitting here is carried out iteratively, first assuming the coeficients of ψ_t given and estimating those of ν_t, then assuming the coeficients of ν_t given and estimating those of ψ_t. In both cases, the estimation procedures are probit. The second panel gives the estimate of c_u with two standard error limits set about 0. There is evidence for the presence of a second state variable, although in the case of the present computations it does not have the appearance of the derivative of the first. The estimate of a_u given in the first panel shows how the noise current is exciting the neuron.

The problem of assessing goodness of fit has not yet been commented on. Figure 14 provides an informal procedure for the model (5.11). The top panel is a plot of (5.11), the bottom panel gives the empirical firing probability as a function of the first and second predictors. To obtain this, one bins the values of the predictors and computes the corresponding proportion of firing occurrences. The agreement does seem reasonable. One could proceed to formal goodness-of-fit tests based on the quantities just graphed, such as chi-squared statistic, but this seems premature because the temporal dependency leaves the sampling properties in doubt.

Brillinger and Segundo (1979) fit the threshold model to some *Aplysia* data by maximum likelihood. Brillinger
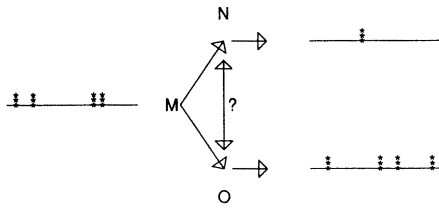
Network connections / causal models

Figure 15. A Network of Three Neurons. Neuron M influences neurons N and O, but one wonders if there is a direct connection from N to O or vice versa.

(1988b) provides a number of references to the threshold modeling of nerve cells' actions and presents further empirical examples.

6. NETWORKS—3 CELL

Suppose one has three neurons, M, N, O, which may be influencing each other. In the experiment analyzed below (see Brillinger, Bryant, and Segundo 1976), it was understood that neuron M was driving both neurons N and O, but it was not known if there were direct connections from N to
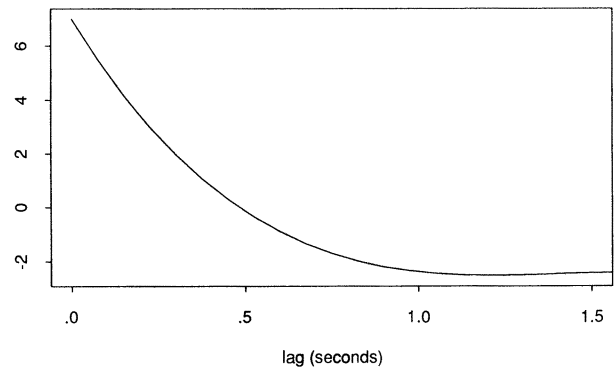
m(.) and o(.)

lag (seconds)

Decay function

lag (seconds)
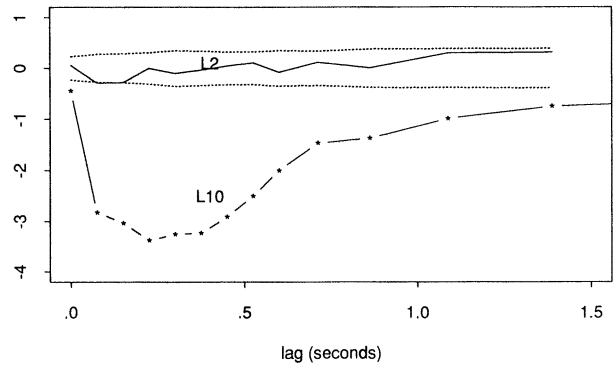
Figure 17. Estimates of m($\cdot$) and o($\cdot$) of (6.2) and of the Cubic Decay Function of (6.3).

Coherence N and O   Coherence M and N

frequency (Hz)   frequency (Hz)

Coherence M and O   Partial coherence N & O

frequency (Hz)   frequency (Hz)

Figure 16. Coherences of Three Neurons. The first three panels provide estimates of the indicated coherences. The final panel is an estimate of the partial coherence of N and O "removing" the effects of the input M. The dashed line gives the upper 95% point of the null distribution.

O or vice versa. The scheme of the situation has been illustrated in Figure 15. One tool for addressing questions of connectivity is partial coherence. The partial coherency at frequency λ of point processes M and N, given the point process O, is defined as

$$R_{NO|M} = \frac{R_{NO} - R_{NM}R_{MO}}{\sqrt{(1 - |R_{NM}|^2)(1 - |R_{OM}|^2)}}. \quad (6.1)$$

Here, R_{NO} denotes a coherency of two stationary point processes as before. Dependence on λ has been surpressed to simplify the display (6.1). The partial coherency may be interpreted via

$$R_{NO|M} = \lim_{T\to\infty} \text{corr}\{d_N^T - \alpha d_M^T, d_O^T - \beta d_M^T\},$$

with

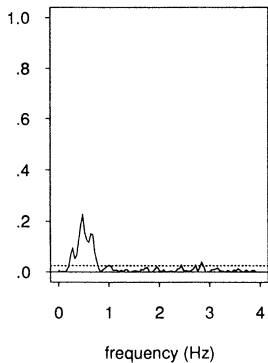$$d_M^T(\lambda) = \sum_j e^{-i\lambda\sigma_j},$$

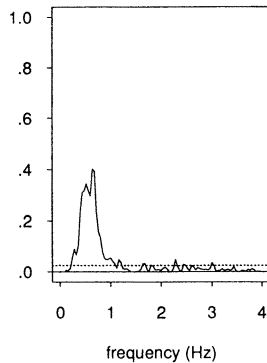for example, as before. Here α and β are the regression coefficients of d_N^T on d_M^T and of d_O^T on d_M^T. The intent of their inclusion is to remove the (linear) effects of the Fourier transform of M from those of N and O.

Figure 16 provides the results of such computations for data on a network of cells $O = L2$, $N = L3$, and $M = L10$ of *Aplysia*. The particular experiments are discussed in Brillinger et al. (1976). The effect of the analysis is quite dramatic.
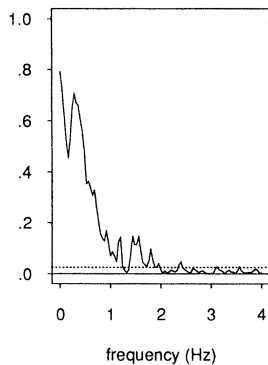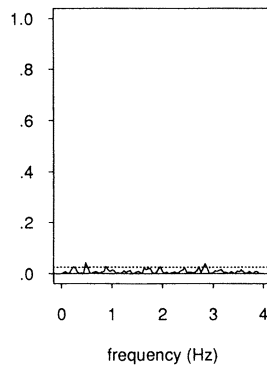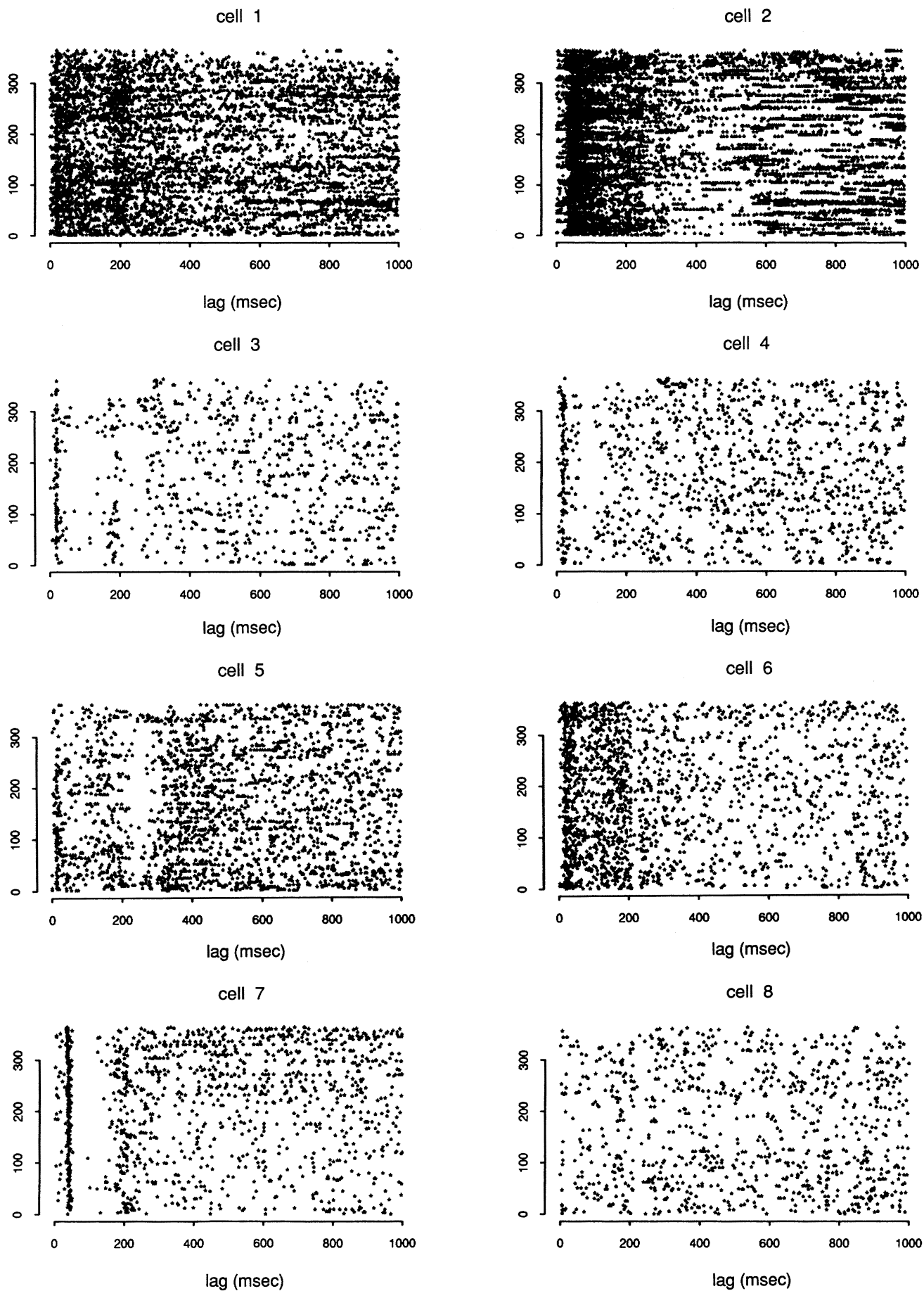
Figure 18. Rastor Plots of the Firings of Eight Cells Following Application of a Noise Stimulus Every 1,000 msec.

From the fourth panel, one can infer that the apparent association of cells N and O, as shown in the first panel, is due to their common association with cell M.

This problem can also be addressed from a likelihood approach by employing a threshold model. Suppose the firing times of cell M are denoted by $\{\sigma_j\}$ and those of cell O by $\{\rho_l\}$. Consider the membrane potential of cell N at time t to be given by

$$U(t) = \sum_j m(t - \sigma_j) + \sum_l o(t - \rho_l), \qquad (6.2)$$

and suppose

$$\Pr\{N_t = 1 \,|\, \text{the past}\} = \Phi(U_t - d - e\gamma_t - f\gamma_t^2 - g\gamma_t^3), \qquad (6.3)$$

γ_t being the elapsed time since N last fired. Here, $m(\cdot)$ and $o(\cdot)$ are summation functions associated with the effects of neurons M and O. One wonders if the function $o(\cdot) \equiv 0$.

Figure 17 on page 267 gives the maximum likelihood estimates of m_u, o_u, and the decay function. The two standard error limits for the cell $O = L2$, set at about 0, suggest an insignificant effect. This is consistent with the results of the coherence analysis. One could do a similar analysis relating O to M and N and achieve the same result.

Various references relating to network analysis are given in Brillinger (1988a), as are further examples. Tick (1963) is an early reference to partial coherence analysis. Gersch (1972)
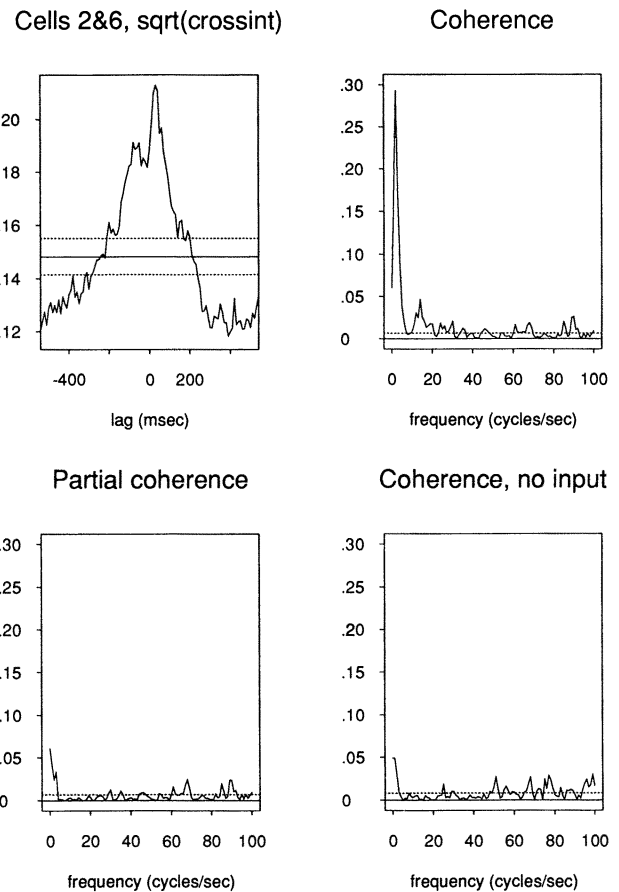

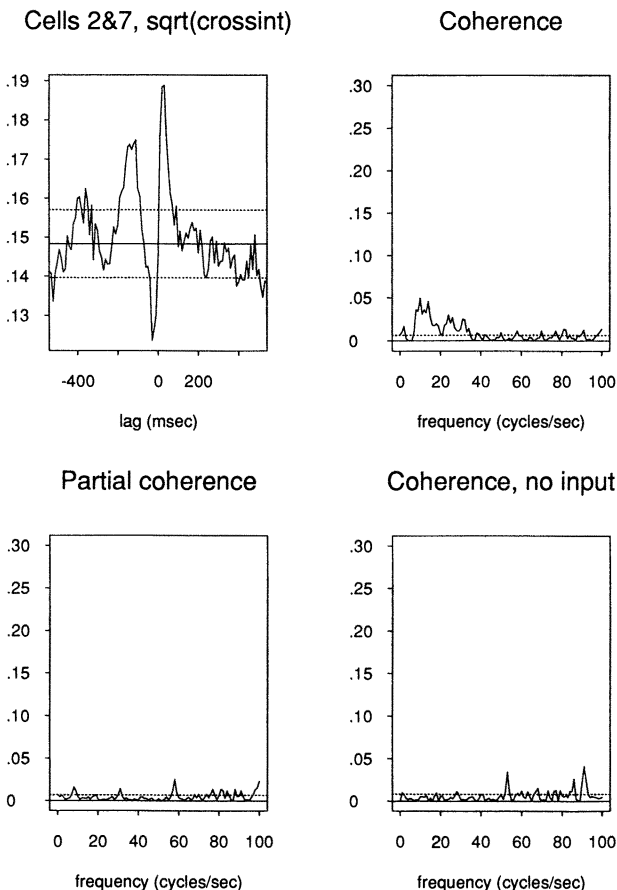Figure 20. Statistics to Investigate the Association of Cells 2 and 6.

Figure 19. Statistics to Investigate the Association of Cells 2 and 7.

discusses empirical partial coherence analysis as a tool to study causality in electrophysiological signal analysis. More examples are provided in Rosenberg, Amjad, Breeze, Brillinger, and Haliday (1989).

7. NETWORKS—8 CELL

In the next analyses presented (albeit preliminarily, as this is work in progress), data were collected in an attempt to understand the auditory pathways of the cat. Microelectrodes were inserted with location tuned to an apparent response to sound and to anatomical knowledge, and responding neurons were located.

The animal was stimulated by white noise bursts of 200 msec duration and at the rate of one per second, through speakers inserted in the ears. The stimulus was applied 364 times. For the eight cells located, Figure 18 on preceding page provides rastor displays of firing times for lags up to 1,000 msec following stimulus application. Various behaviors are exhibited, ranging from the strong association of cells 1, 2, and 6 to the weak association of cell 8. One sees excitation, inhibition, and rebounding.

This work has defined various measures of association of point processes. Figures 19–21 provide them for a selected three of the 28 possible cell pairs. In Figure 19, concerning cells 2 and 7, the cross-intensity and coherence show association. Not much is present, however, when the stimulus is "removed" by partial coherence analysis. This inference is

Cells 2&8, sqrt(crossint)

Coherence

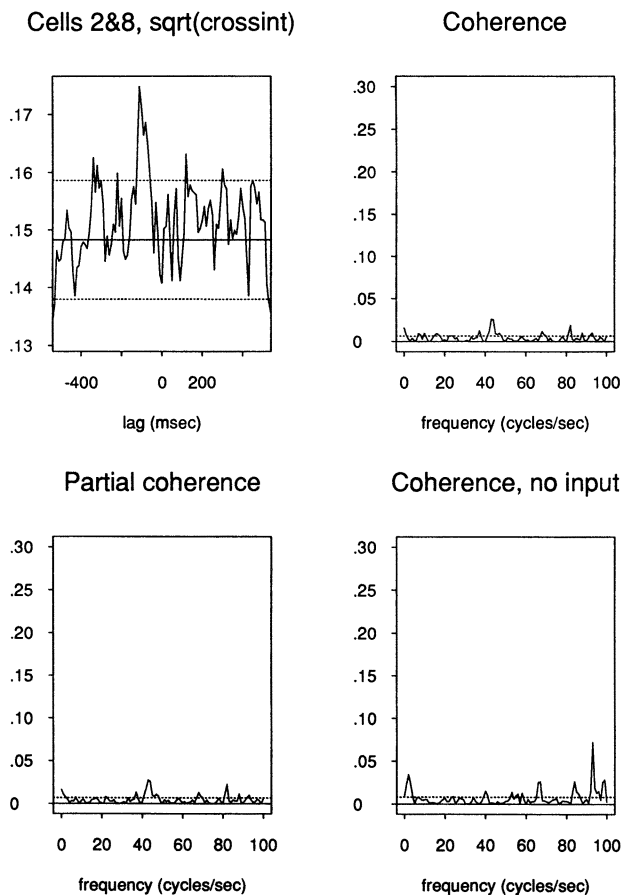## Partial coherence

Coherence, no input

Figure 21. Statistics to Investigate the Association of Cells 2 and 8.

confirmed by the directly measured coherence between the two cells in the case of no applied experimental noise stimulus. Figure 20 provides the same information for cells 2 and 6. Again, the cross-intensity and coherence estimates show association. In this case, however, the partial coherence does suggest that the cells are related beyond the dependence introduced by the common noise stimulus. This inference is again confirmed by the coherence for the case of no experimental stimulus. Figure 21, based on cells 2 and 8, suggests little, if any, connection for these cells. This is consistent with the apparent weak dependence of cell 8 on the stimulus, as shown in Figure 18.

8. DISCUSSION AND SUMMARY

The article has sought to follow the historical statistical progression of description, association, regression, and likelihood analysis. It then continues to the contemporary topics of semiparametric maximum likelihood and causal structure recognition. The data is of a particular type—point process—and is taken from the field of neurophysiology. The paper has illustrated that a calculus is available for point process data analysis and that the calculus allows the computation of standard errors to provide uncertainty measures.

It has been seen that linear techniques—specifically coherence analysis—can elucidate highly nonlinear situations. It has also been seen that stochastic models incorporating basic features of neuron firing and network connections can be set down.

Work ahead includes inferring causal connections for the 8-cell cat network (taking note of the issues and techniques mentioned in Wold 1956, for example), maximum likelihood analysis of the cat data, modeling at the ionic level and, as is topical in contemporary statistical work, improving estimates by borrowing strength (e.g., via random effects models).

REFERENCES

Almudevar, T., and Tibshirani, R. (1990), GAIM, Personal Computer Software for Generalized Additive Interactive Modelling, Toronto: S. N. Tibshirani Enterprises.

Becker, R. A., Chambers, J. M., and Wilks, A. R. (1988), The New S Language, Pacific Grove, CA: Wadsworth.

Brillinger, D. R. (1974), "Cross-Spectral Analysis of Processes With Stationary Increments Including the $G/G/\infty$ Queue," The Annals of Probability 2, 815–827.

—— (1976), "Estimation of the Second-Order Intensities of a Bivariate Stationary Point Process," Journal of the Royal Statistical Society, Ser. B, 38, 60–66.

—— (1977), "Consistent Nonparametric Regression," Discussion of Stone, The Annals of Statistics 5, 622–623.

—— (1978), "Comparative Aspects of the Study of Ordinary Time Series and of Point Processes," in Developments in Statistics 1, ed. P. R. Krishnaiah, New York: Academic Press, pp. 33–133.

—— (1988a), "Maximum Likelihood Analysis of Spike Trains of Interacting Nerve Cells," Biological Cybernetics, 59, 189–200.

—— (1988b), "The Maximum Likelihood Approach to the Identification of Neuronal Firing Systems," Annals of Biomedical Engineering, 16, 3–16.

Brillinger, D. R., and Segundo, J. P. (1979), "Empirical Examination of the Threshold Model of Neuron Firing," Biological Cybernetics, 35, 213–220.

Brillinger, D. R., Bryant, H. L., and Segundo, J. P. (1976), "Identification of Synaptic Interactions," Biological Cybernetics, 22, 213–228.

Bryant, H. L., Ruiz Marcos, A., and Segundo, J. P. (1973), "Correlations of Neuronal Spike Discharges Produced by Monosynaptic Connexions and by Common Inputs," Journal of Neurophysiology, 36, 205–225.

Bryant, H. L., and Segundo, J. P. (1976), "Spike Initiation by Transmembrane Current: A White-Noise Analysis," Journal of Physiology, 260, 279–314.

Cajal, S. Ramon y (1895), Les Nouvelles Idées sur la Structure du Système Nerveux Chez L'Homme et Chez les Vertébrés, Paris: Reinwald.

Cleveland, W. S. (1979), "Robust Locally Weighted Regression and Smoothing Scatterplots," Journal of the American Statistical Association, 74, 829–836.

Cleveland, W. S., and Devlin, S. J. (1988), "Locally Weighted Regression: An Approach to Regression Analysis by local Fitting," Journal of the American Statistical Association, 83, 596–610.

Cleveland, W. S., and Kleiner, B. (1975), "A Graphical Technique for Enhancing Scatterplots With Moving Statistics," Technometrics, 17, 447–454.

Cox, D. R., and Lewis, P. A. W. (1966), The Statistical Analysis of Series of Events, London: Methuen.

Fisher, R. A. (1956), Statistical Methods and Scientific Inference, London: Oliver and Boyd.

Gersch, W. (1972), "Causality or Driving in Electrophysiological Signal Analysis," Mathematical Biosciences, 14, 177–196.

Gilchrist, W. G. (1967), "Methods of Estimation Involving Discounting," Journal of the Royal Statistical Society, Ser. B, 29, 355–369.

Hastie, T., and Tibshirani, R. (1984), Generalized Additive Models, Technical Report 2, Laboratory for Computational Statistics, Stanford University.

Koch, C., and Segev, I. (1989), Methods in Neuronal Modelling, Cambridge, MA: MIT Press.

McCullagh, P., and Nelder, J. A. (1989), Generalized Linear Models (2nd ed.), London: Chapman and Hall.

Mosteller, F., and Tukey, J. W. (1977), Data Analysis and Regression, Reading, MA: Addison-Wesley.

Rosenberg, J. R., Amjad, A. M., Breeze, P., Brillinger, D. R., and Haliday, D. M. (1989), "The Fourier Approach to the Identification of Functional Coupling Between Neuronal Spike Trains," Progress in Biophysics and Molecular Biology, 53, 1–31.