SRC TR 89-36

Identification of Connectivity in Neural Networks

By

X. Yang & S. A. Shamma

IDENTIFICATION OF CONNECTIVITY IN NEURAL NETWORKS

Xiaowei Yang and Shihab A. Shamma

Electrical Engineering Department, Systems Research Center,

and the University of Maryland Institute for Advanced Computer Studies

University of Maryland, College Park, Maryland 20742, USA

Abstract. *Analytical and experimental methods are provided for estimating synaptic connectivities from simultaneous recordings of multiple neurons. The results are based on detailed, yet flexible neuron models in which spike trains are modeled as general doubly stochastic point processes. The expressions derived can be used with non-stationary or stationary records, and can be readily extended from pair-wise to multi-neuron estimates. Furthermore, we show analytically how the estimates are improved as more neurons are sampled, and derive the appropriate normalizations to eliminate stimulus-related correlations. Finally, we illustrate the use and interpretation of the analytical expressions on simulated spike trains and neural networks, and give explicit confidence measures on the estimates.*

1 Introduction

Most functions of the mammalian nervous system are performed by networks of highly interconnected neurons. In the experimental study of these networks, extracellular recording are often employed to sample the patterns of action potentials simultaneously generated by several neurons (Abeles and Goldstein 1977; Gerstein and Perkel 1969; Yang and Shamma 1988). The correlations among the recorded firings of the different cells are then used as measures of the type and strength of their interconnections. Many such measures have been proposed to accomplish the latter task; they include the cross-interval histograms, the cross-correlation histograms, the cross-covariance histogram, and the joint peri stimulus time (PST) histogram (the scatter diagram) (Gerstein 1970; Gerstein and Perkel 1969). In all cases, the histograms provide statistical measures in support of various hypotheses such as whether the two (or more) neurons under study directly influence each other or simply share common inputs, and whether the influences are excitatory or inhibitory.

There are three basic difficulties with these methods that we tackle in this report. The first concerns the lack of flexible general analytical treatments that outline the relations between the synaptic connectivities and the correlation measures that are used to estimate them. Thus, while various features in the above mentioned histograms may reflect qualitatively the underlying connections, several parameters and conditions can render these measures inadequate. Examples of such difficulties are the differing integrating dynamics of different cell types, and the potentially severe errors due to stimulus-induced (rather than synaptic) correlations. Attempts to overcome these problems, as in the use of the *shuffling method* to reduce stimulus effects, are shown here to be largely inadequate.

The second basic shortcoming of the above correlation methods stems from the nonstationarity of the neural records. In constructing cross-interval and cross-

correlation histograms, counts are usually obtained not only by averaging over different stimulus presentation but also by averaging over the time duration of each presentation period. This makes these two estimates inadequate when working with non-stationary records and, instead, measures based on time-dependent histograms such as the joint PST scatter diagram should be used for the analysis (Habib and Sen 1985; Stokkum et al. 1986).

Finally, it is unclear in many existing methods how to extend the analysis to more than two neurons, and how to evaluate the degree to which a pair-wise estimate is improved when the records from many other neurons are included. This is a particularly important criterion as progress in multiunit recording technologies promises to increase significantly the number of records of simultaneously active neurons.

To summarize, the objectives of this report are (1) to provide rigorous analytical and experimental methods to estimate synaptic connectivities from simultaneous recordings of multiple neurons that are based on accurate and flexible neuron models; (2) to express synaptic connectivity in terms of probability densities of joint neuronal firings and individual neuronal firings that can be used with non-stationary (or stationary) records; (3) to extend these methods from pair-wise to multiunit correlations.

The paper is organized as follows. In the next section, a stochastic nonlinear neuron model is proposed, and the spike train generated by the model is expressed by a doublely stochastic process. This model will serve as the fundamental tool upon which the analytical results are based. In section 3, quantitative analyses of neuronal connectivities are carried out through the model. These include derivations of the relations between the synaptic connectivity and the firing probability densities, and extending the pair-wise correlations to the multi-neuron case. In section 4, the results are summarized and discussed in the context of practical implementations

and considerations of the accuracy of the estimates. Finally, the analytical results are simulated and discussed in section 5. The proofs of lemmas and theorems are given in the Appendix.

All the analytical treatments are contained within sections 2 and 3. For the reader interested only in using the final expressions, section 4 outlines the results and is sufficient as a guide for their experimental applications.

2 The Neuron Model

The basic unit of the nervous system which receives and transmits neural signals is the neuron. The interactions of neurons in a network occur in most cases through synaptic connections between them. Most synapses are found between the axon terminals of a presynaptic neuron and the soma or dentritic tree of a post-synaptic neuron. Since there can be many synapses between any two neurons, it is impractical in modeling the neural network to account for individual synapses; rather, it is more fruitful both for experimental investigation and mathematical description to consider the *total effective influence* of one cell on another.

A synaptic connection from a presynaptic neuron B to post-synaptic neuron A is said to be excitatory (inhibitory) if the firing rate of neuron A increases (decreases) when neuron B fires. For the purposes of the model, we assume that the post-synaptic potentials due to many presynaptic inputs are continuously integrated to produce a post-synaptic *membrane potential.* A neuron fires an action potential when its membrane potential exceeds a *threshold* level. After each action potential, there is a period during which the probability of firing is reduced. This period is divided into two intervals: The first is the *absolute refractory period*, in which the neuron cannot fire again; the second is the *relative refractory period* where the neuron may

generate a spike only when the stimulus is fairly strong.

Since the action potentials of a given neuron are similar in shape, we assume that the transmitted information is carried only through the temporal patterns of the spike trains, and hence we use a sequence of impulses to abstract a train of action potentials. Because the instantaneous firings of a neuron are not deterministic, a stochastic point process is adopted to model the firings (Perkel et al. 1967). All the stochastic processes and random variables to be discussed are defined on some probability space $(\Omega, \mathcal{F}, P)$. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $\{\mathcal{F}_t : t \geq 0\}$ be a non-decreasing family of sub σ-field of $\mathcal{F}$ (i.e., $\mathcal{F}_s \subset \mathcal{F}_t$, for every $s < t$). The family $\{\mathcal{F}_t\}$ is called a *history*, and $\mathcal{F}_t$ represents the information collected during $[0, t]$. Let $\{V_t\}_{t\geq 0}$ be a stochastic process (representing the semi-membrane potential process) defined on $(\Omega, \mathcal{F}, P)$. The family of the sub σ-field generated by $\{V_t\}_{t\geq 0}$, $\mathcal{H}_t = \sigma(V_s : 0 \leq s \leq t)$, is called a history of $\{V_t\}_{t\geq 0}$ if $\mathcal{H}_t \subset \mathcal{F}_t$, for $t \geq 0$. And in this case, $\{V_t\}_{t\geq 0}$ is said to be *adapted* to $\{\mathcal{F}_t\}$. Let $\mathcal{R}^+$ be the σ-field generated by set $[0, \infty)$. A function f from $(\Omega, \mathcal{H}_t)$ into $([0, \infty), \mathcal{R}^+)$ is measurable if for every $S \in \mathcal{R}^+$, $f^{-1}(S) \in \mathcal{H}_t$.

Consider that neuron A is influenced by a family of neurons B_i, $i = 1, 2, \cdots, n$. The model we use is depicted in Fig. 1; it is similar in many respects to that studied by Knox (1974) and by Van Den Boogaared et al. (1986). A sequence of impulses from neuron B_i is transformed into a membrane potential in neuron A. If the integrated membrane potential exceeds a threshold value $\theta(t)$, an impulse (spike) is generated while the membrane potential discharges to a resting level v_o. $h_i(t, s)$ is the impulse response (not necessarily time-invariant) which describes the total temporal influence of neuron B_i on neuron A from past up to present, including the conduction and transmission delay. A synaptic connection is said to be excitatory if $h(t, s) \geq 0$ for all t, all s in $\mathbf{R}$; it is said to be inhibitory if $h(t, s) \leq 0$ for all t, all s

in **R**.

The somatic (membrane) potential (W_t^A) of neuron A is represented by a linear spatial-temporal superposition of all input action potentials of neurons $B_1, B_2, \cdots, B_n$ (including self-inhibition and/or self-excitation), and an unknown random potential U_t which represents the influence of all other unobservable neurons and biophysical factors. A sigmoid function g is used to map the somatic potential as follows:

$$W_t^A = g(U_t + \sum_i \int_0^t h(t,\tau) dN_{B_i}(\tau)) = g(U_t + \sum_i \sum_{k=1}^{N_{B_i}(t)} h_i(t, T_k^{B_i})) \qquad (1)$$

where $\{T_k^{B_i}\}_{k\geq 1}$ are the epoch times of spike train from neuron B_i, and $\{N_{B_i}(t)\}_{t>0}$ is the associated counting process, i.e., the number of spikes arriving from neuron B_i in the interval $(0,t]$.

For mathematical simplicity, let us assume that the nonlinearity g has the form of $g(x) = \alpha e^x$, $\alpha > 0$, i.e., that neuron A is operating around threshold and is thus not strongly driven. Suppose further, without loss of generality, that we are interested in finding the connectivity between two neurons (A and B_1). Then we write

$$W_t^A = \frac{1}{\alpha}\, g(U_t + \sum_{i\neq 1} \sum_{k=1}^{N_{B_i}(t)} h(t, T_k^{B_i}))\, V_t^A \qquad (2)$$

where V_t^A is called here the *semi-membrane potential* and is defined as:

$$V_t^A = g(\sum_{k=1}^{N_{B_1}(t)} h(t, T_k^{B_1})). \qquad (3)$$

In order to account for the firings of neuron A that are due to V_t^A, we can think of the factor $\frac{1}{\alpha} g(U_t + \sum_{i\neq 1} \sum_{k=1}^{N_{B_i}(t)} h(t, T_k^{B_i}))$ as a continuous random variable Z, such

that a random threshold $\theta(t)$ is formed, which is defined as:

$$\theta(t) = Z\theta_o(t) \tag{4}$$

where $\theta_o(t)$ is a time function. Due to the refractory period r during which a neuron is unable to produce a successive spike, the time function can be taken as simple as

$$\theta_o(t) = \begin{cases} \infty, & T_k \leq t < T_k + r \\ \theta_o, & T_k + r \leq t < T_{k+1} \end{cases} \tag{5}$$

where $\theta_o > 0$ is a constant, T_k and T_{k+1} are the times at which the k-th and $(k+1)$-st spike occur, respectively.

A spike occurs whenever the threshold is exceeded by the accumulated semi-membrane potential, i.e.,

$$\int_{t_0}^{t} V_\tau^A d\tau > \theta(t) \tag{6}$$

where t_0 is the instant of the preceding spike. Denoting by $N_A(t)$ the number of spikes in train A during time interval $(0, t]$, a stochastic counting process $\{N_A(t)\}_{t\geq 0}$ is associated with spike train A with $N_A(0) = 0$. Let $\Delta N_A(t) = N_A(t + \Delta t) - N_A(t)$ be the number of spikes in an infinitesimal duration Δt. We say that a process is orderly if $P_r(\Delta N_A(T) > 1) = o(\Delta t)$.

Armed with this general neuron model, we are ready for the analysis of the interneuronal connectivities deduced from the stochastic firing of several neurons.

3 Analytical Results

In this section, we shall derive and elaborate on three basic results. We will first consider the simpler case of two observable neurons, and show how the connectivity

between them can be expressed analytically in terms of the neuron model outlined above. We then consider the sources of uncertainty in this estimate and how they can be reduced through added information from neighboring neurons. Finally, we will comment on the critical normalization procedures used to remove the confounding effects of stimulus artifacts. In the following discussion, assume that $A_t = \{$ neuron A firing at instant $t\}$, and $B_s = \{$ neuron B firing at instant $s\}$. The three basic results derived are:

Result 1. The joint firing probability of a pair of presynaptic and post-synaptic neurons can be expressed as the product of individual firing probabilities and the pairwise connectivity, and a corrupting (uncertainty) factor due to other unobservable influences on the firing of A:

$$P_r(A_t, B_s) = P_r(A_t)P_r(B_s)\gamma(t,s)e^{h(t,s)} \tag{7}$$

where $\gamma(t,s)$ is the corrupting factor ($\gamma \geq 0$) given by:

$$\gamma(t,s) = \frac{E[f_A(t,\theta_t)P_B(s)]}{E[f_A(t,\theta_t)]\,E[P_B(V_s^B)]}. \tag{8}$$

where $f_A(t,\theta_t)$ is measurable with respect to $\sigma(V_\tau^A, N_A(\tau); \tau < t)$ and $P_B(s)$ is measurable with respect to $\sigma(V_\tau^B, N_B(\tau); \tau < s)$.

Result 2. The uncertainty can be reduced (i.e., the corrupting factor made closer to 1) if more interacting neurons are observed simultaneously. The pairwise connectivity becomes:

$$h(t,s) = \log \frac{P_r(A_t, B_s, C_t)P_r(C_t)}{P_r(A_t, C_t)P_r(B_s, C_t)} - \log \gamma^* \tag{9}$$

with $C_t = \{\text{neurons } C_1, C_2, \cdots, C_m \text{ firing at instant } t\}$, where

$$\gamma^*(t,s) = \frac{E[f_A(t,\theta_t)P_B(s)|C_t]}{E[f_A(t,\theta_t)|C_t]\ E[P_B(s)|C_t]} \tag{10}$$

is a quantity satisfying $|\gamma^* - 1| \leq |\gamma - 1|$. If γ^* is very close to 1, then $\log \gamma^*$ may be negligible.

Result 3. In order to minimize the effects of the stimulus on the estimators of the connectivity, the normalized joint firing probability given by :

$$N_p(t,s) = P_r(A_t, B_s)/P_r(A_t)P_r(B_s) \tag{11}$$

leads to estimators superior to those produced by the often employed *shuffle* method:

$$N_s(t,s) = P_r(A_t, B_s) - P_r(A_t)P_r(B_s), \tag{12}$$

which is the quantity that the cross-covariance histogram estimates.

3.1 Definitions and Further Relationships

In order to discuss the derivation of the above stated results, we will need to utilize a few more definitions and relationships. In particular, we will make use of the PST histogram of a single cell spike train which measures the firing rate of a neuron with respect to the stimulus onset. Each bin of the PST histogram is an unbiased estimator for the probability density of the average neuron firing over a short period Δt at instant t corresponding to that bin. The firing probability density of a neuron (A) is defined as

$$E[P_A(t)] \ = \ \lim_{\Delta t \to 0} \frac{P_r(\Delta N_A(t) = 1)}{\Delta t} \tag{13}$$

where $P_A(t)$ is the conditional firing probability density of the *post-synaptic* neuron given the history of the intensity process of the presynaptic neuron and the history

of spike train A denoted by $\mathcal{N}_t^A = \sigma(N_A(s) : s \leq t)$, that is,

$$P_A(t) = \lim_{\Delta t \to 0} \frac{P_r(\Delta N_A(t) = 1 | \mathcal{H}_t^B, \mathcal{N}_t^A)}{\Delta t}. \tag{14}$$

Therefore, the PST histogram of a neuron A estimates $E[P_A(t)]\Delta t$.

Likewise, we have

$$E[P_B(s)] = \lim_{\Delta t \to 0} \frac{P_r(\Delta N_B(s) = 1)}{\Delta s} \tag{15}$$

where $P_B(s)$ is the conditional firing probability density of the *presynaptic* neuron given the history of the intensity process of the presynaptic neuron and the history of spike train B, that is,

$$P_B(s) = \lim_{\Delta s \to 0} \frac{P_r(\Delta N_B(s) = 1 | \mathcal{H}_s^B, \mathcal{N}_s^B)}{\Delta s}. \tag{16}$$

Therefore, the individual PST histograms of neurons A and B estimate $E[P_A(t)]\Delta t$ and $E[P_B(s)]\Delta s$, respectively.

Moreover, the joint PST histogram of the two neurons estimates $E[P_{AB}(t,s)]\Delta t \Delta s$ where

$$E[P_{AB}(t,s)] = \lim_{\Delta t, \Delta s \to 0} \frac{P_r(\Delta N_A(t) = 1, \Delta N_B(s) = 1)}{\Delta t \ \Delta s} \tag{17}$$

and

$$P_{AB}(t,s) = \lim_{\Delta t, \Delta s \to 0} \frac{P_r(\Delta N_A(t) = 1, \Delta N_B(s) = 1 | \mathcal{H}^B_{\max(t,s)}, \mathcal{N}_t^A, \mathcal{N}_s^B)}{\Delta t \ \Delta s}. \tag{18}$$

Given a pair of interacting neurons (A and B), the following lemmas will play an important role in the analysis below.

Lemma 1. $P_B(t)$ can be expressed as a map from the semi-membrane potential space of neuron B onto $[0, \infty)$,

$$P_B(t) = \frac{f_B(x_t)}{1 - F_B(x_t)} V_t^B \tag{19}$$

where $x_t = \int_{t_0}^{t} V_\tau^B d\tau$, and $f_B(\cdot)$, $F_B(\cdot)$ are the density and the distribution functions of the threshold of neuron B, respectively. The function $P_B(\cdot)$ may have a very simple form. For example, if the threshold is an exponentially distributed independent random variable with mean λ, then $P_B(t) = \lambda V_t^B$. And in this case, $\{N_B(t)\}_{t \geq 0}$ is a doubly stochastic Poisson process.

Lemma 2. The conditional expectation of the product of the semi-membrane potential of neuron A and the firing rate of neuron B can be expressed as

$$E[V_t^A \frac{dN_B(s)}{ds}] = e^{h(t,s)} E[V_t^A P_B(s)] \tag{20}$$

The proofs of the lemmas are given in the Appendix. Lemma 1 gives the expression of the conditional firing probability density of the presynaptic neuron given the history of the intensity process of that neuron. Lemma 2 relates the connectivity, the membrane potential of the post-synaptic neuron, and the firing rate of the presynaptic neuron.

3.2 Discussion of Result 1

We will first need to derive an expression for the firing rate of the post-synaptic neuron (A). In general, the threshold θ_t of this neuron is not an independent variable, since it depends on all other unobservable inputs to the neuron. Given an arbitrary value for θ_t, we can write

$$P_r(\Delta N_A(t) = 1 | \theta_t) \simeq$$

$$P_r(\int_t^{t+\Delta t} V_\tau^A d\tau \geq \theta_t|\theta_t) \simeq P_r(V_t^A \geq \frac{\theta_t}{\Delta t}|\theta_t). \tag{21}$$

Since θ_t is assumed to be a positive threshold, by the Markov inequality we have

$$P_r(\Delta N_A(t) = 1|\theta_t) \leq \frac{\Delta t}{\theta_t} E[V_t^A|\theta_t]. \tag{22}$$

Averaging for all possible θ_t and taking limit as Δt goes to zero result

$$P_r(A_t) = P_r(dN_A(t) = 1) \leq E[\frac{1}{\theta_t} V_t^A]dt. \tag{23}$$

In fact, it can be shown (see the Appendix) that there exists a measurable function $f_A(t, \theta_t)$ with respect to $\sigma(V_\tau^A, N_A(\tau); \tau \leq t)$ such that

$$P_r(A_t) = E[f_A(t, \theta_t)]dt. \tag{24}$$

Similarly, the conditional joint firing probability can be expressed as

$$P_r(\Delta N_A(t) = 1, \Delta N_B(s) = 1|\theta_t) \simeq$$

$$P_r(\int_t^{t+\Delta t} V_\tau^A d\tau \geq \theta_t, \Delta N_B(s) = 1|\theta_t) \leq \frac{\Delta t}{\theta_t} E[V_t^A \Delta N_B(s)|\theta_t]. \tag{25}$$

By Lemmas 1 and 2, we therefore have

$$P_r(A_t, B_s) = P_r(dN_A(t) = 1, dN_B(s) = 1) \leq e^{h(t,s)} E[\frac{1}{\theta_t} V_t^A P_B(s)]dtds. \tag{26}$$

As in Eq.(24) above, a more precise expression can be derived as

$$P_r(A_t, B_s) = e^{h(t,s)} E[f_A(t, \theta_t) P_B(s)]dtds. \tag{27}$$

Since the firing probability of the presynaptic neuron is

$$P_r(B_s) = E[P_B(s)]ds, \tag{28}$$

then combining equations (27), (24) and (28) gives Result 1 with

$$\gamma(t,s) = \frac{E[f_A(t,\theta_t)P_B(s)]}{E[f_A(t,\theta_t)]\,E[P_B(s)]}. \tag{29}$$

3.3 Discussion of Result 2

The factor $\gamma(t,s)$ reflects our ignorance of the input to neuron B, or that of the knowledge of the threshold θ_t. For a completely known input $\{V_s^B\}$ (hence $P_B(s)$ is determined), $\gamma(t,s) = 1$; for a completely known threshold, $\gamma(t,s)$ is a constant. When the activity of more neurons are known, the uncertainty in the input and/or threshold decreases, and $\gamma(t,s)$ approaches 1. For instance, if the activities of more interacting neurons $(C_1, C_2, \cdots, C_m)$ are available, we can use a multiunit PST histogram in addition to the conventional joint and individual histograms to estimate

$$\frac{P_r(A_t,B_s,C_t)P_r(C_t)}{P_r(A_t,C_t)P_r(B_s,C_t)} = \frac{P_r(A_t,B_s|C_t)}{P_r(A_t|C_t)P_r(B_s|C_t)} = \gamma^*(t,s)e^{h(t,s)} \tag{30}$$

where $\gamma^*(t,s)$ is defined in Eq.(10).

Note that observing more neurons makes γ^* closer to one than γ is in Eq.(29), and hence the estimator for $h(t,s)$ is more reliable. This is because neurons C_t may contain information about $P_B(s)$ and/or θ_t (for instance, if neurons C_t influence the activity of either or both neurons A and B), hence making $f_A(t,\theta)$ less correlated with $P_B(s)$, and γ^* closer to one.

3.4 Discussion of Result 3

An important factor in correctly interpreting the correlations among the activities of different cells concerns the effects of the stimulus. Specifically, this refers to the fact that unconnected cells may exhibit strong correlations in their firings purely

due to the fact that they are driven by the same stimulus. In order to eliminate these effects, some form of normalization is necessary. In Result 3 we show how the stimulus shuffle alone fails to accomplish this task. In order to illustrate this with explicit analytic expressions, three simplifying assumptions will be adopted concerning the properties of the post-synaptic neuron threshold θ_t (used in Theorem 1 below) and the distribution of the presynaptic potential (used in Theorem 3). We start by stating two of these assumptions and the theorems associated with them, and then proceed to relate the correlation functions explicitly to the inter-neuronal connectivity ($h(t,s)$) in a pair of neurons (A and B).

Assumption 1. The random variable Z of the threshold in Eq.(4) is independent of V_t^A, and has an exponential pdf:

$$p_{\mathbf{Z}}(z) = \begin{cases} \theta_o e^{-(\theta_o z - v_o)}, & v_o/\theta_o \leq z < \infty \\ 0, & z < v_o/\theta_o \, . \end{cases} \tag{31}$$

This assumption is typically valid in cases where Neuron B is only related to neuron A, i.e., it is weakly related to any other neuron. It can be verified that under Assumption 1, the output spike train of the neuron A is a doubly stochastic *Poisson* process $\{N_A(t)\}_{t\geq 0}$ with the intensity process $\{\Lambda_t^A\}_{t\geq 0}$, where

$$\Lambda_t^A = \begin{cases} 0, & T_k \leq t < T_k + r \\ V_t^A, & T_k + r \leq t < T_{k+1} \end{cases} \tag{32}$$

where r is the refractory period.

For a doubly stochastic Poisson process $\{N_A(t)\}$, we have

$$P_r(\Delta N_A(t) = 0 | \mathcal{H}_t^A, \mathcal{N}_t^A) = 1 - \Lambda_t^A \Delta t + o(\Delta t), \tag{33}$$

$$P_r(\Delta N_A(t) = 1 | \mathcal{H}_t^A, \mathcal{N}_t^A) = \Lambda_t^A \Delta t + o(\Delta t), \tag{34}$$

$$P_r(\Delta N_A(t) > 1 | \mathcal{H}_t^A, \mathcal{N}_t^A) \;=\; o(\Delta t) \tag{35}$$

Note that Λ_t^A depends on $\mathcal{N}_t^A$, and hence $\{N_A(t)\}_{t\geq 0}$ is a self-exciting process with the intensity function $E[\Lambda_t^A | \mathcal{N}_t^A]$.

Assumption 2. The refractory period is much smaller than any interspike interval and hence is negligible.

Under this assumption Λ_t^A does not depend on $\mathcal{N}_t^A$, and hence the intensity process becomes the membrane potential process.

Theorem 1. Under Assumptions 1 and 2, $\{N_A(t)\}_{t\geq 0}$ is a doubly stochastic Poisson process with the intensity process $\{V_t^A\}_{t\geq 0}$. Furthermore, the conditional joint firing probability density of neurons A and B can be expressed as

$$P_{AB}(t,s) \;=\; P_A(t) P_B(s) e^{h(t,s)} \tag{36}$$

for all t and all s, where $h(t,s)$ is the inter-neuronal connectivity with non-zero transmission delay.

Theorem 1 states that the joint firing probability density can be expressed as the product of the individual firing densities and the connectivity. Thus the inter-neuronal connectivity $h(t,s)$ can be directly identified by

$$h(t,s) = \log P_{AB}(t,s) - \log P_A(t) - \log P_B(s). \tag{37}$$

This is an ideal case.

Experimently, if only neurons A and B are recorded, the semi-membrane potential of the presynaptic neuron B is generally unknown (hence P_{AB}, P_A, and P_B are unknown), because the membrane potentials are unobservable in extracellular recordings. Therefore, one must instead evaluate the normalized *unconditional* joint probability density $N_p(t,s)$.

Let us recall that the PST histogram of neuron A estimates $E[P_A(t)]\Delta t$ and that of neuron B estimates $E[P_B(t)]\Delta t$, and the joint PST histogram estimates $E[P_{AB}(t,s)]\Delta t\Delta s$. Therefore, $N_p(t,s)$ can be formed by these three histograms:

$$N_p(t,s) = \frac{E[P_{AB}(t,s)]}{E[P_A(t)]E[P_B(s)]}. \tag{38}$$

Lemma 3. Under Assumptions 1 and 2, the conditional firing probability of the post-synaptic neuron can be expressed as

$$P_A(t) = E[V_t^A|\mathcal{H}_t^B] \;=\; \alpha \exp\{\int_0^t (e^{h(t,\tau)} - 1)V_\tau^B d\tau\}. \tag{39}$$

Theorem 2. If Assumptions 1 and 2 hold, the uncertainty γ in Result 1 can be expressed as

$$\gamma(t,s) = \frac{E[V_s^B \exp\{\int_0^t (e^{h(t,\tau)} - 1)V_\tau^B d\tau\}]}{E[V_s^B]\; E[\exp\{\int_0^t (e^{h(t,\tau)} - 1)V_\tau^B d\tau\}]}. \tag{40}$$

Theorem 3. If the semi-membrane potential process of the presynaptic neuron can be decomposed in the form of $V_t^B = Xf(t)$ where X is a random variable, and $f(t)$ is a deterministic time function, then

$$\gamma(t,s) = \frac{M'(\eta_t)}{E[X]\, M(\eta_t)} \tag{41}$$

where $M(\cdot)$ is denoted as the moment generating function of X, and $M'(\eta_t)$ is the first derivative of $M(\cdot)$ with respect to η_t which is expressed by

$$\eta_t = \int_0^t (e^{h(t,\tau)} - 1)f(\tau)d\tau. \tag{42}$$

Furthermore, $\gamma(t,s) \to 1$ when $Var(X) \to 0$.

Assumption 3. The presynaptic membrane potential X is Gamma distributed with parameters (λ, ν).

One consequence of Theorem 3 is that if Assumption 3 holds — a relatively common occurrence (Bishop et al. 1964; Nakahama et al. 1968; Correia and Landolt 1977) — the normalized unconditional joint probability density $N_p(t,s)$ can be explicitly evaluated in terms of these parameters as

$$N_p(t,s) = \frac{\lambda}{\lambda - \eta_t} e^{h(t,s)}, \quad \eta_t < \lambda. \tag{43}$$

Comparing this expression with that of Theorem 2 suggests that $\gamma(t,s) = \lambda/(\lambda - \eta_t)$. Therefore, for a given Gamma distribution (of degree ν), as the variance of X ($= \nu/\lambda^2$) becomes smaller, λ increases, and $\gamma(t,s) \to 1$. In other words, the more is known about V_s^B (e.g., from recordings of additional neurons), the more accurate is the estimate of the connectivity between neurons A and B. We will illustrate these result through simulations later in section 5 (see Fig. 3).

If the membrane potential does not vary much for different stimulus presentation (small variance of X), then $\lambda \gg \eta_t$. Consequently, we have

$$N_p(t,s) \simeq e^{h(t,s)}. \tag{44}$$

This confirms the conclusions established in Result 2 earlier.

In contrast to the normalization used in Eq.(11), the conventional cross-covariance histogram (which is the modified joint PST diagram using the shuffling method) uses a *difference* normalization which estimates (Gerstein 1970; Habib and Sen 1985)

$$N_d(t,s) = E[P_{AB}(t,s)] - E[P_A(t)]\, E[P_B(s)]. \tag{45}$$

In general, this expression is very complicated. However, if we make use of the assumptions in Theorem 3 for the intensity process of the presynaptic neuron (i.e.,

a Gamma distribution), it reduces to

$$N_d(t,s) = \alpha\nu f(s) M(\eta_t)(\frac{e^{h(t,s)}}{\lambda - \eta_t} - \frac{1}{\lambda}). \tag{46}$$

If the membrane potential is not varying too much for different stimulus presentation ($\lambda \gg \eta_t$), then $N_d(t,s)$ can be approximately written as

$$N_d(t,s) \simeq \frac{\alpha\nu f(s)}{\lambda}(e^{h(t,s)} - 1). \tag{47}$$

This expression suggests that identifying the connectivity here is considerably more difficult than that of the normalization $N_p(t,s)$ used earlier, since quantities α, ν, λ and function $f(s)$ are generally unknown. Nevertheless, Eq.(47) suggests that the shuffling method remains effective in indicating *the absence of a direct connection* (i.e., when $h(t,s)$ is very small), since in that case $N_d(t,s)$ is approximately zero regardless of the confounding terms (α, ν, λ and function $f(s)$). We will illustrate this in simulations in section 5 (see Figs. 4).

4 Experimental Considerations

In the analysis of multi-neuronal connectivities, spike trains from several neurons are recorded in response to the repeated presentation (e.g., R times) of a stimulus. Spikes are usually sampled and parsed into (i.e., labeled by) small time bins, using the onset of the stimulus as the initial bin. The bin width Δt is always chosen to be so small that at most one spike may occur in each bin (which corresponds to the orderliness of the point process). Thus each spike train is converted into a discrete 0-1 process, and is further segmented into R segments, each for one stimulus presentation.

Let A_{rn} be the time bin corresponding to the n-th bin associated with the r-th stimulus presentation. A spike train can then be represented by a $R \times N$ random matrix **A** with elements $(A_{rn},\ r = 1, 2, \cdots, R;\ n = 1, 2, \cdots, N)$ — called here a *spike matrix*. Let us assume that the firing activity during each stimulus presentation is statistically independent. Therefore, each element is a random variable taking values $\{0, 1\}$, and the elements in the same column are independent and identically distributed.

The PST histogram $(H_n,\ n = 1, 2, \cdots, N)$ reflects the stimulus-locked firing rate of each single neuron, and it is formed by taking average over every column of the spike matrix,

$$H_n^A = \frac{1}{R}\sum_{r=1}^{R} A_{rn},\ \ n = 1, 2, \cdots, N. \tag{48}$$

The value of H_n^A counts the average spikes over R stimulus presentations in the n-th bin in a spike train A.

The joint PST scatter diagram of two neurons A and B $(H_{mn}^{AB},\ m = 1, 2 \cdots, N;\ n = 1, 2, \cdots, N)$ measures the coincidence spikes in train A and in train B relative to stimulus onset. It is a two-dimensional histogram with one axis (m) for train A and the other axis (n) for train B, and hence it is an N square matrix **H**. Element H_{mn}^{AB} represents the average count for coincidence of a spike in the m-th bin of train A and a spike in the n-th bin of train B over R stimulus presentations, that is,

$$H_{mn}^{AB} = \frac{1}{R}\sum_{r=1}^{R} A_{rm}B_{rn},\ m = 1, 2 \cdots, N;\ n = 1, 2, \cdots, N \tag{49}$$

where A_{rm} and B_{rn} are the elements of spike matrices for trains A and B, respectively. Therefore, the matrix presentation of the joint PST scatter diagram is

$$\mathbf{H} = \frac{1}{R}\mathbf{A}^T\mathbf{B} \tag{50}$$

where T denotes transposition. The expanded joint PST histogram for multiunit recordings (of M neurons) is then

$$H^{C_1C_2\cdots C_M}_{n_1n_2\cdots n_M} = \frac{1}{R}\sum_{r=1}^{R} C^1_{rn_1}C^2_{rn_2}\cdots C^M_{rn_M},\ n_i = 1,2\cdots,N;\ i = 1,2,\cdots,M \qquad (51)$$

where $C^i_{rn_i}$ is the element of the spike matrix for the i-th neuron.

4.1 Using the Scatter Plot to determine neuronal connectivities

The correlations between a pair of recorded neurons (A and B) can be computed from the experimental estimate of the expression of Result 1, i.e.,

$$h(t,s) = \log(\frac{N_p(t,s)}{\gamma(t,s)}) = \log(\frac{E[P_{AB}(t,s)]}{E[P_A(t)]\ E[P_B(s)]}) - \log(\gamma(t,s))$$

where $E[P_A(t)]$ and $E[P_B(s)]$ represent the PST histograms of firings of the neuron pair, $E[P_{AB}(t,s)]$ is their scatter plot, and γ (≥ 0) is the corrupting factor representing the uncertainty in the estimate due to the influences of other unobserved neurons and biophysical factors. Thus in terms of bin numbers m and n, the above equation can be written as

$$\log(\frac{H^{AB}_{mn}}{H^A_m\ H^B_n}) = h(m\Delta t, n\Delta t) + \log(\gamma(m\Delta t, n\Delta t)). \qquad (52)$$

In the case of time invariant connectivities, $h(t,s)$ becomes $h(t-s)$, and the correlation peak becomes a band that runs parallel to the principal diagonal ($t-s=0$).[1]

In the practical application of Eq.(52), the confounding $\gamma(t,s)$ contributions are not known. However, the analysis shows that additional simultaneous recordings can

[1]Note that one can detect further correlations in the *unnormalized* scatter plot, such as the more diffuse bands of time-invariant common inputs (Yang 1989). Of course, these features are intentionally removed by the normalization since they do not reflect direct connectivities between the neuron pair.

be used to reduce these uncertainties. Therefore, by using the additional data, the improved estimator for $h(t,s)$ becomes

$$\log(\frac{H^{ABC_3\cdots C_M}_{mnm\cdots m}\ H^{C_3\cdots C_M}_{m\cdots m}}{H^{AC_3\cdots C_M}_{mm\cdots m}\ H^{BC_3\cdots C_M}_{nm\cdots m}}) = h(m\Delta t, n\Delta t) + \log(\gamma^*(m\Delta t, n\Delta t)) \qquad (53)$$

where $H^{C_1C_2\cdots C_M}_{n_1n_2\cdots n_M}$ are simply the joint multi-dimensional scatter plots defined in Eq.(51), and the uncertainty factor γ^* ($\leq \gamma$) is defined in Eq. (10). The estimates of Eqs.(52) and (53) are illustrated in network simulations in section 5.

4.2 Establishing confidence measures on the estimates

The histograms are random variables subject to fluctuations. Hence, it is important to determine upper and lower bounds such that we assume a connection between neurons A and B whenever these bounds are surpassed. By the law of large numbers, H^{AB}_{mn} converges to $E[P_{AB}(t,s)]$, so does H^A_m to $E[P_A(t)]$ and H^B_n to $E[P_B(s)]$ almost surely as $R \to \infty$. Therefore, if neurons A and B are independent, by theorems on limiting distributions,

$$\frac{H^{AB}_{mn}}{H^A_m H^B_n} \to 1 \ \ as \ \ R \to \infty \qquad (54)$$

almost surely.

The hypothesis $\mathcal{H}_0$ is that the two neurons are statistically independent, which is supported by

$$E[P_{AB}(t,s)] = E[P_A(t)]\ E[P_B(s)]. \qquad (55)$$

And the alternative hypothesis $\mathcal{H}_1$ is that the two neurons depend, which is described by

$$E[P_{AB}(t,s)] \neq E[P_A(t)]\ E[P_B(s)]. \qquad (56)$$

One expects $H^{AB}_{mn}/H^A_m H^B_n$ to be close to 1 if hypothesis $\mathcal{H}_0$ is true. Conversely, if the amount it deviates from 1 exceeds a bound b, one accepts hypothesis $\mathcal{H}_1$.

Now for a given significance level α, we need to find the bound b satisfying

$$P_r(|\frac{H_{mn}^{AB}}{H_m^A H_n^B} - 1| > b \mid \mathcal{H}_0) = \alpha. \tag{57}$$

The hypothesis testing is stated as the following theorem.

Theorem 4. Let b be a bound which divides a critical region for the hypothesis testing. One announces that there is a dependence between the two observed neurons if

$$|\frac{H_{mn}^{AB}}{H_m^A H_n^B} - 1| > b. \tag{58}$$

For the given significance level α of false announcement of dependence, the bound can be calculated by

$$b = \varepsilon_b \sqrt{\frac{1 - H_m^A H_n^B}{R H_m^A H_n^B}} \tag{59}$$

where the value of ε_b is determined from

$$\Phi(\varepsilon_b) = 1 - \frac{\alpha}{2} \tag{60}$$

and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-x^2/2} dx$.

The function $\Phi(x)$ is usually available as the standard normal distribution table. For example, $\alpha = 0.05$ gives $\varepsilon_b = 1.96$.

The above theorem implies that element $H_{mn}^{AB}/(H_m^A H_n^B)$ of the normalized joint PST diagram has a conditional expectation value 1 and an approximate conditional variance

$$\sigma_{mn}^2 \simeq \frac{1 - H_m^A H_n^B}{R H_m^A H_n^B} \tag{61}$$

given the values of H_m^A and H_n^B under hypothesis $\mathcal{H}_0$. Since H_m^A and H_n^B are usually very small and R is fairly large, this approximation is close to a recent result by Palm et al. (1988) where their conditional variance is

$$\sigma_{mn}^2 = \frac{(1 - H_m^A)(1 - H_n^B)}{(R-1)H_m^A H_n^B} \tag{62}$$

under hypothesis $\mathcal{H}_0$.

The bound dividing the hypothesis regions can be made more useful in neural networks with time invariant connectivities. Let w_m reflect the fluctuation in the normalized joint PST diagrams such that

$$\frac{H_{mn}^{AB}}{H_m^A H_n^B} = \gamma(m\Delta t, n\Delta t)e^{h((m-n)\Delta t)} + w_m \ , \tag{63}$$

and the mean of w_n is zero. Let $k = m - n$. A collapsed version can be generated by averaging over diagonals of the normalized joint PST diagram. This collapsed version is a 1-dimensional histogram G_k expressed by

$$G_k = \frac{1}{N - |k|} \sum_{n=\max(1,1-k)}^{\min(N,N-k)} \frac{H_{n+k,n}^{AB}}{H_{n+k}^A H_n^B}$$

$$k = -N+1, \cdots, -1, 0, 1, \cdots, N-1 \tag{64}$$

where $k = 0$ is the collapsed point of the principal diagonal.

Since averaging reduces the fluctuations (the average of w_m has a smaller variance), G_k is a better estimator for the time invariant connectivity $h(t,s) = h(t-s)$. This enables us to establish a bound such that

$$P_r(|G_k - 1| > b_k \mid \mathcal{H}_0) = \alpha. \tag{65}$$

Theorem 5. Given a significance level α, let b_k be a bound of critical region satisfying the above equation, then b_k may be approximately written as

$$b_k \simeq \frac{\sqrt{\sum_{n=\max(1,1-k)}^{\min(N,N-k)} \sigma_{n+k,n}^2}}{N-|k|} \varepsilon_b \tag{66}$$

where ε_b is the the same as in Theorem 4, and σ_{mn}^2 is given by (61). Furthermore, b_k will reduce to

$$b_k \simeq \frac{b}{\sqrt{N-|k|}} \tag{67}$$

when σ_{mn}^2's are taken as constants.

This theorem indicates that the critical region is enlarged (the bound value decreases) when the collapsed version of the normalized joint PST histogram is used.

5 Simulations and Discussion

In order to illustrate the nature of the estimates, uncertainties, and bounds derived earlier, we show the results from simulations of networks of excitatory and inhibitory neurons. The neuron model used for the simulations is depicted in Fig. 1(c) where the nonlinearity $g(x) = \alpha e^x$ and the random threshold has an exponential distribution with mean 1.

In the first case (Fig. 2), pair-wise excitatory and inhibitory, time-invariant connections are estimated using the normalized scatter plots; the uncertainty factor (γ) is equal to 1. The upper plots show the two-dimensional normalized scatter plots. The correlations appear as bands along the principal diagonal because $h(t,s)$ is time-invariant. Hence, the scatter plot can be collapsed along this axis to produce the lower histograms. Note that time-variations in $h(t,s)$ (e.g., due to post-stimulus

adaptation) do not allow this reduction. Consequently, it should only be performed on the portions of the neural record that display obvious stationary behavior. In both simulations of Fig. 2, the predicted analytical estimates are also plotted for comparison, together with the bound lines for the confidence measures (determined by Theorem 5).

In order to illustrate the effects of the uncertainty factor γ, we examine in Fig. 3 the interactions among three neurons with time-invariant connectivities. Here, neuron A is inhibited by neuron B and excited by neuron C, and neuron C is in turn excited by neuron B. Because of the interactions between B and C, the threshold in neuron A is no longer independent of the firings of B. Thus, if we attempt to identify the connectivity between neurons A and B from pair-wise recordings, the estimates will be contaminated by the γ uncertainty factor. The top curve in Fig. 3 first shows the "target" theoretical connectivity obtained from the multi-recording estimate given by formula (30) with $\gamma^*(t,s) = 1$ (i.e., $e^{h_{AB}(t-s)}$). If neuron C is ignored, the pair-wise estimate of $e^{h_{AB}(t-s)}$ is shown as the middle curve in Fig. 3 (corresponding to formula (52)). The correlation is so distorted that actual inhibition becomes false excitation because of the strong excitatory activity from neuron C. To correct the erroneous correlation, we have to use the information from the third neuron. The tripartite correlation according to formula (53) is displayed at bottom of Fig. 3, which is much closer to the analytical estimate.

Fig. 4 compares the preferred normalization with the difference normalization (shuffle method) under two situations. In the absence of a direct connection, the shuffle method provides accurate indication of the lack of synaptic inputs between the two neurons. However, in the presence of a direct connection, the shuffle method fails to remove completely the stimulus correlations as indicated by the deviation from the analytical results. Instead, the normalization suggested in this paper performs

well in both cases.

In conclusion, the above simulations confirm the proposed theory. The neuron model adopted is quite general because (1) the synaptic connectivity $h(t, s)$ represents a time-varying system; (2) the processes representing spike trains are not necessarily Poisson processes, and (3) the nonlinear function $g(x) = \alpha e^x$ is an approximation of $\alpha e^x / 1 + \alpha e^x$ when $\alpha e^x \ll 1$, meaning that the neuron is operating at low firing rates. Moreover, our analytical Results 1 and 2 do not dependent on any further assumptions. Although the three simplifying assumptions were made in order to see Result 3 more clearly, we did not use these assumptions in the simulations of Fig. 4.

The analysis presented in this paper also points to the following sobering conclusion: For multiunit correlation analysis to play a useful role in establishing the basic circuitry of the nervous system, new technologies have to be developed for *stable, multi-unit* recordings. These requirements stem from the need for extended simultaneous recordings from many cells in order to construct adequate scatter histograms and to minimize inherent uncertainty due to unobserved but related activities. Unfortunately, neither of these requirements are easily met at present, although extensive efforts towards this goal are underway through the use of silicon-based microelectrode arrays (Krüger 1983).

Acknowledgment

This work was supported in part by grants from the Whitaker Foundation and the Naval Research Laboratory.

References

[1] Abeles M, Goldstein MH (1977) Multispike train analysis. Proc. IEEE 65:762–772

[2] Bishop PB, Levick WR, Williams WO (1964) Statistical analysis of the dark discharge of lateral geniculate neurones. J Physiol 170:598–612

[3] Boogaard H van den, Hesselmans G, Johannesma P (1986) System identification based on point processes and correlation densities. I. The nonrefractory neuron model. Math Biosci 80:143–171

[4] Correia MJ, Landolt JP (1977) A point process analysis of the spontaneous activity of anterior semicircular canal units in the anesthetized pigeon. Biol Cybern 27:199–213

[5] Gerstein GL (1970) Functional association of neurons: Detection and interpretation. In: Schmitt FO (ed) The Neurosciences Second Study Program. Rockefeller Univ. Press, New York, pp 648–661

[6] Gerstein GL, Perkel DH (1969) Simultanously recorded trains of action potentials: Analysis and functional interpretation. Science (Washington) 148:828–830

[7] Habib MK, Sen PK (1985) Non-stationary stochastic point-process models in neurophysiology with applications to learning. In: Sen PK (ed) Biostatistics: statistics in biomedical, public health and enviromental sciences. Elsevier/North-Holland, Amsterdam, pp 481–509

[8] Knox CK (1974) Cross-correlation functions for a neuronal model. Biophys J 14:567–582

[9] Krüger J (1983) Simultaneous individual recordings from many cerebral neurons: techniques and results. Rev Physiol Biochem Pharmacol 98:177-233

[10] Larson HJ, Shubert BO (1979) Probabilistic models in engineering sciences. (Vol. II) Random noise, signals, and dynamic systems. John Wiley, New York, pp 591–592

[11] Nakahama H, Suzuki H, Yamamoto H, Aikawa S, Nishioka S (1968) A statistical analysis of spontaneous activity of central single neurons. Physiol Behav 3:745–752

[12] Palm G, Aertsen AMHJ, Gerstein GL (1988) On the significance of correlations among neuronal spike trains. Biol Cybern 59:1–11

[13] Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes, I. The single spike train. Biophys J 7:391–418

[14] Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes, II. Simultaneous spike train. Biophys J 7;419–440

[15] Stokkum IHM van, Johannesma PIM, Eggermont JJ (1986) Representation of time-dependent correlation and recurrence time functions. Biol Cyber 55:17–24

[16] Yang X, Shamma SA (1988) A totally automated system for the detection and classification of neural spikes. IEEE Trans BME-35:806–816

[17] Yang X (1989) Detection and classification of neural signals and identification of neural networks. Ph D Dissertation, Electrical Engineering Department, University of Maryland, College Park

Appendix

Proof of Lemma 1: Since the threshold of neuron B is a continuous random variable, let its probability density function and distribution function be $f_B(x)$ and $F_B(x)$, respectively. Let $x_t = \int_{t_0}^{t} V_\tau^B d\tau$, where t_0 is the occurrence instant of the previous spike. From definition (16) we have

$$P_B(t) = \lim_{\Delta t \to 0} \frac{P_r(x_{t+\Delta t} \geq \theta_t | x_t < \theta_t; \mathcal{H}_t^B, \mathcal{N}_t^B)}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{P_r(x_t < \theta_t \leq x_{t+\Delta t} | \mathcal{H}_t^B, \mathcal{N}_t^B)}{\Delta t \; P_r(x_t < \theta_t | \mathcal{H}_t^B, \mathcal{N}_t^B)} = \frac{V_t^B P_{\theta_t}(x_t)}{1 - F_{\theta_t}(x_t)}. \tag{68}$$

Furthermore, if the threshold is exponentially distributed with mean λ, then $f_B(x_t)/1 - F_B(x_t) = \lambda$, and hence $P_B(t) = \lambda V_t^B$. In this case, $P_B(t)$ does not depend on $\{\mathcal{N}_B(t)\}$, and $\{N_B(t)\}_{t \geq 0}$ evolves without aftereffects.

Proof of Lemma 2: Because $\Delta N_B(t)$ can take values 0 and 1 only, by Eq.(3) we have

$$E[V_t^A \frac{\Delta N_B(s)}{\Delta s} | \mathcal{H}_{t \vee s}^B, \mathcal{N}_s^B]$$

$$= E[\alpha \exp\{ \sum_{k=1}^{N_B(t)} h(t, T_k^B)\} | \Delta N_B(s) = 1; \mathcal{H}_{t \vee s}^B, \mathcal{N}_s^B] \frac{P_r(\Delta N_B(s) = 1 | \mathcal{H}_s^B, \mathcal{N}_s^B)}{\Delta s}. \tag{69}$$

For $t > s$, the conditional expectation in the above equation can be written as

$$E[\alpha \exp\{ \sum_{k=1}^{N_B(t)} h(t, T_k^B)\} | \Delta N_B(s) = 1; \mathcal{H}_t^B, \mathcal{H}_s^B]$$

$$= E[\alpha \exp\{ \sum_{k=1}^{N_B(s)} h(t, T_k^B)\} \exp\{h(t, s + \Delta s)\} \exp\{ \sum_{k=N_B(s+\Delta s)+1}^{N_B(t)} h(t, T_k^B)\} | \mathcal{H}_t^B, \mathcal{N}_s^B], \tag{70}$$

which becomes

$$e^{h(t,s)}E[V_t^A|\mathcal{H}_t^B,\mathcal{N}_s^B] \tag{71}$$

as Δs goes to 0. Since $P_B(s)$ is a measurable function with respect to $\sigma(\mathcal{H}_s^B \times \mathcal{N}_s^B)$, hence

$$E[V_t^A \frac{dN_B(s)}{ds}] = e^{h(t,s)}E[E[V_t^A|\mathcal{H}_t^B,\mathcal{N}_s^B]P_B(s)] = e^{h(t,s)}E[V_t^A P_B(s)]. \tag{72}$$

For $t \leq s$, we have

$$E[V_t^A \frac{\Delta N_B(s)}{\Delta s}|\mathcal{H}_s^B,\mathcal{N}_s^B] = E[V_t^A|\mathcal{H}_s^B,\mathcal{N}_s^B]\ E[\frac{\Delta N_B(s)}{\Delta s}|\mathcal{H}_s^B,\mathcal{N}_s^B]$$

$$= E[V_t^A|\mathcal{H}_s^B,\mathcal{N}_s^B]\ \frac{P_r(\Delta N_B(s)=1|\mathcal{H}_s^B,\mathcal{N}_s^B)}{\Delta s}, \tag{73}$$

hence

$$E[V_t^A \frac{dN_B(s)}{ds}] = E[E[V_t^A|\mathcal{H}_s^B,\mathcal{N}_s^B]P_B(s)] = E[V_t^A P_B(s)]. \tag{74}$$

Since $h(t,s)$ represents a synaptic connectivity, which is a causal system with non-zero transmission delay, $h(t,s) = 0$ for $t \leq s$. Thus Eq.(71) holds for all t and s.

Proof of Equation (24): Replace neuron A for neuron B in the proof of Lemma 1. Here we emphasis the threshold θ_t of neuron A so that $f_{\theta_t}(\cdot)$ and $F_{\theta_t}(\cdot)$ replace $f_B(\cdot)$ and $F_B(\cdot)$, respectively. Thus $\lim_{\Delta t\to 0} P_r(\Delta N_A(t)=1|\mathcal{H}_t^A,\mathcal{N}_t^A)/\Delta t = V_t^A P_{\theta_t}(x_t)/1-F_{\theta_t}(x_t)$, where $x_t = \int_{t_0}^t V_\tau^A d\tau$. Although the joint distribution function of θ_t and x_t is unknown in general, we can define a measurable function

$$f_A(t,\theta_t) = \frac{V_t^A f_{\theta_t}(x_t)}{1-F_{\theta_t}(x_t)}. \tag{75}$$

Therefore, we have

$$P_r(A_t) = E[P_r(dN_A(t)=1|\mathcal{H}_t^A,\mathcal{N}_t^A)] = E[f_A(t,\theta_t)]dt. \tag{76}$$

Proof of Theorem 1: Suppose that Assumptions 1 and 2 hold, and that the threshold θ_t has an exponential distribution with mean 1. By the proof of equation (24) and Lemma 1,

$$\lim_{\Delta t \to 0} P_r(\Delta N_A(t) = 1 | \mathcal{H}_t^A, \mathcal{N}_t^A)/\Delta t = f_A(t, \theta_t) = V_t^A, \tag{77}$$

hence spike train $N_A(t)$ is represented by a doubly stochastic Poisson process. Therefore, by Eqs.(32) and (34) we have

$$P_A(t) = E[V_t^A | \mathcal{H}_t^B]. \tag{78}$$

Because Poisson process evolves without aftereffects, the conditional probability given the firing histories of neurons A and B can be split into

$$P_r(\Delta N_A(t) = 1, \Delta N_B(s) = 1 | \mathcal{H}_t^A, \mathcal{H}_{t \vee s}^B)$$

$$= P_r(\Delta N_A(t) = 1 | \mathcal{H}_t^A)\ P_r(\Delta N_B(s) = 1 | \mathcal{H}_t^A, \mathcal{H}_{t \vee s}^B). \tag{79}$$

By Eqs.(32) and (34), the first factor is

$$P_r(\Delta N_A(t) = 1 | \mathcal{H}_t^A) = V_t^A \Delta t + o(\Delta t). \tag{80}$$

We write the second factor as

$$P_r(\Delta N_B(s) = 1 | \mathcal{H}_t^A, \mathcal{H}_{t \vee s}^B) = E[\Delta N_B(s) | \mathcal{H}_t^A, \mathcal{H}_{t \vee s}^B], \tag{81}$$

and we have

$$P_r(\Delta N_A(t) = 1, \Delta N_B(s) = 1 | \mathcal{H}_t^A, \mathcal{H}_{t \vee s}^B) = E[V_t^A \Delta N_B(s) | \mathcal{H}_t^A, \mathcal{H}_{t \vee s}^B] + o(\Delta t). \tag{82}$$

By taking average over the σ-field $\mathcal{H}_t^A$, we obtain

$$P_{AB}(t, s) = E[V_t^A \frac{dN_B(s)}{ds} | \mathcal{H}_{t \vee s}^B], \tag{83}$$

which is, by the proof of Lemma 2,

$$P_{AB}(t, s) = e^{h(t,s)} E[V_t^A | \mathcal{H}_t^B] P_B(s) = e^{h(t,s)} P_A(t) P_B(s). \tag{84}$$

Proof of Lemma 3: Under Assumptions 1 and 2, $P_A(t)$ becomes the conditional expectation of the post-synaptic membrane potential, which can be expressed as

$$E[V_t^A|\mathcal{H}_t^B] = E[\alpha \exp\{\sum_{k=1}^{N_B(t)} h(t, T_k^B)\}|\mathcal{H}_t^B] \tag{85}$$

Define an event $D_n = \{N_B(t) = n\}$, which has a conditional Poisson distribution

$$P_r(D_n|\mathcal{H}_t^B) = \frac{[\int_0^t V_\tau^B d\tau]^n}{n!} e^{-\int_0^t V_\tau^B d\tau}. \tag{86}$$

Because $\{N_B(t)\}_{t\geq 0}$ is an inhomogeneous Poisson process with the associated point process $\{T_k\}_{k>0}$ for the given realization of the intensity process $\{V_t^B\}_{t\geq 0}$, it can be shown that (Larson and Shubert 1979)

$$E[\exp\{\sum_{k=1}^{N_B(t)} h(t, T_k^B)\}|N_B(t) = n, \mathcal{H}_t^B] = [\frac{\int_0^t e^{h(t,\tau)} V_\tau^B d\tau}{\int_0^t V_\tau^B d\tau}]^n. \tag{87}$$

Therefore, we have

$$E[\exp\{\sum_{k=1}^{N_B(t)} h(t, T_k^B)\}|N_B(t) = n, \mathcal{H}_t^B]$$

$$= \sum_{n=0}^{\infty} E[\exp\{\sum_{k=1}^{N_B(t)} h(t, T_k^B)\}|N_B(t) = n, \mathcal{H}_t^B] P_r(D_n|\mathcal{H}_t^B)$$

$$= \sum_{n=0}^{\infty} \frac{[\int_0^t V_\tau^B d\tau]^n}{n!} e^{-\int_0^t V_\tau^B d\tau}, \tag{88}$$

and the summation in the above equation is a series expression for an exponential function. Thus rewriting it proves the lemma.

Proof of Theorem 2: By Lemma 1, $P_B(s) = V_s^B$; by Theorem 1, $f_A(t, \theta_t) = V_t^A$. We write

$$\gamma(t, s) = \frac{E[V_t^A V_s^B]}{E[V_t^A] E[V_s^B]}. \tag{89}$$

Note that $E[V_t^A V_s^B] = E[E[V_t^A|\mathcal{H}_t^B]V_s^B] = E[P_A(t)V_s^B]$ and $E[V_t^A] = E[E[V_t^A|\mathcal{H}_t^B]] = E[P_A(t)]$. Hence, applying Lemma 3 completes the proof.

Proof of Theorem 3: In the expression (40), the numerator of $\gamma(t,s)$ can be written as $E[Xf(s)e^{X\eta_t}]$; the denominator can be written as $E[Xf(s)]E[e^{X\eta_t}]$. Eq.(41) is true by canceling $f(s)$. Furthermore, as the variance of X decays to zero, $M'(\eta_t) \to \mu e^{\mu\eta_t}$ and $M(\eta_t) \to e^{\mu\eta_t}$, consequently $\gamma(t,s) \to 1$.

Proof of Theorem 4: For a given significance level α, we need to find a bound b satisfying

$$P_r(|\frac{H_{mn}^{AB}}{H_m^A H_n^B} - 1| > b \mid \mathcal{H}_0) = \alpha. \tag{90}$$

Let us remember that RH_{mn}^{AB} is binomially distributed with parameters $(R, E[P_{AB}(t,s)])$ and that $H_{mn}^{AB} \to H_m^A H_n^B$ almost surely under $\mathcal{H}_0$. Hence by the central limiting theorem,

$$\frac{RH_{mn}^{AB} - RH_m^A H_n^B}{\sqrt{RH_m^A H_n^B(1 - H_m^A H_n^B)}} \to N(0,1) \ \ as \ \ R \to \infty \tag{91}$$

where $N(0,1)$ is denoted as a standard Gaussian random variable. This means that Eq.(56) can be approximately written as

$$P_r(|N(0,1)| > \varepsilon_b \mid \mathcal{H}_0) = \alpha \tag{92}$$

where

$$\varepsilon_b = \frac{bRH_m^A H_n^B}{\sqrt{RH_m^A H_n^B(1 - H_m^A H_n^B)}} \tag{93}$$

which results in an expression of the bound as

$$b = \varepsilon_b \sqrt{\frac{1 - H_m^A H_n^B}{RH_m^A H_n^B}}. \tag{94}$$

The value of ε_b is determined by

$$\Phi(\varepsilon_b) = 1 - \frac{\alpha}{2} \tag{95}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-x^2/2} dx$.

The above arguments imply that element $H_{mn}^{AB}/(H_m^A H_n^B)$ of the normalized joint PST diagram has a conditional expectation value 1 and an approximate conditional variance

$$\sigma_{mn}^2 \simeq \frac{1 - H_m^A H_n^B}{R H_m^A H_n^B} \tag{96}$$

under hypothesis $\mathcal{H}_0$.

Proof of Theorem 5: Let us note that under hypothesis $\mathcal{H}_0$, $H_{mn}^{AB}/H_m^A H_n^B$ is approximately Gaussian distributed with mean 1 and variance σ_{mn}^2. Hence

$$|G_k - 1| \simeq |\frac{1}{N - |k|} \sum_{n=\max(1,1-k)}^{\min(N,N-k)} \sigma_{n+k,n}\ N_n(0,1)| \tag{97}$$

where each $N_n(0,1)$ approximately has a standard Gaussian distribution expressed by

$$N_n(0,1) = \frac{R H_{n+k,n}^{AB} - R H_{n+k}^A H_n^B}{\sqrt{R H_{n+k}^A H_n^B (1 - H_{n+k}^A H_n^B)}}, \tag{98}$$

and σ_{mn}^2 is given in Eq.(61). Therefore, $G_k - 1$ is approximately Gaussian distributed with zero-mean and variance

$$Var(G_k - 1) = \frac{1}{(N - |k|)^2} \sum_{n=\max(1,1-k)}^{\min(N,N-k)} \sigma_{n+k,n}^2 \tag{99}$$

where mutual independence of $N_n(0,1)$'s is assumed. Let

$$\varepsilon_b = \frac{b_k}{\sqrt{Var(G_k - 1)}}, \tag{100}$$

we obtain

$$P_r(|G_k - 1| > b_k \mid \mathcal{H}_0) = 2(1 - \Phi(\varepsilon_b)). \tag{101}$$

If all σ_{mn}^2's are the same, observing the bound b in Theorem 4 completes the proof.

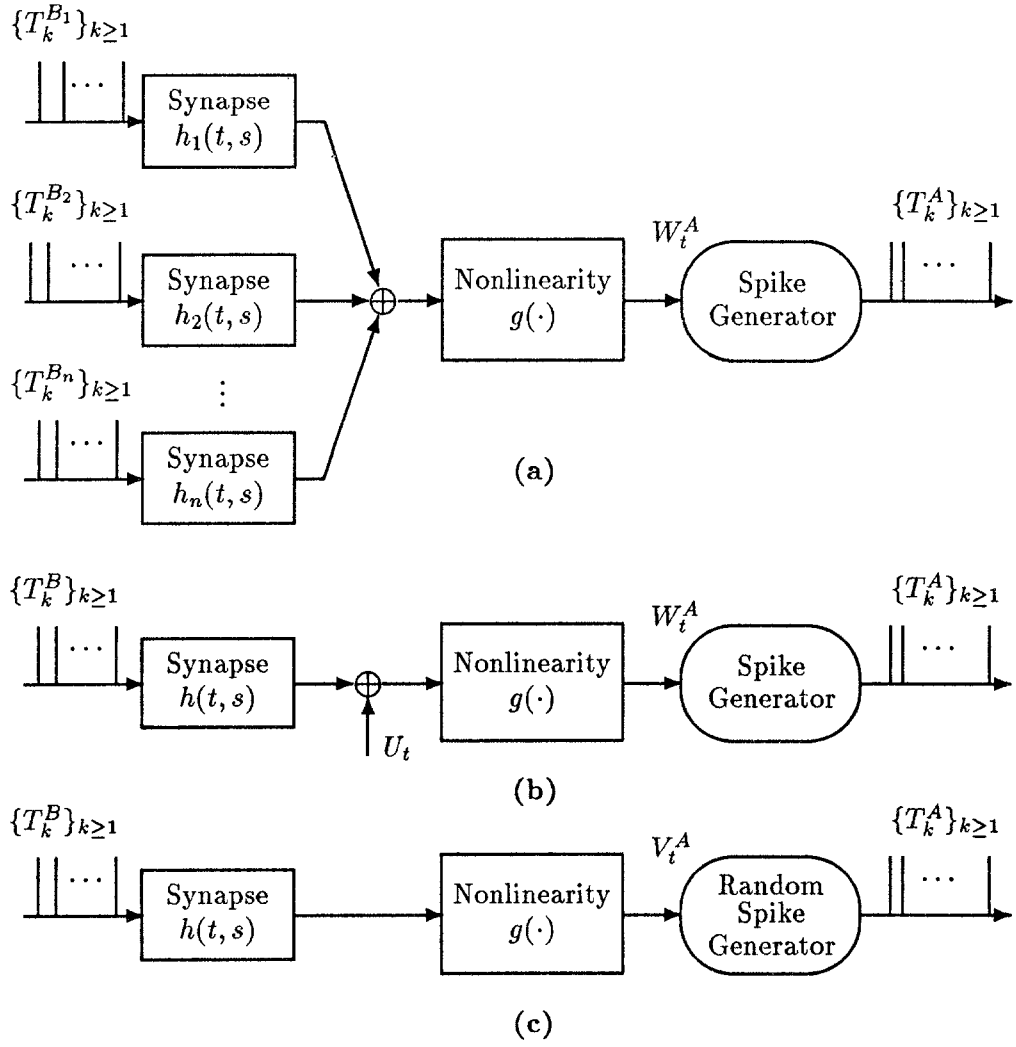

Figure 1: A dynamical nonlinear neuron model, where neuron A is considered as the post-synaptic neuron.

(a) Neuron A is influenced by presynaptic neurons $B_1, B_2, \cdots, B_n$.

(b) A synaptic connection between neurons A and B; the influences of other neurons on neuron A are summarized by U_t.

(c) An equivalent probabilistic version of the neuron model. The impact of the random input U_t is now moved to the spike generator where the threshold becomes random.

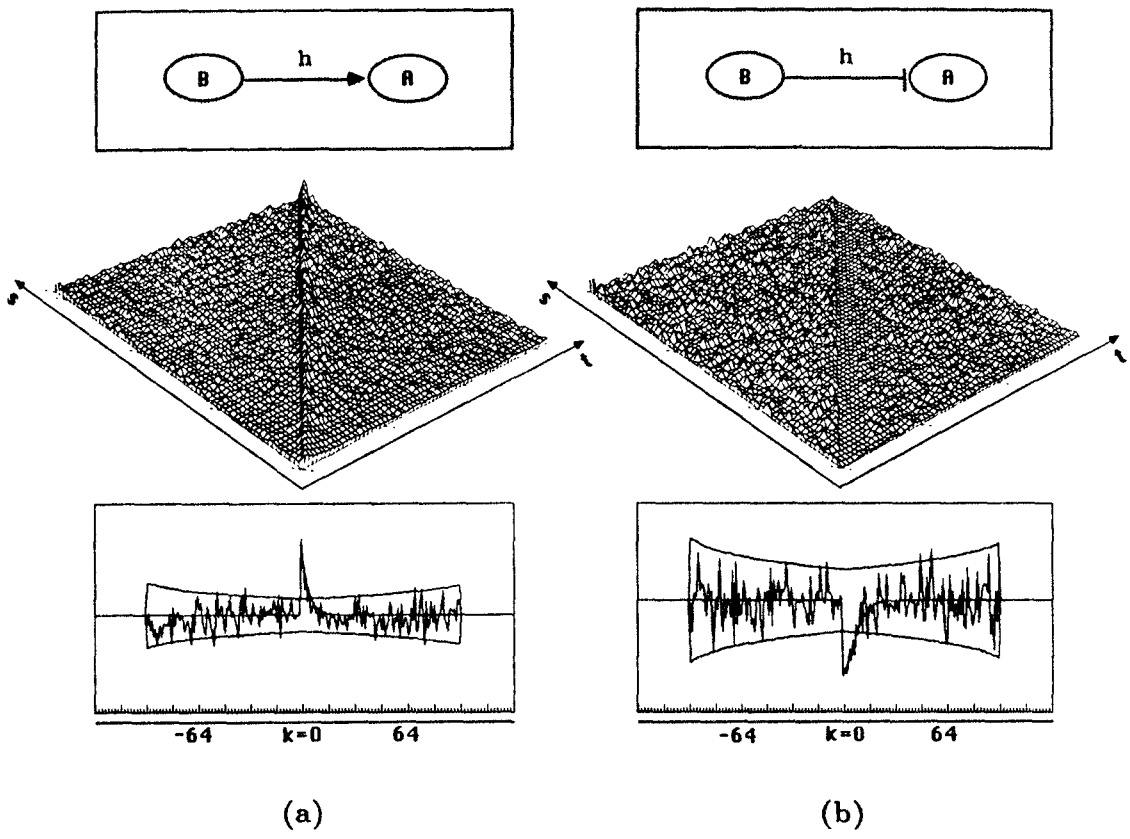

Figure 2: Simulations for pair-wise excitatory and inhibitory correlations.

(a) Excitatory coupling $h(t,s) = 0.8e^{-20(t-s)}$, $t > s$. Shown is the two-dimensional normalized scatter plot generated by the spike trains of the two neurons; below it is the histogram G_k that results from collapsing the scatter plot along the principal diagonal. It corresponds to the function $N_p(k) = e^{h(k)}$. The upper and lower bound lines represent the 95 % confidence measure.

(b) Inhibitory coupling, similar to (a) for $h(t,s) = -3.0e^{-20(t-s)}$, $t > s$.

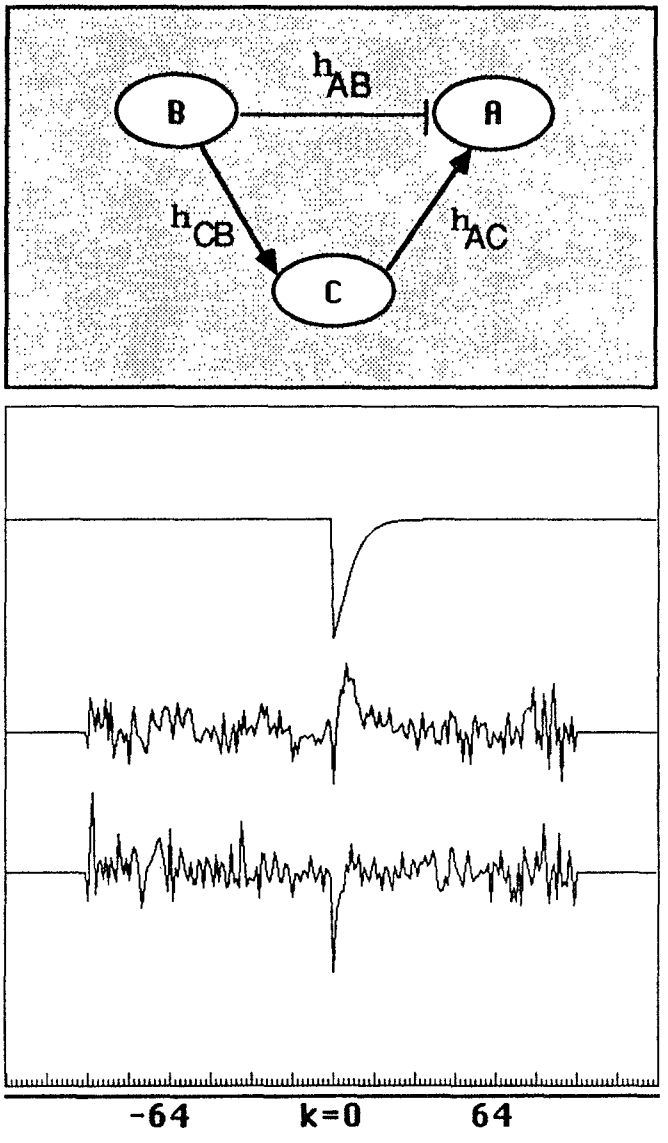

Figure 3: Interaction among three neurons. The network structure is displayed on the top graph: neuron B inhibits neuron A and excites neuron C, and neuron C excites neuron A. $h_{AB}(t) = -1.8e^{-20t}$, $h_{AC}(t) = 3.6e^{-20t}$, and $h_{CB}(t) = 2.0e^{-20t}$. The top curve gives the theoretical connectivity from formula (30) with $\gamma^*(t,s)$. The middle one is the correlation curve corresponding to formula (52) generated from spike trains A and B only. The correlation is so distorted that actual inhibition becomes a false excitation (which is actually due to a strong excitatory input from neuron C). The bottom curve shows the tripartite correlation according to formula (53), which displays the correct inhibitory sign for the connectivity.

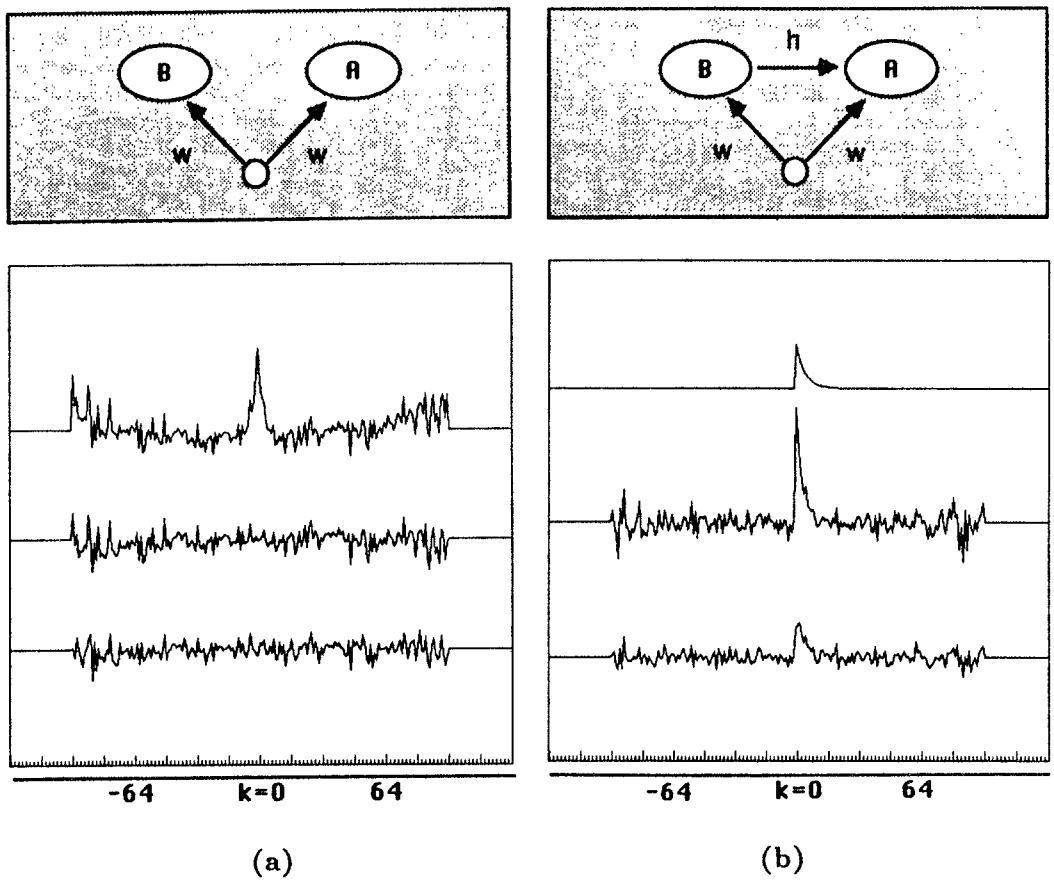

Figure 4: Comparison of the preferred normalization with the difference normalization (shuffle method).

(a) The absence of a direct connection case ($h = 0$): neurons A and B have a common input source — a neuron driven by a stimulus. The connection strength from the common input is $w = 1$. The top curve gives the collapsed version of the joint PST histogram without any normalization. The correlation peak is purely due to stimulus effects. The middle curve represents the difference normalized correlation. The bottom curve shows the preferred normalized correlation curve. Both methods perform well in indicating the absence of connection between A and B.

(b) The presence of a direct connection case ($h \neq 0$): neurons A and B have a common input source as in (a), and in addition, a direct synaptic connectivity from B to A, $h_{AB}(t) = 0.4e^{-20t}$. The top curve gives the theoretical correlation predicted from $N_p(k) = e^{h(k)}$. The middle curve shows the difference normalized correlation. Although the connectivity is weak (only 0.4), the large sharp peak in the correlation leads to a false impression of high excitatory connectivity, which is in fact due to stimulus effects. The bottom curve shows the preferred normalized correlation, which is very close to the theoretical function $0.4e^{-20t}$.