SOME GRAPHICAL METHODS FOR ASSESSING THE DEPENDENCE
STRUCTURE BETWEEN NEURONAL SPIKE TRAINS

BY

HANI DOSS   and   JOSEPH MARHOUL

TECHNICAL REPORT NO. 426

FEBRUARY 13, 1990

DEPARTMENT OF STATISTICS
STANFORD   UNIVERSITY
STANFORD, CALIFORNIA

SOME GRAPHICAL METHODS FOR ASSESSING THE DEPENDENCE

STRUCTURE BETWEEN NEURONAL SPIKE TRAINS

HANI DOSS   and   JOSEPH MARHOUL

BY

TECHNICAL REPORT NO. 426

FEBRUARY 13, 1990

DEPARTMENT OF STATISTICS

STANFORD   UNIVERSITY

STANFORD, CALIFORNIA

1. INTRODUCTION.

Suppose that two (possibly dependent) point processes are observed simultaneously over a period of time, yielding observations at times $A_1 < A_2 < \cdots < A_{N_A}$ for the first process, and at times $B_1 < B_2 < \cdots < B_{N_B}$ for the second. Such data arises in many contexts, and it is often of interest to discover and quantify the association between the two processes. Two fields in which this situation occurs are neurophysiology and reliability theory, from which the following four examples are drawn.

In neurophysiology, point processes arise as the impulse times of neurons. An example of a hypothesis testing problem arises when one wants to determine whether or not the impulse times of two neurons, say A and B, are associated. Of course, if an association is discovered, it is usually of interest to identify the nature of the relationship between the two neurons; e.g. is A driving (inhibiting) B? If so, by how much? This is typically a harder problem.

A second example, one in which estimation is necessary, is given by the following situation. An animal is to be taught (or trained) to perform a certain task. Now consider two connected neurons which are essential in the performance of this task, and simultaneously record their firing patterns. Before the learning has taken place, the neurons will in general affect each other. The problem is to determine the way in which this dependence is altered after the learning has taken place.

1

The next two examples are taken from reliability theory. Consider a system which contains two components, A and B, and for which A is replaced according to an age replacement policy: the unit is replaced at failure or at time T, whichever comes first. Let $A_1 < A_2 < \cdots$ be the successive times at which the component is replaced without having failed, and let $B_1 < B_2 < \cdots$ be the times of failure of component B. It may be of interest to determine if the point processes $\{A_i\}$ and $\{B_j\}$ are dependent, and in particular to determine if the age replacement policy is beneficial (or harmful) to component B.

As a second example, consider a series system of two subsystems A and B, where A and B are parallel structures of k_A and k_B components, respectively $(k_A \geq 2, k_B \geq 2)$. Assume that any failed component is repaired, and that repair time is negligible. Thus, if any component fails the system continues to operate. Let $A_1 < A_2 < \cdots$ and $B_1 < B_2 < \cdots$ denote the times at which failures occur in subsystems A and B, respectively. Then, these form a bivariate point process, and it may be of interest to determine if the failure times of subsystems A and B are dependent.

In this article, we describe and discuss certain graphs, plots, as well as more formal methods that can assess the dependence between point processes. Specifically, these methods indicate whether or not the likelihood of an A-point is increased (decreased) after the occurrence of a B-point. The techniques are illustrated on simulated data. Although bivariate point processes arise in many fields, we emphasize the applications in neurophysiology.

We hope to illustrate the techniques described here on data drawn from real neurons later.

2. THE FUNCTION $\lambda_{A|B}(\cdot)$ AND THE CROSS-CORRELATION HISTOGRAM.

We proceed at a heuristic level. A rigorous approach and formal definitions of the terms used below may be found in Daley and Vere-Jones (1972). The book by Cox and Isham (1980) is a good guide to the literature and contains all the necessary information to formalize what we do here.

We assume that the points $\{A_i\}$ and $\{B_j\}$ form a bivariate point process that is stationary and orderly. Intuitively, the stationarity assumption is that the process has been going on for a long time and is in steady state. The orderliness condition is that each univariate process has no multiple occurrences. For a Borel subset of the real line S, let $N_A(S)$ and $N_B(S)$ denote the number of points A_i and B_j, respectively, that lie in S. We will use $N_i(s,t)$ to denote $N_i((s,t])$, for $i = A,B$. We may view $N_A(\cdot)$ and $N_B(\cdot)$ as random Borel measures.

We define the rates λ_A and λ_B as follows:

$$(1) \qquad \lambda_A = \lim_{h \to 0} \frac{1}{h} P\{N_A(t,t+h) > 0\} ,$$

with a similar definiton for λ_B. By the stationarity assumption, λ_A and λ_B are independent of t. The existence of the limit in (1) was first proved by Khintchine (1960). Korolyuk's Theorem (see Leadbetter, 1968) states that if $t_1 < t_2$, then

$$(2) \qquad E \ N_A(t_1,t_2) = \lambda_A(t_2 - t_1) .$$

Similarly for the B process. A simple consequence of the ergodic theorem is that with probability one,

$$(3) \qquad \frac{N_i(0,T)}{T} \to \lambda_i \quad \text{as} \quad T \to \infty \ ,$$

for $i = A, B$. This gives a third way of thinking about the rates λ_A and λ_B.

The discussion so far relates only to the processes $N_A(\cdot)$ and $N_B(\cdot)$, taken one at a time. To see how the processes affect each other, we define the following quantities:

$$(4) \qquad \lambda_{A|B}(u) = \lim_{h \to 0} \frac{1}{h} \, P\{N_A(t+u, t+u+h) \,|\, N_B(\{t\})=1\} \quad \text{for} \quad -\infty < u < \infty \ .$$

Thus, roughly speaking, $\lambda_{A|B}(u)$ gives the infinitesimal probability of an A-point u units after a B-point. The stationarity assumption implies that $\lambda_{A|B}(u)$ is independent of t. We may also define the function $\lambda_{B|A}(\cdot)$, but it is simple to see that

$$\lambda_B \, \lambda_{A|B}(u) = \lambda_A \, \lambda_{B|A}(-u) \quad \text{for} \quad -\infty < u < \infty \ ,$$

so that it suffices to consider $\lambda_{A|B}(u)$, as long as we consider both positive and negative values of u.

The function $\lambda_{A|B}(\cdot)$ has been considered in different forms and contexts by Cox (1965), Cox and Lewis (1972), Brillinger (1976) and Griffith and Horn (1963), among others. If the processes $N_A(\cdot)$ and

4

$N_B(\cdot)$ are independent, we clearly have

$$\lambda_{A|B}(u) = \lambda_A \quad \text{for all } u ,$$

so that $\lambda_{A|B}(\cdot)$ can indicate deviations from independence. The following two examples of very simple neuronal networks serve to illustrate this point.

Example 1: Two independent neurons.

$N_A(\cdot)$ and $N_B(\cdot)$ are independent. Then, as was just mentioned,

$$(5) \qquad\qquad \lambda_{A|B}(u) = \lambda_A \quad \text{for all } n ,$$

so that in particular, $\lambda_{AB}(\cdot)$ is constant.

Example 2: A network of three neurons.

The spontaneous firings of neurons A and B form independent processes $N_1(\cdot)$ and $N_2(\cdot)$. A stimulus neuron (neuron S) has spikes which form the process $N_3(\cdot)$. Let λ_i equal the rate of $N_i(\cdot)$, for i=1,2,3. Suppose that every spike from S deterministically gives rise to a spike from A and a spike from B ℓ_A and ℓ_B units of time later, where ℓ_A and ℓ_B are fixed constants. Figure 1 gives a diagram depicting this situation. The overall spike trains of neurons A is therefore the superposition of $N_1(\cdot)$ and $N_3(\cdot)$ delayed by ℓ_A, and similarly for neuron B. We therefore have

$$\lambda_A = \lambda_1 + \lambda_3$$

(6)

$$\lambda_B = \lambda_2 + \lambda_3 \ .$$

If in addition $N_3(\cdot)$ is a Poisson process, a simple calculation gives

(7)
$$\lambda_{A|B}(u) = \begin{cases} \lambda_A & \text{if} \quad u \neq \ell_A - \ell_B \\ \\ \infty & \text{if} \quad u = \ell_A - \ell_B \end{cases}$$

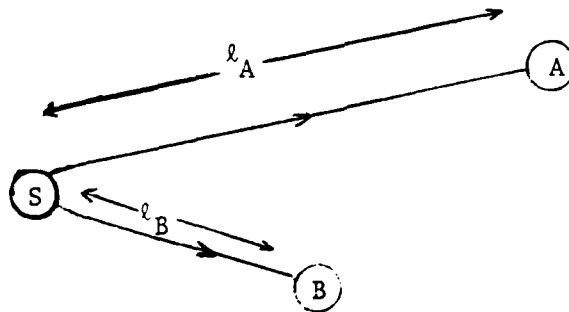Figure 1: Neuron S excites A and B with latencies ℓ_A and ℓ_B, respectively.

When $\lambda_{A|B}(\cdot)$ is not constant, the processes $N_A(\cdot)$ and $N_B(\cdot)$ are dependent. It should be stressed that if this is the case, we cannot infer causality: neither A nor B need be affecting the other directly. Instead they may both be affected by a third source, as Example 2 illustrates.

6

Since $\lambda_{A|B}(\cdot)$ can indicate deviations from independence, it is important to be able to estimate it from the data. Cox (1965), Cox and Lewis (1972) and Brillinger (1976) have proposed estimates of the function $\lambda_{A|B}(\cdot)$, or what is essentially equivalent, $\lambda_B \cdot \lambda_{A|B}(\cdot)$. Their estimate is formed as follows: From the processes $A_1 < A_2 < \cdots < A_{N_A}$, $B_1 < B_2 < \cdots < B_{N_B}$, compute all the differences $A_i - B_j$, for $1 \le i \le N_A$, $i \le j \le N_B$. Then form a histogram (or more generally any density estimate) of these $N_A \cdot N_B$ points. This histogram is called a cross-correlation histogram (CCH).

Brillinger (1976) showed that if the processes $\{A_i\}$ and $\{B_j\}$ are observed over a period of length T, then as $T \to \infty$ very roughly, the suitably normalized CCH resembles the function $\lambda_{A|B}(\cdot)$. In a little more detail, suppose that the bin width used to form the CCH is b. His result is that under some regularity conditions, if $T \to \infty$ and $b \to 0$ in such a way that bT remains constant, then the height of the normalized CCH at a fixed point u has mean $\lambda_{A|B}(u)$ and variance that is of the order of $\frac{1}{bT}$. His result applies to individual points u, and does not imply that the normalized CCH as a whole resembles $\lambda_{A|B}(\cdot)$. Furthermore, the variance of the height of the normalized CCH is of a larger order of magnitude than $\frac{1}{T}$. Nevertheless, his result aids greatly in understanding the way in which the normalized CCH estimates $\lambda_{A|B}(\cdot)$.

We now illustrate the use of the CCH's in the situations given by the two examples above.

Example A: Two independent neurons.

This is the situation described in Example 1. The parameters are

$$\lambda_1 = \lambda_2 = 100 \ ,$$

and

$$T = 1$$

(recall that T is the length of the period of observation). Two features of the CCH are apparent. First, as expected, the CCH is generally flat, or at least there are no apparent peaks or valleys. Second, although the heights of the different bins in the CCH all have the same distribution, they clearly display a lot of fluctuation, in accordance with the remark made concerning the variance.

Example B: Two neurons simultaneously excited by a third source.

This is the situation described in Example 2. The parameters are

$$\lambda_1 = \lambda_2 = \lambda_3 = 100$$

$$T = 1$$

$$\ell_A = .04$$

$$\ell_B = .03 \ .$$

In this situation we have (see equation (7))

$$
(8) \qquad \lambda_{A|B}(u) = \begin{cases} \lambda_A & \text{if } u \neq .01 \\ \\ \infty & \text{if } u = .01 \ . \end{cases}
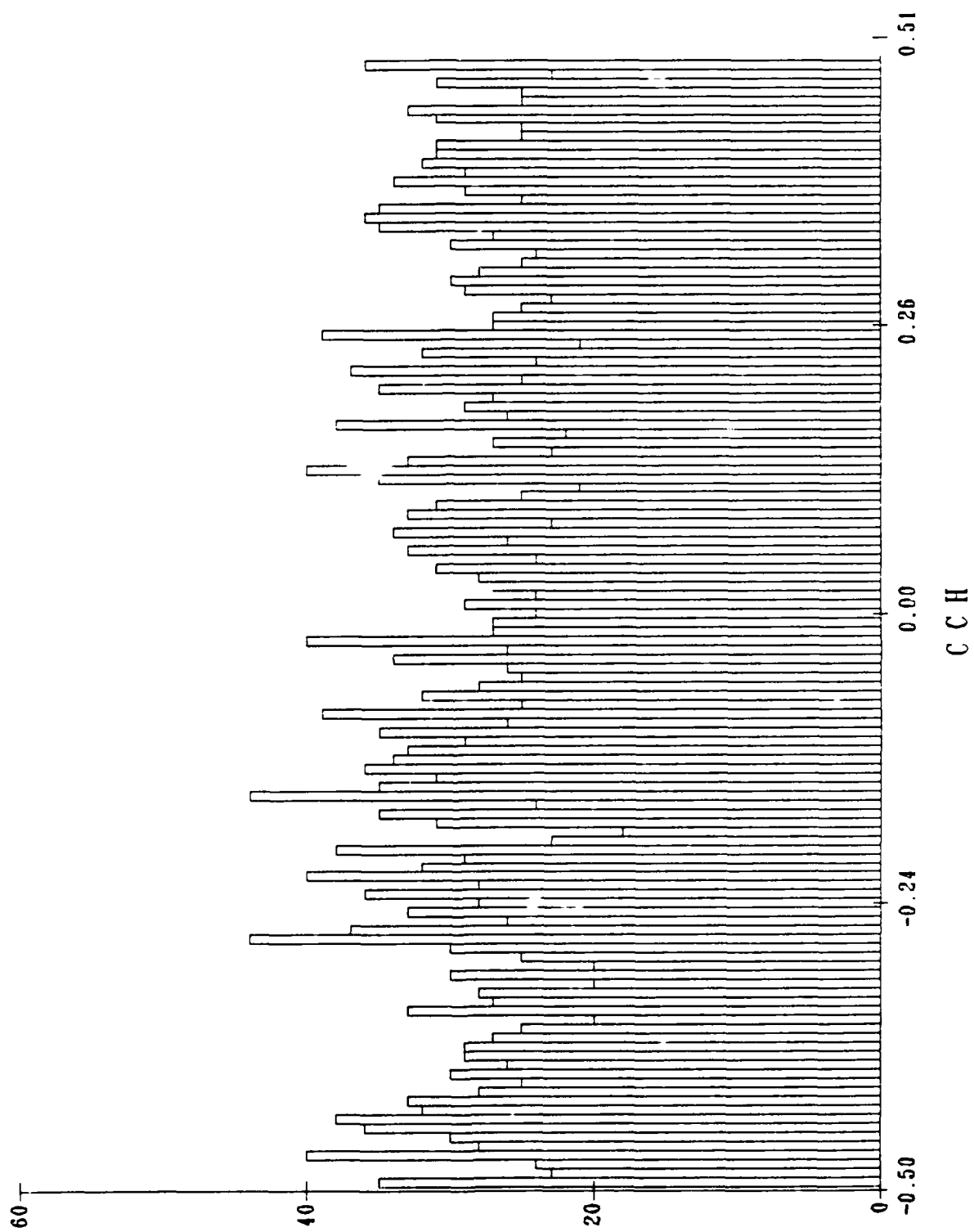$$

Figure 2: CCH for two neurons firing independently.

Figure 3 gives the CCH for this setup. The location of the spike is at $u = .01$, in accordance with (8).

Example C: Two neurons simultaneously excited by a third source.

This is the same as Example B, except that $T = 10$. The function $\lambda_{A|B}(\cdot)$ is still given by (8). Figure 4 gives the CCH for this situation. The spike is much sharper than in Figure 3, due to the fact that T is much larger.
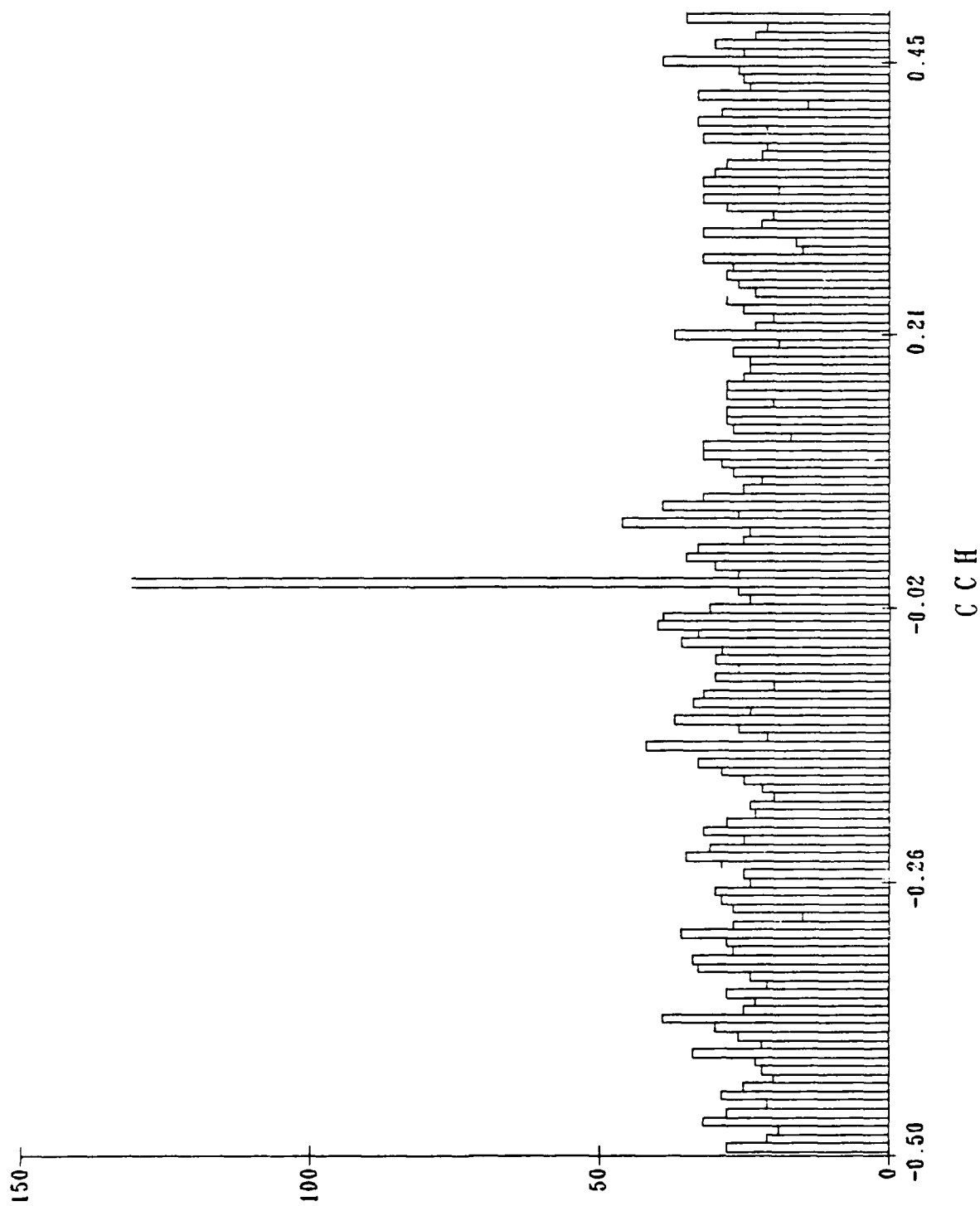
Figure 3: CCH for two neurons simultaneously excited by a third source.
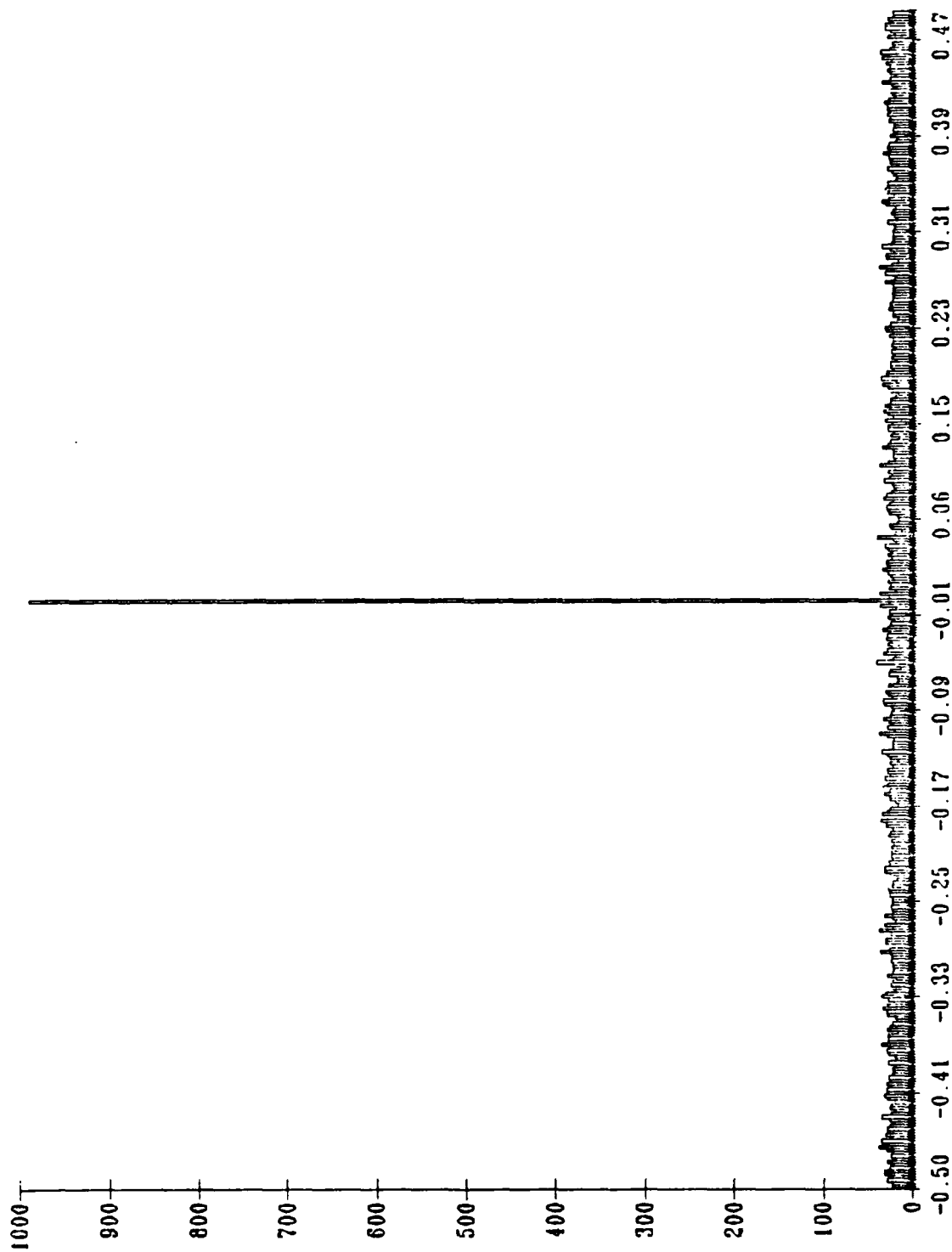
Figure 4: CCH for two neurons simultaneously excited by a third source. The period of observation is longer than in Figure 3.

3.   ESTIMATES AND CONFIDENCE INTERVALS FOR THE NORMALIZED RENEWAL FUNCTION.

In Example 2, the lags ℓ_A and ℓ_B were taken to be fixed constants, and this led to the highly discontinuous form of $\lambda_{A|B}(\cdot)$ given by (8). We saw that the CCH was particularly suited for revealing the dependence structure in this situation. In practice however, the lags will never be fixed constants; this will result in $\lambda_{A|B}(\cdot)$ being a smooth function. The normalized CCH will then not be well suited for estimating it: it is more suitable as a pointwise rather than as a global estimate of $\lambda_{A|B}(\cdot)$.

Example 3: A two-neuron network.

Consider the network described in Example 2, except that $\ell_A \equiv 0$ and ℓ_B is random. More specifically, if $\ell_B^{(i)}$ denotes the lag between impulse i of the stimulus neuron and the induced impulse from neuron B, assume that the $\ell_B^{(i)}$ are iid with a density f. Then it is simple to see that

$$(9) \qquad \lambda_{A|B}(u) = (\lambda_1+\lambda_3) + \frac{\lambda_3}{(\lambda_2+\lambda_3)} \, f(u) \ .$$

In this example, suppose for simplicity that f was the standard uniform distribution on [0,1]. Then, if a CCH was formed it would be likely to have a peak between 0 and 1. If we knew the distribution of the area under the histogram, we would be able to determine whether the peak was significant or rather was due to random fluctuation.

Let $AR(t_1,t_2)$ denote the area under the histogram between the points t_1 and t_2. Here t_1 and t_2 are arbitrary points satisfying $t_1 < t_2$. We may then write

13

$$\text{(10)} \qquad AR(t_1, t_2) = b \sum_{i=1}^{N_B} \sum_{j=1}^{N_A} I\{A_j - B_i \in (t_1, t_2)\} \; ,$$

where   b   is the   bin   width.

Let us now define for $t_1 < t_2$

$$\text{(11)} \qquad U_{A|B}(t_1, t_2) = E\{N_A(t_1, t_2) \,\Big|\, \begin{matrix} \text{There is a B point} \\ \text{at } t = 0 \end{matrix}\} \; .$$

$U_{A|B}(\cdot, \cdot)$ is called the conditional renewal function or renewal function, for short. Note that if $N_A(\cdot)$ and $N_B(\cdot)$ are independent processes, then $U_{A|B}(t_1, t_2) = \lambda_A(t_2 - t_1)$ by Korolyuk's theorem. Defining the normalized renewal function $W(\cdot, \cdot)$ by

$$\text{(12)} \qquad W(t_1, t_2) = \frac{1}{\lambda_A} U_{A|B}(t_1, t_2) \; ,$$

we then have

$$\text{(13)} \qquad W(t_1, t_2) = \begin{cases} = t_2 - t_1 & \text{(independence)} \\ > t_2 - t_1 & \text{(excitation)} \\ < t_2 - t_1 & \text{(inhibition)} \end{cases}$$

with the words "excitation" and "inhibition" suitably referring to the interval (t_1, t_2). Thus, the family of parameters $\{W(t_1, t_2); \, t_1 < t_2\}$ is useful in describing the dependence structure between $N_A(\cdot)$ and $N_B(\cdot)$.

An appropriate way to search for an excitatory or an inhibitory effect is to proceed as follows. Fix some constant $\Delta > 0$, and consider $W(t - \frac{\Delta}{2}, t + \frac{\Delta}{2})$ as a function of t. Here, Δ is determined by the experimenter as the likely duration of an interaction, and is determined by physiological considerations. Under independence this function is constantly equal to Δ, so that deviations from Δ indicate a dependence structure. It is therefore necessary to estimate $W(t - \frac{\Delta}{2}, t + \frac{\Delta}{2})$.

In Doss (1983) it is shown that under some regularity conditions, $AR(t - \frac{\Delta}{2}, t + \frac{\Delta}{2})$, suitably normalized, is asymptotically normal, with mean $W(t - \frac{\Delta}{2}, t + \frac{\Delta}{2})$, and variance of the order of $\frac{1}{T}$. More specifically, the result is that for large N_B,

(i) $AR(t - \frac{\Delta}{2}, t + \frac{\Delta}{2})(\frac{T}{bN_B N_A})$ is approximately $N(W(T - \frac{\Delta}{2}, t + \frac{\Delta}{2}), \frac{\sigma^2(t)}{N_B})$.

Furthermore,

(ii) $\sigma^2(t)$ can be estimated consistently from the data.

The main feature of the result is (ii), which allows confidence intervals to be formed. This makes possible a formal analysis. The entire function $AR(t - \frac{\Delta}{2}, t + \frac{\Delta}{2}) \frac{T}{bN_A N_B}$ can be plotted, and a confidence band can be put around it. Under the hypothesis of independence, the function is essentially flat, at height Δ. Upwards or downwards deviations from Δ (dependence) can therefore be discerned at a glance from the plot.

Example D: Two neurons simultaneously excited by a third source.

This is the situation described in Example 2. The parameters are:

$$\lambda_1 = \lambda_2 = \lambda_3 = 100, \quad t = 3 \ ,$$

$$\ell_A = .1, \quad \ell_B = .05, \quad (\text{thus } \ell_A - \ell_B \text{ is negative})$$

$$\text{and} \quad \Delta = .1 \ .$$

Figure 5 shows the estimate of the normalized renewal function

$$W(t - \frac{\Delta}{2}, \ t + \frac{\Delta}{2}) \quad \text{for} \quad -.5 \le t \le .5 \ .$$

The peak in the diagram clearly shows the dependence structure.

In Figure 5, the estimate of the normalized renewal function appears without the confidence band. Also, the diagram was constructed from data where the lags were fixed and not random. In a later paper we hope to carry out the analysis further by

(i) constructing the confidence bands

(ii) using data generated as in Example 3, i.e., with random lags, and

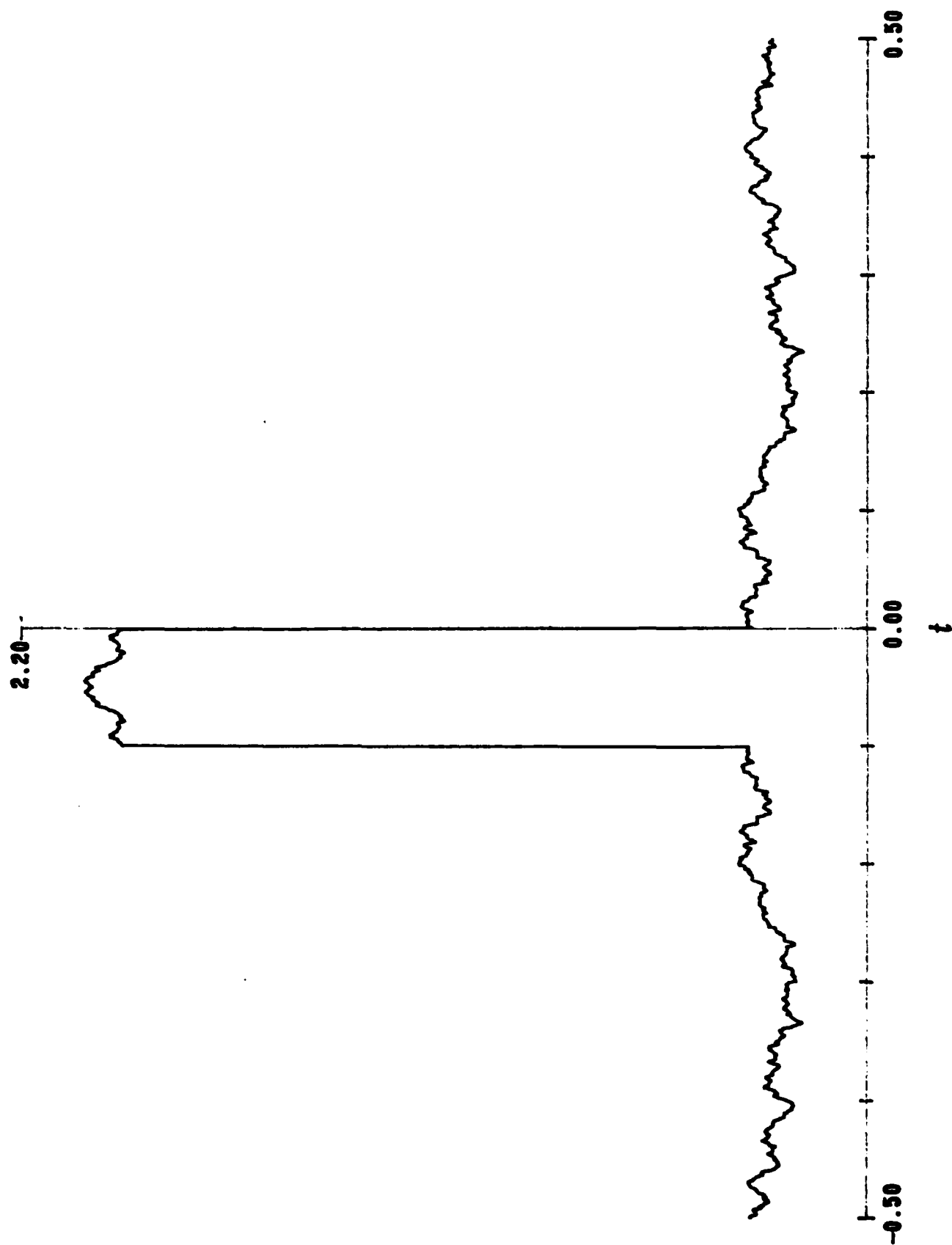(iii) illustrate all the procedures discussed here on real data as well.

Figure 5: Estimate of the normalized renewal function for two neurons simultaneously excited by a third source.

17

REFERENCES

Brillinger, David R. (1976). Estimation of the second-order intensities of a bivariate stationary point process. J. Roy. Statist. Soc. Series B, 38, 60-66.

Cox, David R. (1965). On the estimation of the intensity function of a stationary point process. J. Roy. Statist. Soc. Series B, 27, 332-337.

Cox, David R. and Isham, Valerie (1980). Point Processes. Chapman and Hall, London.

Cox, David R. and Lewis, P.A.W. (1972). Multivariate Point Processes. Proc. 6th Berkeley Symp., 3, 401-448.

Daley, D.J. and Vere-Jones, D. (1972). A summary of the theory of point processes, in Stochastic Point Processes, ed. P.A.W. Lewis, pp. 299-383. New York, Wiley.

Doss, Hani (1983). On estimating the dependence between two point processes. Stanford University Technical Report. (submitted).

Griffith, J.S. and Horn, G. (1963). Functional coupling between cells in the visual cortex of the unrestrained cat. Nature, Lond. 199, 876; 893-895.

Khintchine, A. Ya. (1960). Mathematical Methods in the Theory of Queueing. London: Griffin.

Leadbetter, M.R. (1968). On three basic results in the theory of stationary point processes. Proc. Amer. Math. Soc., 19, 115-117.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 426	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Some Graphical Methods For Assessing The Dependence Structure Between Neuronal Spike Trains		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Hani Doss and Joseph Marhoul		8. CONTRACT OR GRANT NUMBER(s) N00014-89-J-1627
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 1111		12. REPORT DATE February 13, 1990
		13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.		
17. DISTRIBUTION STATEMENT (of the abstract entered In Block 20, If different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Graphical methods, Neurophysiology, Reliability theory.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) PLEASE SEE FOLLOWING PAGE.		

TECHNICAL REPORT NO. 426

20. ABSTRACT

Suppose that two (possibly dependent) point processes are observed simultaneously over a period of time, yielding observations at times $A_1 < A_2 < \cdots < A_{N_A}$ for the first process, and at times $B_1 < B_2 < \cdots B_{N_B}$ for the second. Such data arises in many contexts, and it is often of interest to discover and quantify the association between the two processes. Two fields in which this situation occurs are neurophysiology and reliability theory. In this article, we describe and discuss certain graphs, plots, as well as more formal methods that can assess the dependence between point processes. Specifically, these methods indicate whether or not the likelihood of an A-point is increased (decreased) after the occurrence of a B-point. The techniques are illustrated on simulated data. Although bivariate point processes arise in many fields, we emphasize the applications in neurophysiology. Keywords: Reliability Theory. (Fw)