Nerve Cell Spike Train Data Analysis: A Progression of Technique

By

David R. Brillinger*

Technical Report No. 323
September 1991

*Research partially supported by NSF Grant DMS-8900613.

Department of Statistics
University of California
Berkeley, California 94720

$2\pi \neq 1$

$2\pi \neq 1$

$2\pi \neq 1$

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 data type that is fairly common, but not broadly discussed. This work is concerned with the particular circumstance of the firing times of nerve cells. The nerve cells studied interact and form networks.

An intention of the work is to review a progression of statistical analysis technique. The techniques begin with description, next association as measured by moments and correlation, then onto 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.

* David R. Brillinger is Professor of Statistics at the University of California, Berkeley. This article formed the R. A. Fisher Lecture given at Atlanta, Georgia in August 1991. The research was partially supported by the National Science Foundation Grant DMS-8900613. The computations were carried out at the Statistical Computing Facility, University of California, Berkeley, Director Leo Breiman. The figures were prepared employing S, see Becker, Chambers and Wilks (1988). The author thanks the many individuals who gave him help and advice with the computing and with the presentation of the material. The author particularly thanks Professor J. P. Segundo, Department of Anatomy & Cell Biology, UCLA who for almost 20 years has helped him with the intricacies of the pertinent neurophysiology.

KEY WORDS: Association; Binary data; Coherence; Cross-intensity function; Descriptive statistics; Impulse response; Likelihood; Nerve cell; Network; Nonlinear system; Partial Coherence; Point process; Probit analysis; Semiparametric model; Threshold phenomenon

1. INTRODUCTION

"*... 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 the present work is the better understanding of the nerve cell system and of the construction of 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 historical development of the field of statistics. In particular he was central to the progression of data analysis techniques from description and simple measures of association past the tools of association and regression analysis and onto likelihood analysis. A principal goal of this work is 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 work is concerned with a particular biological system, that of 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. These were collected by A. E. P. Villa at Lausanne, Switzerland. In his experiments, cats were subjected to sound stimuli and data for 8 nerve cells recorded simultaneously. The experiments are described in Villa (1988, 1990). Also studied are data of *Aplysia californica* (the sea hare) collected by J. P. Segundo at the University of California, Los Angeles. This is simultaneous data for networks of two and three identified nerve cells, in particular the cells L2, L3, L5, L10 of *Aplysia*. The experiments in which those data were collected are described in detail in Bryant, Ruiz Marcos and Segundo (1973) and in Bryant and Segundo (1976). *Aplysia* is often 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 figures are presented. These figures are central to the analysis.

The layout of the paper is the following: What is a nerve cell? What are point process data? Association - second-order moments; Regression - a linear model; Likelihood - conceptual modelling; Networks - 3 cell, 8 cell; Discussion and summary.

Important aspects of nerve cell firing that are not addressed in this present work include spatial effects and intracellular data collection and analysis.

2. 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 to be directly related. The functions include the accumulation, processing and transmission of 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 synapses are the junctions of neural networks. Figure 1, taken from Cajal (1894), is a drawing of a collection of neighboring neurons. The arrows are meant to indicate the flow of information. The cell bodies are the five blobs, 4 of which are labelled A, B, C, 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 tree-like 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 in turn affect other neurons. The action potentials are of near identical size and shape, see the spikes in Figure 2. This figure provides measured voltage fluctuations within cell R2 of *Aplysia,* see Bryant and Segundo (1976). It may be argued that, because of reduced sensitivity to noise, the firing times are the crucial variates in communication amongst neurons.

Some discussion of the reduction to point processes is given in Segundo et al. (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 more likely or less likely to fire. Neurons also may fire spontaneously in the case of no outside stimulus. Further there is the phenomenon of *refractoriness* wherein after a neuron has fired, the chance of it firing again is reduced (perhaps to zero) for a time period.

Questions of interest include: Can an analytic model that incorporates the basic features of neuron behavior be developed and fit? and Given the firing times of a network of neurons, can one infer their causal connections?

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

3. 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 occured in some time interval, say $(0,T]$. Usual examples are the times of telephone calls or the times of emission of particles by some radioactive material. A naive descriptive statistic derived from such data is the observed *rate,* given by K/T here. It has dimensions of counts per unit time and is useful in elementary comparisons of point process behavior. For the data studied in this paspike per, the rates range from about 1 per second to about 20 per second. In the case of Figure 2 there were 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, that 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 of length 1000 milliseconds. Let τ_{kl} refer to the time elapsed since the start of the l-th segment, of the k-th 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 now a noise stimulus was introduced into the ears of the cat every 1000 milliseconds. The points are plotted as before with τ refering to the time elapsed since each presentation of the stimulus. What this picture shows is that this neuron fires a short time after the application of the stimulus. Then there is a time period during which the neuron is unlikely to fire and perhaps then a rebound period when the cell is more likely to fire. Plots, such as those of 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 $\{\tau_k\}$ those 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). Figure 4 is a plot of the points $\{(\tau_k - \sigma_j, j), k = 1,2,\ldots,K_j\}$ for $j = 1,2,3 \cdots$ This plot is consistent with the idea that the firing of L10 tends to inhibit the 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 to be randomly distributed.

In order 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 $N(I_1),...,N(I_J)$ where I_j is a Borel set and $N(I_j)$ is the number of points falling in I_j for $j=1,...,J$ and $J = 1,2,....$ 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) for some discussion. The classic reference to the analysis of point process data is Cox and Lewis (1966).

4. 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*

$$\lim_{h \rightarrow 0} Prob\{N \text{ point in } (u+t,u+t+h] \mid M \text{ point at } t\}/h$$

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

$$\frac{\#\{u+\sigma_j < \tau_k \leq u+\sigma_j+h\}}{\#\{\sigma_j\}h} \tag{4.1}$$

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 it is often more convenient to plot the square root of the estimate, see Brillinger (1976), and this was done here. The horizontal dashed lines provide ±2 standard error limits set about 0. The diagram shows a period initial inhibition after L10's firing followed by a rebound at about .3 second. In some sense Figure 5 is not adding new information to that of Figure 4, but it does provide a specific way with which to interpret and assess the phenomena occurring. 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

$$\sum_{0<\sigma_j \le T} e^{-i\lambda\sigma_j} \quad , \quad \sum_{0<\tau_k \le T} e^{-i\lambda\tau_k}$$

for $0 \le \lambda < \infty$. The quantity

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

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, see 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 that 1

cycle/second. The dashed line in the figure gives the (approximate) 95% upper point of the null distribution of the estimate. In view of the essential nonlinear relationship of the data being studied, the magnitude here of the coherence estimate at the low frequencies is surprising.

5. REGRESSION - A LINEAR MODEL

Consider next a model

$$\lim_{h \to 0} Prob\{N \text{ point in } (t, t+h] \mid M\}/h = \mu + \sum_j a(t - \sigma_j) \qquad (5.1)$$

This model is linear and time invariant. The function $a(.)$ is meant to represent the various chemical, electrical, spatial and temporal delay processes involved in neuron M's firing influencing the firing of neuron N. For example were the τ's given by $\tau_j = \sigma_j + Y_j$, with the Y's independent and of density function $a(.)$, then the result (5.1) would hold with $\mu = 0$, see Brillinger (1974). The model (5.1) may be fit by cross-spectral analysis, *ibid.* The resulting estimate of $a(.)$ is given in Figure 7 for the *Aplysia* data addressed in the preceding section. The estimate is seen to mimic that of Figure 5. The distinction is that, as is the case in ordinary regression analysis one is nearer to a system invariant. This analysis for this particular data set is not dramatically enlightening, but interesting examples may be found in Brillinger, Bryant and Segundo (1976). The next section presents a more satisfying analysis of the present data in any case.

6. LIKELIHOOD - CONCEPTUAL MODELLING

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 subject to slow decay (although the decay is not always monotonic). The effect of the reset is to prevent firing recurring immediately and hence 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 $a(.)$, consider the following time varying state variable

$$U(t) = \sum_{\sigma_j \leq t} a(t-\sigma_j) \tag{6.1}$$

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 $a(.)$ is referred to as a *summation function.* As in the previous section $a(.)$ is meant to represent the various processes involved in M's firing influencing 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 will be represented by a function $b(.)$.

Figure 8 is a layout of the situation. The bottom two panels give hypothetical $a(.)$ and $b(.)$ for the case of an inhibitory synapse. (Shortly empirical estimates of $a(.)$ and $b(.)$ will be provided.) The vertical asterisks of the top plot are the firing times of the input neuron. The hook-shaped curves are the translates of the function $b(.)$, with a new translate introduced with each firing of the principal neuron. If γ_t denotes the time elapsed since 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 $U(t)$ crossing $\theta(t)$.

Consideration turns to developing a stochastic version of this model and of a corresponding likelihood function to employ in analysing available data. Suppose first that the point processes are simplified to discrete

time ($t = 0,\pm1,\pm2,...$) 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 $t = 0,\pm1,\pm2,...$ Corresponding discrete versions of N and $a(.)$ are similarly defined. Now

$$U(t) = \sum a(t-\sigma_j) \simeq \sum a_{t-u} M_u \tag{6.2}$$

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

$$\theta_t = \sum_{v=1}^{\gamma_t} b_v N_{t-v} \tag{6.3}$$

with γ_t the time elapsed since the last N-firing.

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

$$P_t = Prob\{N_t = 1 \mid the\ past\} = P(\psi_t) \tag{6.4}$$

where

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

The loglikelihood is

$$\sum [N_t \log P_t + (1-N_t)\log (1-P_t)] \tag{6.5}$$

Estimates of the a's and b's may now be determined by the maximization of (6.5), employing iteratively reweighted least squares algorithms such as those described in McCullagh and Nelder (1989).

Figure 9 presents the results of these computations taking $P(.)$ to be $\Phi(.)$, 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 (or L10)'s firing inhibiting the firing of N (or L3). This effect of L10 appears to last for approximately a second. There is no apparent rebound effect 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 seconds 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

$$P(\psi) = \Phi(g(\psi)) \tag{6.6}$$

with the consequence that $g(.)$ will be linear if the noise is in fact normal.

The estimation procedure employed in this case involves computations carried out recursively. To begin set $\hat{g}(\psi) = \psi$ and $\hat{g}'(\psi) = 1$.

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

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

$$\sum w(\psi - \hat{\psi}_t)[N_t \log P_t + (1-N_t)\log(1-P_t)] \tag{6.7}$$

with $w(.)$ 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 estimate of $g(\psi)$ is now taken to be $\hat{\alpha}_\psi + \hat{\beta}_\psi \psi$ and of the derivative taken to be $\hat{\beta}_\psi$.

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), Staniswalis (1989). An early version of GAIM, see Almudevar and Tibshirani (1990), gave the author confidence that this procedure was feasible for the present situation. The weight function of (6.7) 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 $\hat{g}'(.)$ they are placed about the level 1.0. The derivative estimate $\hat{g}'(.)$ 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 $g(.)$ and hence of normality. This is further reflected in 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, see Mosteller and Tukey (1977). In the present case replicates were based on 99% of the data and 20 replicates were formed.

Consideration next turns to an alternate type of experiment with *Aplysia*. There is a different stimulus and a correspondingly altered state variable. In the experiment, noise current is fed directly into the neuron L5 and once again evoked spike times recorded. Some input and corresponding output are provided in Figure 12. It is to be mentioned that 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, see Segundo (1968). For example, the speed of the crossing appears to be pertient as well. The preceding threshold model suggests consideration of the state variable

$$U(t) \approx \sum_{u \leq t} a_{t-u} X_u \tag{6.8}$$

with X the input noise and ψ_t the corresponding linear predictor

$$\psi_t = U_t - d - e\gamma_t - f\gamma_t^2 - g\gamma_t^3 \tag{6.9}$$

(In these computations it was convenient to take the threshold decay function to be cubic in order to avoid an excess of computations.) Consider also a second state variable

$$\nu_t = \sum c_u X_{t-u} \tag{6.10}$$

Suppose further that

$$Prob\{N_t = 1 \mid the\ past\} = \Phi(\psi_t)\Phi(\nu_t) \tag{6.11}$$

as an extension of (6.4). 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. The estimation procedures in these cases are both probit. The second panel gives the estimate of c_u with 2 standard error limits set about 0. There is evidence for the presence of a second state variable, although is 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 (6.11). The top panel is a plot of (6.11). 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 since 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 (1988b) provides a number of references to the threshold modelling of nerve cells' action and presents further empirical examples.

7. NETWORKS - 3 CELL

Suppose one has 3 neurons, M, N, O, which may be influencing each other. In the experiment whose data analysis is about to be presented, (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 O or vice versa. The scheme of the situation is 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 to be

$$R_{NO\,|M} = \frac{R_{NO} - R_{NM}R_{MO}}{\sqrt{(1 - |R_{NM}|^2)(1 - |R_{OM}|^2)}} \tag{7.1}$$

Here R_{NO} denotes a coherency of two stationary point processes as before. Dependence on λ has been surpressed to simplify the display (7.1). The partial coherency may be interpreted via

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

with

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

for example, as before. Here α, β are the regression coefficients of d_N^T on d_M^T and of d_O^T on d_M^T respectively. 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 $O = L2$, $N = L3$, $M = L10$ of *Aplysia*. The particular experiments are discussed in Brillinger, Bryant and Segundo (1976). The effect of the analysis is quite dramatic. 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 present 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

$$U(t) = \sum_j m(t - \sigma_j) + \sum_l o(t - \rho_l) \tag{7.2}$$

and suppose

$$Prob\{N_t = 1 \mid the\ past\} = \Phi(U_t - d - e\gamma_t - f\gamma_t^2 - g\gamma_t^3) \tag{7.3}$$

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

Figure 17 gives the maximum likelihood estimates of m_u, o_u and the decay function. The 2 standard error limits for the cell O = L2, set 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. Here the result is the same.

A variety of 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) discusses empirical partial coherence analysis as a tool to study causality in electrophysiological signal analysis. More examples are provided in Rosenberg *et al.* (1989).

8. NETWORKS - 8 CELL

In the next analyses to be presented, albethey preliminary 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. Responding neurons were located.

The animal was stimulated by white noise bursts, of duration 200 msec., at the rate of 1 per second, through speakers inserted in the ears. For the 8 cells located, Figure 18 provides rastor displays of firing times for lags up to 1000 milliseconds following the stimulus application. The stimulus was applied 364 times. A variety of behaviors show themselves 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 evidence association. Not much is present however when the stimulus is "removed" by partial coherence analysis. This inference is confirmed by the directly measured coherence between the two cells in the case of no applied experimental noise stimulus. Figure 20 provides the same for cells 2 and 6. Again the cross-intensity and coherence estimates evidence the presence of association. The partial coherence in this case however 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 there is not much connection, if any for these cells. This is consistent with the apparent weak dependence

of cell 8 on the stimulus, as seen in Figure 18.

9. DISCUSSION AND SUMMARY

The paper has sought to follow the historical statistical progression of description, association, regression and then likelihood analysis.The work 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. Amongst other things 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 can be set down that incorporate basic features of neuron firing and network connections.

Work lying 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, modelling at the ionic level and, as is topical in contemporary statistical work, improving estimates by borrowing strength, eg. 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: Wadsworth.

BRILLINGER, D. R. (1974), Cross-spectral analysis of processes with stationary increments including the $G/G/\infty$ queue, *Ann. Prob.* 2, 815-827.

BRILLINGER, D.R. (1976), Estimation of the second-order intensities of a bivariate stationary point process, *J. Royal Stat. Soc. B* 38, 60-66.

BRILLINGER, D.R. (1977), Discussion of Stone (1977), *Ann. Statist. 5, 622-623.*

BRILLINGER, D. R. (1978), Comparative aspects of the study of ordinary time series and of point processes, pp. 33-133 in *Developments in Statistics 1* (ed. P. R. Krishnaiah), New York: Academic.

BRILLINGER, D.R. (1988a), Maximum likelihood analysis of spike trains of interacting nerve cells, *Biological Cybernetics* 59, 189-200.

BRILLINGER, D. R. (1988b), The maximum likelihood approach to the identification of neuronal firing systems, *Ann. 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, *J. Neurophysiol.,* 36, 205-225.

BRYANT, H.L. and SEGUNDO, J. P. (1976), Spike initiation by transmembrane current: a white-noise analysis, *J. Physiol.,* 260, 279-

314.

CAJAL, S. Ramon y (1894), *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, *J. Amer. Statist. Assoc.*, 74, 829-836.

CLEVELAND, W.S. and KLEINER, B. (1975), A graphical technique for enhancing scatterplots with moving statistics, *Technometrics* 17, 447-454.

CLEVELAND, W.S. and DEVLIN, S.J. (1988), Locally weighted regression: an approach to regression analysis by local fitting, *J. Amer. Statist. Assoc.*, 83, 596-610.

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, *Math. Biosciences*, 14, 177-196.

GILCHRIST, W.G. (1967), Methods of estimation involving discounting, *J. Royal Statistical Society B* 29, 355-369.

HASTIE, T. and TIBSHIRANI, R. (1984), *Generalized additive models*, Tech. Report 2, Laboratory for Computational Statistics, Stanford University.

KOCH, C. and SEGEV, I. (1989), *Methods in Neuronal Modelling*, Cambridge: MIT Press.

McCULLAGH, P. and NELDER, J.A. (1989), *Generalized Linear Models*, Second Edition, London: Chapman and Hall.

MOSTELLER, F. and TUKEY, J.W. (1977), *Data Analysis and Regression*, Reading: 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 Moleculat Biology*, 53, 1-31.

SEGUNDO, J.P. (1968), Functional possibilities for communication and coding of neuronal properties and interactions, *Acta Neurol. Latinoamer.*, 14, 340-344.

SEGUNDO, J.P. (1984), *La Neurofisiologia: Alguno Supuestos y Bases, Recovecos e Implicaciones*, Mexico: SECEP.

SEGUNDO, J. P. (1986), What can neurons do to serve as integrating devices?, *J. Theoret. Neurobiol.*, 5, 1-59.

SEGUNDO, J.P., ALTSHULER, E., STIBER, M. and GARFINKEL, A. (1991), Periodic inhibition of living pacemaker neurons: I. locked, intermittent, messy and hopping behaviors, *International J. Bifurcation and Chaos*, 1, in press.

STANISWALIS, J.G. (1989), The kernel estimate of a regression function in likelihood-based models, *J. Amer. Statist. Assoc.*, 84, 276-283.

STEIN, R. B. (1972), The stochastic properties of spike trains recorded from nerve cells, pp. 700-731 in *Stochastic Point Processes* (ed. P. A. W. Lewis), New York: Wiley.

TIBSHIRANI, R. and HASTIE, T. (1987), Local likelihood estimation, *J. Amer. Statist. Assoc.*, 82, 559-567.

TICK, L.J. (1963), Conditional spectra, linear systems and coherency, pp. 197-203 in *Time Series Analysis* (Ed. M. Rosenblatt), New York:

Wiley.

VILLA, A.E.P. (1988), *Influence de l'Ecorce Cérébrale sur l'Activité Spontanée et Evoquée du Thalamus Auditif du Chat,* Thesis, Faculty of Sciences, University of Lausanne.

VILLA, A.E.P. (1990), Physiological differentiation within the auditory part of the thalamic reticular nucleus of the cat, *Brain Research Reviews,* 15, 25-40.

WOLD, H. (1956), Causal inference from observational data a review of ends and means, *J. Royal Statistical Societ* 119, 28-61.

Captions

Figure 1. Drawing of Cajal (1894) illustrating a network of 5 cells. The arrows suggest that the input arrives along the fibre E and progresses from it both directly and indirectly to the cells A, B, C, D.

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 the times at which nerve cell 7 fires in successive time segments of length 1 second. In the lefthand panel there was no experimental stimulus. In the case of the second panel, a noise stimulus was applied at the beginning of each time segment.

Figure 4. Times of neuron L3's firings relative to those of L10.

Figure 5. The square root of the cross-intensity statistic (4.1). The dashed lines give upper and lower two standard error limits placed about 0 level.

Figure 6. An estimate of the coherence of neurons L10 and L3 obtained in the fashion described in Brillinger, Bryant and Segundo (1976).

Figure 7. An estimate of the function $a(.)$ of (5.1) obtained in the fashion described in Brillinger, Bryant and Segundo (1976).

Figure 8. The lower curve of the top panel gives $U(t)$ of (6.1) 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.

Figure 9. Estimates of the functions a_u and b_u of (6.2) and (6.3). The dashed lines provide two standard error limits.

Figure 10. Estimates of the functions $g(.)$ and $P(.)$ of (6.6) and of the derivative of $g(.)$.

Figure 11. Estimates of a_u and b_u for the case of unknown $P(.)$.

Figure 12. The neuron L5 of *Aplysia* is stimulated directly by the Gaussian noise of the lower panel and fires as in the upper panel.

Figure 13. Estimates of a_u and c_u of (6.8) and (6.10) and of the cubic decay function of (6.9).

Figure 14. The top panel gives the right-hand side of (6.11). The bottom panel provides the observed proportion of times the neuron fires as a function of the first and second linear predictor values.

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

Figure 16. 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.

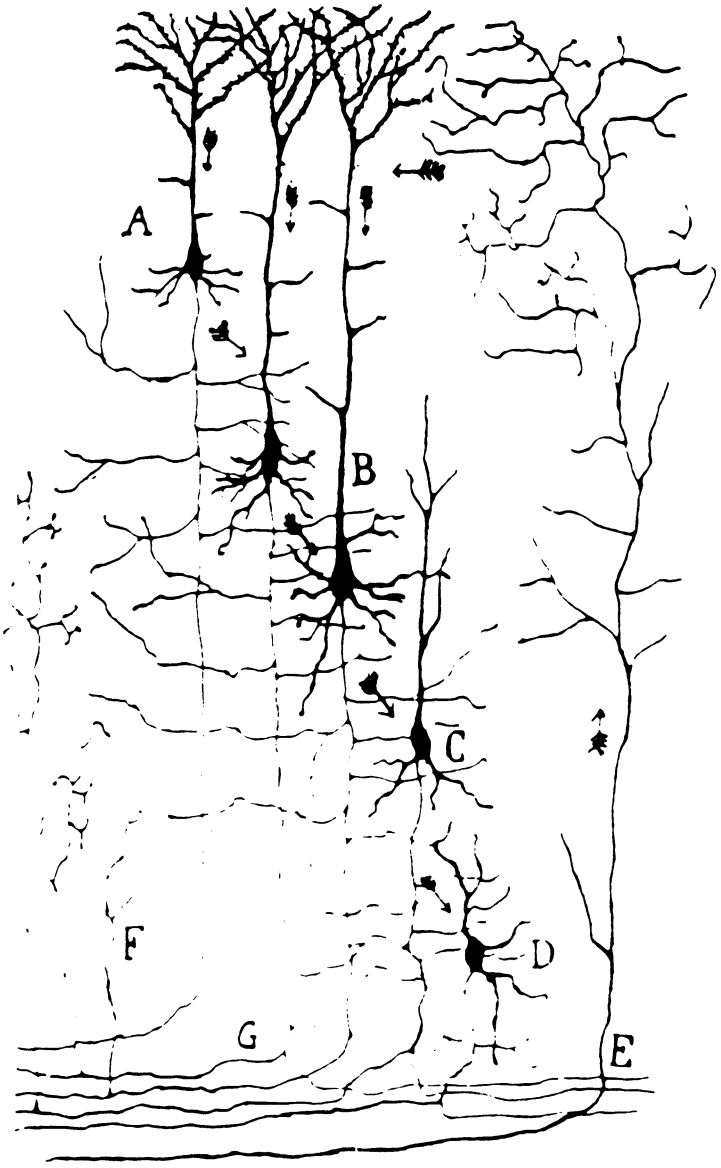
Figure 17. Estimates of $m(.)$ and $o(.)$ of (7.2) and of the cubic decay function of (7.3).

Figure 18. Rastor plots, as in Figure 3, of the firings of 8 cells following periodic application of a noise stimulus. The stimulus is applied every 1000 milliseconds.

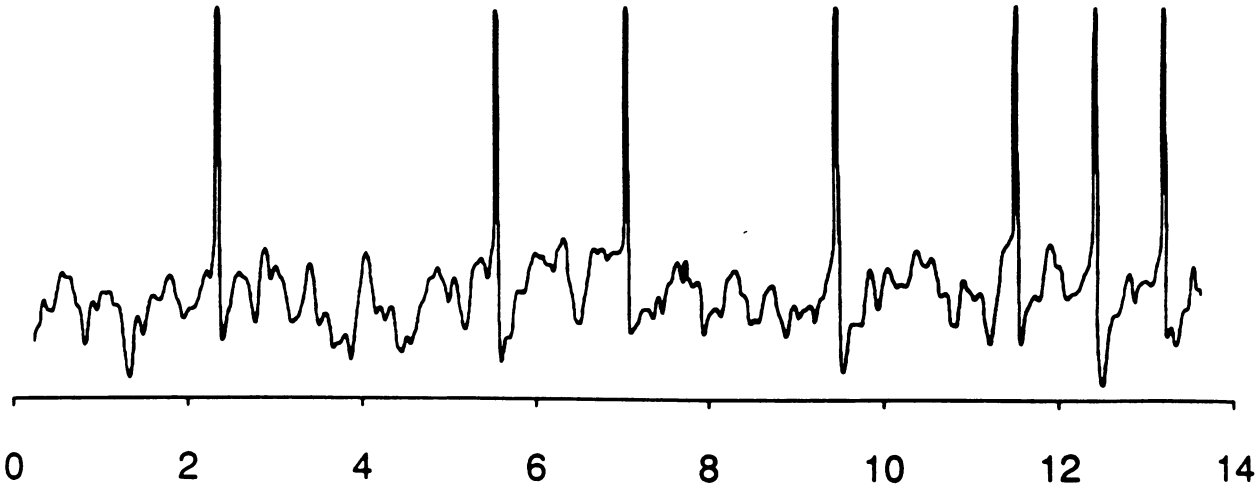
Figure 19. Statistics to investigate the association of cells 2 and 7.

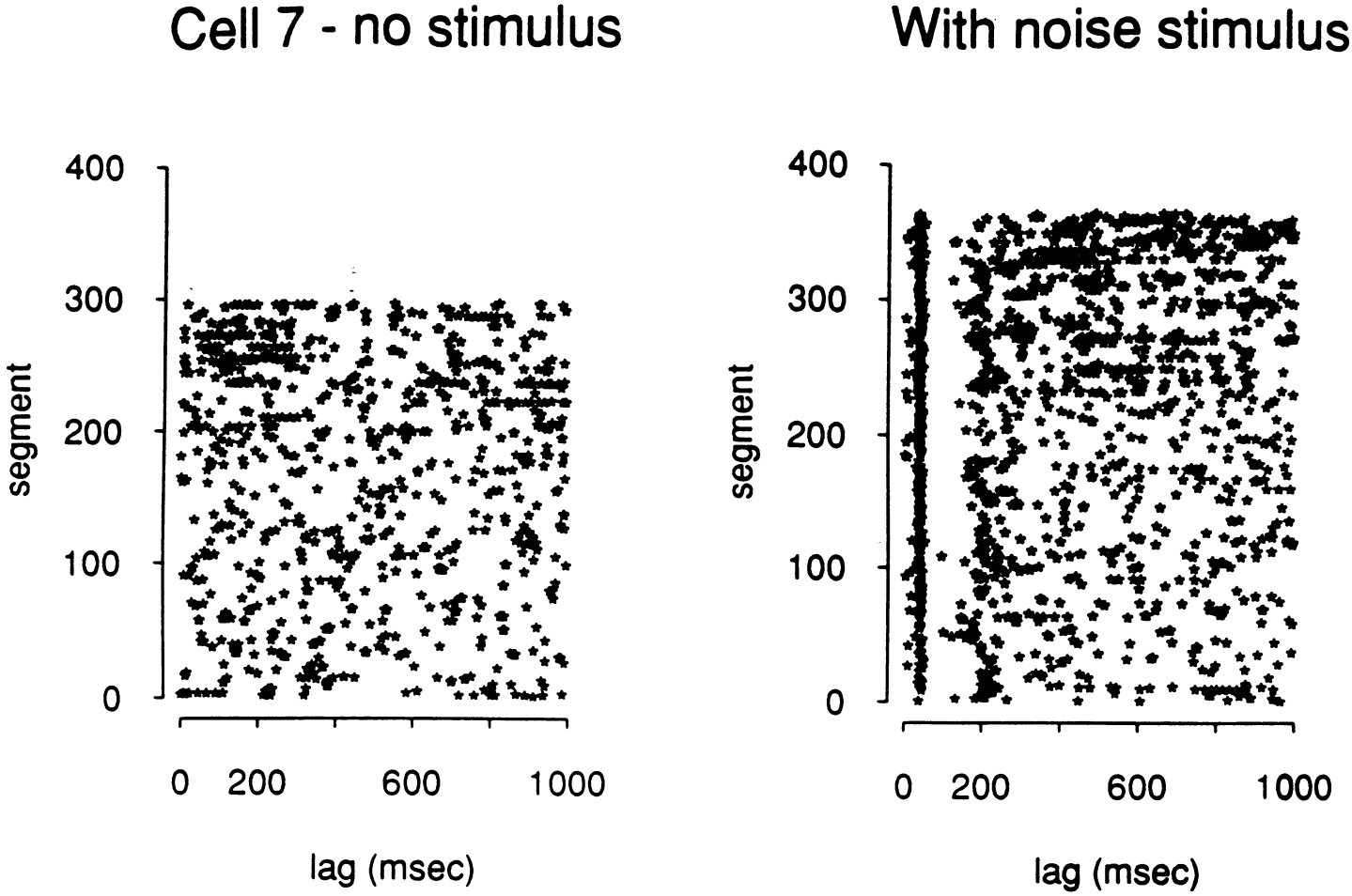
Figure 20. Statistics to investigate the association of cells 2 and 6.

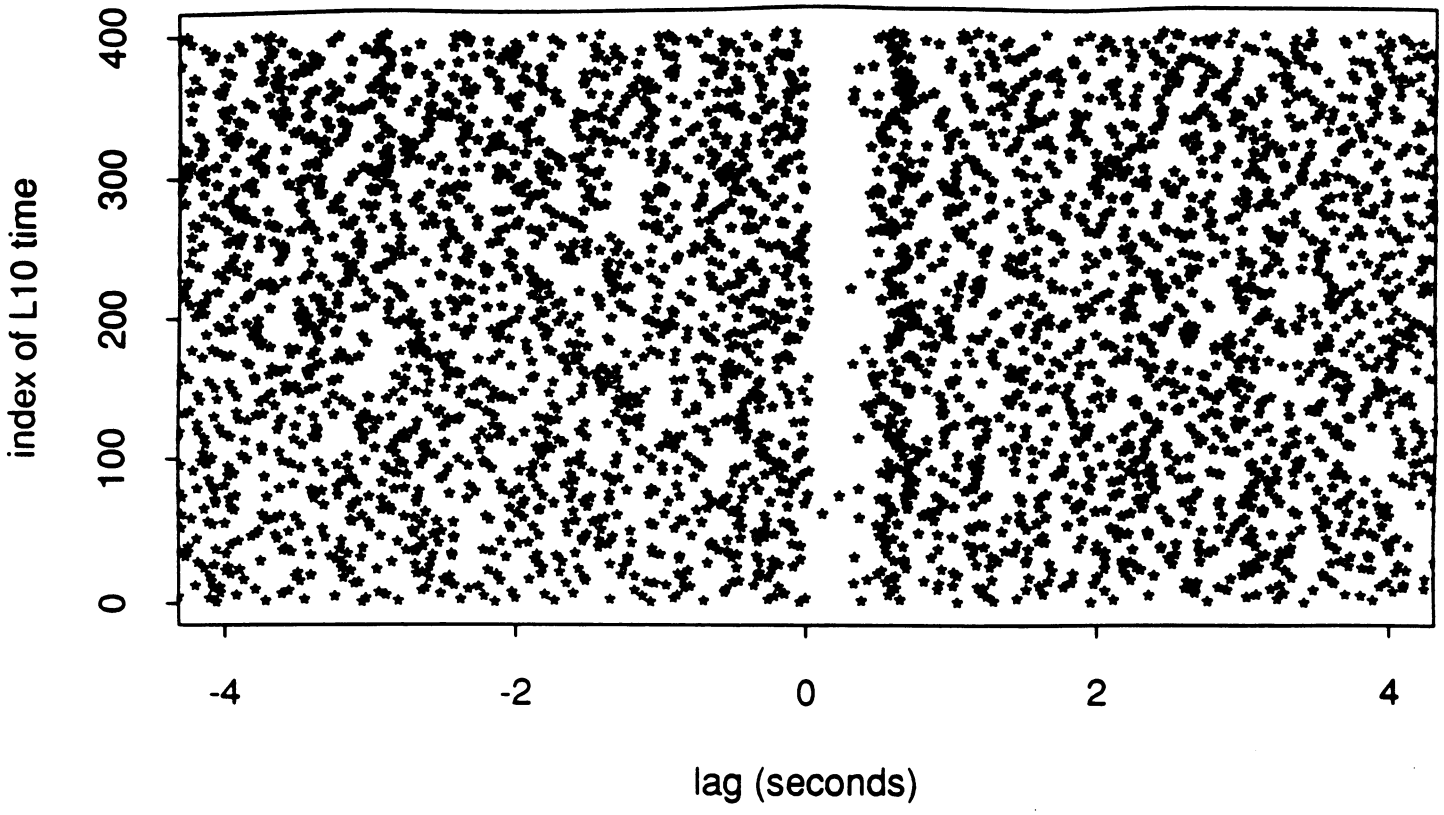
Figure 21. Statistics to investigate the association of cells 2 and 8.


A
B
C
D
E
F
G

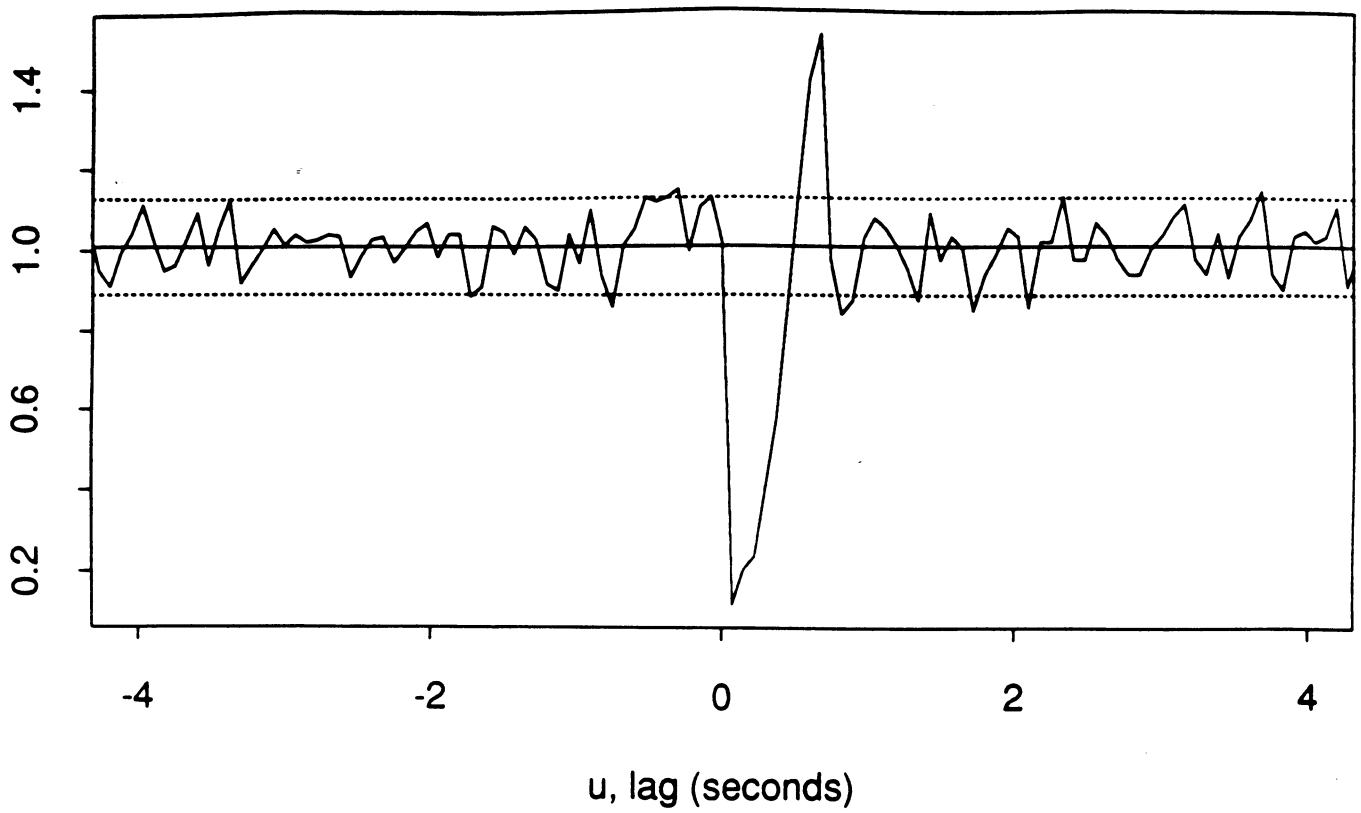
Transmembrane Potential



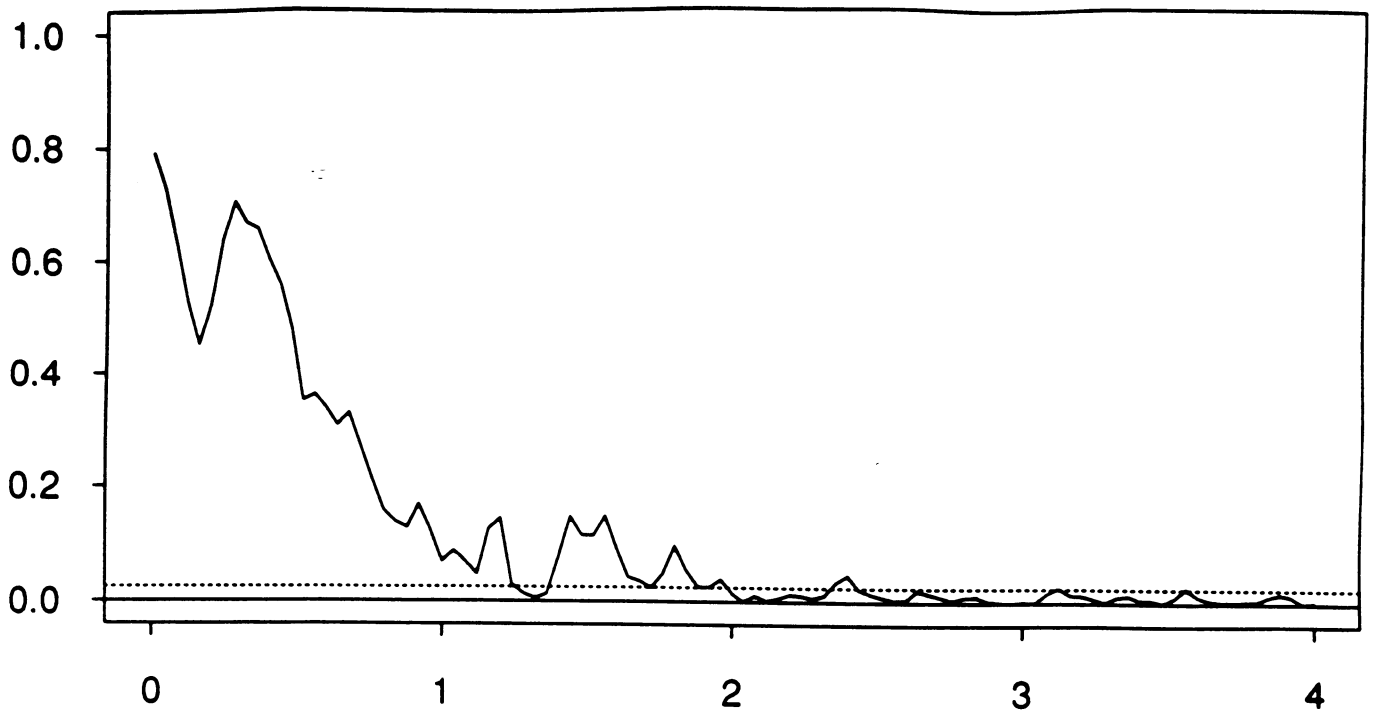

Cell 7 - no stimulus
With noise stimulus
400
300
200
100
0
segment
0
200
600
1000
lag (msec)


L3 spikes following L10's
index of L10 time
0
100
200
300
400
-4
-2
0
2
4
lag (seconds)

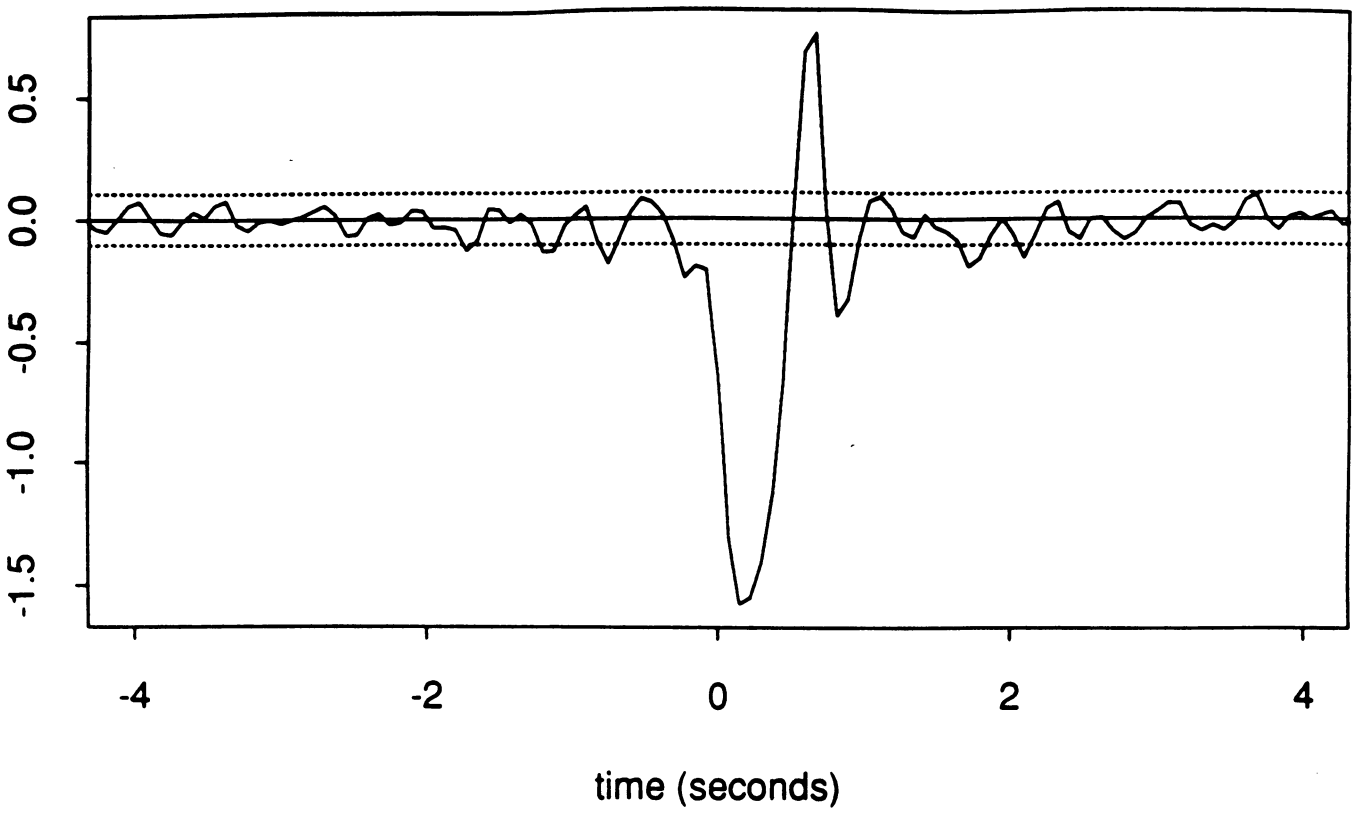
Square root crossintensity L3 given L10



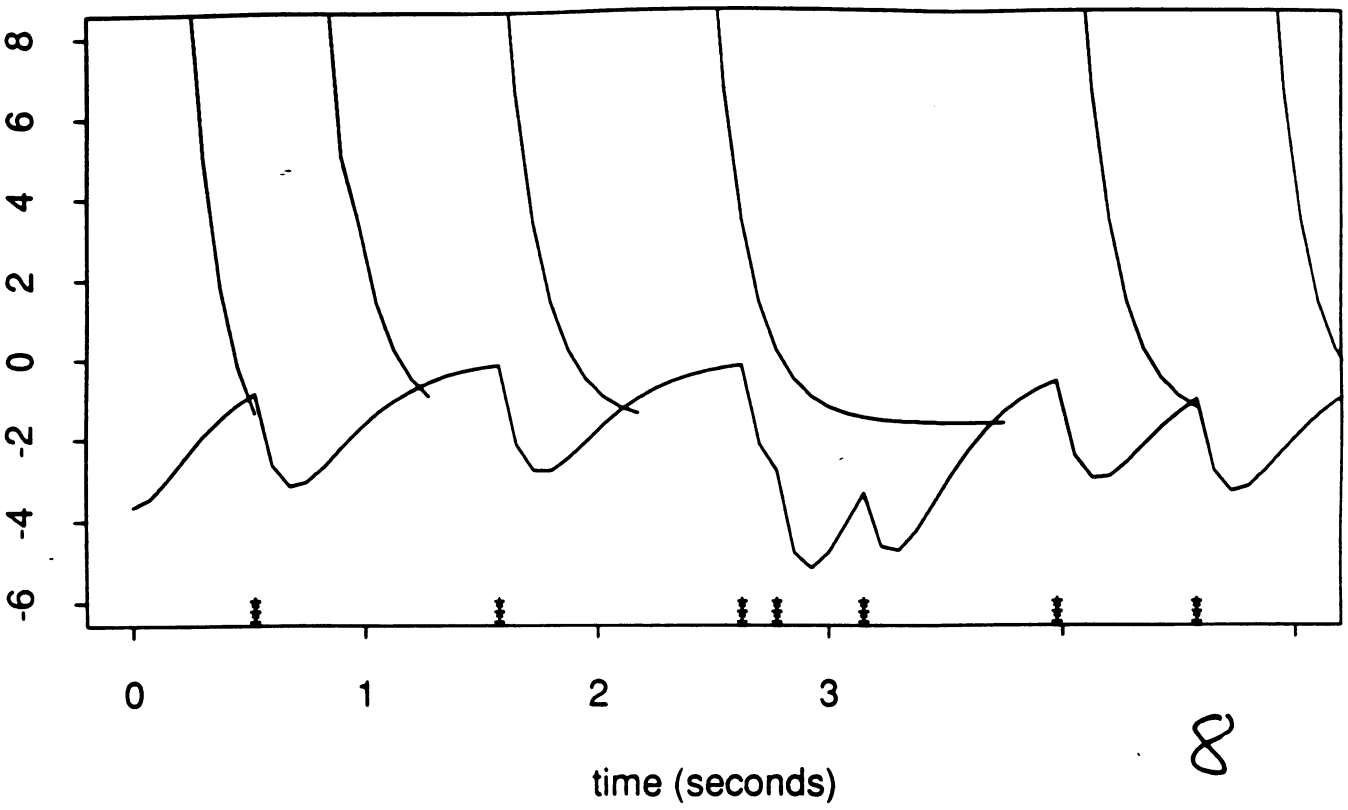
Coherence L10 and L3



frequency (cycles/second)
..... line gives 95% null point

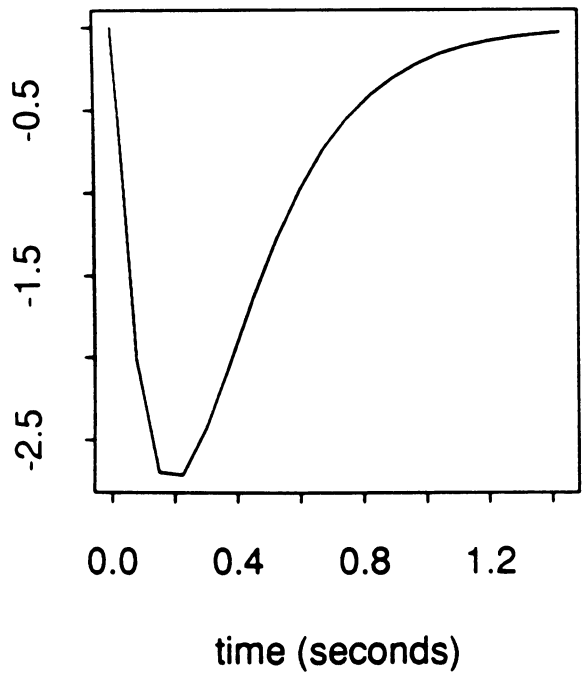

Impulse response of L3 given L10
0.5
0.0
-0.5
-1.0
-1.5
-4
-2
0
2
4
time (seconds)

Membrane potential and threshold function

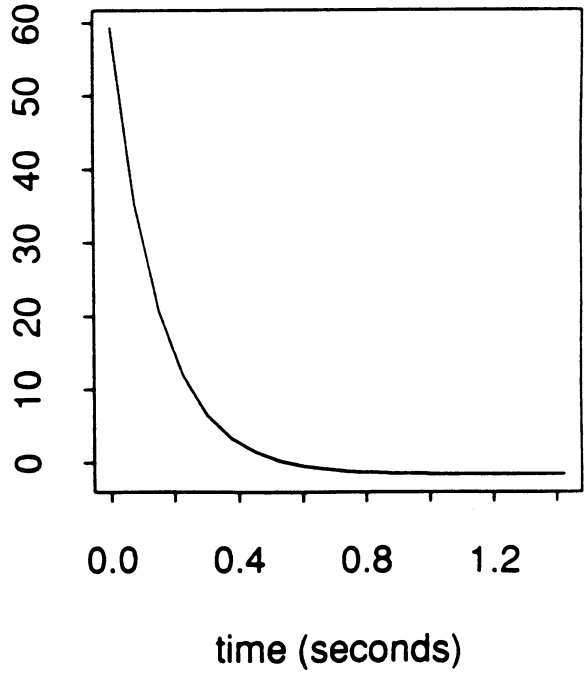


8

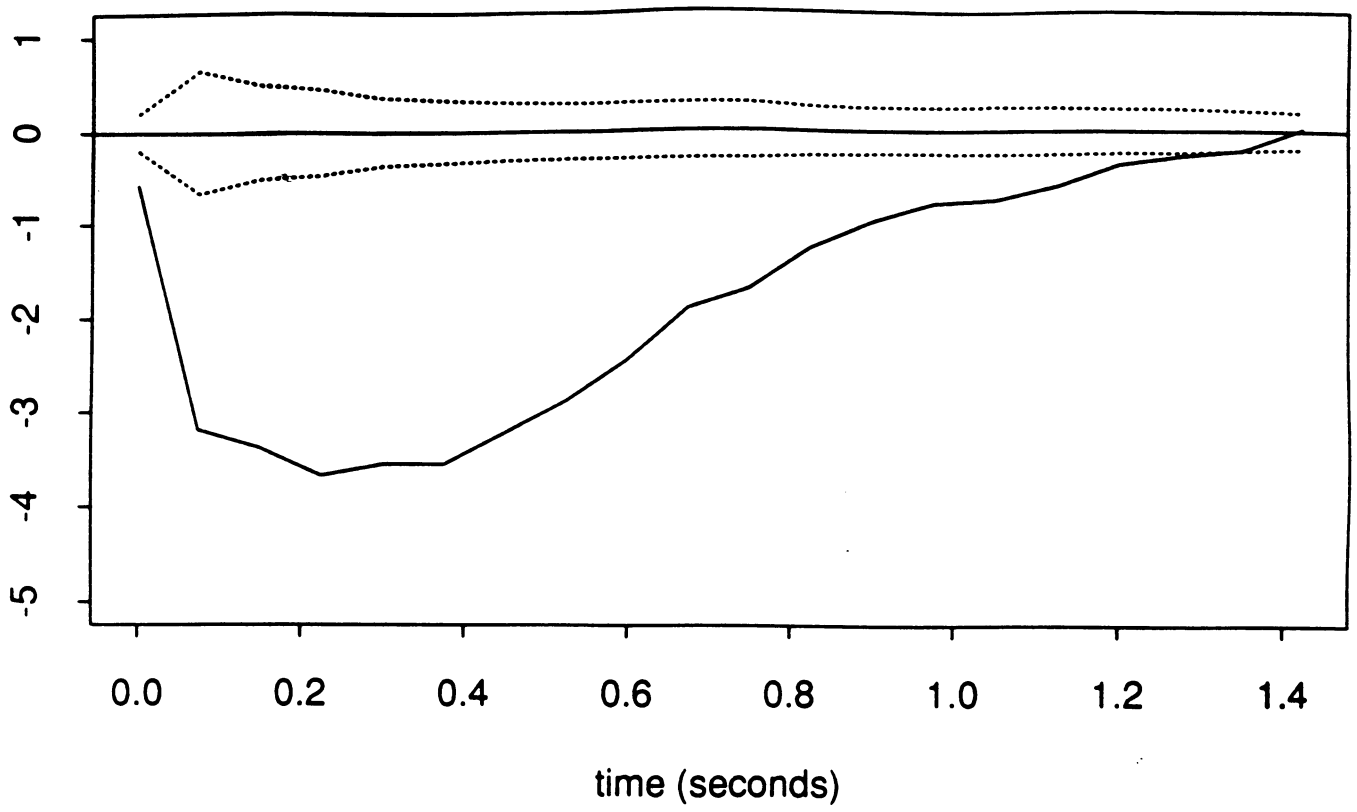
Summation function



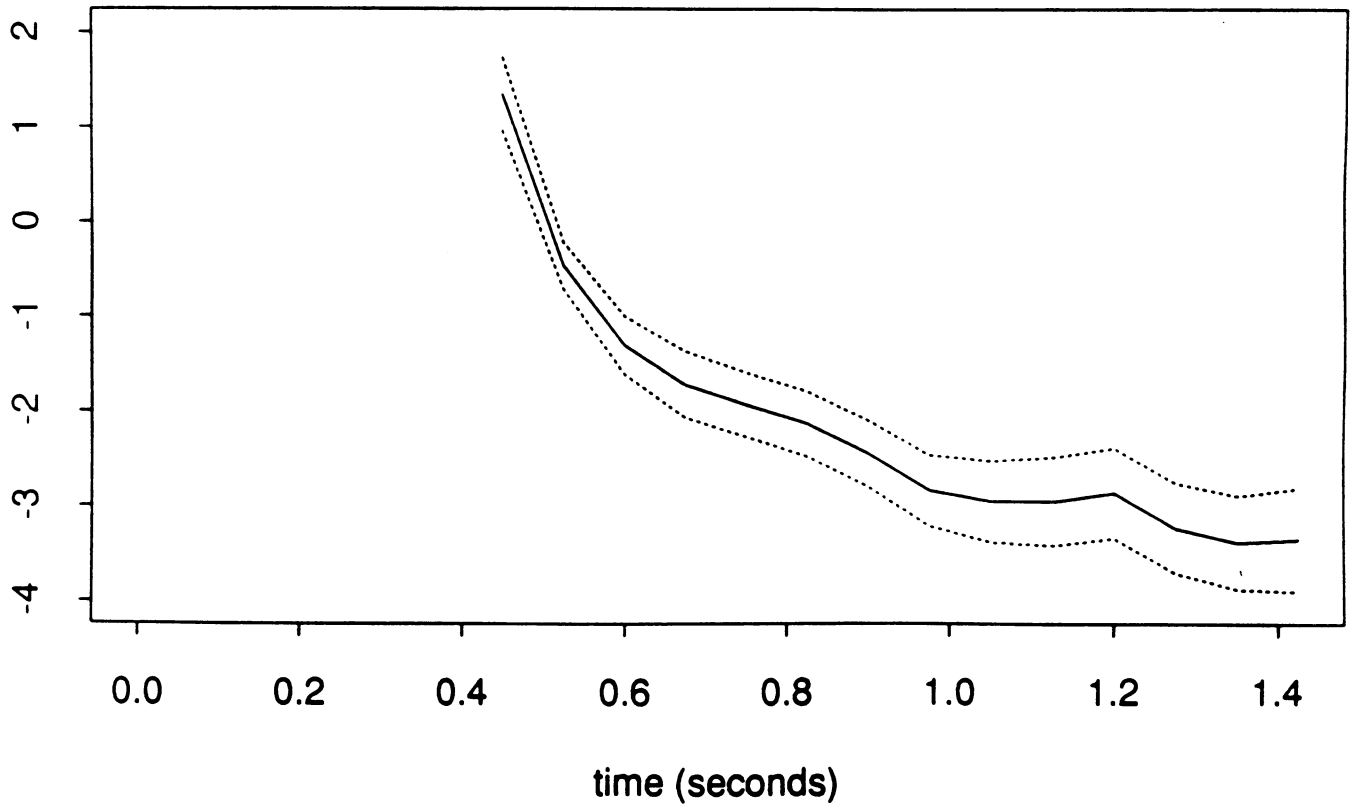
Decay function



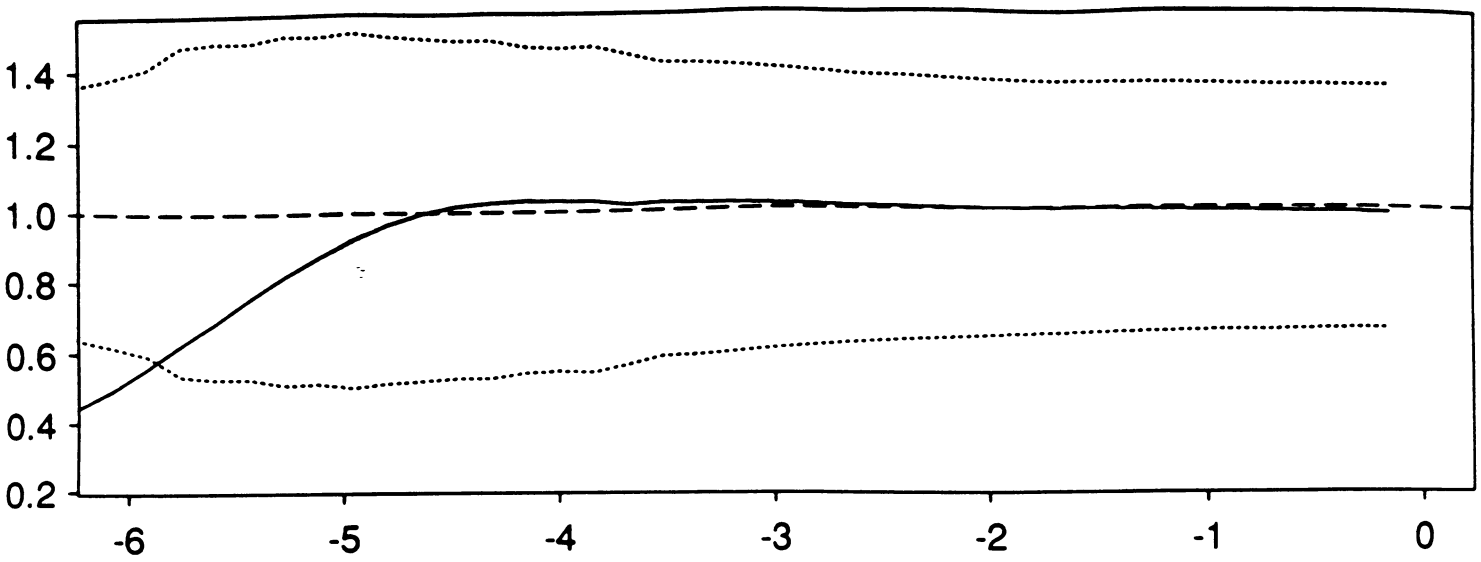
Summation function



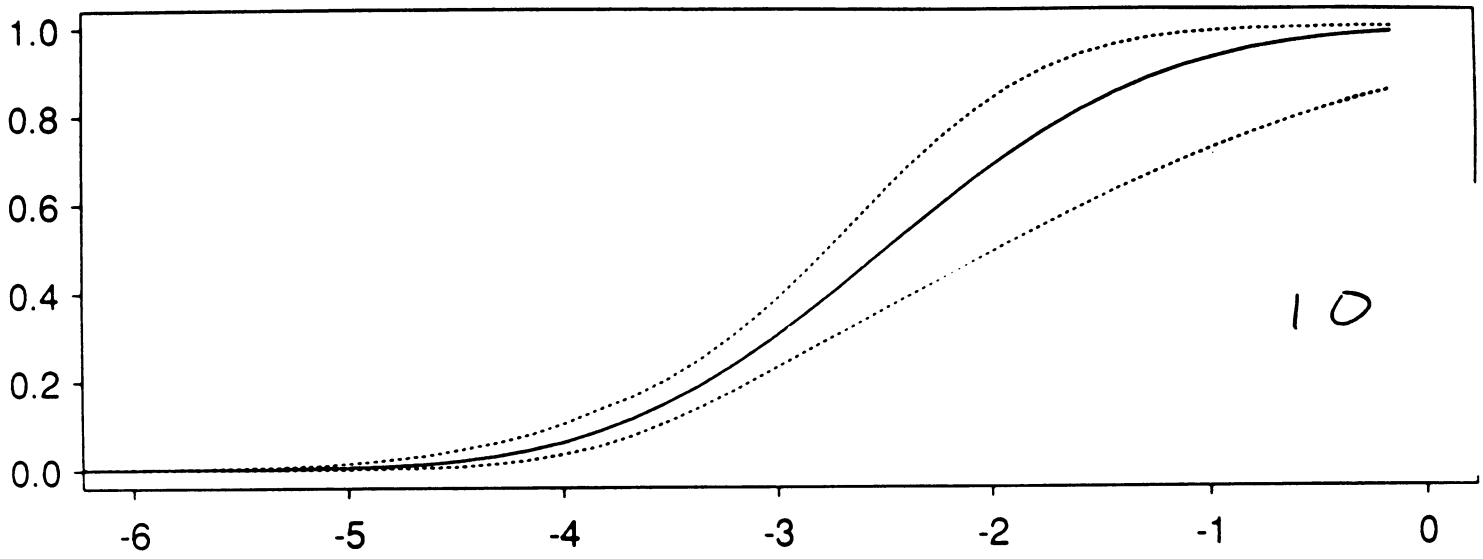
Decay function



Derivative, g'(.)

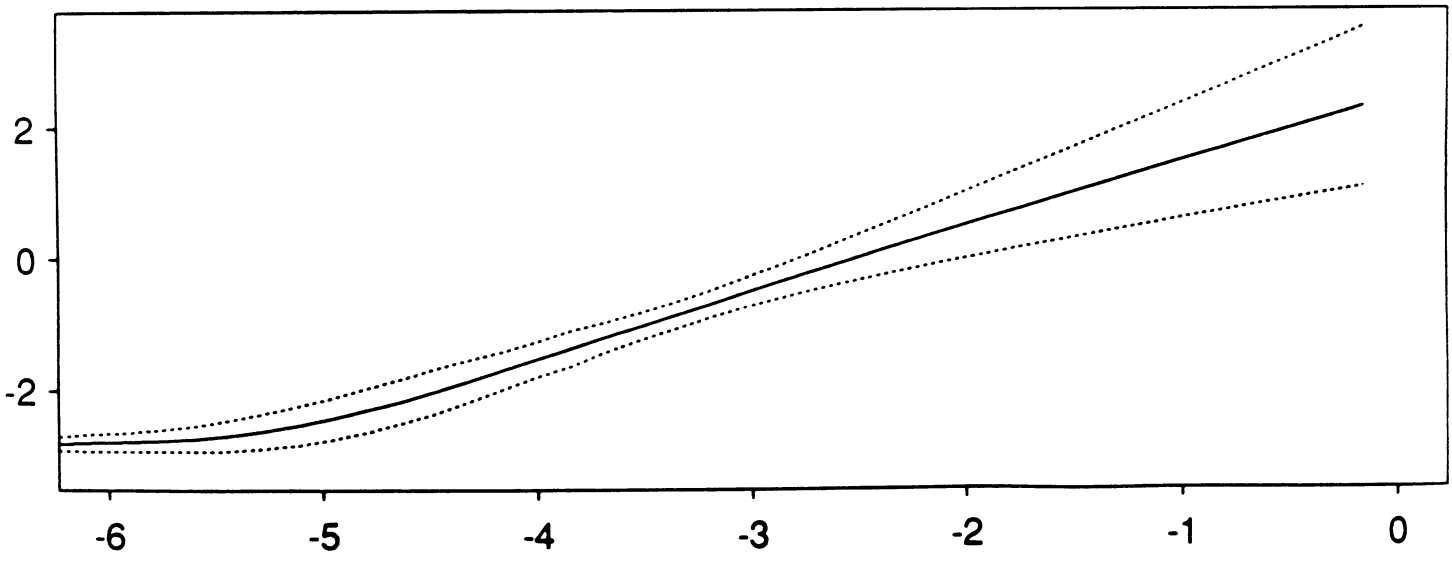


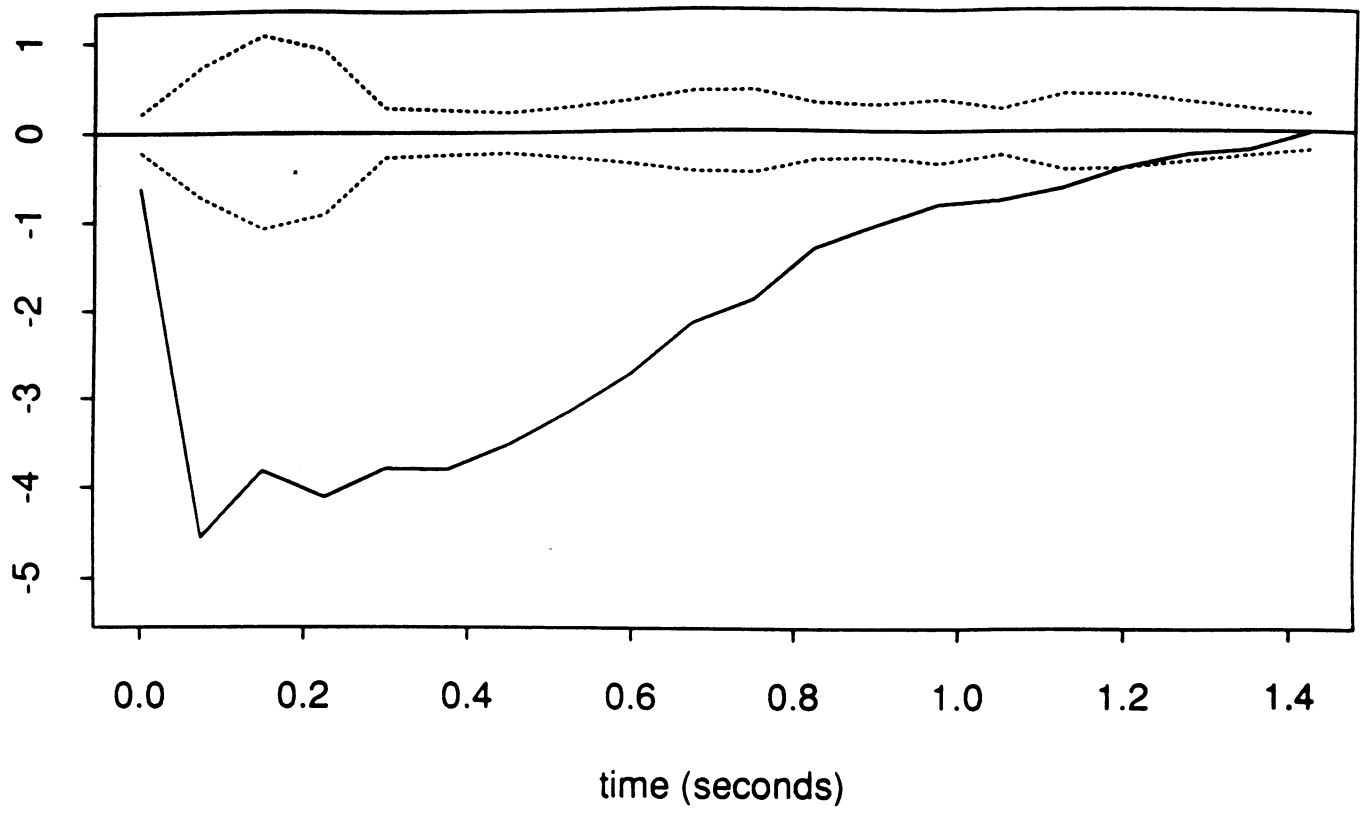
Probability function, P

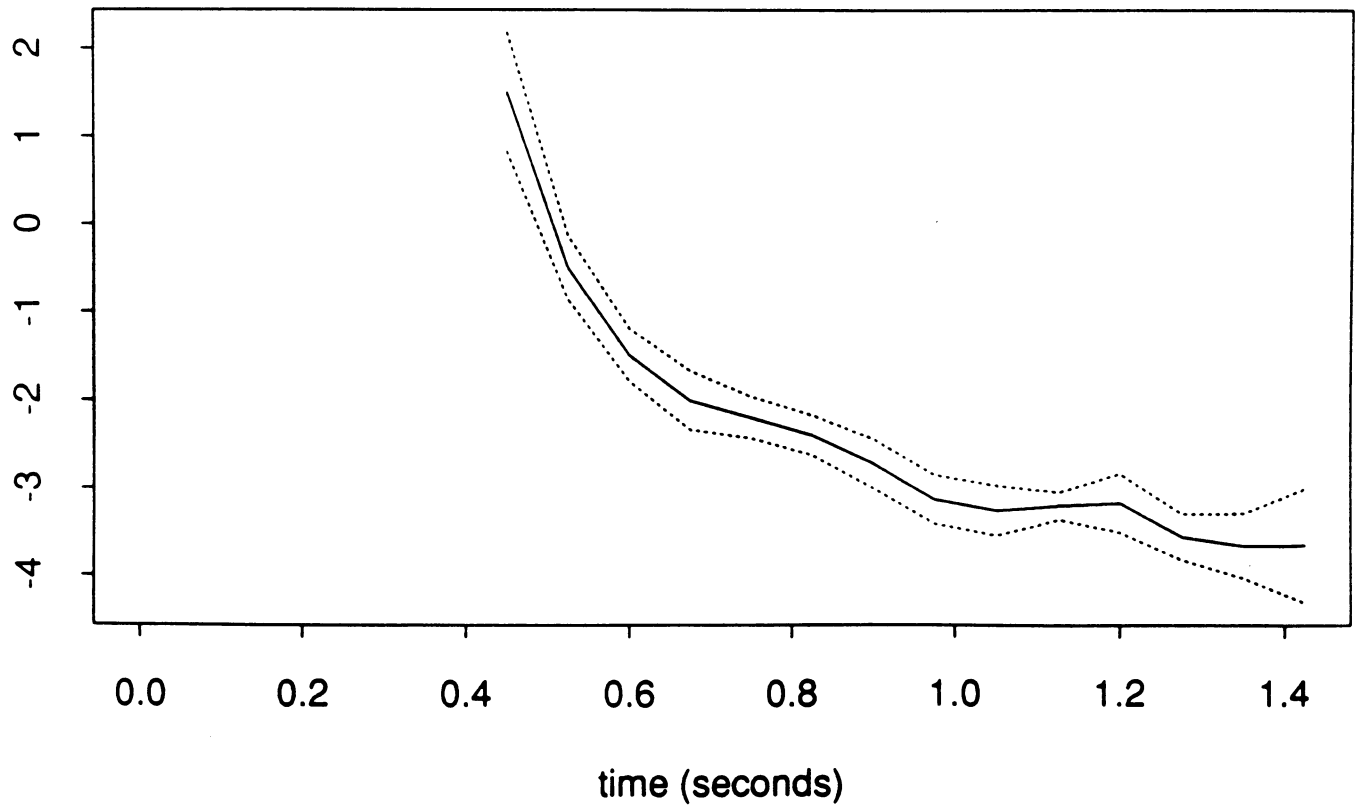


10

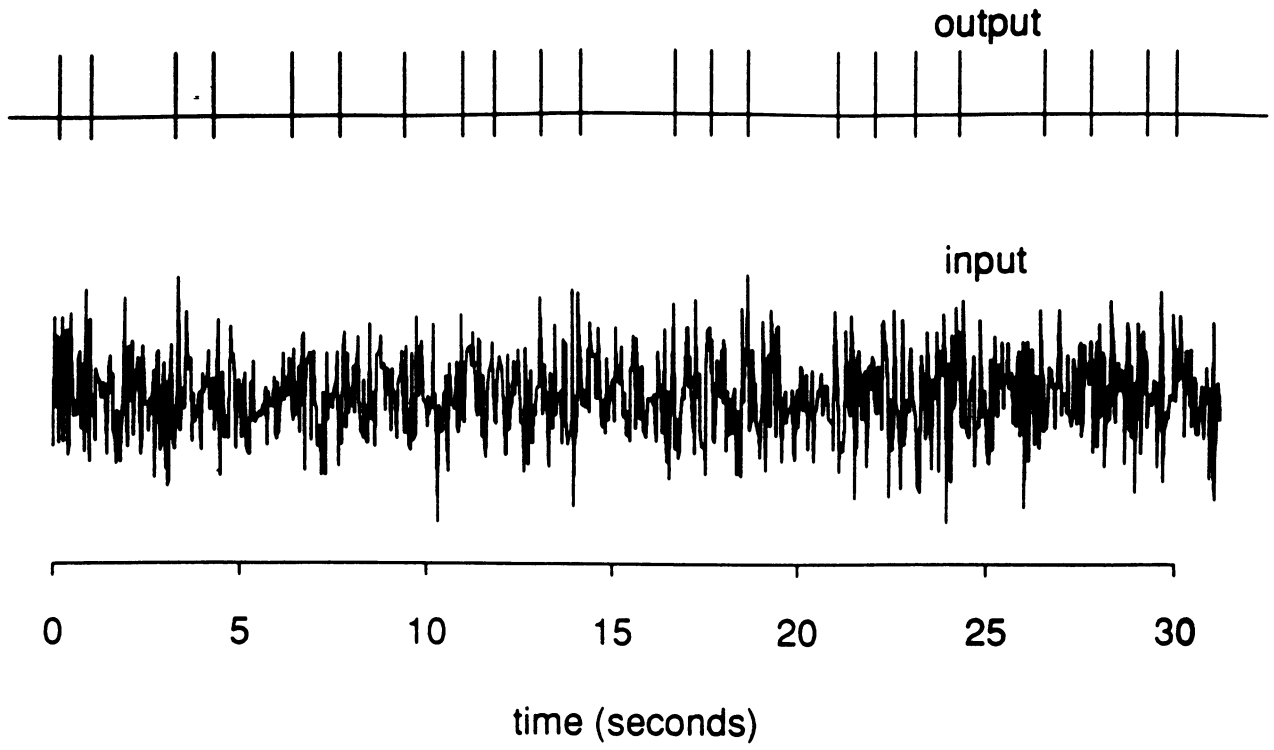
Transform, g(.)

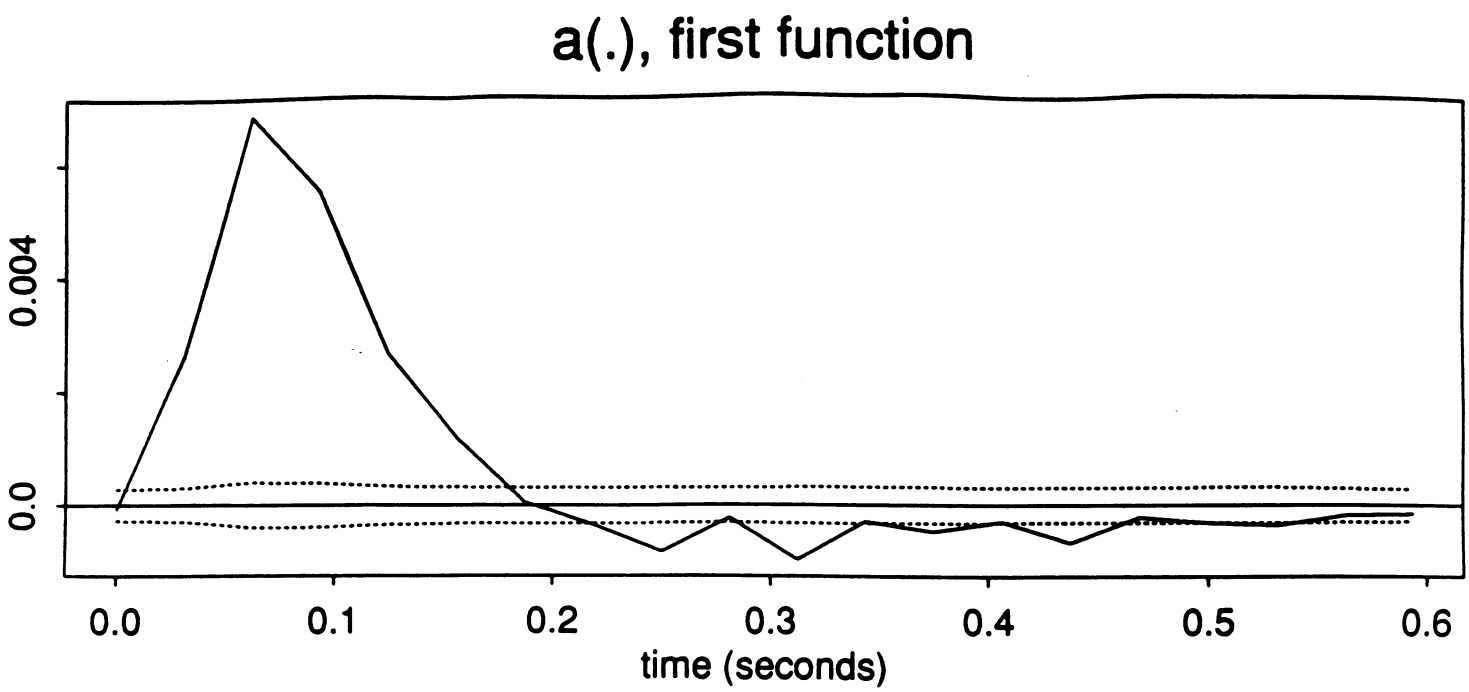


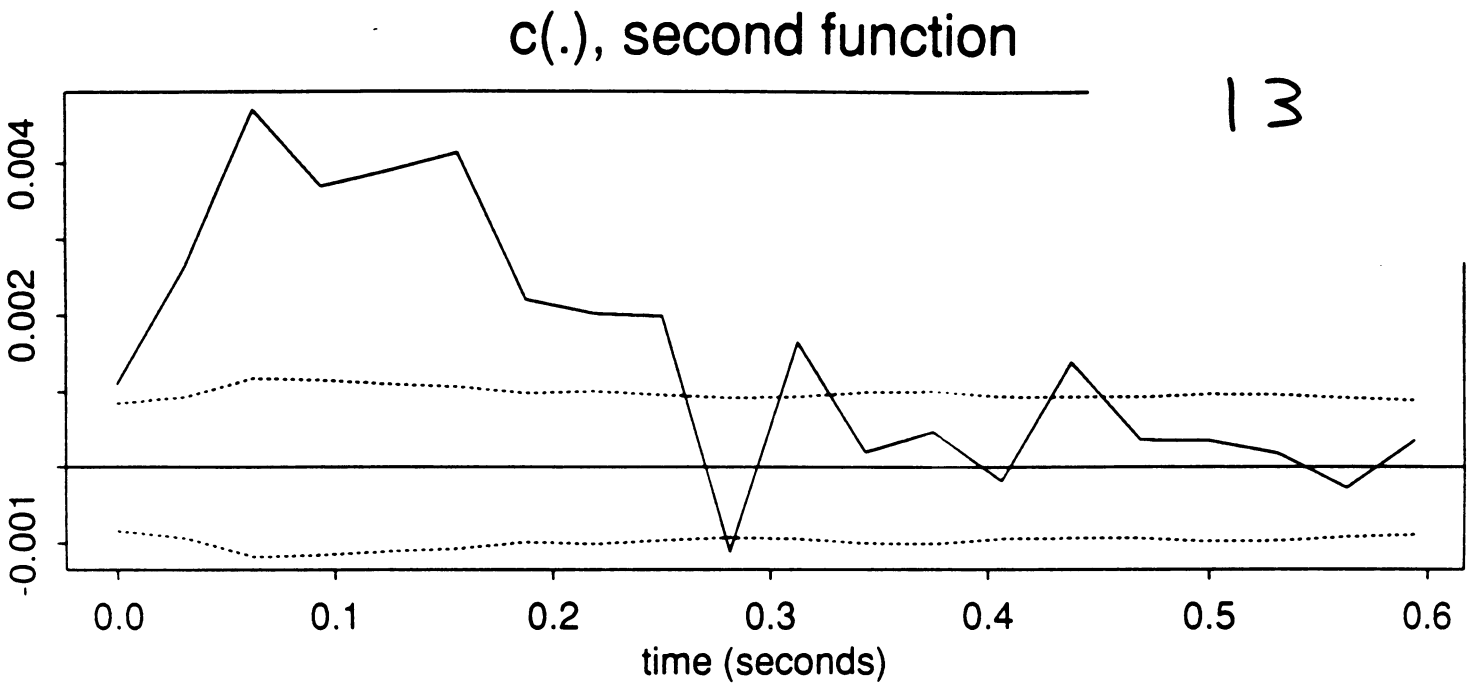

Summation function
1
0
-1
-2
-3
-4
-5
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
time (seconds)

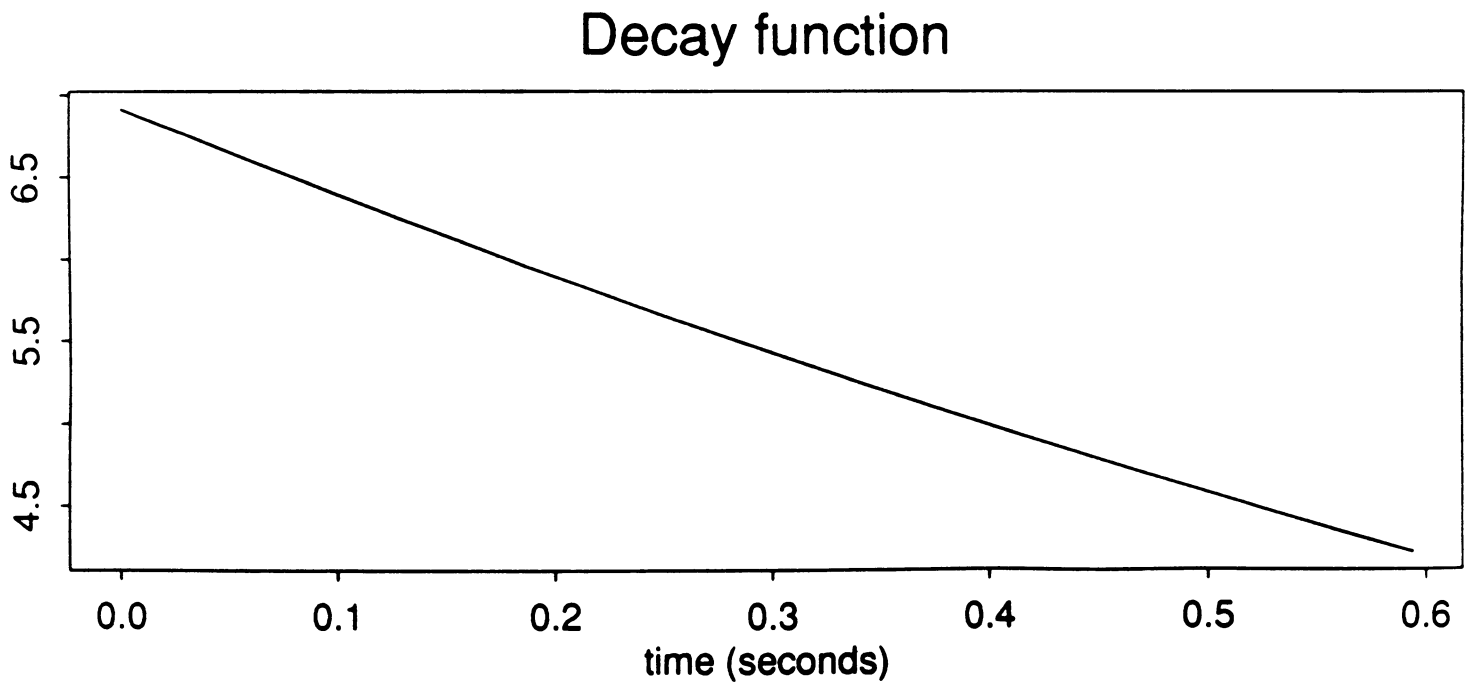

Decay function
2
1
0
-1
-2
-3
-4
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
time (seconds)

Neuron L5 - Noise Driven

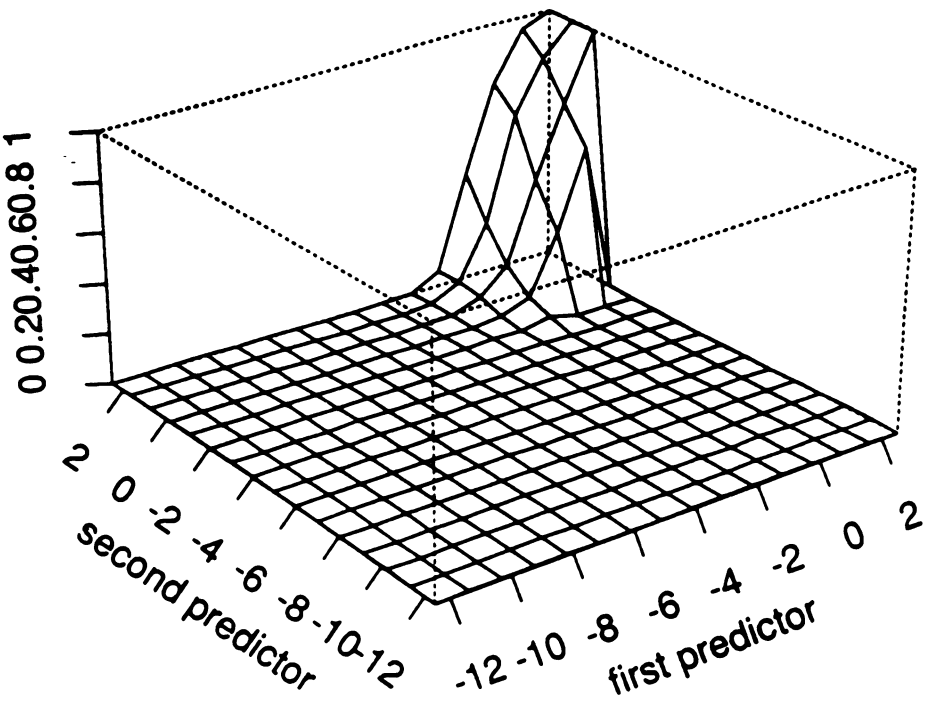



a(.), first function
0.004
0.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
time (seconds)

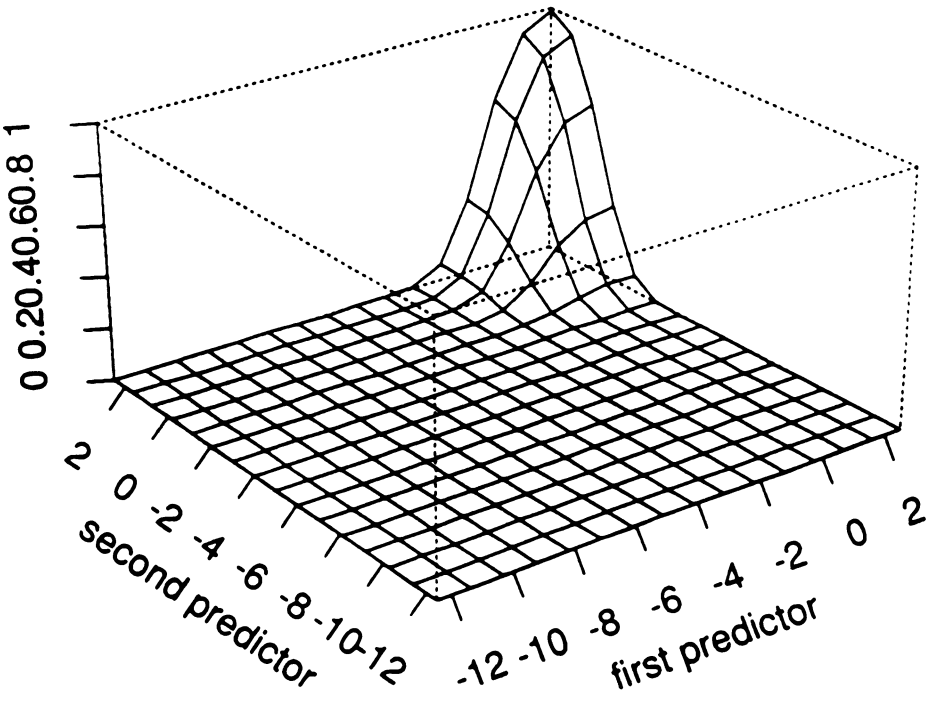

c(.), second function
13
0.004
0.002
-0.001
0.0
0.1
0.2
0.3
0.4
0.5
0.6
time (seconds)


Decay function
6.5
5.5
4.5
0.0
0.1
0.2
0.3
0.4
0.5
0.6
time (seconds)

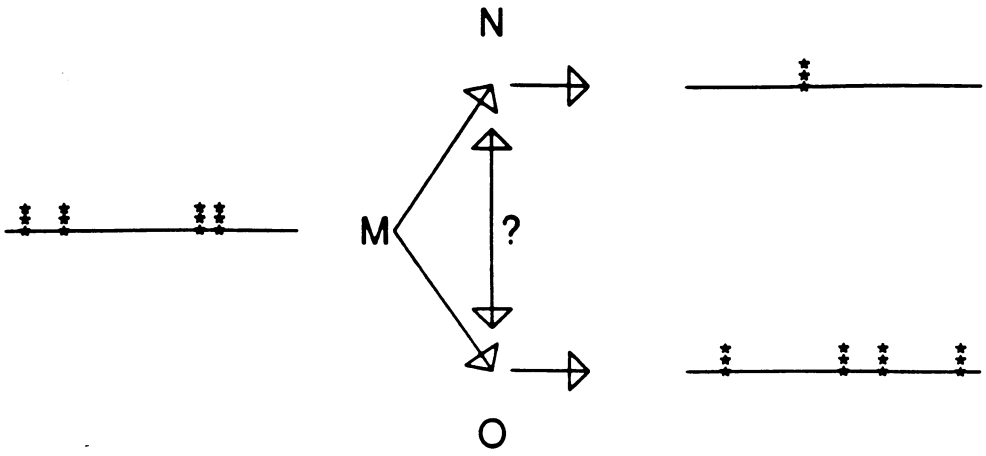
Empirical Firing Probability



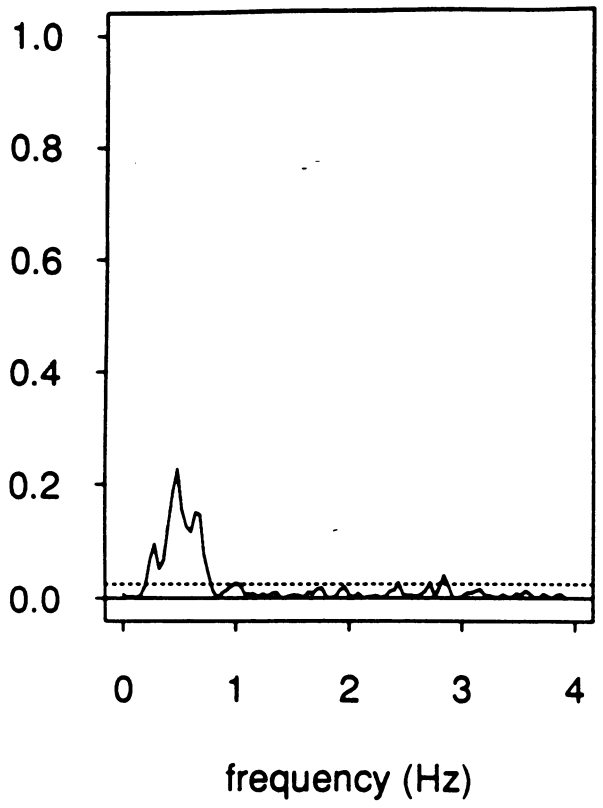
Theoretical Firing Probability



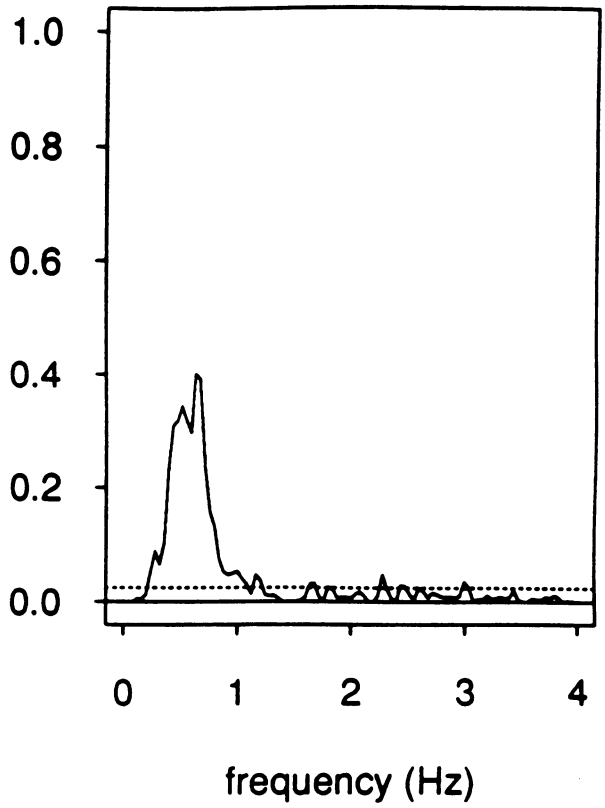
Network connections / causal models



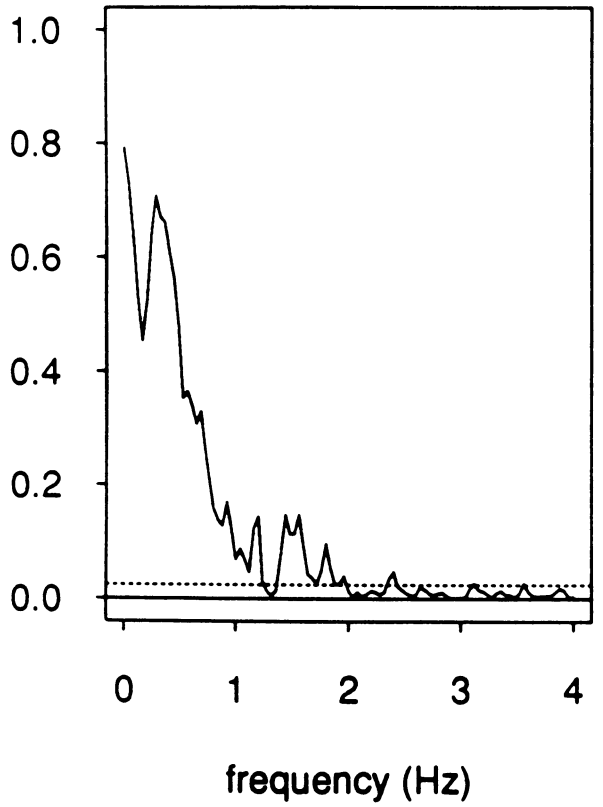
Coherence N and O



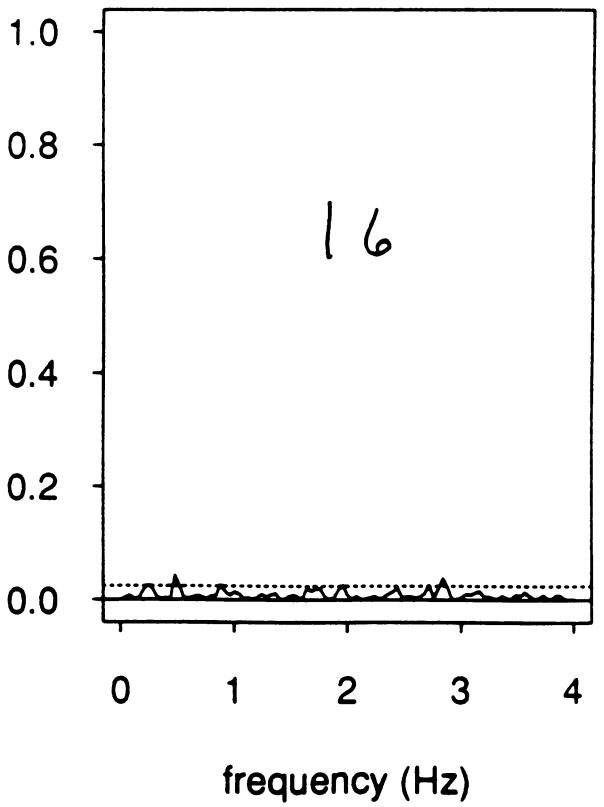
Coherence M and N

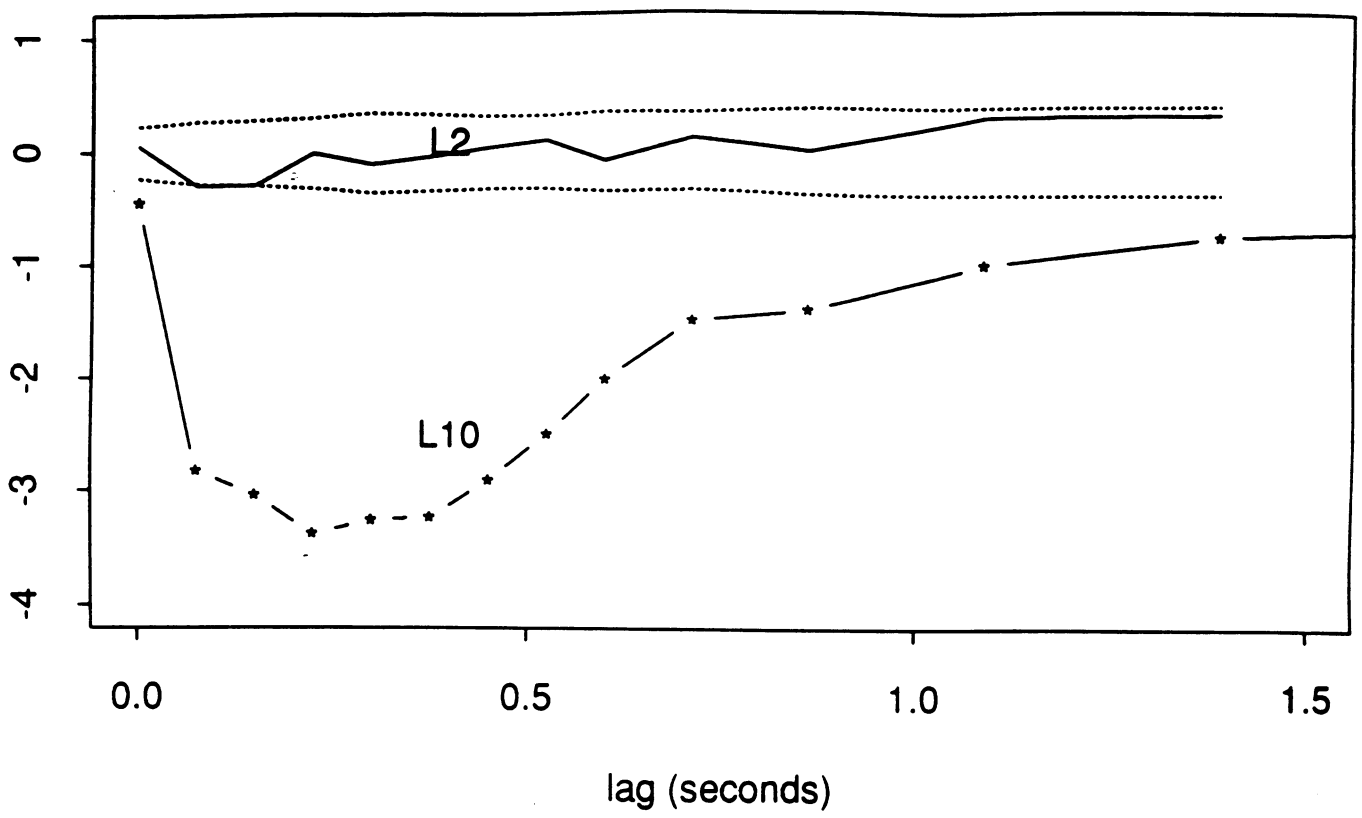


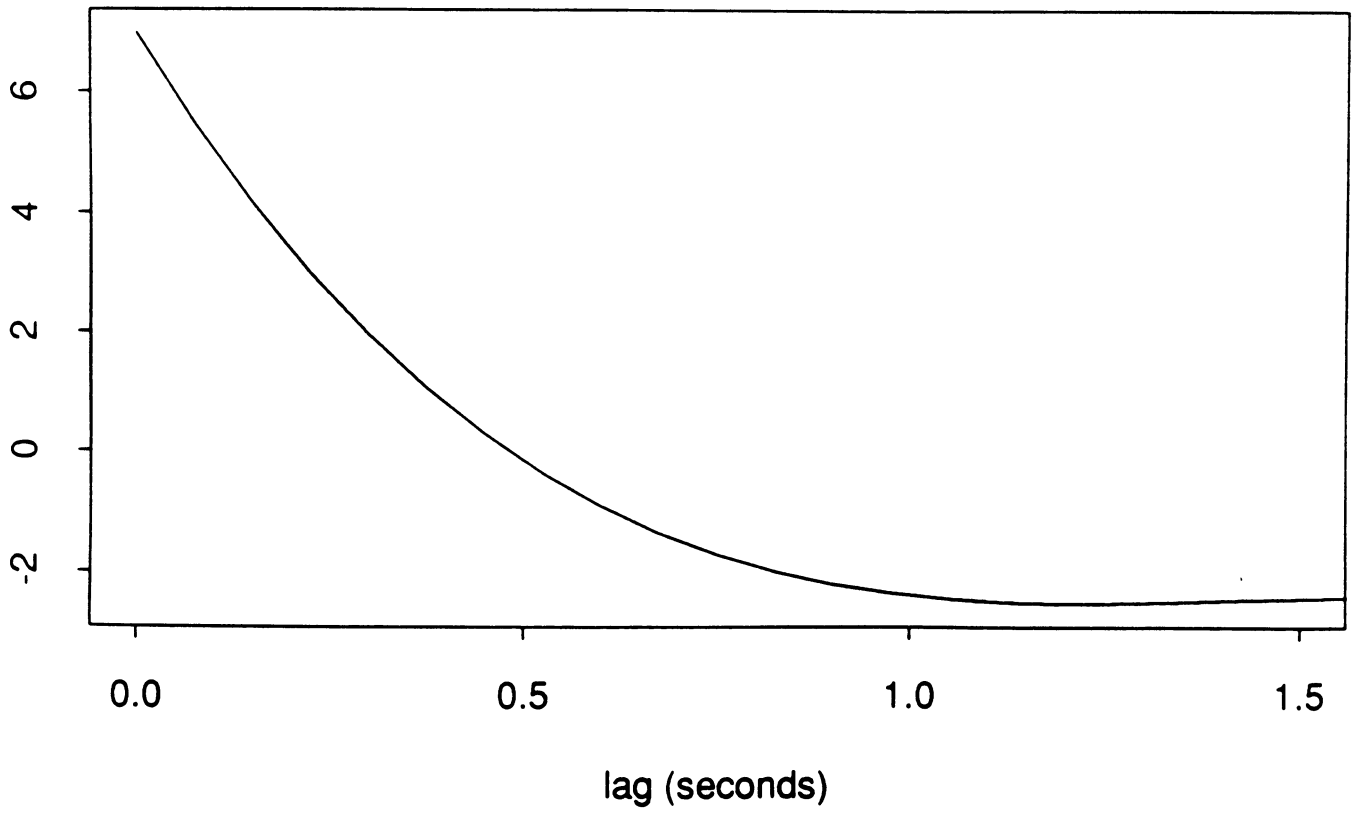
Coherence M and O

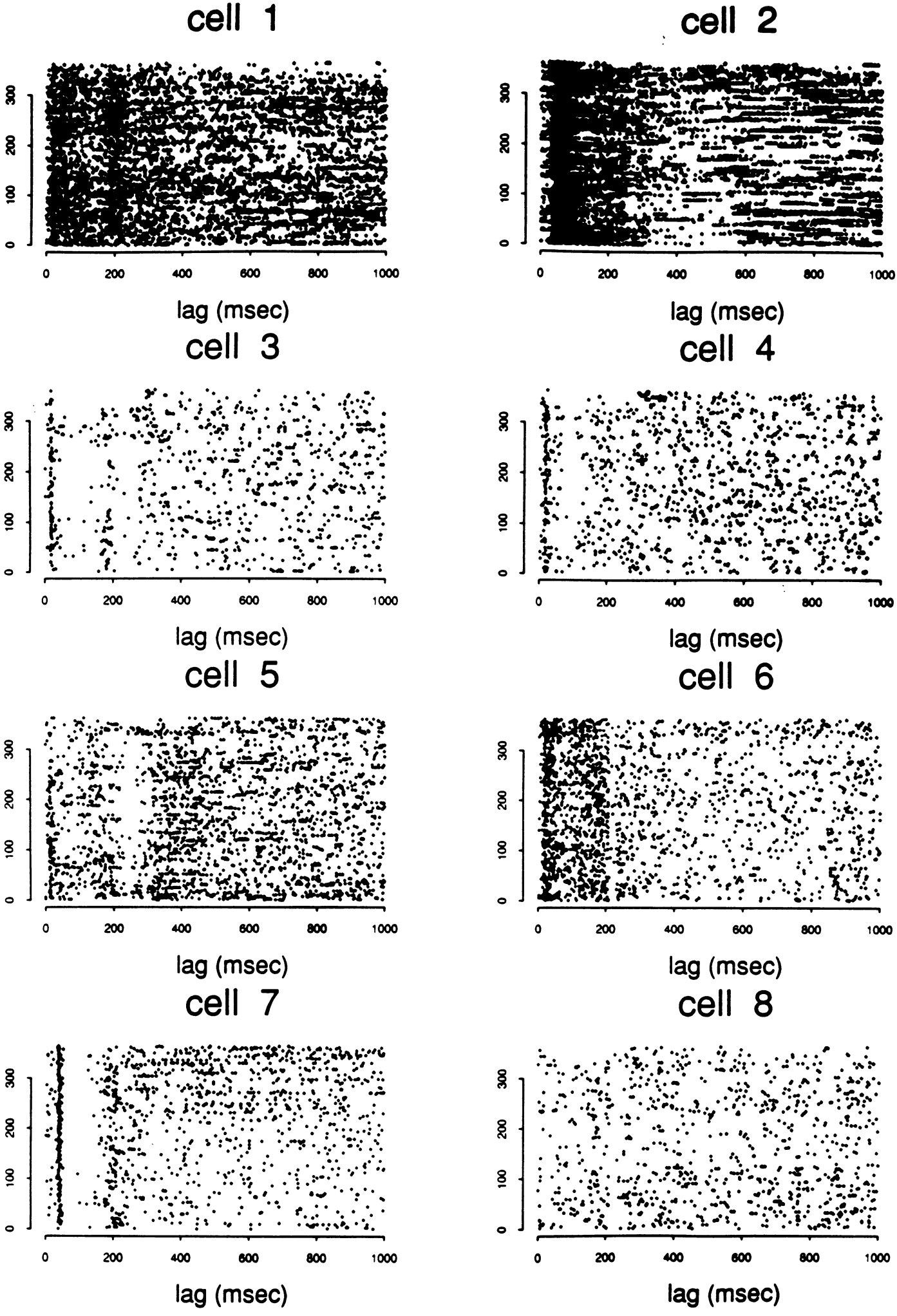


Partial coherence N & O

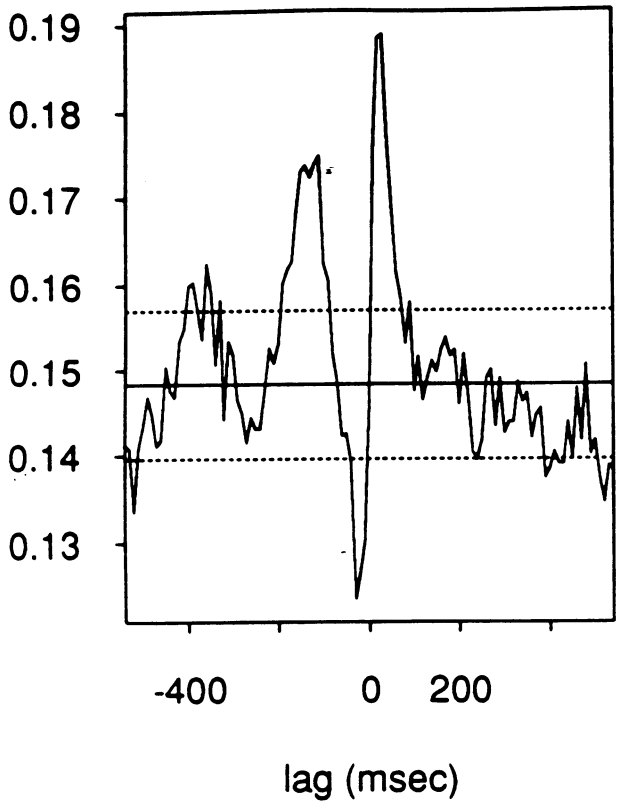



m(.) and o(.)
1
0
-1
-2
-3
-4
L2
L10
0.0
0.5
1.0
1.5
lag (seconds)

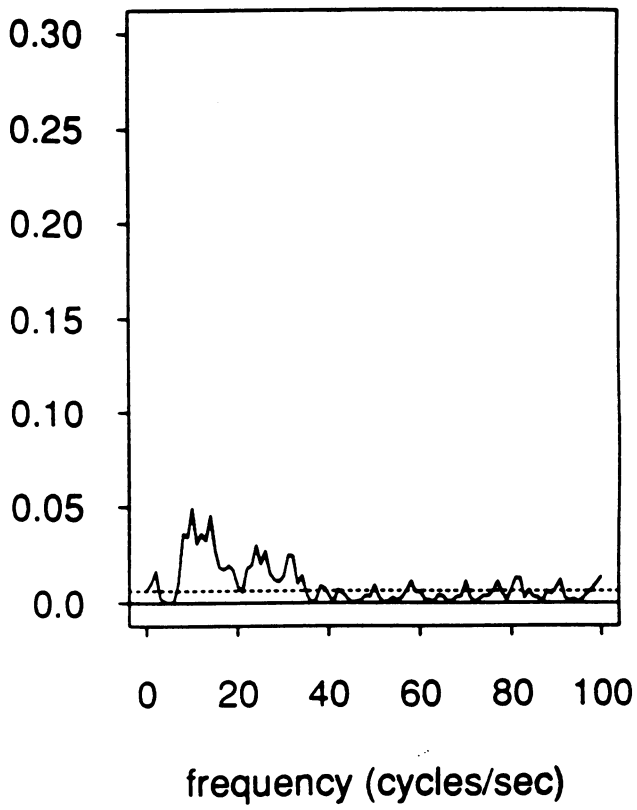

Decay function
6
4
2
0
-2
0.0
0.5
1.0
1.5
lag (seconds)


cell 1
cell 2
cell 3
cell 4
cell 5
cell 6
cell 7
cell 8
0
100
200
300
0
200
400
600
800
1000
lag (msec)

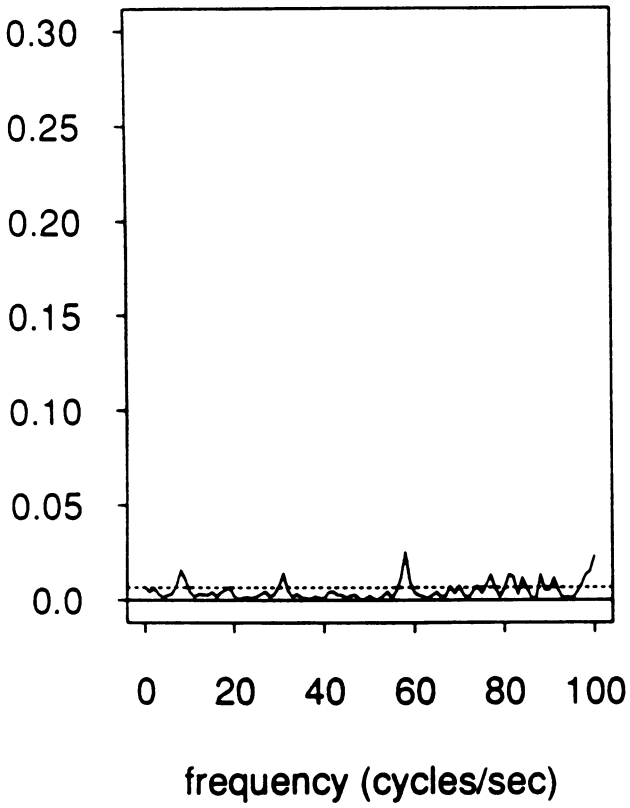
Cells 2&7, sqrt(crossint)



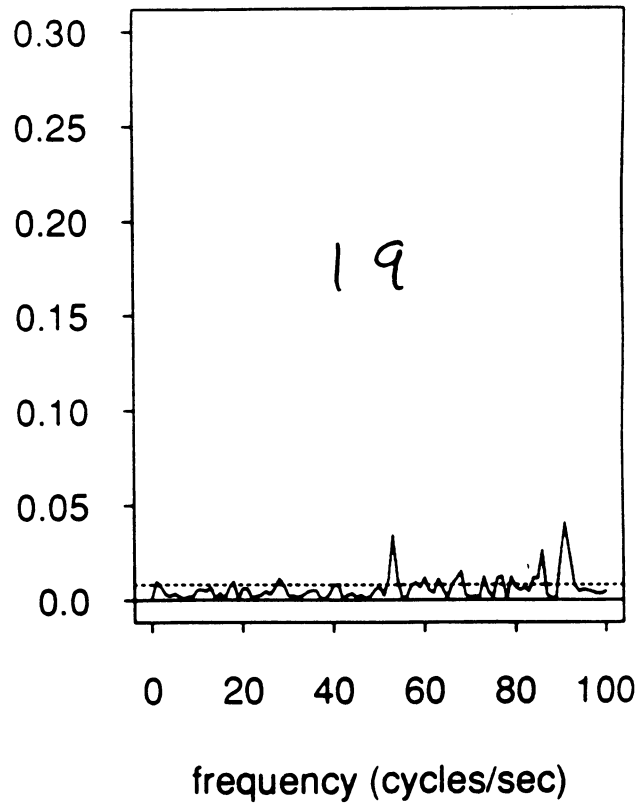
Coherence



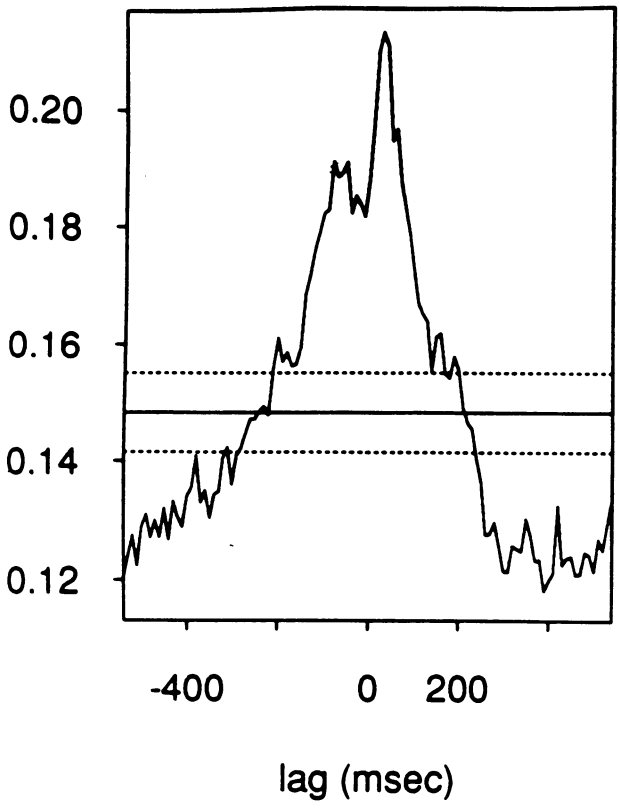
Partial coherence



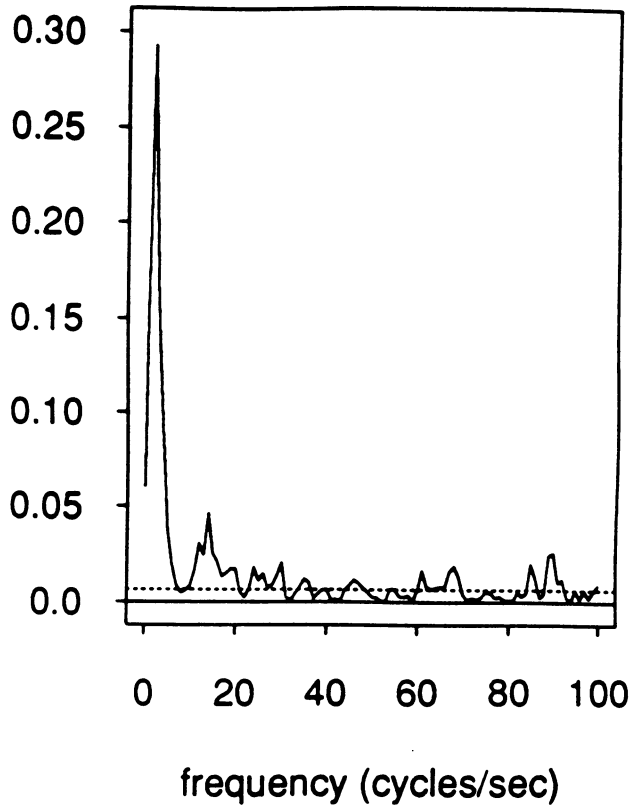
Coherence, no input



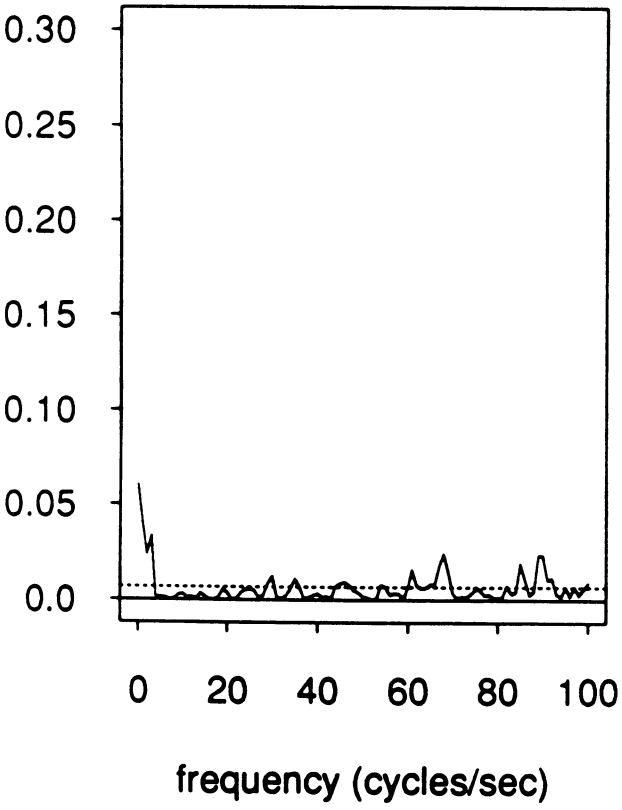
Cells 2&6, sqrt(crossint)



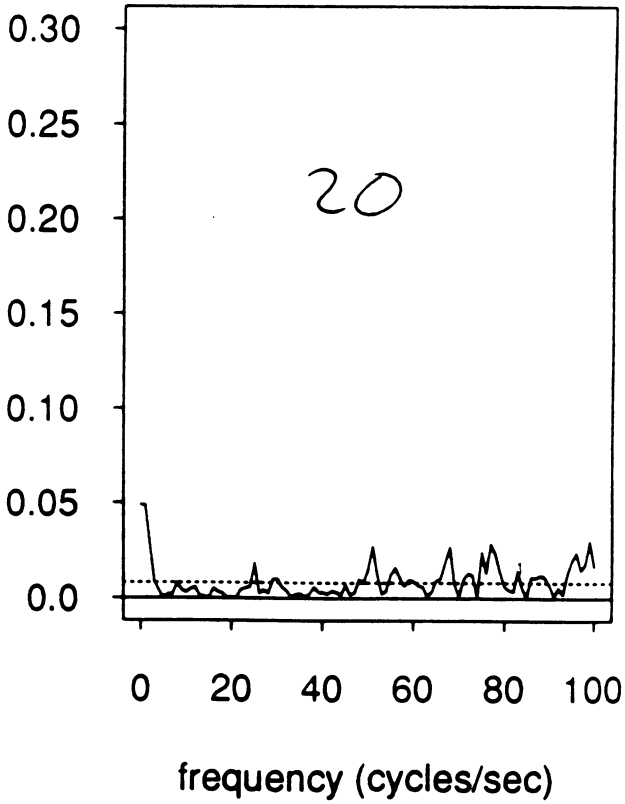
Coherence



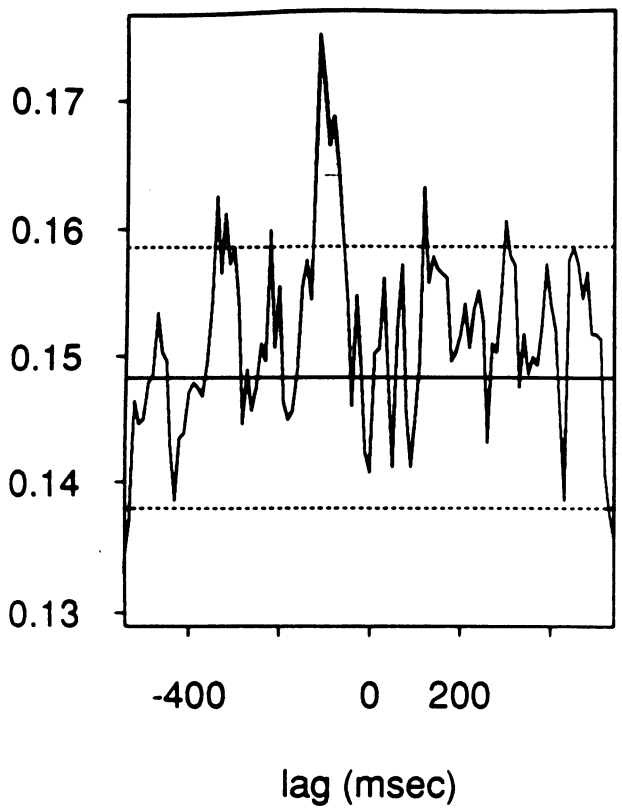
Partial coherence



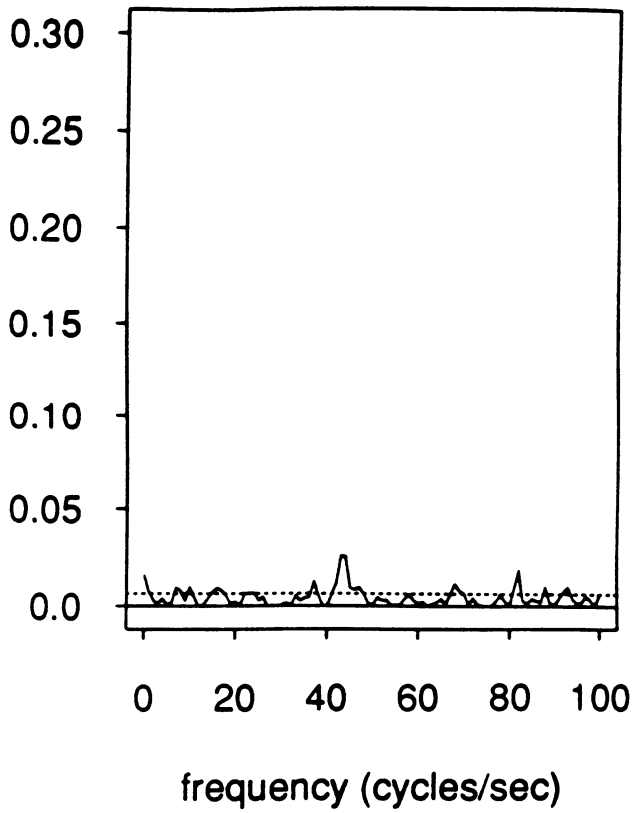
Coherence, no input



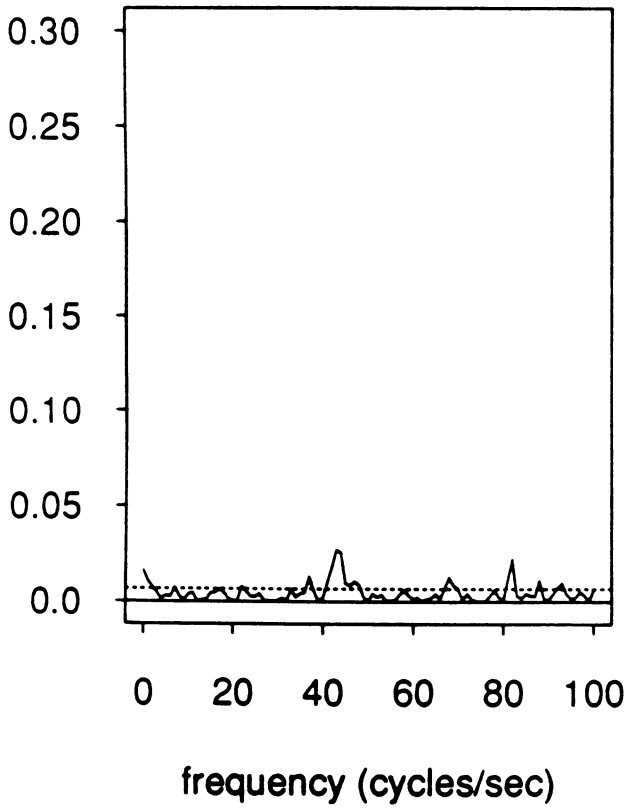
Cells 2&8, sqrt(crossint)



Coherence



Partial coherence



Coherence, no input