MAXIMUM LIKELIHOOD ANALYSIS OF NEURONAL SPIKE TRAINS

BY

YOUSEF MOHAMMED EMHEMMED

A Thesis Submitted to the
University of Glasgow
for the Degree of
Doctor of Philosophy (Ph.D.)

Faculty of Science
Department of Statistics

June, 1995

© 1995, Yousef, M. E.

ProQuest Number: 13818457

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 13818457

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Thesis
10191
Copy 1

GLASGOW
UNIVERSITY
LIBRARY

In the name of Allah
the merciful the compassionate

Table of Contents

Acknowledgements

First, I owe a debt of gratitude to many people for providing me with every possible assistance, Statistical, Physiological or otherwise, during the period of my research.

On the top of the list, I would like to express my deep appreciation and gratitude to my supervisors Mr. **Peter Breeze** and Professor **Jay Rosenberg** of the Statistics and Physiology Departments at the University of Glasgow for suggesting the topic of this research and for their advice, continual encouragement and patience throughout the period of this study.

I would also like to thank Dr. **David Halliday** of the Physiology Department for kindly providing the data analysed in this thesis and for many useful discussions.

I wish to extend my thanks to Professor **Ian Ford**; Head of the department of Statistics, the staff members and to my fellow research students for providing me with very comfortable working environment.

I am indebted to many friends in **Glasgow** and fellow research students for making my stay here so enjoyable.

I am grateful to the Department of Statistics, El. Fateh University in **Tripoli** for their financial support.

My final acknowledgement is a word of special thanks reserved for my wife for her magnificent devotion to her family and for her continuous understanding, encouragement, moral support and patience without which this research would have been made infinitely more difficult.

Glasgow, June, 1995.

To the memory of my brother Issa;

To those who sacrifice for me;

To my Mother;

To my Father;

To my Wife;

and To my daughter Mauada Yousef,

To you all, I would like to dedicate this thesis.

Notation

Table 1: Operators and symbols and their meanings.

Operators and Symbols	
$Pr\{.\}$	Means "Probability of".
$E(.)$	Means "Expected value of".
$Var(.)$	Means "Variance of".
$Cov(.)$	Means "Covariance of".
M, N and L	Point processes M, N and L.
X_{t-u}	Observed input at time $t-u$ (i.e., at lag u).
Z_{t-u}	"Unobservable" input at time $t-u$ (i.e., at lag u).
Y_t	Observed output at time t.
$\sum_{i=1}^{k} s_i$	Summation of s_i from i=1 to $i=k$.
$\prod_{i=1}^{k} s_i$	Product of s_i from i=1 to $i=k$.
$\Phi(.)$	Standard Normal cumulative distribution function.
χ^2_m	Chi square random variable with degrees of freedom "m".

γ_t	Time elapsed since the time of the last output spike at any given time t.
ζ_1	Minimum of the output inter-spike intervals.

Table 2: Parameters for different functions and their corresponding estimates.

Function	Parameter	Estimate
Counting measure that counts the number of events in the interval (0, t].	$N(t)$	——
Mean intensity function of point process N.	P_N	$\hat{P}_N$
Cross-intensity function between point processes N and M at lag v.	$m_{NM}(v)$	$\hat{m}_{NM}(v)$
Auto-spectrum of point process M at frequency λ.	$f_{MM}(\lambda)$	$f_{MM}^{(T)}(\lambda)$
Cross-spectrum between two point processes N and M at frequency λ.	$f_{NM}(\lambda)$	$f_{NM}^{(T)}(\lambda)$
Ordinary coherence function between two point processes N and M at frequency λ.	$\lvert R_{NM}(\lambda)\rvert^2$	$\lvert R_{NM}^{(T)}(\lambda)\rvert^2$
Partial coherence function of order 1 between two point processes N and M after removing the effects of point process L.	$\lvert R_{NM.L}(\lambda)\rvert^2$	$\lvert R_{NM.L}^{(T)}(\lambda)\rvert^2$
Multiple coherence function between an output point process N and input processes M and L.	$\lvert R_{N.ML}(\lambda)\rvert^2$	$\lvert R_{N.ML}^{(T)}(\lambda)\rvert^2$

Phase-spectrum function between point processes N and M at frequency λ.	$\phi_{NM}(\lambda)$	$\phi_{NM}^{(T)}(\lambda)$
The time delay between two point processes N and M.	τ	$\hat{\tau}$
The likelihood function for parameter $\underline{\theta}$ and data $\underline{y}$.	$l_0(\underline{\theta};\ \underline{y})$	$l_0(\hat{\underline{\theta}};\ \underline{y})$
The natural log of the likelihood function for parameter $\underline{\theta}$ and data $\underline{y}$.	$l(\underline{\theta};\ \underline{y})$	$l(\hat{\underline{\theta}};\ \underline{y})$
Membrane potential on the trigger zone of the neurone at time t.	U_t	$\hat{U}_t$
Linear predictor of the model at any given time t.	η_t	$\hat{\eta}_t$
Conditional probability of neurone firing at any given time t.	P_t	$\hat{P}_t$
Empirical probability of neurone firing at any given value of the linear predictor η_t.	$P(\eta_t)$	$\hat{P}(\eta_t)$
Deviance function for parameter $\underline{\theta}$ and data $\underline{y}$.	$D(\underline{\theta};\ \underline{y})$	$D(\hat{\underline{\theta}};\ \underline{y})$
Link function.	$H(P_t)$	$H(\hat{P}_t)$
Inverse of the link function.	$H^{-1}(\eta_t)$	$H^{-1}(\hat{\eta}_t)$
Summation function at lag u.	$\{a_u\}$	$\{\hat{a}_u\}$
Carry-over effect function at lag w.	$\{c_w\}$	$\{\hat{c}_w\}$
Recovery function at lag α where $\alpha=(\gamma_t-\zeta_1-1)$.	V_α	$\hat{V}_\alpha$
Summation function for the first observed input at lag u.	$\{_1a_u\}$	$\{_1\hat{a}_u\}$

Summation function for the second observed input at lag u.	$\{{}_2a_u\}$	$\{{}_2\hat{a}_u\}$
Carry-over effect function for the first observed input at lag w.	$\{{}_1c_w\}$	$\{{}_1\hat{c}_w\}$
Carry-over effect function for the second observed input at lag w.	$\{{}_2c_w\}$	$\{{}_2\hat{c}_w\}$
Summation function for the "unobservable" input at lag u.	$\{b_u\}$	$\{\hat{b}_u\}$
Carry-over effect function for the "unobservable" input at lag w.	$\{d_w\}$	$\{\hat{d}_w\}$

Abstract

The main aim of this thesis is to introduce and develop a very powerful statistical technique, maximum likelihood estimation, to show how best this approach can be used in analysing neuronal spike train data. We then compare some of the likelihood results with those obtained via stochastic point processes techniques which will highlight the advantages of using the likelihood approach.

Chapter 1 is aimed to give the physiological background and to provide a brief description of some aspects of neurophysiology which are relevant to the discussions that follow throughout this thesis. A brief description of the neuromuscular control system is followed by a more detailed one of the structure and operation of one of its components, the muscle spindle. Some basic statistics and a brief summary of a real data set obtained from a mammalian muscle spindle are also presented. The last part of this introductory chapter describes the simulation procedure used to generate the data sets for this thesis.

Chapter 2 presents some brief historical notes of point process theory, followed by a definition of point process and some of the standard assumptions. The final part of this chapter is a review of the stochastic point processes techniques in both time and frequency domains followed by a demonstration of the uses of the square root of the cross intensity function, the ordinary coherence and the phase function with two simple examples from simulated neuronal spike train data.

In chapter 3 we introduce the maximum likelihood estimation procedure as an alternative technique to the point process techniques used in the analysis of neuronal spike train data followed by a definition of the likelihood function and the maximum likelihood estimator (m.l.e.). An analytic likelihood model is introduced. The model is based on two underlying processes, the linear summation of the effects of the input spike train on the membrane potential and a recovery process, which, among other things, represents intrinsic properties of the neurone. The link function, the log likelihood for binomial data and a computational procedure are also discussed. The analysis of deviance, which highlights the difficulty in the goodness of fit assessment for models used to analyse binary data, is discussed. A linearisation technique for estimating non-linear parameters, which is used to estimate the non-linear parameters in the case of an exponentially decaying threshold, is introduced. In the final part of this chapter the summation, recovery and threshold functions are estimated using the same two sets of data considered in chapter 2, where a comparison of the summation function with the corresponding cross intensity function in each example indicates that the cross intensity function is underestimating the underlying excitatory effects of a synaptic input and may be misleading.

We start chapter 4 with a discussion of certain issues concerning the likelihood approach, in particular an assessment of goodness of fit. We introduce a method of checking the validity of the model based on a graphical

comparison between estimated and corresponding theoretical probabilities. We follow this with a discussion of the choice of link function. The second part of this chapter then applies the likelihood procedure to some simulated data sets. We start the analysis with a spontaneous discharge data set using three different link functions. This is a case where only threshold and recovery functions can be estimated, and also where the traditional stochastic point process techniques do not provide an analogous measure for the spontaneous behaviour of the cell. This gives the likelihood approach a further advantage over time and frequency domain analyses. Also we consider the case of a single input and single output neuronal spike train data set, where we introduce the idea of a carry-over effect of the synaptic inputs on the firing of a neurone. The likelihood approach is able, to some extent, to separate aspects of the relationship between spike trains through the threshold, recovery, summation and carry over effect functions and such ability and flexibility are not provided by other techniques. The demonstrations again suggest that the cross intensity function is difficult to interpret and may be misleading and underestimate the underlying excitatory and inhibitory effects of a synaptic input. In each case the improvement of the model is assessed at each stage of complexity by constructing a table of deviances. A sufficient reduction in deviance when proceeding to higher levels of complexity reflects a significant improvement in the model.

In chapter 5 we extend the application of the likelihood procedure by taking advantage of its flexibility in the cases of one and two observed inputs and a single observed output of neuronal spike train data both in the absence and presence of "unobservable" inputs. This is to show that the approach is sufficiently flexible, and it may further be extended in principle to the case of an arbitrary number of neurone inputs. We start the analyses with one simulated data set with one observed (spike train) input and one "unobservable" (continuous) input, and one observed (spike train) output followed by two

examples where two observed spike train inputs and one "unobservable" input, and one observed output have been considered. The two observed inputs are uncorrelated with each other in the first set of data and correlated with each other in the second set of data. The likelihood approach again shows the ability to separate the aspects of the relationships between spike trains. The demonstrations again suggest that the square root of the cross intensity function is underestimating the underlying excitatory effects of a synaptic input compared to the corresponding summation function. In the second part of this chapter we have discussed a real set of data obtained from a mammalian muscle spindle where the relationships between the two fusimotor γ_0 and γ_b inputs and each of the sensory Ia and II outputs are investigated.

As a final chapter of this thesis, chapter 6 presents a summary and general conclusion of the present work and also indicates some possible areas in which the work of this thesis may be extended.

Chapter 1

1 Neurophysological Background

1.1 Introduction

Neurophysiology is a branch of science that concerned with how the elements of the nervous system function and how they communicate with each other. The function of the nervous system at all levels is seen to involve chemical mechanisms, electrical mechanisms and physical arrangements.

The goals of neurophysiologists range from understanding the properties of neurone to the heroic: how to explain things like memory, emotion, learning, sleep, expectation, behaviour, etc. At a less ambitious level neurophysiologists are concerned with how a single nerve cell responds to stimuli, transmits information and changes with alterations of the environment. The main aim of this introductory chapter is to give the necessary background and also to provide brief descriptions of some aspects of neurophysiology which are relevant to the discussions that follow throughout this thesis.

Many biological systems have the important feature that under normal operating conditions they are acted upon by several inputs simultaneously, and in response may give rise to several outputs. This common feature of biological systems plays a crucial role in its function. The field of neurophysiology provides many examples of such systems where continuous signals (e.g. change in muscle length) can interact with neuronal signals to produce other neuronal signals. These neuronal signals consists of a series of spikes known as a spike-train (they are referred to as spikes because of their relatively short duration which is about 1 msec and fixed amplitudes), and can be represented as a series of isolated events by considering the spike-train as a point-process. Any element of the neuromuscular system which is acted upon or generates spike-trains can then be considered as a point-process system. One obvious example of such a system is the muscle spindle. The muscle spindle is an important component of the neuromuscular system which is thought to provide information to the other parts of the nervous system that is important in the control of movement and maintenance of posture. During the course of a movement the muscle spindle is acted upon by a continuous change in length, which occurs as a consequence of the movement. In addition to the length change, the output activity from the spindle is further modified by several point process inputs that are initiated within the central nervous system.

The spike trains of the central nervous system reveal a certain degree of randomness associated with its activity. Moreover, it has been recognised by many statisticians that the field of neurophysiology provides a rich source of problems and data relating to stochastic process systems (Brillinger, 1975a; Brillinger et al, 1976; Feinberg, 1974; Sampath and Srinivasan, 1977). First we present a brief description of the neuromuscular control system followed by a more detailed one of the structure and operating of one of its components, the muscle spindle. Secondly we provide a brief summary of the basic data sets

which will be used as experimental material throughout the subsequent chapters.

1.2 Neuromuscular Control System

The neuromuscular control system may be thought of as all those parts of the nervous and muscular systems concerned with the initiation and control of movement and maintenance of posture. Anatomically and functionally this system has further been divided into the central nervous system and the peripheral nervous system.

1.2.1 Central Nervous System (Brain)

The central nervous system consists of the brain and the spinal cord. The human brain contains a very complicated network of perhaps as many as ten billion highly specialised cells called neurones (or nerve cells) which are the basic building blocks of an animal's central communication system.

Fig.1.2.1 illustrates a sample of four kinds of neurones with different shapes and sizes.

The nerve cells may be considered as input-output systems of a particular structure having important functions. The nerve cells are not isolated, but communicate in a very characteristic way using both chemical and electrical mechanisms. The output voltage from a neurone is a pulse that travels along a neurone's output fibre. It is of near constant shape, and its generation depends on broad variety of non-linear phenomena.

It is pertinent to discuss both structure and function of these cells, because in biology often the two seem directly related. Fig.1.2.2 illustrates

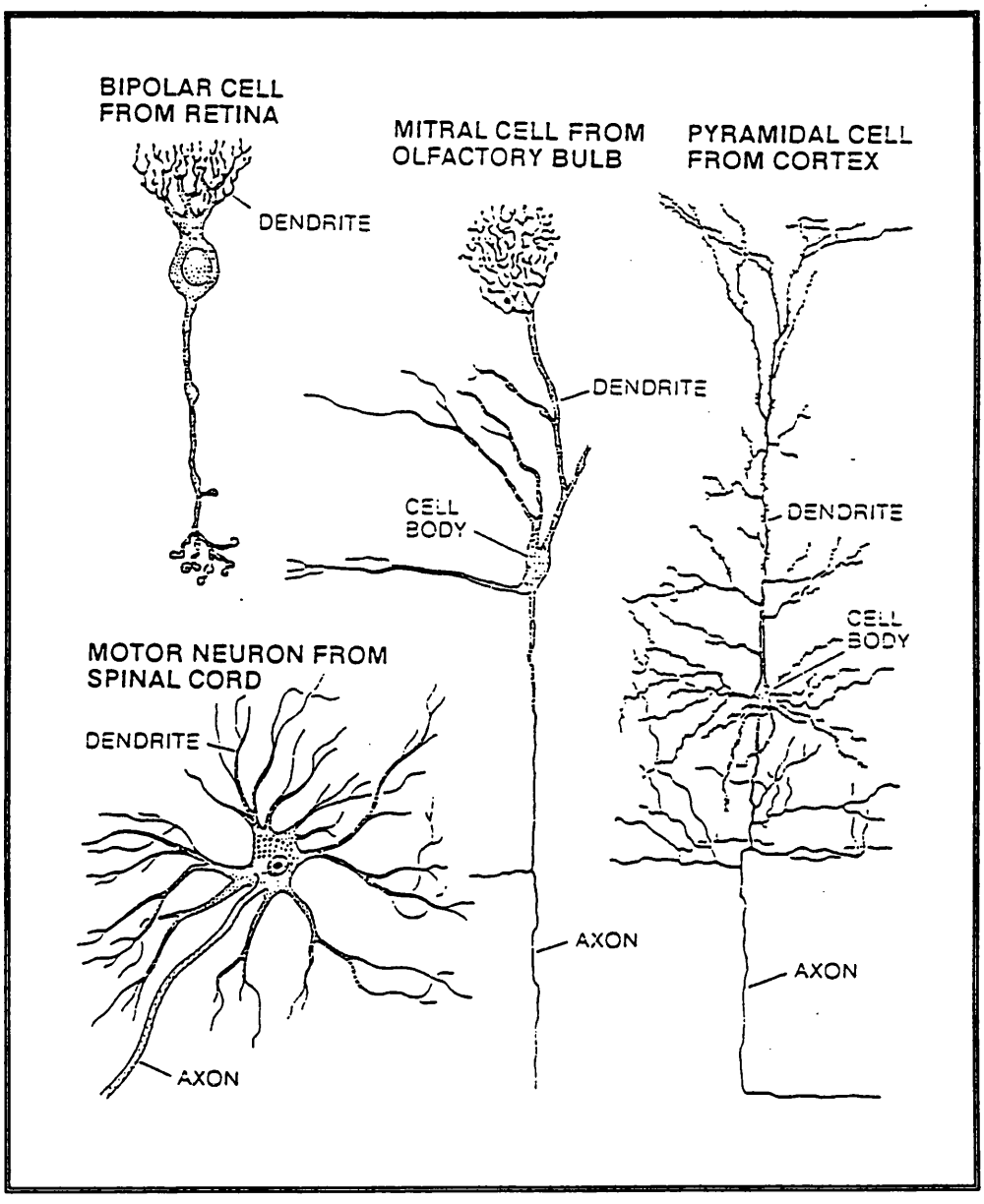


Fig.1.2.1 Shapes and sizes of four selected neurones.

The motor neurone was dissected from a mammalian spinal cord, the bipolar cell is from the retina of a dog, the pyramidal cell from the cortex of a mouse, and the mitral cell from the olfactory bulb (a relay station in the pathway concerned with smell) of a cat (after Kuffler and Nicholls, 1976).

a schematic diagram of small network of three neurones. Although there are many kinds of neurones with structural and functional differences, the basic components of a neurone can be identified as follows:

a. The Axon

The axon is the element that links the neurone to a neighbouring neurone. Normally a cell gives rise to a single axon but this axon may give off side branches and characteristically divides up into a number of smaller branches just before terminating. The internal and external fluids of the axon are composed mainly of ionised potassium chloride and sodium chloride, with the concentration of potassium ions inside the axon much higher than that outside. This results in a movement of the potassium ions to the outside, and hence an electric field which opposes the chemical field. Equilibrium is attained when the two forces are nearly equal, resulting in a potential difference across the membrane, the inside being more negative. This negative potential compared to the surrounding fluid is called the resting potential. If an ionic change occurs that causes the inside of the axon to be more negative, it is called hyper-polarisation, otherwise it is called depolarisation. If the depolarisation is so large that the membrane potential exceeds a certain value, called the threshold, then a spike will occur and propagates along the axon. The direction of propagation is usually away from the neurone cell body. Once the wave-form or what is called the action potential has been generated, the region of stimulation cannot be excited for a short time thereafter, and this property is referred to as refractoriness.

b. The Synapse

The synapses are the points at which the neurones are interconnected with each other, and they are of particular importance, since at these points the information flows from one cell into another, and interactions between neurones take place. The terminal end of an axon broadens into a bulge called the bouton and lies adjacent to the cell membrane or dendrites of the cell body of another neurone cell. The bouton has small packets called vesicles which contain a chemical known as the transmitter, the type of transmitter depends on the kind of junction. Fig.1.2.3 illustrates the various regions and points of contact between one neurone and three other neurones. When an action potential arrives at the bouton, a transmitter is released and its molecules modify the permeability of the membrane of the next neurone to different ions. If the resulting change is positive it is called an excitatory post-synaptic potential (EPSP), otherwise it is called an inhibitory post-synaptic potential (IPSP).

The post synaptic potentials are induced at several points along the soma, which is due to successive arrivals of action potentials, the resulting membrane potential may be the linear sum of the individual potentials. As in the axon, recovery effects keep taking place in the absence of inputs, i.e. the membrane is leaky. The neurone tends to fire when the integrative effect exceeds the threshold.

c. The Soma

The soma is the cell body of the neurone which is about 30 - 100 micra across. The cell bodies are clustered in certain areas in the brain while other regions consist of axons running from one group of neurones to another. In these regions large numbers of closely packed axons run parallel to form

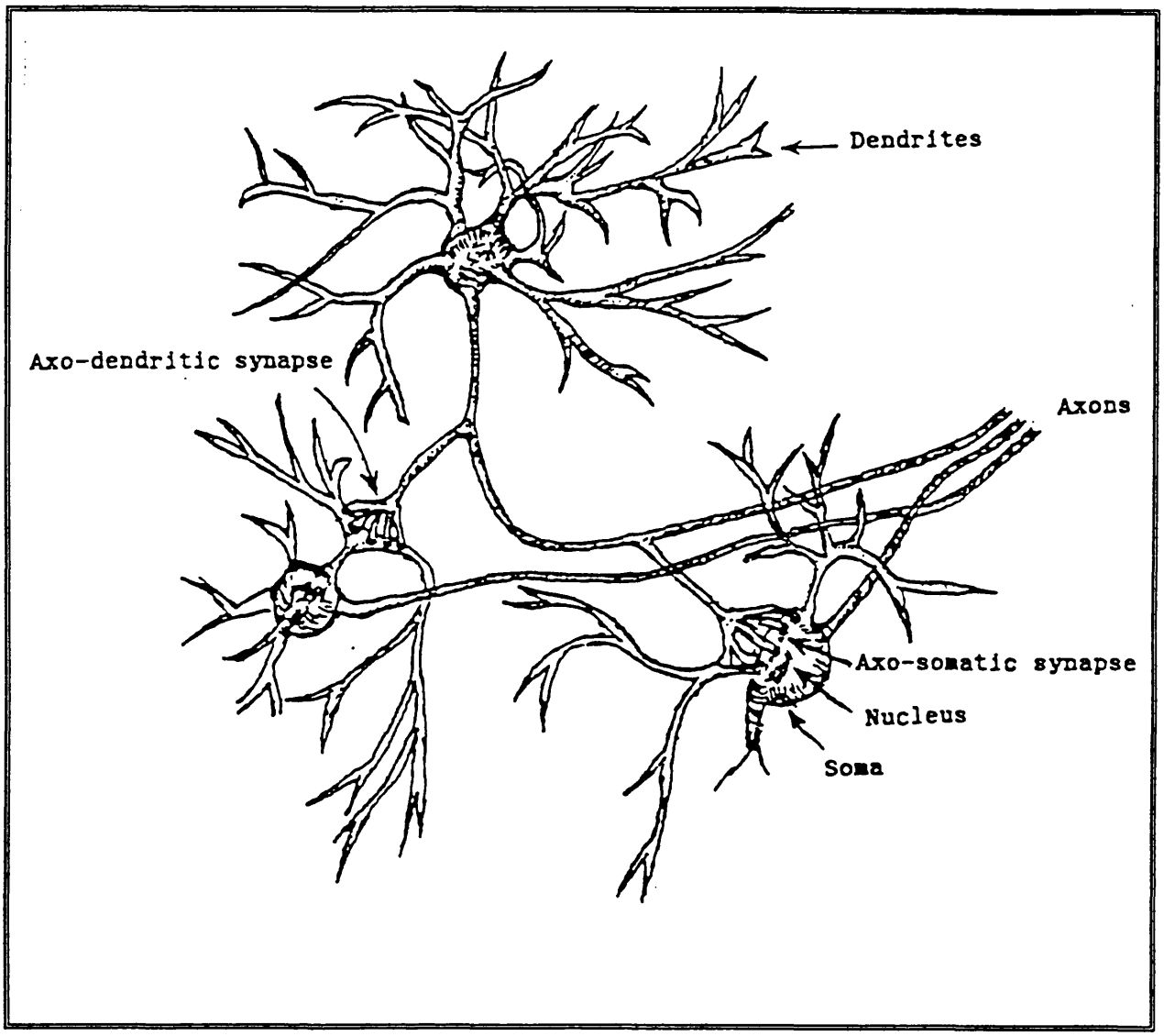


Fig 1.2.2 Schematic diagram showing a small network of three neurones, and the points of connection "synapses" (after Amjad, 1989).

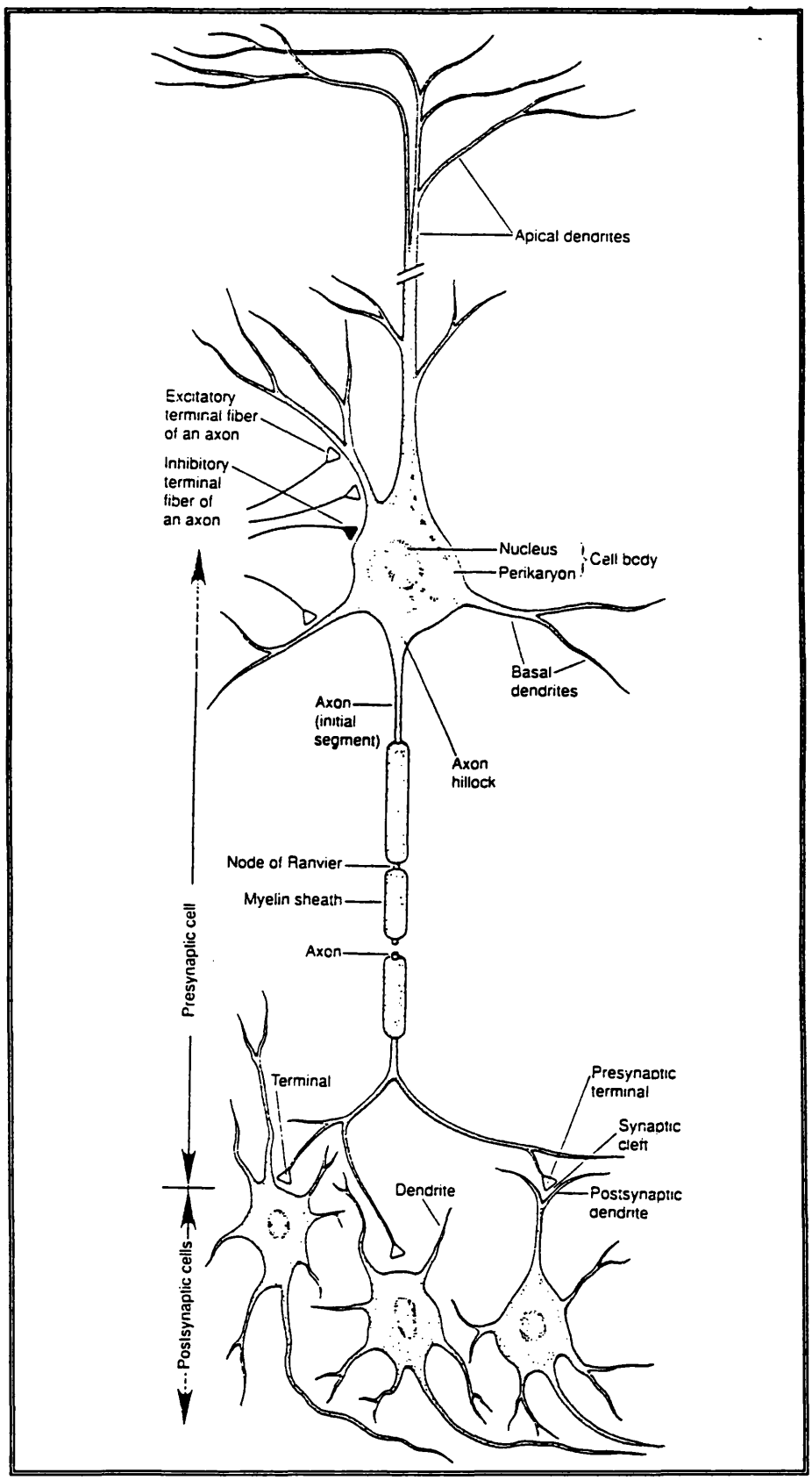


Fig.1.2.3 The various regions and points of contact between one neurone "at the top" and three other neurones "at the bottom" (after Kandel, Schwartz and Jessell,1991).

structures called fibre tracks. Similar tracks run outside the brain to muscles where they are referred to as nerves.

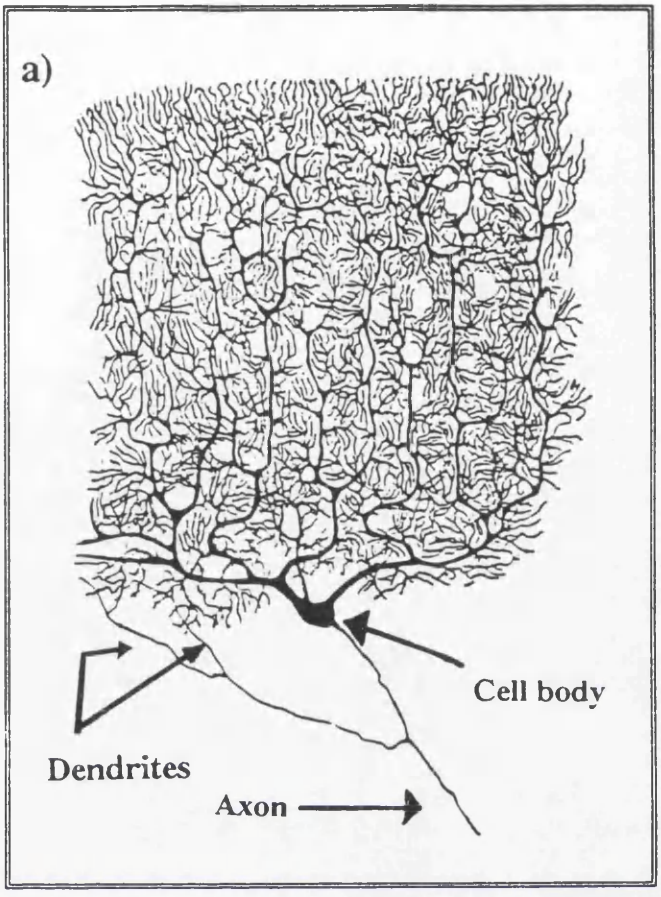
Like the axon, the soma also has the property of decreased excitability following the generation of an action potential. In the case of an alpha-motoneurone this property is known as the after hyper polarisation, and it is thought to play an important role in the control of repetitive firing events and in characterising the function of the alpha-motoneurone.

d. The Dendrites

The dendrites are hair-like extensions projected from the cell body (soma) which are about 200 to 300 micra long. Along with the other components of the neurone, the dendrites play an important role of producing the spike activity of the cell. Fig.1.2.4 is an example of one of the most common types of neurone in the mammalian nervous system, the multipolar cell called the Purkinje cell which is characterised by its rich and extensive dendritic tree. The figure also illustrates the connection between the cell bodies of several Purkinje cells and the dendrites of one basket cell.

1.2.2 Peripheral Nervous System

At the level of the spinal cord, the peripheral nervous system is arranged in a sequence of identically organised repeating segmental layers called segmental levels of the spinal cord. Fig.1.3.1a illustrates some of the pathways connecting a muscle spindle to one segmental level in the spinal cord. It also outlines one of these segmental levels along with the other components which form the peripheral neuromuscular system at this level.



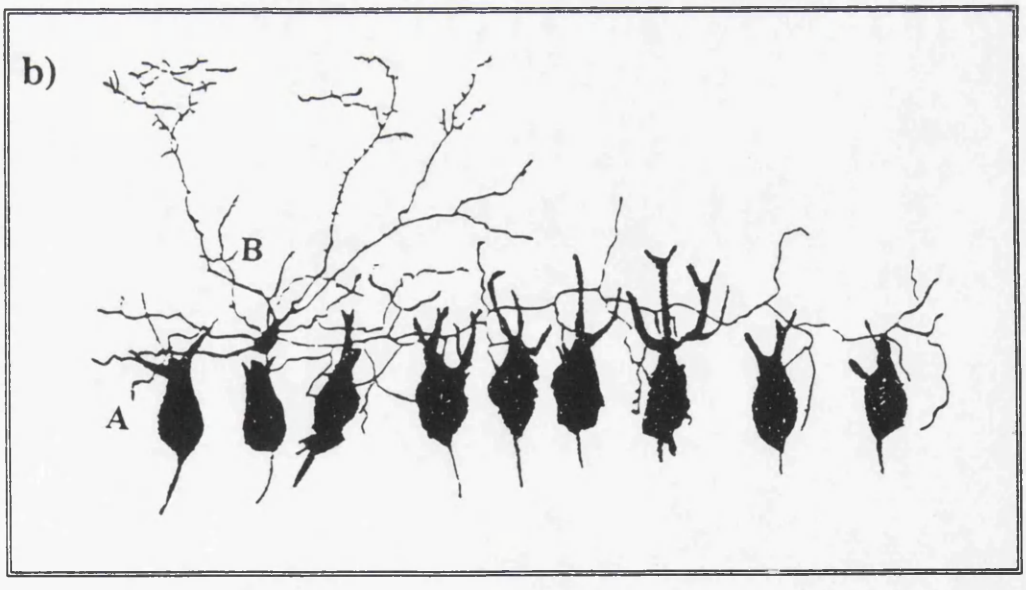


Fig.1.2.4 **a)** Purkinje cell of the cerebellum with an axon and many dendritic processes (after Kuffler and Nicholls, 1976).

b) Schematic diagram showing the connection between one basket cell "**B**" and the cell bodies of several Purkinje cells "**A**" (after Cajal, 1955).

There are several classes of nerve cells which lie within the spinal cord in groups called nuclei some of which may contain as many as 2000 cells. One group is the alpha-motoneurones whose axons innervats the load-bearing, or extrafusal muscle fibres responsible for generating forces or changes of length. The axons of the alpha-motoneurones normally conduct nerve impulses from the cell body to the extrafusal muscle fibres. The cell bodies of the alpha-motoneurones have diameters ranging from 25 to 100 micra and the axons are from 8 to 20 micra in diameter. The axons conduct nerve impulses which travel at velocities in the range 50 to 120 m/sec. from the bodies to the extrafusal muscle fibres. A neurone cell can generate propagated impulses repetitively to produce a train of spikes with mean frequency which may vary from one pulse every few seconds to several hundred pulses per second. Once the axon of an alpha-motoneurone reaches a muscle, it divides into a number of fine branches, which end on specialised areas of the extrafusal muscle called the motor end plate. When a nerve impulse reaches the junction between the axon and the muscle fibre, a sequence of electro-chemical events occurs which leads to the contraction of the muscle fibres. Each terminal branch of the alpha-motoneurones innervats a single extrafusal fibre of one muscle, and all of the extrafusal fibres innervated by one alpha-motoneurone lie within the same muscle. The alpha-motoneurone together with all the extrafusal fibres that it may innervate are called a motor unit, the function of a particular muscle is closely related to the number of motor units and their sizes within that muscle.

Associated with the extrafusal muscle fibres and the tendons which attach the muscles to bone are a number of physiological transducers called muscle receptors, which are very sensitive to imposed length changes or force acting on the parent muscle. The nerves attached to these receptors called sensory nerves normally transmit pulse-coded information in the form of spike trains to the groups of nerve cells lying within the spinal cord. After entering

the spinal cord each sensory axon divides into a number of branches which make synaptic-contact with a large number of nerve cells over several segmental levels of the spinal cord. Each cell within the spinal cord receives input information from a large number of sensory axons from different receptors in the same muscle as well as from the receptors attached to different muscles. The train of action potentials along the axon releases a sequence of electro-chemical events, which occur at the point of contact synapses, then modify the on-going activity of these inter-related cells. An introduction to the organisation of the spinal cord can be found in Shepherd (1974) and a detailed review of this along with the properties of the spinal cord and its interconnections is given in Burke and Rudomin (1977).

1.3 The Transmission of Information

One way the information is transmitted through the dendrites and axon is via changes in electrical activity. An abrupt pulse-like change in the membrane potential is usually called a nerve impulse or action potential. The nerve impulse is approximately 100 mV in amplitude and 1 msec in duration. Because of this relatively short duration impulses are often referred to as spikes or spike trains. Spikes are propagated along the axon with a velocity which depends partly on the axon's diameter. Fig.1.3.1b summarises the sequence of signals that produces a reflex action. A neurone cell can generate propagated impulses repetitively to produce a train of spikes. When a pulse reaches the synapse it provokes the release of a transmitter substance which alters the permeability of the dendrites of the next cell to certain ions. The resulting flow of ions generates a small electric current which moves down the dendrites to the soma. If the synapse is excitatory the probability of spike activity of the second cell is increased, if inhibitory it is decreased.

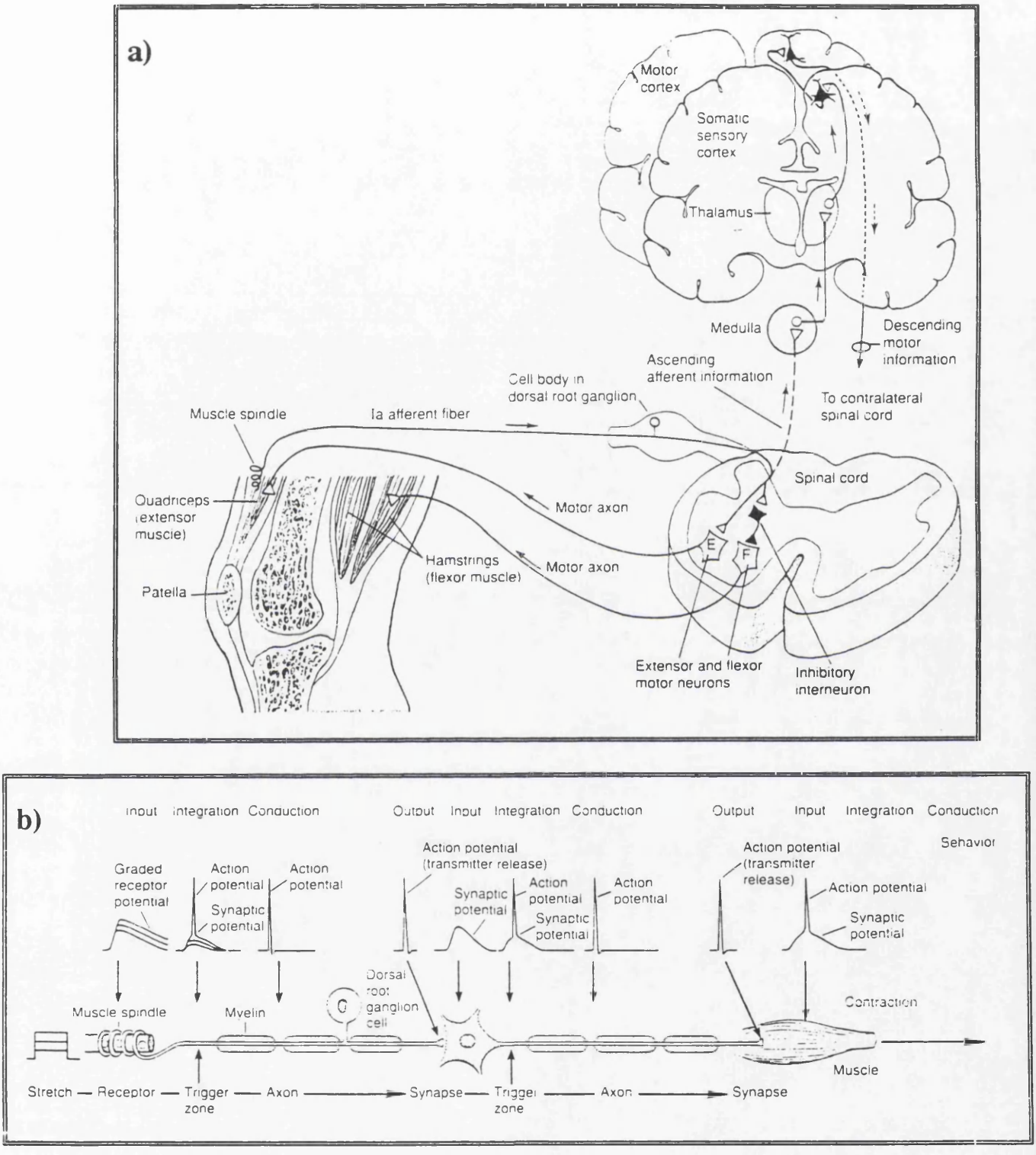


Fig.1.3.1 **a)** Diagram of some of the pathways connecting a muscle spindle to the spinal cord. **b)** Diagrammatical summary of the sequence of signals that produces a reflex action (after Kandel, Schwartz and Jessell,1991).

1.4 The Muscle Spindle

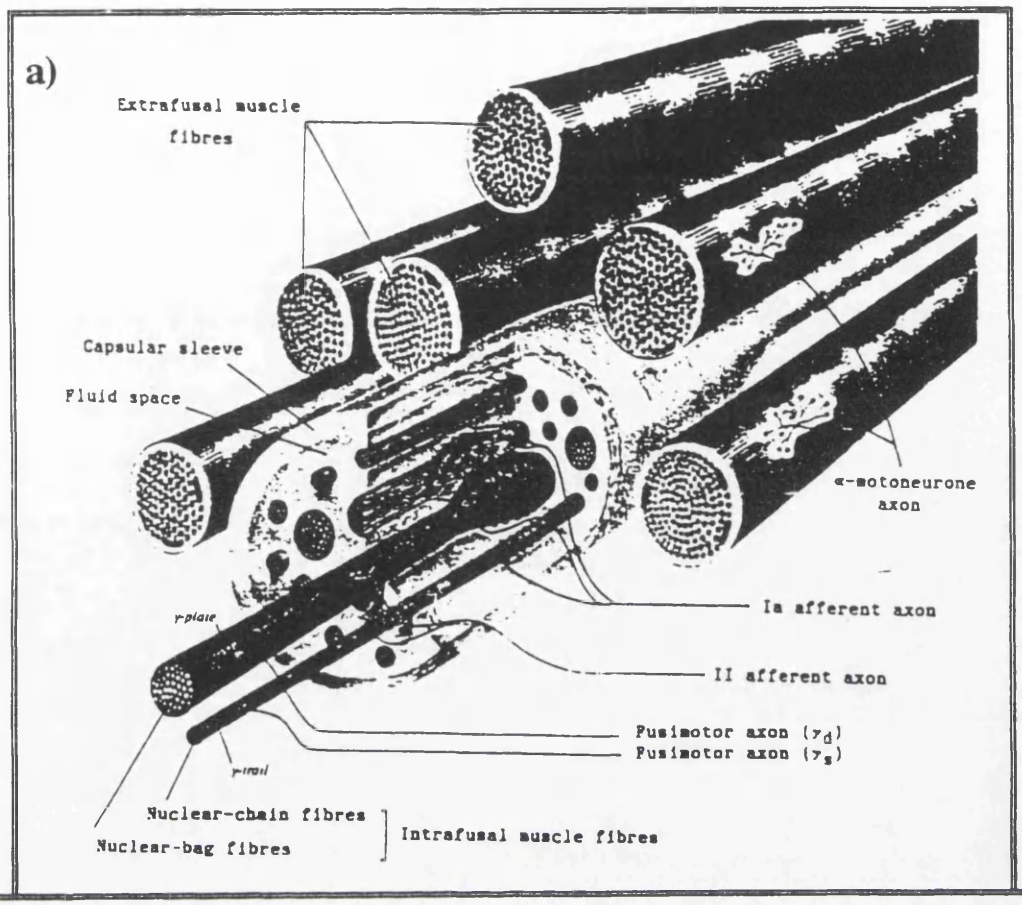
One class of muscle receptors with particular importance is the muscle spindle, which is thought to have a critical role in initiating and controlling movements and maintaining posture. The muscle spindle is a transducer which responds to length changes imposed on the parent muscle. Most skeletal muscles contain a number of muscle spindles lying in parallel with the extrafusal fibres. These spindle fibres are much shorter than the extrafusal fibres and are partially surrounded by a fluid-filled capsule of connective tissue. These intrafusal fibres have been divided into three different types, namely, the dynamic nuclear-bag fibres (Db1), the static nuclear-bag fibres (Sb2), and the nuclear-chain fibres (C).

The properties of the three types of intrafusal fibres are different, and hence they might respond distinctly to length change imposed on the parent muscle (Bessou and Pages, 1975; Boyd, 1980). The effects of the imposed length changes on the intrafusal muscle fibres are transmitted to the spinal cord by the two types of sensory axons closely associated with the muscle spindle. These are the primary or group Ia, and the secondary or group II axons, each spindle having one primary and several secondary axons associated with it. Action potentials in primary sensory axons have a conduction velocity in the range 72 to 120 m/sec, whereas in the secondary sensory axons it is in the range 24 to 72 m/sec.

The sensory axons from the muscle spindle generate action-potentials at a constant rate, when the parent muscle is held at a fixed length, and this constant rate depends upon the muscle length (Matthews and Stein, 1969). The rate, however, is increased with an increase in the muscle length. Each muscle spindle is innervated by a single Ia sensory axon, but may have several group II axons. The changes in activity in the Ia sensory axon in part reflect the responses to imposed length changes in all the three types of intrafusal fibres,

whereas the activity in the II sensory axons reflects, mainly, changes in the muscle-chain fibres. It has been shown that the Ia and II sensory axons project largely to different groups of cells within the spinal cord and therefore may be associated with quite different functions (Johansson, 1981).

In addition to Ia and II sensory axons, the intrafusal muscle fibres are innervated by the axons of a group of cells lying within the spinal cord in the neighbourhood of the alpha-motoneurones. These cells have cell bodies smaller than those of the alpha-motoneurones. These motoneurones generate impulses at a velocity ranging from 10 to 50 m/sec. These gamma motoneurones innervate only the intrafusal fibres. Each gamma motoneurone may innervate intrafusal fibres lying in different muscle spindles within the same muscle. These motoneurones have been further divided into two broad categories, the gamma dynamic and gamma static axons (Matthews, 1962; Emonet-Denand et al, 1977). The gamma dynamic axons innervate the dynamic nuclear bag fibres, whereas the gamma static axons innervate either the nuclear chain fibres or the static nuclear-bag fibres, or both (Boyd, 1980; Matthews, 1981). A single muscle spindle may be innervated by as many as six fusimotor neurones. Fig.1.4.1 summarises the main features of the muscle spindle. Fig.1.4.1a is a schematic diagram of the sensory and motor innervation of the muscle spindle. Fig.1.4.1b is a naive description of the normal innervation of the muscle spindle.



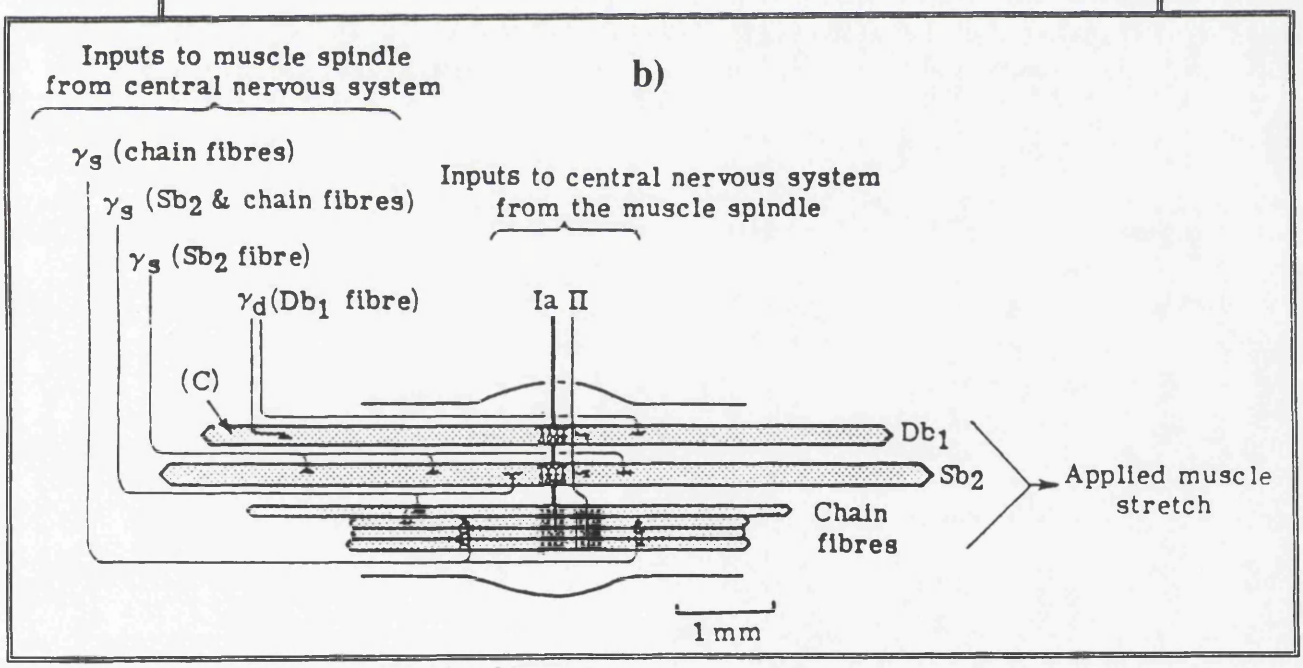


Fig.1.4.1 **a)** Schematic diagram of the sensory and motor innervation of the muscle spindle. The spirals of the primary sensory ending are seen wrapped round all the intrafusal fibres, whereas the terminals of the secondary sensory ending are restricted to the chain fibres. **b)** Simplified diagram of the normal innervation of the muscle spindle (after Amjad, 1989).

1.5 Basic Data Sets and Recording Procedures

The example data sets shown in this chapter were obtained mainly from a muscle lying within the tenuissimus muscle in the hind limb of anaesthetised adult cats by isolating a muscle spindle within the parent muscle and dissecting the selected fusimotor nerves from the spinal cord and the primary and sensory nerves to the spinal cord. The parent muscle was clamped in a muscle puller to keep the muscle length under control during the recording.

The recording technique used here is what is called recording in which fine silver wire electrodes insulated except for the tip were attached to the cut nerve endings and the fusimotor ending stimulated with voltage pulses. This sequence of pulses of the primary and secondary responses was recorded to form the data sets.

1.5.1 Discharges of the Ia and II Sensory Axons in Different Conditions

The Ia and II sensory axons of the muscle spindle in the absence of any fusimotor activity may still be able to generate nerve impulses at nearly constant rate depending on the given length at which the parent muscle is held (Matthews, and Stein, 1969). The rate, however, is increased when the fusimotor axon inputs γ_0 and γ_b are applied. The discharge activity of the sensory axons when no fusimotor inputs are applied is referred to as the spontaneous discharge of the spindle.

Fig.1.5.1 is an example which illustrates the output point process activity (discharge) obtained from the Ia and II sensory axons of the muscle spindle. Fig.1.5.1a is the inter-spike interval histogram (h=1 msec) of the II spontaneous discharge. Fig.1.5.1b is the inter-spike interval histogram (h=1 msec) of the Ia spontaneous discharge.

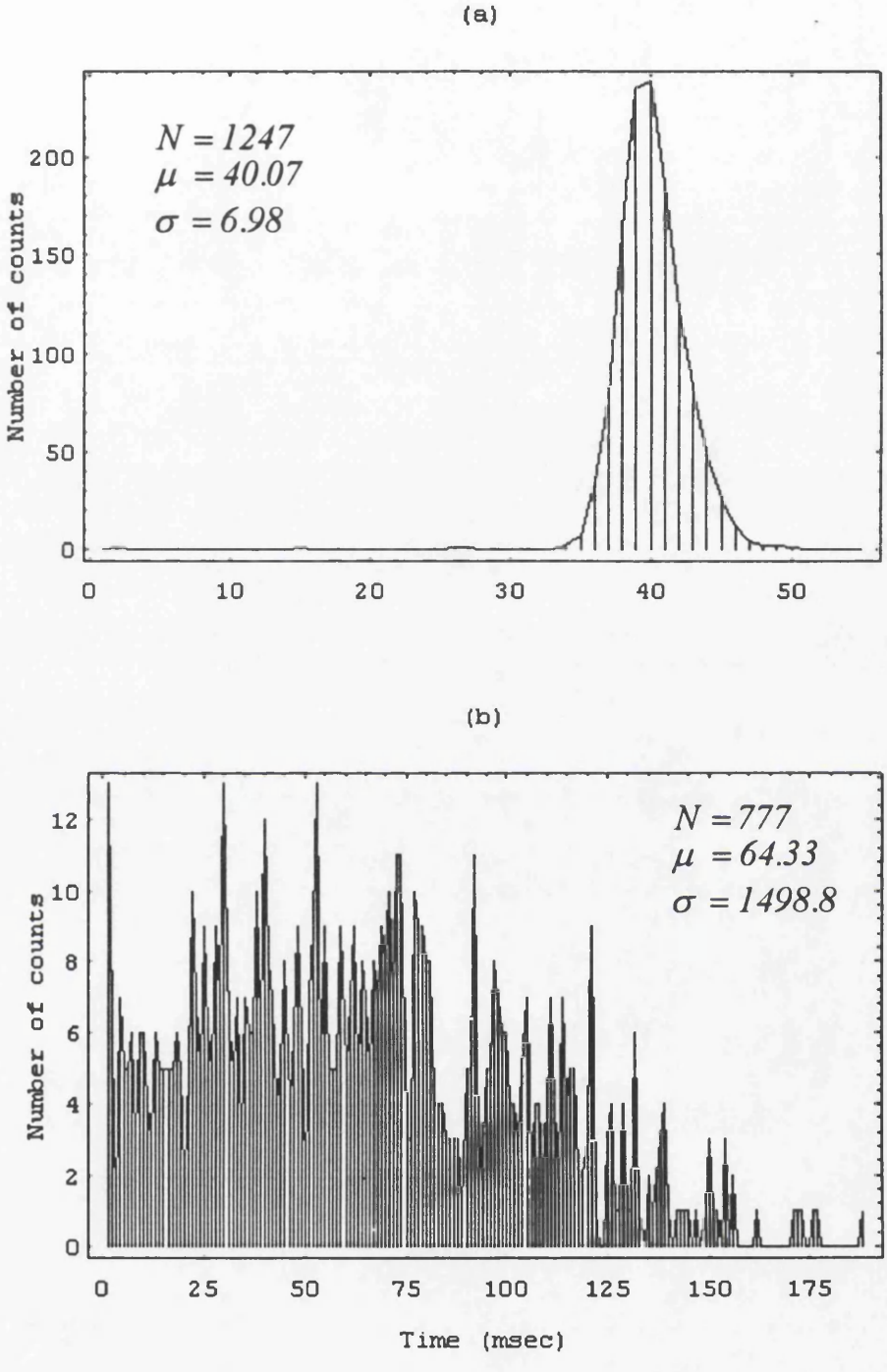


Fig.1.5.1 Basic statistics of the Ia and II spontaneous discharges.

a) The inter-spike interval histogram (h=1 msec) of the II spontaneous discharges.

b) The inter-spike interval histogram (h=1 msec) of the Ia spontaneous discharges.

Fig.1.5.2 represents the Ia and II discharge when the fusimotor axon inputs γ_0 and γ_b are applied. Fig.1.5.2a represents the inter-spike interval histogram (h=1 msec) of the II discharge in the presence of both fusimotor axon inputs γ_0 and γ_b whereas Fig.1.5.2b represents the inter-spike interval histogram (h=1 msec) of the Ia discharge in the presence of both fusimotor axon inputs γ_0 and γ_b.

The Ia and II sensory axons discharges may be considerably affected when the parent muscle is acted upon by both fusimotor axon inputs γ_0 and γ_b. Fig.1.5.2a and Fig.1.5.2b are examples which illustrate how the output point process activity (discharge) of a muscle spindle is affected by the two kinds of inputs when compared to those of the corresponding spontaneous discharge shown in Fig.1.5.1a and Fig.1.5.1b respectively when the muscle spindle is assumed to receive no input effects.

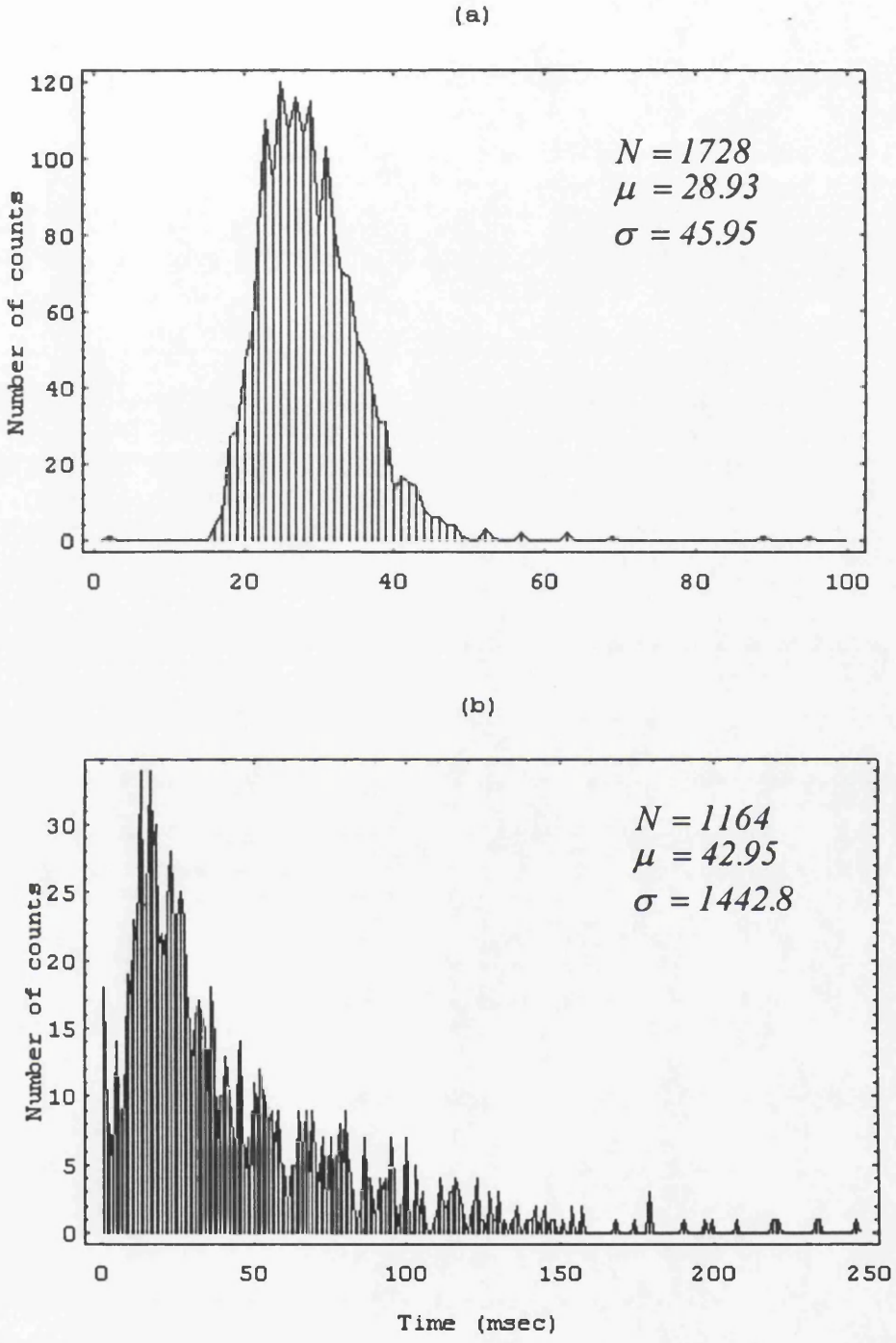


Fig.1.5.2 Basic statistics of the Ia and II discharges in the presence of γ_0 and γ_b fusimotor axon inputs.

a) The inter-spike interval histogram (h=1 msec) of the II discharges in the presence of γ_0 and γ_b fusimotor axon inputs.

b) The inter-spike interval histogram (h=1 msec) of the Ia discharges in the presence of γ_0 and γ_b fusimotor axon inputs.

1.6 Simulation Procedure

Attempts to understand the nervous system and how it performs various tasks require a biological, physical and computational framework in which to develop mathematical models describing how the nervous system, or some parts of it, carry out certain operations. In the last decade, with the remarkable increase in computer power, simulation models of neurones that incorporate as much anatomical and biophysical detail as available have been set down. These mathematical models have been derived from electrical circuits representing the main features used by the nerve cell for electrical signalling. In these models, all of the important functional properties of the neurone are represented by an electrical circuit called an "equivalent circuit" consisting only of conductances, batteries, and capacitors so that the behaviour of these models matches as many real features and as much of the behaviour of the neurone as possible, and can simulate the processes that are used by neurones for signalling (Hodgkin and Huxley, 1952 and Getting, 1989).

In this study we used realistic models of neurone to provide the data sets for analysis.

In the simulations 60 second records of repetitive firing have been generated (see appendix "A"), with the firing times of the input and output spike times recorded as the quantities available for analysis (Halliday, 1994). These record lengths are similar to those used in experimental studies. At each step of the simulation, parameters were chosen so that the behaviour of the model matched comparable experimental data. Once parameters were chosen they could not vary at subsequent steps. This procedure helped to ensure that the final set of parameters matched experimental data at the cellular and synaptic levels. Appendix (A) illustrates a full description of the model and its electrical equivalent circuit.

Chapter 2

2 Review of Stochastic Point Processes Techniques

2.1 Introduction

Neurophysiology is an area where both ordinary time series and point processes may play an important role. The study of the behaviour of small networks of neurones is needed to determine the strength of association between component neurones, or an assessment of the timing relations between them.

The short duration of an action potential or spike compared with the random intervals between successive pulses provides the basis for considering the spike train as a realisation of a point process along a line. This process is described fully by an ordered sequence of the realised times

$$\ldots\tau_{-3} < \tau_{-2} < \tau_{-1} < \tau_0 < \tau_1 < \tau_2 < \tau_3 \ldots,$$

of occurrence of the spikes (Cox and Lewis 1966).

The main object of this chapter is to review some existing techniques applied to the study of neuronal networks, namely time and frequency domain

analyses. Such approaches involve both point process and time series methods. The reason for presenting these (already) existing techniques is to demonstrate their points of weakness and their limitations, if any. It has been recognised that in some situations, the results obtained by using these techniques are difficult to explain and interpret, and in some other situations are found to give results which are inconsistent with the underlying biological concepts and interpretations. This is on the one hand, and on the other it is important to highlight the necessity to introduce other alternative techniques, such as the one we are about to introduce and discuss throughout this thesis, namely the likelihood approach, in which some extra information could be provided. Also other physiological behaviour of the nerve cell can be studied more precisely, for instance the recovery process of the cell, an estimate of the firing probability at any given time and the general effect of the inputs on subsequent outputs. Another reason for presenting this chapter is to provide the basis needed for comparison between these techniques and the likelihood technique.

A wide variety of examples of point processes are discussed in Lewis (1972a); and Snyler (1975) and more recently in Brillinger (1978a); Cox and Isham (1980); and Daley (1988).

2.2 Brief History of Point Processes

A stochastic point process is defined as a random non-negative integer-valued measure. In practice this gives the times of occurrence of the events. Examples of stochastic point processes are vast : queues, neuronal electrical activity, heartbeats, radio-activity, seismology, population growth and many others. For more details see Lewis (1972).

The frequency domain analysis of signals is thought to have been introduced as early as the middle of the 17th century when Isaac Newton decomposed a light signal into separate components by passing the light

through a glass prism. The earliest study of point processes was begun by Graunt, (1620-1674) when constructing life tables in the study of populations, such tables correspond to the superposition of many independent point processes each with a single point at the time of death of an individual. For an early history see Westgaard (1968).

The Poisson point process was introduced early in the 18th century with the discovery of the Poisson distribution credited to DeMoivre in 1718 and then to the French mathematician and physicist Simeon Denis Poisson (1781-1840) who published it in 1837. Another era for Poisson point processes started when Boltzman in 1868 introduced the expression $exp(-\mu t)$ for the probability of no events in an interval of length t. Bateman (1910) in the analysis of radio-active problems noticed that the number of particles recorded in fixed time intervals satisfies a simple set of differential equations, and the solution of those equations were Poisson probabilities Height (1967). Erlang (1909) made extensive use of such processes in the study of traffic systems, then to telephone systems to determine the optimal number of circuits and also the study of queuing systems with input-output point processes corresponding to times of arrival and departure of customers. The latter was further developed by Khinchine et al (Bhat, 1969). For a historical review, see Height (1967).

Another class of point process is the renewal process, in which the intervals between successive events are independently distributed. Substantial developments of point processes took place in the late 1930's by physicists in the study of the decay of radio-activity.

More recently point process analysis techniques have been used in the fields of seismology and neurophysiology; in particular the study of the behaviour of small networks of neurones. Such processes have been found useful in assessing the degree of association between neuronal signals and in the estimation of biologically meaningful parameters; more details can be

found in Amjad (1989); Brillinger (1988); Halliday (1986); and Rosenberg et al (1989).

2.3 The Basic Assumptions and Notation

A realisation (or a sample) of a point process may be represented by a counting measure denoted as

$$N(t) = \#\left\{\tau_j ;\ 0 < \tau_j \le t\right\};$$

where $\#\{.\}$ indicates the number of events in the interval $(0,\ t]$ and $\{\tau_j\}$ the set of spike times in the sample. Before defining the different parameters of point process it is relevant to introduce differential increments of the process N defined as

$$dN(t) = N(t,\ t+dt],$$

giving the number of events in a small interval $(t,\ t+dt]$ of duration dt.

2.3.1 Stationarity

The point process $N_a\ (a = 1,\ 2,\ \ldots,\ k)$ is said to be first order stationary if the probability distribution of the number of events $N(t,\ t+h]$ is same as that of the number of events $N(t+\tau,\ t+h+\tau]$ $\forall$ t, τ, and $h>0$, and it is said to be second order stationary (or weakly stationary) if the joint probability distribution of the number of events $N(t,\ t+h_1]$ and $N(t+h_2,\ t+h_3]$ is the same as that of the number of events $N(t+\tau,\ t+h_1+\tau]$ and $N(t+h_2+\tau,\ t+h_3+\tau]$. The point process N_a is said to be completely stationary (strong stationarity) if the joint probability distribution of the number of events in any arbitrary number of intervals is invariant under translation.

2.3.2 Orderliness

A point process N_a is said to be orderly if the probability of two or more events occurring in the interval $(t, t+h)$ tends to zero as $h \to 0$. This avoids the occurrence of more than one event in small intervals, in other words that the points of the process N_a do not occur simultaneously.

2.3.3 Strong Mixing

We say that a point process satisfies a (strong) mixing condition if events of the same process well-separated in time are independent. This condition can also be applied to the case of multivariate point processes by assuming that events of one point process separated in time by a distance v from the events of the other point process become independent for large values of v. More details can be found in Cox and Lewis (1972); Srinivasan (1974); Brillinger (1975b); and Cox and Isham (1980).

2.4 Stochastic Point Process Parameters

The field of neurophysiology provides a rich source of data which can be analysed within the framework of point process theory. This theory consists of two parallel approaches, namely, the time domain approach and the frequency domain approach. These two approaches are mathematically equivalent and are assumed to represent the data equivalently and also to contain the same information about the process under investigation (Tukey 1978; and Koopmans 1983). But since the volume of data is finite, mathematical equivalence does not imply equivalent representation. Therefore it is valuable to have both analyses, which illustrate different aspects of the process.

The main aim of this section is to introduce and define certain parameters of point processes in both domains and without any details of properties or estimation procedures. Further details of the estimation procedures, properties and computational procedures with applications can be found in Rigas (1983); Halliday (1986); and Amjad (1989).

2.4.1 Parameters in the Time Domain

It will be assumed that the assumptions of first order stationarity, orderliness and (strong) mixing hold in this section.

(a) The Univariate Point Process

The mean intensity of the process N is defined as

$$P_N = \lim_{h \to 0} Pr\{N \text{ event in } (t, t+h]\} / h \qquad .....(2.4.1)$$

and since the process is orderly, we have

$$P_N = E\{ dN(t) \} / dt \qquad .....(2.4.2)$$

where the mathematical expectation $E\{.\}$ denotes the averaging operator of a random variable.

The product density of order 2 for a stationary point process measures the intensity with which events separated by v time units occur, and is given by

$$P_{NN}(v) = \lim_{h_1, h_2 \to 0} Pr\{ N \text{ event in } (t, t+h_1] \text{ and } N \text{ event in } (t+v, t+v+h_2] \} / h_1 h_2 \qquad .....(2.4.3)$$

and since the process is orderly, we have

$$P_{NN}(v) = E\{dN(t+v)\, dN(t)\} \;/\; dv\, dt \;;\; v \neq 0. \qquad .....(2.4.4)$$

Under the (strong) mixing condition, we have

$$\lim_{v \to \infty} P_{NN}(v) = P_N P_N = \left[P_N\right]^2. \qquad .....(2.4.5)$$

This leads to another function called the second-order cumulant function, given by

$$q_{NN}(v) = P_{NN}(v) - P_N P_N \qquad .....(2.4.6)$$

which has the property that it tends to zero as v increases, i.e.

$$\lim_{v \to \infty} q_{NN}(v) = 0. \qquad .....(2.4.7)$$

A conditional auto-intensity function is defined as

$$m_{NN}(v) = E\{dN(t+v) \mid dN(t) = 1\} \;/\; dv; \quad v \neq 0 \qquad .....(2.4.8)$$

and may be interpreted as

$$m_{NN}(v) = \lim_{h \to 0} Pr\{N \text{ event in } (t+v,\, t+v+h] \mid N \text{ event at } t\} \;/\; h \qquad .....(2.4.9)$$

or, in terms of expected values as

$$\begin{aligned} m_{NN}(v) &= E\{dN(t+v) \mid dN(t) = 1\} / dv \\ &= P_{NN}(v) / P_N \;; \quad v \neq 0 \end{aligned} \qquad .....(2.4.10)$$

and under the (strong) mixing condition, we have

$$\lim_{v \to \infty} m_{NN}(v) = P_N . \qquad \text{.....(2.4.11)}$$

(b) <u>The Bivariate Point Process</u>

Let $\Psi(t) = \{M(t), N(t)\}$ be a real-valued first order stationary bivariate point process which satisfies the assumptions of orderliness and (strong) mixing, then the second-order cross-product density at lag v, $P_{NM}(v)$, is defined as

$$P_{NM}(v) = \lim_{h_1, h_2 \to 0} Pr\{M \text{ event in } (t, t+h_1] \text{ and } N \text{ event in } (t+v, t+v+h_2]\} / h_1 h_2 \qquad \text{.....(2.4.12)}$$

or, in terms of expected values

$$P_{NM}(v) = E\{dN(t+v)\, dM(t)\} / dv\, dt \qquad \text{.....(2.4.13)}$$

Under the assumption of a (strong) mixing condition, and as v becomes large, increments of the process become independent, i.e.

$$\lim_{v \to \infty} P_{NM}(v) = P_N P_M . \qquad \text{.....(2.4.14)}$$

This leads to another function called the second-order cumulant function, given by

$$q_{NM}(v) = P_{NM}(v) - P_N P_M \qquad \text{.....(2.4.15)}$$

which has the property that it tends to zero as v increases, i.e.

$$\lim_{v \to \infty} q_{NM}(v) = 0. \qquad \text{.....(2.4.16)}$$

The cross-intensity function is the traditional method that has been a widely used means for assessing the timing relations between processes. It is also thought to be a good way of characterising the nature of the association between the processes, i.e., the effects of one process on the other are excitatory or inhibitory (Amjad, 1989). The cross-intensity function denoted by $m_{NM}(v)$ is defined as

$$m_{NM}(v) = \frac{P_{NM}(v)}{P_M}. \qquad \text{.....(2.4.17)}$$

Under the assumption of a (strong) mixing condition, and as v becomes large, it also follows that,

$$\lim_{v \to \infty} m_{NM}(v) = P_N. \qquad \text{.....(2.4.18)}$$

The cross-intensity function, $m_{NM}(v)$ can be estimated as

$$\hat{m}_{NM}(v) = \frac{J_{NM}^{(T)}(v)}{h\,M(T)} \qquad \text{.....(2.4.19)}$$

where

$$J_{NM}^{(T)}(v) = \#\left\{(s_k, r_j) : v - (h/2) < r_j - s_k < v + (h/2)\right\} \qquad \text{.....(2.4.20)}$$
$$; \; k = 1, 2, \ldots, M(T) \text{ and } j = 1, 2, \ldots, N(T),$$

and h is a binwidth. The symbol $\#\{W\}$ denotes the number of events in set W, $\{r_j\}$ and $\{s_k\}$ are the observed times of the events of the processes N and M respectively, and $N(T)$ and $M(T)$ are the number of events of processes N and M occurring in time interval (0, T), respectively.

The variables $J_{NM}^{(T)}(v)$ given by expression (2.4.20) are asymptotically independent Poisson random variables (Brillinger, 1976a) with mean

$$(hT)P_{NM}(v) \text{ as } T \to \infty .$$

This implies, for large T, that

$$\hat{m}_{NM}(v) \sim \frac{1}{h T P_M} Po\left[(hT) P_{NM}(v)\right] \qquad .....(2.4.21)$$

where $Po[\lambda]$ denotes a Poisson random variable with mean λ.

Now, for small h and large T, the estimate given in expression (2.4.21) will be approximately normal (Brillinger, 1976a), i.e. ,

$$\hat{m}_{NM}(v) \sim N\left[m_{NM}(v), \frac{m_{NM}(v)}{h T P_M}\right]. \qquad .....(2.4.22)$$

The variance of $\hat{m}_{NM}(v)$ may be stabilised by applying a square root transformation (Kendall & Stuart, 1983), and the distribution of the square root of the estimated cross-intensity function will take the form,

$$\left[\hat{m}_{NM}(v)\right]^{1/2} \sim N\left[\left(m_{NM}(v)\right)^{1/2}, \frac{1}{4 h T P_M}\right]. \qquad .....(2.4.23)$$

An approximate confidence interval for the cross-intensity function estimated using (2.4.23) at a given lag value v under the hypothesis that the two processes are independent can then easily be constructed, with an approximate 95% confidence interval is given by

$$[\hat{P}_N]^{1/2} \pm 1.96\ [4 hT \hat{P}_M]^{-1/2}. \qquad .....(2.4.24)$$

2.4.2 Parameters in the Frequency Domain

It has been mentioned earlier that both time and frequency domain analyses are in some sense mathematically equivalent procedures. The use of the frequency domain has been emphasised (Amjad, 1989) because it reveals more features about the process and might give better understanding to physiologists. But it is valuable, however, to have both analyses.

(a) The Univariate Point Process

Suppose $N(t)$ is a first order stationary point process satisfying the conditions of orderliness and (strong) mixing. Let P_N be the mean rate of process N. Further, suppose that the second order cumulant function $q_{NN}(v)$ as defined in (2.4.6) exists (Amjad, 1989) and satisfies the condition

$$\int_{-\infty}^{\infty} \left| q_{NN}(v) \right| dv < \infty . \qquad \text{.....(2.4.25)}$$

Then the power spectrum of the point process N is defined as

$$f_{NN}(\lambda) = \frac{P_N}{2\pi} + \frac{1}{2\pi} \int_{-\infty}^{\infty} exp(-i\lambda v)\, q_{NN}(v)\, dv \; ; \; -\infty < \lambda < \infty \qquad \text{.....(2.4.26)}$$

and may be interpreted as reflecting the power in each frequency λ of the process (Brillinger et al, 1976).

Using the condition given in equation (2.4.25) the limiting behaviour of the power spectrum of the point process takes the form

$$\lim_{\lambda \to \infty} f_{NN}(\lambda) = \frac{P_N}{2\pi} . \qquad \text{.....(2.4.27)}$$

(b) The Bivariate Point Process

Let $\Psi(t) = \{M(t), N(t)\}$ be a first order stationary bivariate point process which satisfies the conditions of (strong) mixing and orderliness. Suppose the cross-cumulant function, $q_{NM}(v)$, exists and has the form as given in equation (2.4.15) above. Then the cross-spectrum between the two point processes at frequency λ denoted by $f_{NM}(\lambda)$ is defined as

$$f_{NM}(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} exp(-i\lambda v)\, q_{NM}(v)\, dv \;\; ; \;\; -\infty < \lambda < \infty \qquad .....(2.4.28)$$

which can be written as a complex-valued function of the form

$$f_{NM}(\lambda) = [Re\, f_{NM}(\lambda)] + [Im\, f_{NM}(\lambda)] \qquad .....(2.4.29)$$

where $Re\, f_{NM}(\lambda)$ and $Im\, f_{NM}(\lambda)$ are the real and imaginary parts of $f_{NM}(\lambda)$ respectively.

The cross-spectrum between the two point processes at frequency λ, $f_{NM}(\lambda)$, may be interpreted as measuring the association of the processes N and M at frequency λ (Brillinger et al, 1976), i.e.

$$f_{NM}(\lambda) = 0$$

indicates no relationship at frequency λ. Also it has the property that

$$f_{NM}(\lambda) = f_{NM}(-\lambda) \qquad .....(2.4.30)$$

also,

$$\lim_{\lambda \to \infty} f_{NM}(\lambda) = 0. \qquad .....(2.4.31)$$

Further details can be found in Cox and Lewis (1966); Brillinger et al (1976); Amjad (1989); and Rosenberg et al (1989).

The coherence, $\left| R_{NM}(\lambda) \right|^2$ between two point processes N and M at frequency λ can be defined as

$$\left| R_{NM}(\lambda) \right|^2 = \frac{\left| f_{NM}(\lambda) \right|^2}{f_{NN}(\lambda)\, f_{MM}(\lambda)} \qquad .....(2.4.32)$$

and hence it is a normalised cross-spectrum which provides an absolute measure of association and it can be easily shown to be bounded by 0 and 1, i.e.

$$0 \leq \left| R_{NM}(\lambda) \right|^2 \leq 1,$$

where the value $\left| R_{NM}(\lambda) \right|^2 = 0$ indicates no association between the two processes N and M at frequency λ, whereas $\left| R_{NM}(\lambda) \right|^2 = 1$ corresponds to perfect linear association at frequency λ.

Using the respective estimates of expression (2.4.32) above, an estimate of the coherence may be obtained as

$$\left| R_{NM}^{(T)}(\lambda) \right|^2 = \frac{\left| f_{NM}^{(T)}(\lambda) \right|^2}{f_{NN}^{(T)}(\lambda)\, f_{MM}^{(T)}(\lambda)} \qquad .....(2.4.33)$$

whereas an approximate 95% confidence interval for the coherence at

frequency λ may given as

$$\left|R_{NM}^{(T)}(\lambda)\right|^2 \pm 1.96 \left\{ \frac{\left| R_{NM}^{(T)}(\lambda) \right|^2 \left[1 - \left|R_{NM}^{(T)}(\lambda)\right|^2 \right]^2}{(L/2)} \right\}^{1/2} \quad ...(2.4.34)$$

where L is the number of disjoint periodogram sections from records of duration T.

The phase-spectrum defined as the argument of the cross-spectrum is given by

$$\phi_{NM}(\lambda) = arg\left\{ f_{NM}(\lambda) \right\} = tan^{-1}\left[\frac{Im\, f_{NM}(\lambda)}{Re\, f_{NM}(\lambda)} \right] \quad .....(2.4.35)$$

assuming that $f_{NN}(\lambda)$ and $f_{MM}(\lambda)$ are non-zero. The phase-spectrum may be used to assess the timing relations between the processes N and M.

Suppose that (s_k, r_j) represent the spike times for the bivariate process $\{M(t), N(t)\}$, then the cross-spectrum between N and M is given by

$$f_{NM}(\lambda) = \lim_{T\to\infty} \frac{1}{2\pi T} E\left(\sum_k exp(-i\,\lambda\, s_k) \right) \left(\sum_j exp(i\,\lambda\, r_j) \right). \quad .....(2.4.36)$$

If process N is a lagged version of process M with lag τ, i.e.,

$$s_k = r_k + \tau \;; \quad k = 1,2,...,M(T),$$

then the cross-spectrum can be written as

$$f_{NM}(\lambda) = \lim_{T\to\infty} \frac{1}{2\pi T} E\left(\sum_k exp(-i\,\lambda(r_k+\tau))\right)\left(\sum_j exp(i\,\lambda\, r_j)\right)$$
$$= exp(-i\,\lambda\tau)\, f_{MM}(\lambda) \qquad .....(2.4.37)$$

which implies that

$$\phi_{NM}(\lambda) = -\lambda\tau. \qquad .....(2.4.38)$$

where expression (2.4.38) above shows that in the case of a pure delay, the phase $\phi_{NM}(\lambda)$ is a linear function of frequency λ with $-\tau$ being the slope of the line.

The application of expression (2.4.38) is useful for the large number of cases where the relation between two spike-trains can be assumed to be dominated by a delay which may be estimated as the slope of the least squares line fitted to the estimated phase curve. Whenever there is a delay, the coherence between input and output processes is not constant, and consequently a weighted least squares procedure can be used to estimate the delay and its standard error (Rosenberg et al, 1989).

Let $\phi_{NM}^{(T)}(\lambda_i) = \phi_i$ be the estimated phase evaluated at discrete frequencies of the form $\lambda_i = 2\pi i / T\,; i = 1, 2, ..., n$, we define a regression model through the origin of the following form

$$\phi_i = \beta\lambda_i + \varepsilon_i$$

where $\beta = -\tau$, and ε_i are approximately normally distributed with mean zero

and variance σ_i^2, and covariance σ_{ij} where

$$\sigma_{ij} \cong 0 \text{ for } i \neq j.$$

and

$$\sigma_i^2 = \frac{1}{2L}\left[\frac{1}{\left|R_{NM}(\lambda_i)\right|^2} - 1\right]$$

where L is the number of disjoint periodogram sections from records of duration T.

The weighted least squares estimate of β is given by

$$\hat{\beta} = \frac{\sum w_i \phi_i \lambda_i}{\sum w_i \lambda_i^2}$$

which is an unbiased estimate of β, and an estimate of its variance is given by

$$\hat{var}\left(\hat{\beta}\right) = \frac{\hat{\sigma}^2}{\sum w_i \lambda_i^2}$$

where

$$\hat{\sigma}^2 = \frac{\sum w_i \left(\phi_i - \hat{\beta}\lambda_i\right)^2}{n-1}.$$

A plausible choice for the weights is to take w_i as (e.g. Weisberg, 1985, p85)

$$w_i = \frac{1}{\hat{\sigma}_i^2} = 2L\left[\frac{1}{\left|R_{NM}^{(T)}(\lambda_i)\right|^2} - 1\right]^{-1}.$$

Applying standard regression theory, an approximate 95 % confidence interval for the delay is then

$$-\hat{\beta} \pm 1.96 \left[\hat{var}\left(\hat{\beta}\right) \right]^{1/2}$$

or

$$-\hat{\beta} \pm 1.96 \frac{\hat{\sigma}}{\sqrt{\sum w_i \lambda_i^2}} \qquad \text{.....(2.4.39)}$$

(c) **The Multivariate Point Process**

Let $\Psi(t) = \{M(t), N(t), L(t)\}$ be a first order stationary multivariate point process which satisfies the conditions of (strong) mixing and orderliness. The partial coherence of order 1 between two processes N and M after removing the linear effects of the third process L may be defined as;

$$\left| R_{NM.L}(\lambda) \right|^2 = \frac{\left| R_{NM}(\lambda) - R_{NL}(\lambda) R_{LM}(\lambda) \right|^2}{\left[1 - \left| R_{NL}(\lambda) \right|^2 \right] \left[1 - \left| R_{LM}(\lambda) \right|^2 \right]} \qquad \text{.....(2.4.40)}$$

which satisfies the property that it is bounded by 0 and 1, i.e.

$$0 \leq \left| R_{NM.L}(\lambda) \right|^2 \leq 1$$

and with zero corresponding to the situation where the relation between the processes N and M is entirely accounted for by taking into account their individual dependencies on process L.

The multiple coherence at frequency λ between an output point process N and input processes M and L may be defined as;

$$\left|R_{N.ML}(\lambda)\right|^2 = \left|R_{NL}(\lambda)\right|^2 + \left|R_{NM.L}(\lambda)\right|^2 \left\{1-\left|R_{NL}(\lambda)\right|^2\right\} \quad ...(2.4.41)$$

which satisfies the property that it is bounded by 0 and 1, i.e.

$$0 \leq \left|R_{N.ML}(\lambda)\right|^2 \leq 1 ,$$

also giving an interpretation of $\left|R_{N.ML}(\lambda)\right|^2$ as a measure of the linear predictability of the point process N from the other two processes M and L (Brillinger, 1975b; and Jenkins and Watt, 1968).

2.5 Applications

We now demonstrate some of the above-mentioned procedures with two simple examples using both time and frequency domain measures from simulated neuronal spike train data generated by a conductance based neuronal model (Halliday, 1994). For more simulation details see chapter 1, section 1.6, and further details are present in appendix (A).

Fig.2.5.1a and b are the inter-spike interval histograms of the input and of the output, respectively, and suggest approximately an exponential distribution (i.e. a Poisson process) of the input and normal distribution of the output process. Fig.2.5.1c represents the square root of the estimated cross-intensity function as a time domain measure of the association between the two processes and suggests that an excitatory effect of an input spike lasts about four msec, and then followed by a significant decrease in the probability of an output spike from 5-15 msec. The latter seems to indicate an inhibitory effect centred at around 10 msec. No such inhibition was present in the simulation and this behaviour complicates the interpretation of the cross-intensity function and gives an indication that it may lead to misleading results.

Fig.2.5.2a, b correspond to the estimates of coherence and phase, respectively, obtained from the same data. Both estimates have been plotted against the frequencies in the form (1000 J/R) Hz, with R=1024, J=1,2,..., over the range 0-250 Hz. The dotted line in Fig.2.5.2a at each frequency represents the upper 95% confidence interval of the null distribution under the hypothesis of the processes being independent at that frequency. Fig.2.5.2a gives clear evidence that both processes are coupled over the range of about 0-100 Hz.

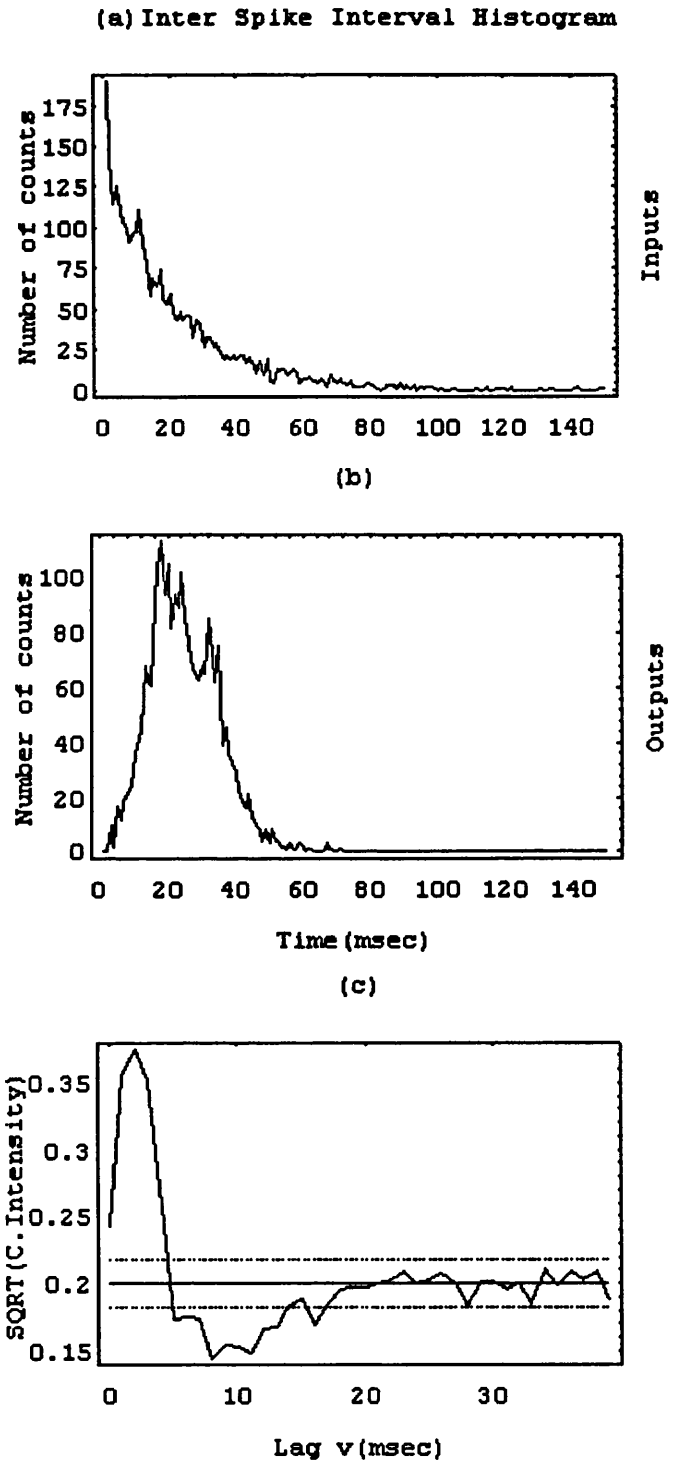


Fig.2.5.1 **a)** and **b)** Input and output inter-spike interval histograms, respectively. **c)** Square root of the estimated cross-intensity function. The horizontal solid line represents the corresponding square root of the estimated output mean rate $\sqrt{\hat{P}_N}$ whereas the dotted lines below and above this line give an approximate 95% confidence interval for the square root of the estimated cross intensity under the hypothesis that the two processes are independent.

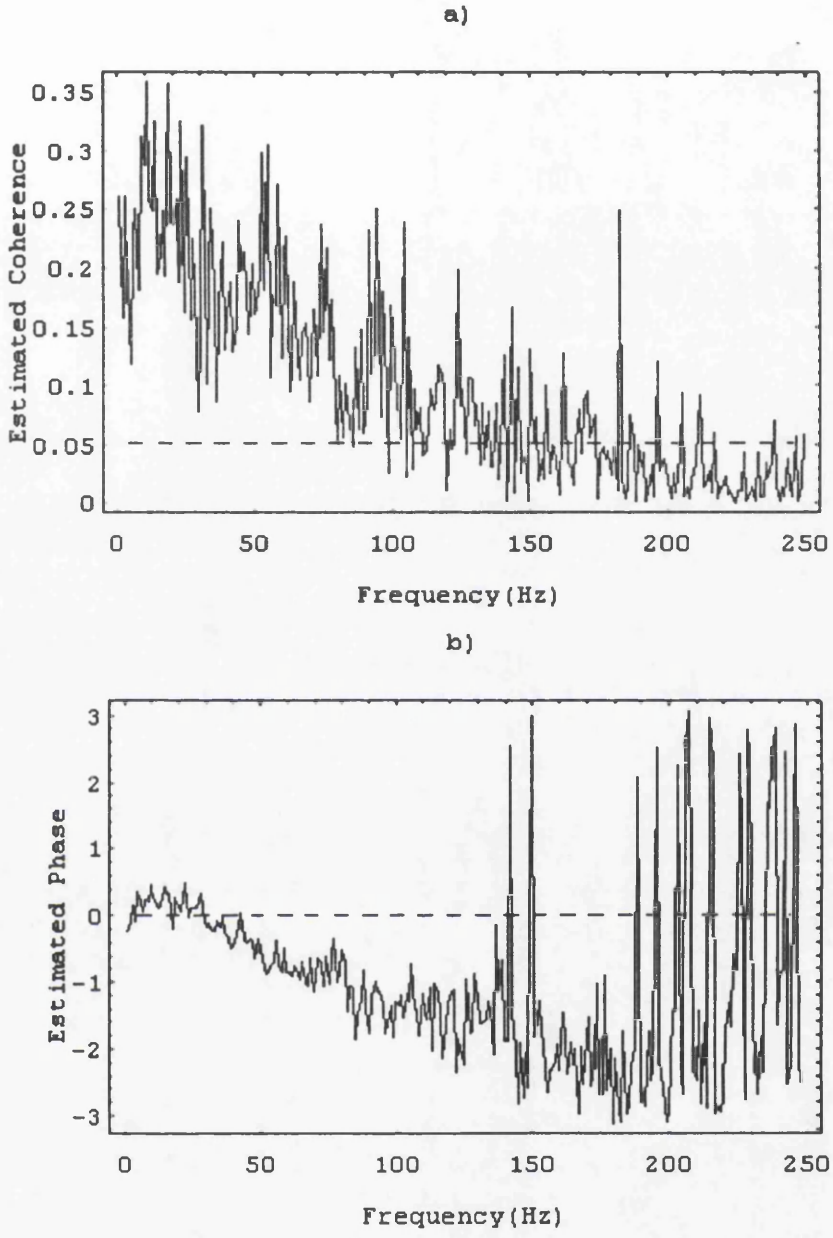


Fig.2.5.2 Illustration of the Coherence and Phase.

a) Estimate of the ordinary coherence; the dotted line represents the upper limit of the 95% confidence interval (marginal) for the coherence under the hypothesis that the two processes are independent.

b) Estimate of the phase.

Fig.2.5.2b gives a clear indication that the estimate of the phase can be approximated by a straight line over a range of frequencies at which the estimated coherence is significant, and thus that a pure time delay is present.

The procedures mentioned above have been applied to another set of simulated neuronal spike train data similar to the previous one except that in these data the input is approximately normally distributed.

Fig.2.5.3a, b are the inter-spike interval histograms of the input and output, respectively, and suggest that both have very approximately a Normal distribution structure. Fig.2.5.3c represents the square root of the estimated cross- intensity function and suggests an excitatory effect of an input spike lasting about 3 msec, followed by a significant decrease in the probability of an output spike from about 4 - 12 msec, and then followed by another excitatory effect from about 16 - 21 msec, and that may suggest a possible periodic response of the output associated with the input. No such feature was present in the simulation, and that again seems to complicate the interpretation of the estimated cross-intensity function and put a question mark on its performance.

Fig.2.5.4a, b correspond to the estimates of coherence and phase, respectively, obtained from the same data set. The coherence indicates clearly that the two processes are strongly associated with each other over the range of 0-200 Hz. Fig.2.5.4b represents the estimated phase which shows that the estimate can be approximated by a straight line over a range of frequencies at which the estimated coherence is apparently significant, and thus that a pure delay is present.

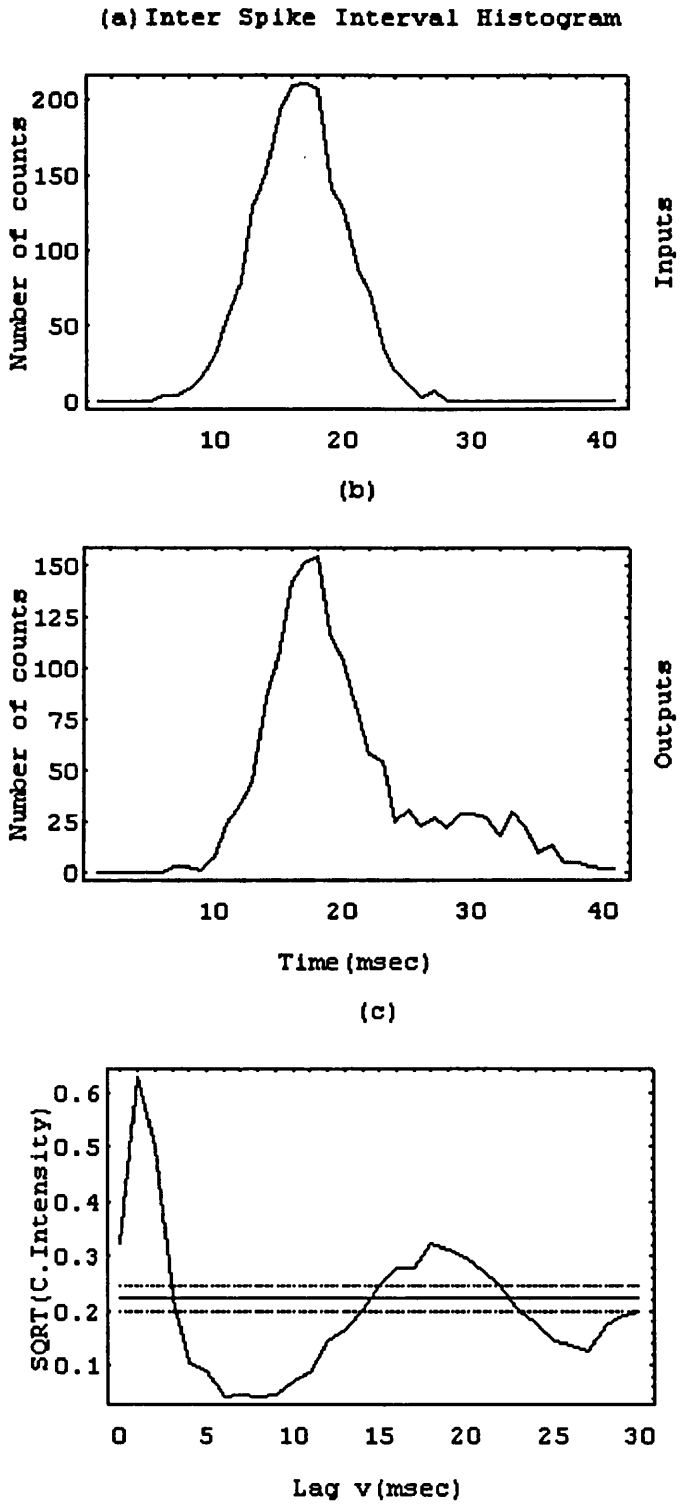


Fig.2.5.3 **a)** and **b)** Input and output inter-spike interval histograms, respectively. **c)** Square root of the estimated cross-intensity function. The horizontal solid line represents the corresponding square root of the estimated output mean rate $\sqrt{\hat{P}_N}$ whereas the dotted lines below and above this line give an approximate 95% confidence interval for the square root of the estimated cross intensity under the hypothesis that the two processes are independent.

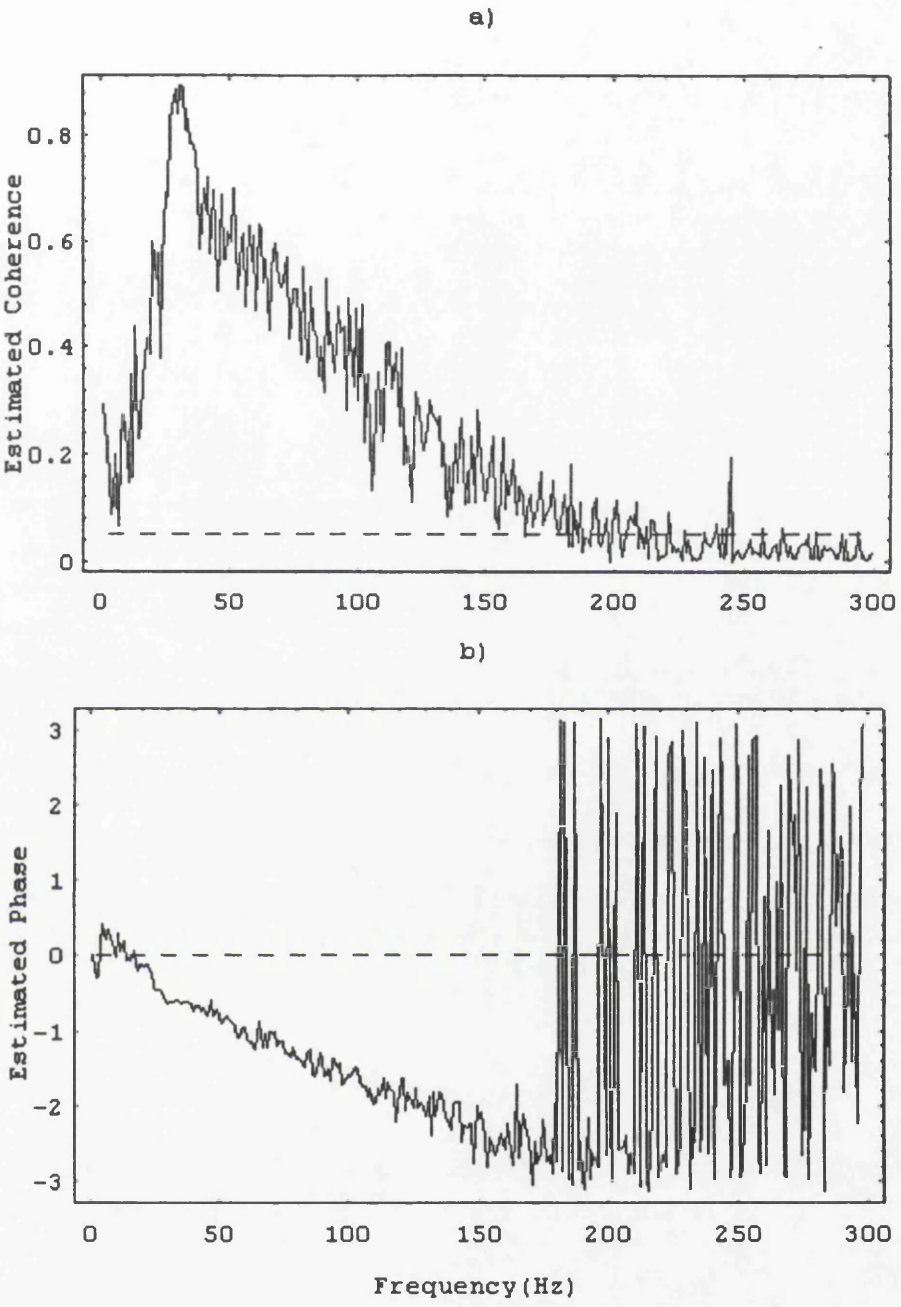


Fig.2.5.4 Illustration of the Coherence and Phase.

a) Estimate of the ordinary coherence, the dotted line represents the upper limit of the 95% confidence interval (marginal) for the coherence under the hypothesis that the two processes are independent.

b) Estimate of the phase.

2.6 Conclusion

From the two sets of data we have seen, it is clear that the square root of the estimated cross-intensity function as a time domain measure of the association between the two processes need not be very helpful, and may indeed be very misleading and fail to reflect the actual properties of the simulation. It is presumably combining information about direct synaptic effects and intrinsic membrane properties, in other words, it seems to combine input information and internal cell recovery information in a way which makes it very difficult to interpret and as a consequence of that it may produce misleading results. This leads us to investigate an alternative approach based on likelihood which will be introduced in chapter 3.

Chapter 3

3 The Likelihood Approach

3.1 Introduction

One of the two most general methods of statistical estimation so far known is the method of maximum likelihood (the other being the method of least squares). The method of maximum likelihood was initially formulated by C. F. Gauss but as a general method of estimation was first introduced by R. A. Fisher and later on developed by him in a series of papers. The historical development of linear models dealing with maximum likelihood and least squares from Gauss and Legendre to Fisher has previously been sketched (McCullagh and Nelder, 1992). For further historical information concerning the development of probability and statistics in general up to the beginning of the twentieth century, see the book by Stigler (1986). We start now by defining the likelihood function.

3.2 Likelihood Function

The likelihood function of n random variables $X_1, X_2, ..., X_n$ is defined to be (proportional to) the joint probability density of the n random variables, say,

$f_{X_1, X_2, ..., X_n}(x_1, x_2, ..., x_n; \underline{\theta})$, which is considered to be a function of unknown parameters, $\underline{\theta}$. In particular, if $X_1, X_2, ..., X_n$ is a random sample of size n from a population with density function $f_X(x; \underline{\theta})$, then the likelihood function denoted by $l_0(\underline{\theta}; \underline{x})$, is their joint density function; given by

$$l_0(\underline{\theta}; \underline{x}) = l_0 = f(x_1; \underline{\theta}).f(x_2; \underline{\theta}) ...f(x_n; \underline{\theta}) = \prod_{i=1}^{n} f(x_i; \underline{\theta}).$$

3.3 Maximum Likelihood Estimator(M. L. E.)

The principle of maximum likelihood estimation consists in finding an estimator for the parameters $\underline{\theta}$ which maximises l_0 for variations in the parameter. Thus if there exists a function $\hat{\underline{\theta}} = \hat{\underline{\theta}}(x_1, x_2, ..., x_n)$ of the sample values which maximises l_0 for variations in the parameter $\underline{\theta}$, then $\hat{\underline{\theta}}$ is to be taken as an estimator of $\underline{\theta}$. And $\hat{\underline{\theta}}$ is usually called the Maximum Likelihood Estimator (M.L.E.). Thus $\hat{\underline{\theta}}$ is the solution if any of

$$\frac{\partial l_0}{\partial \theta_i} = 0 \quad and, \quad \frac{\partial^2 l_0}{\partial \theta_i \, \partial \theta_j} < 0 \qquad .....(3.3.1)$$

where $i \neq j$; $i, j = 1, 2, ..., k$.

The first of the two equations given in (3.3.1) above can be rewritten as

$$\frac{1}{l_0}\frac{\partial l_0}{\partial \theta_i} = 0 \Rightarrow \frac{\partial \log l_0}{\partial \theta_i} = 0 \ ; \ i = 1, 2, ..., k \qquad .....(3.3.2)$$

a form which is much more convenient from a practical point of view. Equations (3.3.2), are usually referred to as the likelihood equations. One reference to the general properties of maximum likelihood estimation is the book by C. R. Rao (1973).

3.4 Likelihood Analysis of Spike Trains

3.4.1 Introduction

The study of the interrelationships within a network of neurones given the individual occurrence times (firing times) has long been of concern (Bryant et al, 1973; Knox et al, 1977; Brillinger et al, 1976; and Borisyuk et al, 1985).

It is known that either a neurone is firing spontaneously or that it is firing under the influence of other neighbouring neurones. Suppose that the firing times of the neurones present are available. A conceptual model that may reflect the main aspects of neurone firing behaviour based on likelihood can then be constructed and studied both theoretically and empirically. This type of model will then enable maximum likelihood estimates of internal quantities to be calculated, such as the post-synaptic potentials of the measured influencing neurones, the membrane potentials, the absolute threshold and the recovery process of the neurone, and may also be able to estimate the effects of unobservable neurones on the firing of others. The power of a likelihood approach is that it allows direct biological interpretation of the results obtained, it is a highly flexible approach in that it allows spontaneous firing and it also extends to the case of an arbitrary number of neurones. Approximate expressions for standard errors of the estimates are always available.

Many stochastic models have been set down to specify the firing of a neurone and to assess the interrelations of firing within moderate sized networks of neurones, but most of these models have only been used extensively theoretically. More details and reviews of these models can be found in Holden (1976); Knight (1972); and more recently Brillinger (1988 and 1992). The only occasion where a specific data set has been used is in Brillinger and Segundo (1979).

Throughout this thesis, the likelihood approach will be discussed and investigated in much greater depth both theoretically and empirically and

applied to many different neuronal spike train and muscle spindle data sets, simulated as well as real, each with different features, in order to demonstrate the informative results the approach can provide. Furthermore, we are introducing new ideas suggested by physiologists on the basis that these may have an important role in describing the firing process of a neurone. For example, one can add the effects of "unobservable" neurones that influence the firing of a particular neurone or a term that allows an exponentially decaying neurone threshold. The flexibility of being able to add new terms to our model shows another advantage of the likelihood approach. The results obtained via likelihood can therefore provide a realistic basis and background and so be helpful to the physiologists who may then better understand the processes involved.

3.4.2 An Analytic Model

The construction of a mechanistic model of the processes governing the firing of a neurone needs an understanding of some of the biological concepts involved. One process by which one neurone influences the firing of a second has been described with more detail in chapter 1.

The biological process may be put into an analytic form and may described formally as follows. Let us consider, for simplicity, the spike trains of two neurones (it may then extend to the case of an arbitrary number of neurones) X and Y which are described by counting measures $X(t)$ and $Y(t)$, respectively. The number of spikes of the neurone X that occur in the time interval $(0, t]$ is given by $X(t)$, and that of the neurone Y is given by $Y(t)$. These are analogous to the counting measure $N(t)$ defined in section (2.3). Now, suppose we are interested in the firing of neurone Y. Suppose that the neurone X fired at time τ. Let $a(t-\tau)$ represent its effect on the potential at time t on

the trigger zone (axon regions which have the lowest threshold for generating an action potential) of the neurone Y. For simplicity, let its evolution be assumed linear (although there may however be non-linear effects present). References suggesting this assumption is reasonable are given in Stevens (1968); Langmoen and Anderson (1983), and more recently; Brillinger (1988). The quantities $\{a(.)\}$, will be called the summation function, which represents the effects of a neurone X (input) on the firing of a neurone Y (output), and describes the course that the potential would follow after a current impulse. The linearity assumption implies that the effects of current pulses at different times are additive. Let $\gamma(t)$ denote the time elapsed at time t since the neurone Y last fired. At this point, let us assume that only inputs occurring after the previous output have any effect, then the membrane potential $U(t)$ at its trigger zone may be represented as

$$U(t) = \int_0^{\gamma(t)} a(u)\, x(t-u)\, du. \qquad \text{.....(3.4.1)}$$

The neurone Y tends to fire when the post-synaptic potential at its trigger zone exceeds an extant level called threshold. Now, let $\theta(t)$ denote the threshold potential level at the trigger zone at time t and assume that it has the form

$$\theta(t) = \theta^*(t) + \varepsilon(t) \qquad \text{.....(3.4.2)}$$

with $\varepsilon(t)$ the noise, which includes contributions of unmeasured neurones that influence the firing of neurone Y, and $\theta^*(t)$ some function of t, representing the underlying form of the threshold at time t.

There is some experimental evidence validating the Gaussian assumption for $\varepsilon(t)$ given in Holden (1976). Two simple forms for $\theta^*(t)$ are

$$\theta^*(t) = \theta_0 \qquad \text{.....(3.4.3)}$$

or,

$$\theta^*(t) = \theta_0 + \mu exp\{-\lambda\gamma(t)\} \; ; \; \mu, \lambda > 0 \qquad \text{.....(3.4.4)}$$

where $\theta^*(t)$ as given in (3.4.3) leads to the assumption of an absolute constant threshold level, whereas that in (3.4.4) leads to the assumption of an exponentially decaying threshold. This is discussed more fully in the chapters that follow.

Furthermore, let $\Omega(t)$ represent the history of a particular process, i.e. those variables determined up to and including time t that are necessary to describe the evolution of the process. For a process $x(t-u)$, we may write

$$\Omega(t) = \{ x(t-u) \; ; \; u \leq t \}. \qquad \text{.....(3.4.5)}$$

In this approach, it will be convenient for computational purposes, specifically in determining the maximum likelihood estimates via standard statistical packages such as GENSTAT or GLIM, to discretise the point processes and record their values only at discrete times t ($t = 0, \pm h, \pm 2h, ...$). If a small sampling interval of length h is selected (i.e., to prevent the occurrence of multiple events in small intervals), then the process will take only the values 0 or 1 (see section 2.3.2).

Thus, for a sampling interval of length h, with h suitably small the point process $Y(t)$ can be replaced by a discrete 0-1 valued series Y_t, such that

$$Y_t = \begin{cases} 1 & ; \textit{ if there is an event in } (t, t+h], \\ 0 & ; \textit{ otherwise.} \end{cases}$$

The discrete approximation of equation (3.4.1) may be written as

$$U_t = \sum_{u=0}^{\gamma_t - 1} a_u x_{t-u} \qquad \text{.....(3.4.6)}$$

where the set of coefficients $\{ a_u \}$ make up the summation function which represents the effects of the neurone X (input) on the potential at time $(t-u)$ on the trigger zone of the neurone Y (output). Similarly, equations (3.4.2; 3.4.3 and 3.4.4) may appear in their discrete approximation form as

$$\theta_t = \theta_t^* + \varepsilon_t \qquad \text{.....(3.4.7)}$$

$$\theta_t^* = \theta_0 \qquad \text{.....(3.4.8)}$$

$$\theta_t^* = \theta_0 + \mu \exp\{-\lambda \gamma_t\}. \qquad \text{.....(3.4.9)}$$

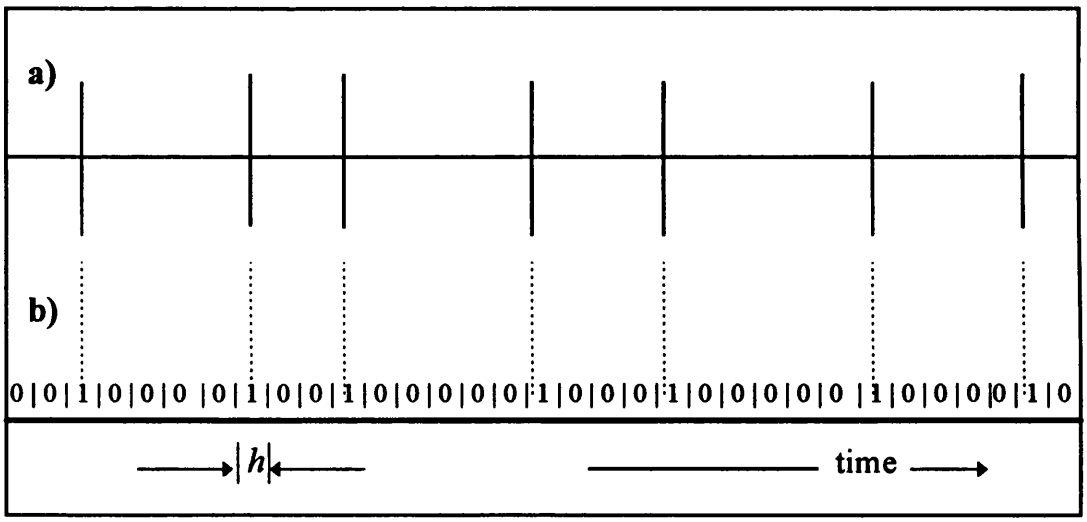


Fig.3.4.1

a) A segment of a neuronal spike train.

b) Neuronal spike train representation as a zero-one valued series.

Now, the conditional probability of the neurone Y firing is given, for h small, by

$$\begin{aligned} P_t &= Pr\{\, Y_t = 1 \mid \Omega_t \,\} \\ &= Pr\{\, U_r \text{ crosses } \theta_r \text{ for some } r \text{ in } (t,\, t+h] \mid \Omega_t \,\} \\ &\cong Pr\{\, U_t \geq \theta_t \mid \Omega_t \,\} \\ &= Pr\{\, U_t \geq \theta^*_t + \varepsilon_t \mid \Omega_t \,\} \\ &= Pr\{\, \varepsilon_t \leq U_t - \theta^*_t \mid \Omega_t \,\} \\ &= F\{\, U_t - \theta^*_t \mid \Omega_t \} \end{aligned} \qquad \text{.....(3.4.10)}$$

where $F(.)$ denotes the cumulative distribution function of the noise ε_t.

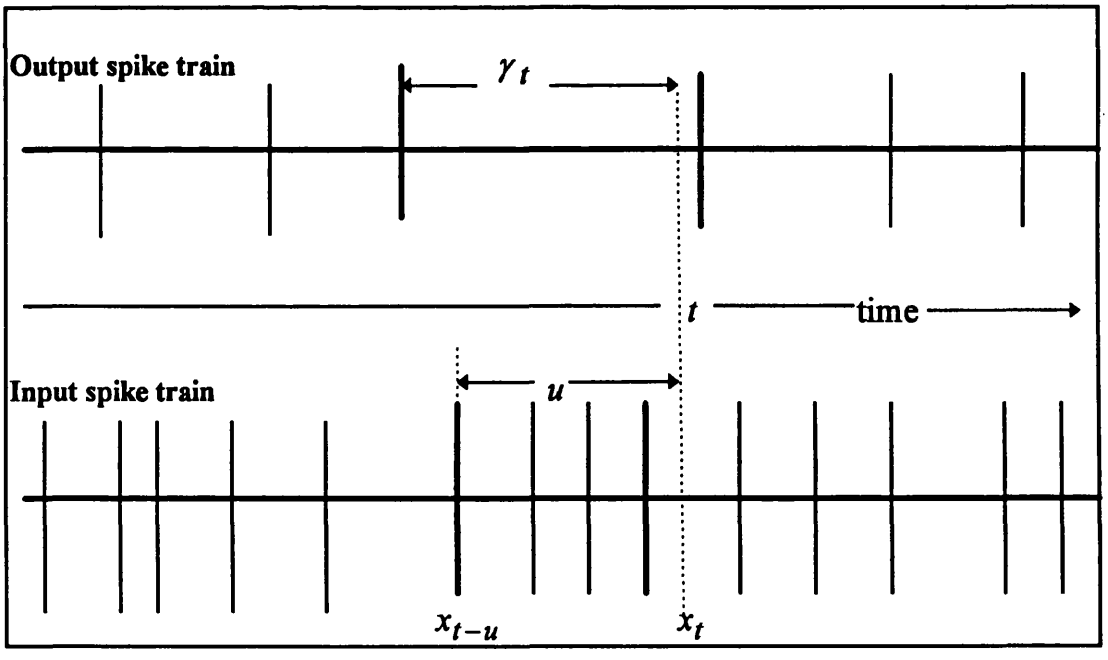


Fig.3.4.2 Diagrammatic representation of the timing convention used for the interval γ_t between an output spike and the time t, u is the interval (starting after the time of the last output spike) between an input spike and time t.

Now, we may write

$$Pr(Y_t = 1 \mid \Omega_t) = P_t \text{ and } Pr(Y_t = 0 \mid \Omega_t) = 1 - P_t \qquad .....(3.4.11)$$

for the probabilities of "success" (or firing) and "failure" (or not firing), respectively. Thus the binomial distribution arises naturally and the probability density function of the response variable Y_t is given by

$$Pr(Y_t = y_t) = P_t^{y_t}(1 - P_t)^{1-y_t} \quad ; y_t = 0, 1. \qquad .....(3.4.12)$$

The likelihood function can be written in the form

$$l_0 = \prod_t P_t^{y_t}(1 - P_t)^{1-y_t} \qquad .....(3.4.13)$$

which can then be maximised.

We also add a recovery term, V_t, to U_t, to allow for spontaneous firings of the neurone and to describe the intrinsic membrane properties of the cell. We shall see later that V_t plays a further role in our model. It seems natural to see if a polynomial form for V_t is adequate.

$$V_t = \begin{cases} \sum_{i=1}^{k} \theta_i (\gamma_t - \zeta_1 - 1)^i & ; \quad \gamma_t \geq \zeta_1 + 1 \\ 0 & ; \quad \gamma_t \leq \zeta_1 + 1 \end{cases} \qquad .....(3.4.14)$$

where γ_t denotes the time elapsed since the time of the last output spike and ζ_1 denotes the minimum of the output inter-spike intervals. We force the recovery function to be zero for $\gamma_t \leq \zeta_1 + 1$ because there will be no output data for smaller values of γ_t. The potential at the trigger zone is reset immediately after the neurone fires and then may rise steadily again on its own without any influence by other neurones. This behaviour suggests adding the

term, V_t, known as the recovery function. References presenting similar analytic models for neurone firing include Johannesma (1968); Knox (1974); and Brillinger (1987).

3.5 Mathematical and Statistical Methodology

3.5.1 Link Function

It is sometimes important to transform data to achieve a specified purpose such as stability of the variance or symmetry. There is a large body of literature concerning transformations for binomial random variables. Throughout this thesis we consider the two most widely used transformations, namely, the logistic and probit transformations. Such transformations are called link functions and will be discussed in the section. More detailed discussions concerning these transformations can be found in McCullagh and Nelder (1992).

To consider the relationship between the response probability, P_t, and the x_{tj} are the values of k variables indicating the presence or absence of an input effect, it is convenient to construct a formal model to describe the effect on P_t of changes in x_{tj} so that the behaviour of the model should be consistent with physical, biological and mathematical laws. One way to investigate this relationship is through the linear combination, usually called the linear predictor

$$\eta_t = \sum_{j=1}^{k} x_{tj}\,\beta_j \;;\; -\infty < \eta_t < \infty \qquad \text{.....(3.5.1)}$$

where β_j $(j = 1, 2, 3, ..., k)$, are the k unknown parameters to be estimated. To consider P_t as a linear combination of the form given in equation (3.5.1) above would inevitably contradict probability laws which require $0 \le P_t \le 1$ and therefore a convenient transformation that maps the unit interval into the

whole real line $(-\infty, \infty)$ is then needed. This leads to the idea of a link function, $H(.)$, such that

$$H(P_t) = \eta_t = \sum_{j=1}^{k} x_{tj}\,\beta_j \;;\; t = 0, \pm h, \pm 2h, \ldots \qquad \ldots\ldots(3.5.2)$$

where $H(P_t)$ is assumed to be a monotonic and differentiable function. Thus

$$P_t = H^{-1}(\eta_t) = H^{-1}\left(\sum_{j=1}^{k} x_{tj}\,\beta_j\right) \qquad \ldots\ldots(3.5.3)$$

where $H^{-1}(.)$ is the inverse of the link function.

The probit link function is one such example commonly used in practice, so that we obtain

$$H(P_t) = \eta_t = \Phi^{-1}(P_t)$$

or,

$$\begin{aligned} P_t &= H^{-1}(\eta_t) = \Phi(\eta_t) \\ &= \Phi\left(\sum_{j=1}^{k} x_{tj}\,\beta_j\right) \end{aligned} \qquad \ldots\ldots(3.5.4)$$

where $\Phi(.)$ is the normal cumulative distribution function.

The identity link function, $P_t = \eta_t = \sum_{j=1}^{k} x_{tj}\,\beta_j$, will not usually be suitable since equal increments of the explanatory variables x_{tj} would imply equal increments in P_t, which would not always be realistic throughout the range of P_t, and indeed lead to values outside its valid range. Cox (1970) has considered other possible link functions, in particular the logistic function,

where

$$H(P_t) = \eta_t = log\left(\frac{P_t}{1-P_t}\right) \text{ or,}$$

$$P_t = H^{-1}(\eta_t)$$

$$= exp(\eta_t) \; / \; (1+exp(\eta_t))$$

$$= exp\left(\sum_{j=1}^{k} x_{tj}\,\beta_j\right) \; / \; \left(1+exp\left(\sum_{j=1}^{k} x_{tj}\,\beta_j\right)\right) \qquad .....(3.5.5)$$

so that equal increments in an explanatory variable produce equal increments in the log-odds ratio and the range of η_t is $(-\infty\,,\,\infty)$.

In fact each distribution has a special link function for which there exists a sufficient statistic equal in dimension to $\underline{\beta}$ for the linear predictor η_t. These special link functions are called canonical link functions. In our case with a binomial distribution the logistic link function is the canonical link, because the log-likelihood depends only on the response variable $\underline{Y}$ through a linear combination $\underline{X}^T\underline{Y}$ which forms a sufficient statistic for the unknown parameter $\underline{\beta}$, as will be seen shortly in section (3.5.2). The logistic link function also has the advantage of having unique estimates for the parameters (Wedderburn 1976 and Haberman 1977). This ensures that the standard statistical packages such as GENSTAT or GLIM will always be able to produce suitable parameter estimates and this may not be true for link functions other than the canonical link. We therefore will invariably investigate the logistic link function first, but will look to examine others if that link function seems in some way inadequate. Cox notes that the logistic and probit links usually produce similar results in practice.

3.5.2 Log Likelihood for Binomial Data

Consider the probability density function of the response variable, Y_t; $t = 0, \pm h, \pm 2h, \pm 3h, \ldots$ given in equation (3.4.12). The likelihood function may be written in the form

$$l_0(\underline{P}; \underline{y}) = \prod_t P_t^{y_t} (1-P_t)^{1-y_t} \qquad \ldots\ldots(3.5.6)$$

where $\underline{P} = \{P_t\}$ and $\underline{y} = \{y_t\}$; $t = 0, \pm h, \pm 2h, \pm 3h, \ldots$.

However, it is easier mathematically to work with the natural logarithm of l_0. The log-likelihood is then

$$\begin{aligned} l(\underline{P}; \underline{y}) &= \log_e \prod_t P_t^{y_t} (1-P_t)^{1-y_t} \\ &= \sum_t \left[y_t \log_e P_t + (1-y_t) \log_e (1-P_t) \right] \\ &= \sum_t \left[y_t \log_e \left(\frac{P_t}{1-P_t} \right) + \log_e (1-P_t) \right]. \end{aligned} \qquad \ldots\ldots(3.5.7)$$

This equation can be expressed as a function of the unknown parameters β_j $(j = 1, 2, 3, \ldots, k)$ through the linear relationship between the response probability P_t and the explanatory variables x_{tj} $(t = 0, \pm h, \pm 2h, \ldots ; j = 1, 2, 3, \ldots, k)$. This relationship takes the form

$$H(P_t) = \eta_t = \sum_{j=1}^{k} x_{tj} \beta_j \quad ; t = 0, \pm h, \pm 2h, \pm 3h, \ldots$$

Now, in order to derive the likelihood equations required for estimating the unknown parameters, we first set down the derivative of the function $l(\underline{P}; \underline{y})$ given in equation (3.5.7), with respect to P_t as

$$\frac{\partial l}{\partial P_t} = \frac{y_t - P_t}{P_t(1-P_t)}.$$

By using the chain rule, we have

$$\frac{\partial l}{\partial \beta_s} = \sum_t \frac{\partial l}{\partial P_t} \frac{\partial P_t}{\partial \beta_s} = \sum_t \frac{y_t - P_t}{P_t(1-P_t)} \frac{\partial P_t}{\partial \beta_s}.$$

Also it is convenient to write $\frac{\partial P_t}{\partial \beta_s}$ in the form

$$\begin{aligned}\frac{\partial P_t}{\partial \beta_s} &= \frac{\partial P_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_s} \\ &= \frac{\partial P_t}{\partial \eta_t} \frac{\partial}{\partial \beta_s} \left(\sum_{j=1}^{k} x_{tj} \beta_j \right) \\ &= \frac{\partial P_t}{\partial \eta_t} x_{ts}.\end{aligned}$$

The derivative of the log-likelihood with respect to β_s then takes the form

$$\frac{\partial l}{\partial \beta_s} = \sum_t \frac{y_t - P_t}{P_t(1-P_t)} \frac{\partial P_t}{\partial \eta_t} x_{ts}. \qquad \text{.....(3.5.8)}$$

Now

$$\begin{aligned}\frac{\partial^2 l}{\partial \beta_r \partial \beta_s} &= \frac{\partial}{\partial \beta_r}\left(\frac{\partial l}{\partial \beta_s}\right) \\ &= \frac{\partial}{\partial \beta_r}\left(\sum_t \frac{y_t - P_t}{P_t(1-P_t)} \frac{\partial P_t}{\partial \eta_t} x_{ts}\right) \\ &= \sum_t \left[(y_t - P_t) \frac{\partial}{\partial \beta_r}\left\{ \frac{1}{P_t(1-P_t)} \frac{\partial P_t}{\partial \eta_t} x_{ts} \right\} + \right. \\ &\qquad \left. \frac{1}{P_t(1-P_t)} \frac{\partial P_t}{\partial \eta_t} x_{ts} \frac{\partial}{\partial \beta_r}(y_t - P_t) \right].\end{aligned}$$

The first term vanishes on taking expectations while the second reduces to

$$E\left\{\frac{\partial^2 l}{\partial\beta_r\,\partial\beta_s}\right\} = -\sum_t \frac{1}{P_t(1-P_t)}\,\frac{\partial P_t}{\partial \eta_t}\,x_{ts}\,\frac{\partial P_t}{\partial\beta_r}$$

$$= -\sum_t \left(\frac{\partial P_t}{\partial \eta_t}\right)^2 \frac{x_{ts}\,x_{tr}}{P_t(1-P_t)}$$

since $\dfrac{\partial P_t}{\partial \beta_r} = \dfrac{\partial P_t}{\partial \eta_t}\,x_{tr}$.

The Fisher information for $\underline{\beta}$ is given by

$$-E\left\{\frac{\partial^2 l}{\partial\beta_r\,\partial\beta_s}\right\} = \sum_t \left(\frac{\partial P_t}{\partial \eta_t}\right)^2 \frac{x_{ts}\,x_{tr}}{P_t(1-P_t)}$$

in the matrix notation this can be represented as

$$-E\left\{\frac{\partial^2 l}{\partial\beta_r\,\partial\beta_s}\right\} = \left\{\underline{X}^T\,\underline{F}\,\underline{X}\right\}_{r\,s}$$

where $\underline{F}$ is a diagonal matrix of weights and has the form

$$\underline{F} = diag\left\{\left(\frac{\partial P_t}{\partial \eta_t}\right)^2 \Big/ P_t\,(1-P_t)\right\}. \qquad \text{.....(3.5.9)}$$

Let us consider the case of the linear logistic model given by

$$\eta_t = log\left(\frac{P_t}{1-P_t}\right)$$

or,

$$P_t = exp(\eta_t) \;/\; (1+exp(\eta_t)).$$

Substituting into the log-likelihood equation given in (3.5.7), we have

$$l(\underline{\beta};\underline{y}) = \sum_t \sum_j y_t x_{tj} \beta_j - \sum_t log[1 + exp(\sum_j x_{tj} \beta_j)] \qquad .....(3.5.10)$$

It has been mentioned earlier that the logistic link function is also the canonical link for the binomial distribution. Because the log-likelihood function given in (3.5.10) above depends on the response variable $\underline{Y}$ only through the linear combination $\underline{X}^T \underline{Y}$ this forms a sufficient statistic for the unknown parameters $\underline{\beta}$.

3.5.3 Computational Procedure

We turn our attention now to the technique used to find the maximum likelihood estimates of the unknown parameters $\underline{\beta}$ in the linear predictor η_t for a given data set. In the case of linear logistic models mentioned above, the derivative given in equation (3.5.8) becomes

$$\frac{\partial l}{\partial \beta_s} = \sum_t (y_t - P_t)\ x_{ts},$$

or in its matrix form

$$\frac{\partial l}{\partial \underline{\beta}} = \underline{X}^T (\underline{Y} - \underline{P}).$$

Also the diagonal matrix of weights $\underline{F}$ given in (3.5.9) reduces for the case of the logistic model to

$$\underline{F} = diag \left\{ P_t (1 - P_t) \right\}.$$

We use the general Newton-Raphson procedure to estimate the unknown parameters $\underline{\beta}$. We define the adjusted dependent variable $\underline{Q}$, with components, q_t $(t = 0, \pm h, \pm 2h, \pm 3h, \ldots)$ as

$$q_t = \hat{\eta}_t + \left(y_t - \hat{P}_t\right)\left\{\frac{\partial \eta_t}{\partial P_t}\right\}_{P_t = \hat{P}_t}$$

$$= \hat{\eta}_t + \left(y_t - \hat{P}_t\right)\left\{\frac{1}{\hat{P}_t\left(1 - \hat{P}_t\right)}\right\}.$$

The maximum likelihood estimates satisfy the equation

$$\underline{X}^T \underline{F}\, \underline{X}\, \underline{\hat{\beta}} = \underline{X}^T \underline{F}\, \underline{Q}$$

or,

$$\underline{\hat{\beta}} = (\underline{X}^T \underline{F}\, \underline{X})^{-1}\, \underline{X}^T \underline{F}\, \underline{Q}. \qquad \ldots\ldots(3.5.11)$$

The procedure underlying the iteration is as follows, beginning with starting values $\hat{P}_t^{(0)}$, say $\hat{P}_t^{(0)} = (y_t + 1/2)/2$ (McCullagh and Nelder, 1992) to calculate $\hat{\eta}_t^{(0)}$ the initial estimate for η_t as

$$\hat{\eta}_t^{(0)} = \log_e\left(\frac{\hat{P}_t^{(0)}}{1 - \hat{P}_t^{(0)}}\right),$$

then evaluating the initial values $q_t^{(0)}$ and $\underline{F}^{(0)}$ for the adjusted dependent variable q_t and the quadratic weights $\underline{F}$, respectively, where

$$q_t^{(0)} = \hat{\eta}_t^{(0)} + \left(y_t - \hat{P}_t^{(0)}\right)\left\{\frac{1}{\hat{P}_t^{(0)}\left(1 - \hat{P}_t^{(0)}\right)}\right\}, \text{ and}$$

$$\underline{F}^{(0)} = \mathrm{diag}\left\{\hat{P}_t^{(0)}(1 - \hat{P}_t^{(0)})\right\}.$$

Calculate $\hat{\underline{\beta}}^{(1)}$, the improved estimate for $\underline{\beta}$ as follows

$$\hat{\underline{\beta}}^{(1)} = (\underline{X}^T \underline{F}^{(0)} \underline{X})^{-1} \underline{X}^T \underline{F}^{(0)} q_t^{(0)}$$

then form new estimates $\hat{\eta}_t^{(1)}$, $\hat{P}_t^{(1)}$, $q_t^{(1)}$ and $\underline{F}^{(1)}$ from $\hat{\underline{\beta}}^{(1)}$ as follows

$$\hat{\eta}_t^{(1)} = \sum_j x_{tj} \hat{\beta}_j^{(1)},$$

$$\hat{P}_t^{(1)} = exp(\hat{\eta}_t^{(1)}) / (1 + exp(\hat{\eta}_t^{(1)})),$$

$$q_t^{(1)} = \hat{\eta}_t^{(1)} + \left(y_t - \hat{P}_t^{(1)}\right) \left\{ \frac{1}{\hat{P}_t^{(1)}(1 - \hat{P}_t^{(1)})} \right\}$$

and

$$\underline{F}^{(1)} = diag\left\{ \hat{P}_t^{(1)}(1 - \hat{P}_t^{(1)}) \right\}.$$

Repeat until changes in deviance in successive iterations are sufficiently small.

Since the response variable Y_t can only take the values 0 or 1, the starting values $\hat{P}_t^{(0)} = (y_t + 1/2)/2$ given above, can take either 1/4 for $Y_t = 0$ or 3/4 for $Y_t = 1$ and these are reasonable initial estimates for the probabilities and may reduce the number of iterations substantially. Another advantage of $\hat{P}_t^{(0)}$ is that it suggests a simple starting procedure by using the data themselves to get the iteration under way.

3.5.4 The Analysis of Deviance

In this section let us consider the general case where Y_t is binomally distributed with index m_t and parameter P_t. Consider the log-likelihood

$$l(\underline{P};\ \underline{y}) = \sum_t \left[y_t \log_e P_t + (m_t - y_t) \log_e (1 - P_t) \right].$$

The maximum achievable value of the log-likelihood for all T observations is obtained when $\underline{P}$ is a vector of T parameters (i.e., the model contains as many parameters as there are data points), in other words the maximum achievable value is attained at the point $\underline{\tilde{P}} = (y_t / m_t)$. Suppose the maximum value for the model under test, M_0, is attained at $\underline{\hat{P}}$. The comparison of observed to predicted values is based on the difference between log likelihoods for the two maximum values, $\underline{\tilde{P}}$ and $\underline{\hat{P}}$ given by the following expression

$$\begin{aligned} D(\underline{\hat{P}};\ \underline{y}) &= -2\left[l(\underline{\hat{P}};\ \underline{y}) - l(\underline{\tilde{P}};\ \underline{y}) \right] \\ &= 2\sum_t \left[y_t \log_e \left\{ \frac{y_t}{\hat{P}_t} \right\} + (m_t - y_t) \log_e \left\{ \frac{m_t - y_t}{m_t - m_t \hat{P}_t} \right\} \right] \end{aligned} \qquad \text{.....(3.5.12)}$$

The statistic, $D(\underline{\hat{P}};\ \underline{y})$, in equation (3.5.12) is usually called the deviance (or the residual deviance) for the model M_0. The deviance for binomial models behaves in much the same way as the residual sum of squares in ordinary linear models (Hosmer and Lemeshow, 1989).

The case we are considering is to analyse neuronal data where y_t is either zero or one and $m_t = 1$. Substituting in (3.5.12), the deviance function will take the form

$$D(\underline{\hat{P}};\ \underline{y}) = 2\sum_t \left[y_t \log_e (y_t / \hat{P}_t) + (1 - y_t) \log_e \left\{ (1 - y_t) / (1 - \hat{P}_t) \right\} \right]$$

$$= -2\sum_t \left[y_t \log_e \left\{ \hat{P}_t / \left(1-\hat{P}_t\right) \right\} + \log_e (1-\hat{P}_t) - \right.$$

$$\left. y_t \log_e(y_t) - (1-y_t)\log_e(1-y_t) \right]$$

$$= -2\sum_t \left[y_t \log_e \left\{ \hat{P}_t / \left(1-\hat{P}_t\right) \right\} + \log_e (1-\hat{P}_t) \right]$$

$$= -2\sum_t \left[y_t \sum_{j=1}^{k} x_{tj}\hat{\beta}_j + \log_e (1-\hat{P}_t) \right],$$

since

$$y_t \log_e(y_t) = (1-y_t)\log_e(1-y_t) = 0; \text{ for } y_t = 0 \text{ or } 1$$

and

$$\log_e \left\{ \hat{P}_t / \left(1-\hat{P}_t\right) \right\} = \hat{\eta}_t = \sum_{j=1}^{k} x_{tj}\hat{\beta}_j.$$

Thus

$$D(\underline{\hat{P}};\ \underline{y}) = -2\underline{\hat{\beta}}^T \underline{X}^T \underline{Y} - 2\sum_t \log_e(1-\hat{P}_t)$$

$$= -2\underline{\hat{\beta}}^T \underline{X}^T \underline{\hat{P}} - 2\sum_t \log_e(1-\hat{P}_t)$$

$$= -2(\underline{X}\underline{\hat{\beta}})^T \underline{\hat{P}} - 2\sum_t \log_e(1-\hat{P}_t)$$

$$= -2\underline{\hat{\eta}}^T \underline{\hat{P}} - 2\sum_t \log_e(1-\hat{P}_t) \qquad \text{.....(3.5.13)}$$

since $\underline{X}^T \underline{Y} = \underline{X}^T \underline{\hat{P}}$ is the maximum likelihood equation. The deviance function as shown in (3.5.13) above compared with that shown in (3.5.12) is clearly a function of $\underline{\hat{\beta}}$ alone and does not depend on $\underline{Y}$. In other words, given $\underline{\hat{\beta}}$, $D(\underline{\hat{P}};\ \underline{y})$ has a conditional distribution which depends only on $\underline{\hat{\beta}}$ and thus cannot be used to test goodness of fit (McCullagh and Nelder, 1992). Furthermore for $D(\underline{\hat{P}};\ \underline{y})$ to be approximately independent of $\underline{\hat{\beta}}$ and therefore

distributed proportional to a $\chi^2_{(T-k)}$ random variable requires the following assumptions

a) $\lim_{\forall t} m_t \rightarrow \infty$, and in fact $\lim_{\forall t} m_t P_t (1-P_t) \rightarrow \infty$ and,

b) The sample size T is large.

This approximate independence is essential for $D(\hat{\underline{P}}; \underline{y})$ to be considered as a goodness of fit statistic. If sample size T is relatively small or the number of binomial trials $m_t = 1$, the approximation no longer holds and hence deviance cannot be used to assess the goodness of fit. An alternative way to test the goodness of fit is discussed in chapter 4.

The deviance in our case (i.e. with $m_t = 1$) is still useful to compare two nested models to test whether the addition of a further term significantly improves the fit. To make this clear, let M_0 represent the model under test and, M_1 the model with an additional term. The reduction in deviance

$$D(\hat{\underline{P}}_0; \underline{y}) - D(\hat{\underline{P}}_1; \underline{y}) = -2\left[l(\hat{\underline{P}}_0; \underline{y}) - l(\hat{\underline{P}}_1; \underline{y})\right],$$

is distributed approximately like $\chi^2_{(1)}$ random variable under the assumptions that the observations are distributed independently according to the binomial distribution and the sample size T is large (neuronal spike train data usually have very large sample sizes), see McCullagh and Nelder (1992).

The χ^2 approximation is therefore usually quite adequate for differences in deviances even though it is inadequate for the deviances themselves (Hosmer and Lemeshow, 1989). For the purpose of comparing nested models, the deviance is a useful measure and deviance tables will be demonstrated in our application sections as a part of the analysis throughout this thesis.

3.5.5 Linearisation Technique for Estimating Non-Linear Parameters

One example where non-linear parameters arise is in the case of modelling the threshold. It is necessary to involve non-linear parameters to describe the threshold accurately. For example an exponentially decaying function of the form $\theta_0 + \mu \exp(-\lambda \gamma_t)$, sometimes has the advantage of making a substantial reduction in the order of the polynomial, V_t, given in (3.4.14) necessary to model the recovery function. We will see this in subsequent chapters.

If we consider a function of an explanatory variable x, such as $f(x;\ \lambda) = \exp(-\lambda x)$, this is an acceptable variable in a linear predictor, $\underline{\eta}$, given in (3.5.1) provided that λ is known since one simply uses the value of $\exp(-\lambda x)$ in place of x in the model matrix. But if λ is to be estimated from the data, then non-linearity arises. Here we describe a fitting procedure by linearisation which allows us to obtain approximate estimates of the non-linear parameters in the linear predictor described in section (3.5.1) above. Consider the term $\delta_1 f(x;\ \lambda)$ to be added to the linear predictor, with δ_1 (linear) and λ (non-linear) unknown parameters to be estimated. For an initial value $\lambda^{(0)}$, $f(x;\ \lambda)$ can be expanded as

$$f(x;\ \lambda) \cong f(x;\ \lambda^{(0)}) + (\lambda - \lambda^{(0)}) \left[\partial f(x;\ \lambda) / \partial \lambda \right]_{\lambda = \lambda^{(0)}}.$$

Then,

$$\delta_1 f(x;\ \lambda) \cong \delta_1 f(x;\ \lambda^{(0)}) + \delta_1 (\lambda - \lambda^{(0)}) \left[\partial f(x;\ \lambda) / \partial \lambda \right]_{\lambda = \lambda^{(0)}}$$

or,

$$\begin{aligned} \delta_1 f(x;\ \lambda) &\cong \delta_1 z_1 + \delta_2 z_2 \\ &= \underline{\delta}^T \underline{Z}. \end{aligned} \qquad \text{.....(3.5.14)}$$

We replace the term $\delta_1 f(x;\ \lambda)$ by two linear terms, $\delta_1 z_1$ and $\delta_2 z_2$, where

$z_1 = f(x; \lambda^{(0)})$, $z_2 = [\, \partial f(x; \lambda) / \partial \lambda]_{\lambda = \lambda^{(0)}}$ *and* $\delta_2 = \delta_1 (\lambda - \lambda^{(0)})$.

The iteration procedure can be described simply as follows. Use the starting value $\lambda^{(0)}$ (say $\lambda^{(0)}$=1/2) to calculate starting values $z_1^{(0)}$ and $z_2^{(0)}$ for z_1 and z_2, respectively, using the expressions for z_1 and z_2 given above.

Now, since all quantities appearing on the right of (3.5.14) above are linear and computed using the starting value $\lambda^{(0)}$, regress $\underline{Z}^{(0)} = (z_1^{(0)}, z_2^{(0)})$ on $\underline{X}$ (McCullagh and Nelder, 1992) to give new estimates $\hat{\delta}_1^{(1)}$ and $\hat{\delta}_2^{(1)}$ for the parameters δ_1 and δ_2, respectively. Calculate $\hat{\lambda}^{(1)}$, the improved estimate for λ as

$$\hat{\lambda}^{(1)} = \lambda^{(0)} + \frac{\hat{\delta}_2^{(1)}}{\hat{\delta}_1^{(1)}} .$$

Use $\hat{\lambda}^{(1)}$ to calculate new values $z_1^{(1)}$ and $z_2^{(1)}$ for z_1 and z_2, as follows

$$z_1 = f(x; \lambda^{(1)}) \text{ and}$$
$$z_2 = [\, \partial f(x; \lambda) / \partial \lambda]_{\lambda = \lambda^{(1)}}$$

Repeat until changes in deviance in successive iterations become sufficiently small. Description of a fitting technique by linearisation can be found in Box and Tidwell (1962).

This linearisation fitting technique is noted here because most of the standard statistical packages such as GENSTAT or GLIM, which are quite suitable for analysing neuronal spike train data, are generally constrained to be linear models packages only. Using the technique described above we will be able to adopt those packages to estimate all the parameters in the model including those which are non-linear. Results obtained by this procedure can be seen at the end of this chapter and in the chapters that follow.

3.6 Application

We now turn to applications of the likelihood method discussed above. First we show how the likelihood approach can be used effectively in analysing neuronal spike train data. This method allows us to estimate internal quantities, such as the summation function which represents the effects of the input on the output, the absolute threshold and the recovery process of the neurone. Secondly, we will be able to compare the results obtained via likelihood with those obtained by the other approaches discussed in chapter 2; for this purpose the same two sets of simulated data considered in the time domain and frequency domain given in that chapter have been used. The simulation details are given in chapter 1 (section 1.6) and further details are given in appendix (A).

The spike trains in the first illustration were replaced by zero-one valued series taking a sampling interval, h, of 1 msec. This led to time series of approximately 60000 points. The unknown parameters were estimated by maximising the likelihood given in (3.5.10), employing the logistic link function given in (3.5.5). Fig. 3.6.1 illustrates the deviance table and suggests that both the summation function and the recovery and (exponentially decaying) threshold functions separately lead to a large reduction in deviance and a model with both leads to a further large reduction in deviance. Both functions contribute information about the process, and so the full model with both functions is required. Further explanation of various features will follow in chapter 4.

Fig.3.6.3a represents the estimated summation function, $\{\hat{a}_u\}$ (and estimated ± two standard error lines plotted around 0), as a measure of the effects of the input on the output, and suggests that the duration of an excitatory input is about 7 msec, i.e., $\{\hat{a}_u\}$; $u = 0, 1, \ldots, 7$, are statistically

significant. We keep adding new parameters into the summation function until they stop being statistically significant. The square root of the estimated cross-intensity function given in Fig.2.5.1c applied to the same set of data, suggests

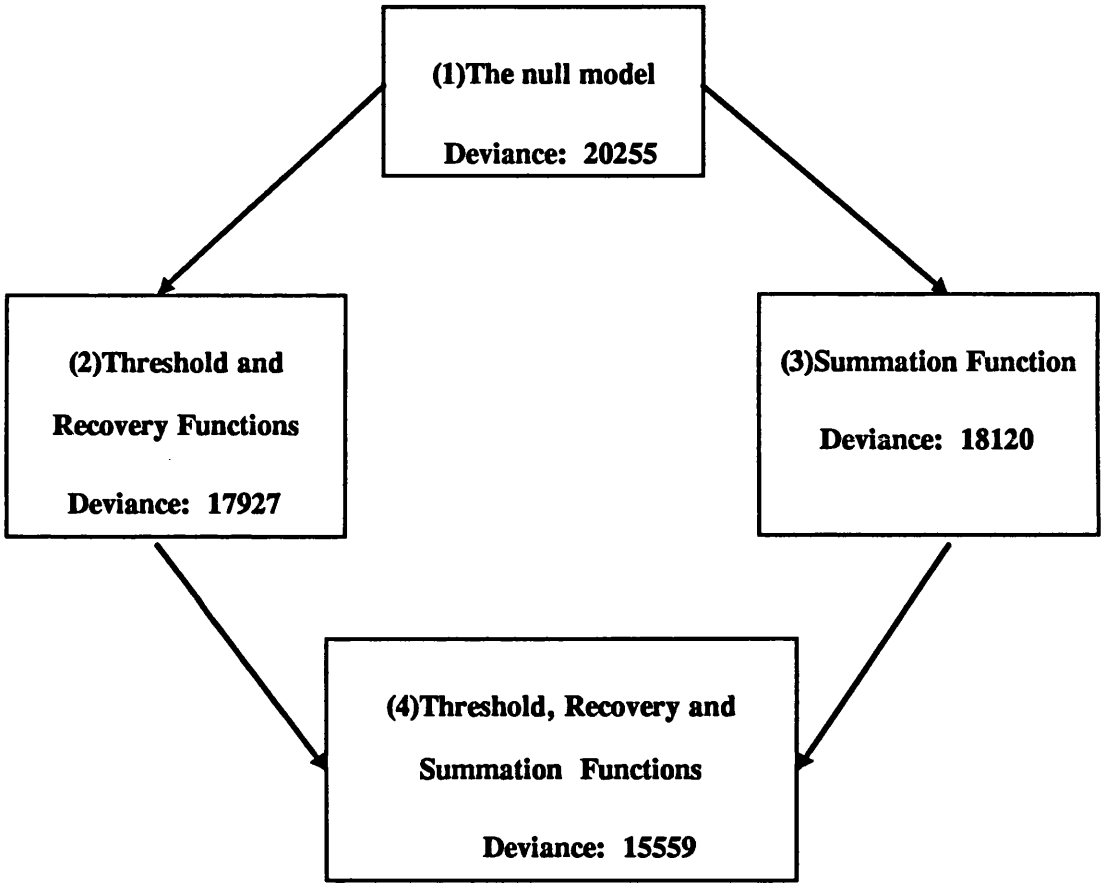


Fig.3.6.1 Diagrammatic representation of the deviance table.

(1) Represents the null model.

(2) A model with only the recovery and (exponentially decaying) threshold functions.

(3) A model with only the summation function.

(4) The full model, i.e., a model with both the summation function and the recovery and (exponentially decaying) threshold functions.

an excitatory effect lasting about 4 msec only, then seems to indicate an inhibition at around 10 msec and that contradicts the way the data was simulated, since there was no such inhibition present in the simulation. Furthermore, the significant duration of the estimated summation function compared with that of the square root of the cross-intensity function given in Fig.2.5.1c corresponds better with the structure of the neuronal model used in the simulation in which the half-width of the estimated excitatory post synaptic potential (EPSP) was about 9.8 msec as shown in Fig.3.6.2.

The recovery and threshold functions taken together (Fig.3.6.3b) indicate that the probability of an output spike is very small up to about 40 msec after the previous output spike, but then it increases rapidly. Using an exponential instead of a constant threshold reduced the order k of the polynomial used in the recovery function, V_t given in (3.4.14), from order 3 to order 1, since $\{\hat{\theta}_i\}$; $i \geq 2$ are not statistically significant. It also improved the fit of the model a little, since it has slightly larger reduction in deviance.

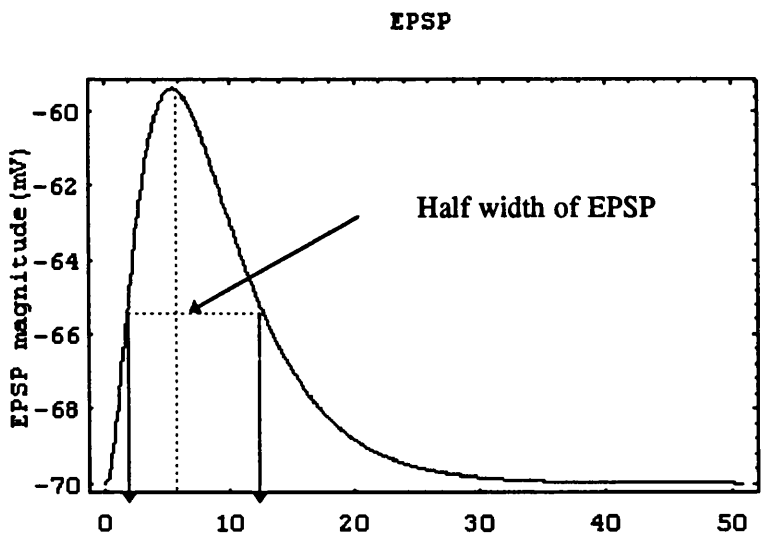


Fig.3.6.2 Illustration of the excitatory postsynaptic potential (EPSP).

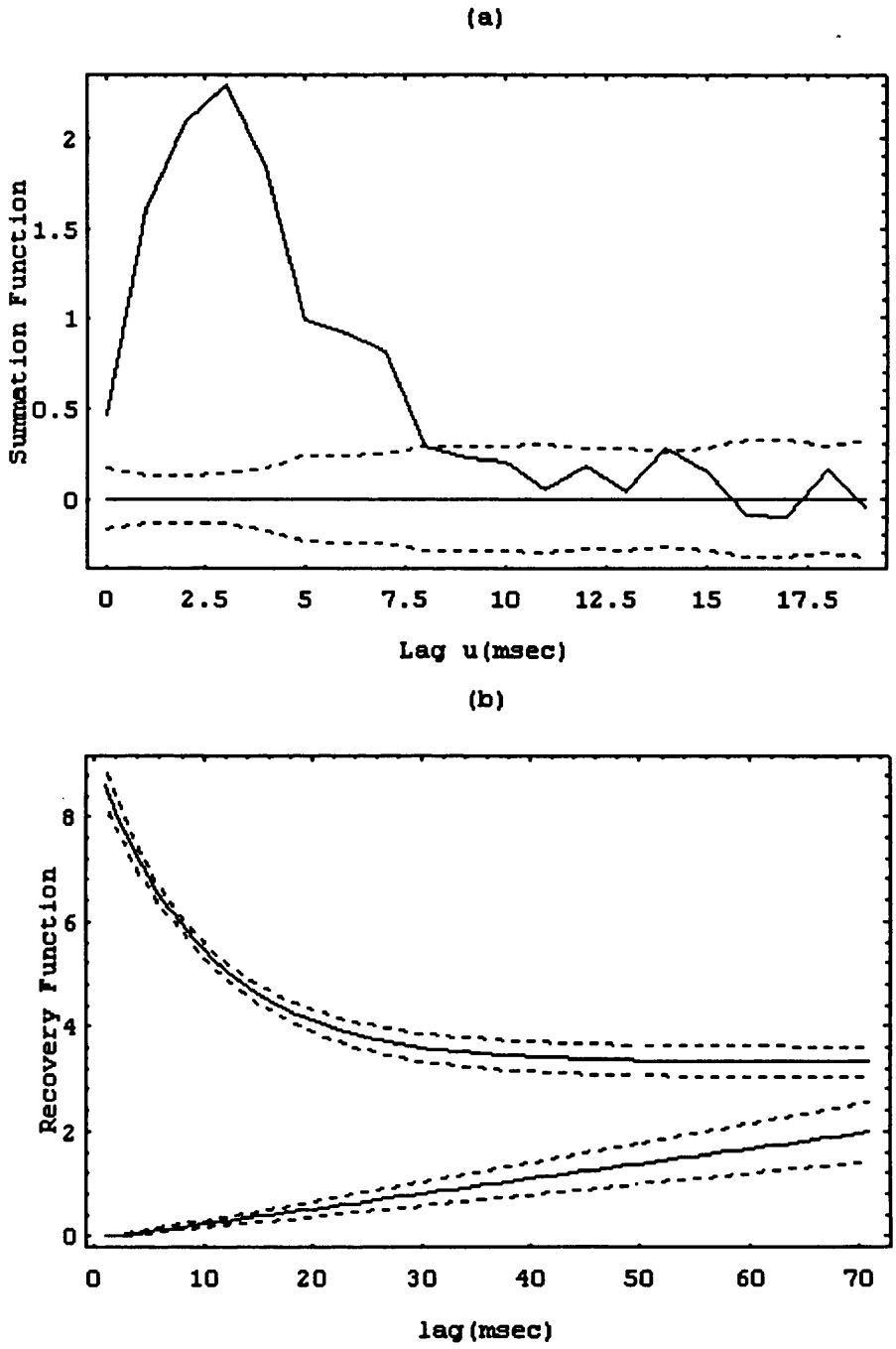


Fig.3.6.3

a) Estimated summation function, the dotted lines provide ± 1.96 standard error limits plotted around zero.

b) Estimated recovery (lower curve) and threshold (upper curve) functions, the dotted lines present the $\pm$ two standard error limits plotted around each curve.

The likelihood procedure was then applied to another set of simulated neuronal spike train data with known features. The spike trains were replaced by zero-one valued series taking a sampling interval, h, of 2 msec. This led to a 0-1 valued series of approximately 30000 points. The unknown parameters were again estimated by maximising the likelihood given in (3.5.10), employing the logistic link function given in (3.5.5). Fig. 3.6.4 illustrates the deviance table. A model with a summation function alone leads to a large reduction in the deviance and a model with recovery and threshold functions reduces the deviance but not as much. This suggests that the data are input-dominated, the summation function is very informative compared to the recovery function. A full model with summation, recovery and threshold functions leads to a further worthwhile reduction in deviance and therefore is the best model.

The estimated summation function given in Fig.3.6.6a is similar to the one shown in Fig. 3.6.3a above, except that the excitatory effect of an input lasts much longer, about 16 msec, though declining steadily after about 3 msec and again the significant duration of the estimated summation function compared with that of the cross-intensity function given in (2.5.3c) corresponds better with the structure of the neuronal model in which the half-width of the estimated excitatory post synaptic potential (EPSP) was about 19.5 msec as shown in Fig.3.6.5 below.

In this example, the minimum output inter-spike interval is 6 msec (i.e., $\zeta_1 = 6$ msec) and the order of the polynomial recovery function needed is one (i.e. $k = 1$), since $\{\hat{\theta}_i\}$; $i \geq 2$ are not statistically significant. We force the recovery function to be zero for $\gamma_t \leq \zeta_1 + 1$ (as given in 3.4.14) ,because there were no output data for smaller values of γ_t. A constant threshold as well as an exponentially decaying threshold were tried. The order of the

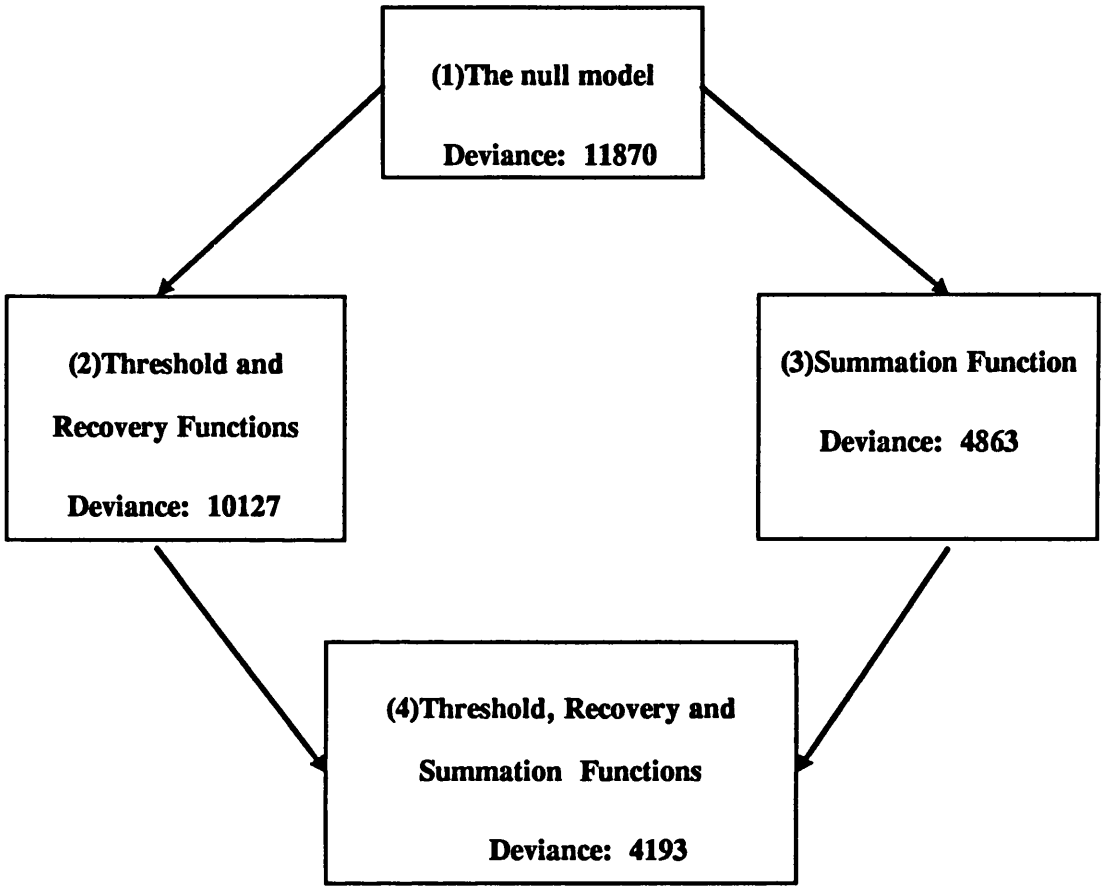


Fig.3.6.4 Diagrammatic representation of the deviance table.

(1) Represents the null model.

(2) A model with only the recovery and (constant) threshold functions.

(3) A model with only the summation function.

(4) The full model, i.e., a model with both the summation function and the recovery and (constant) threshold functions.

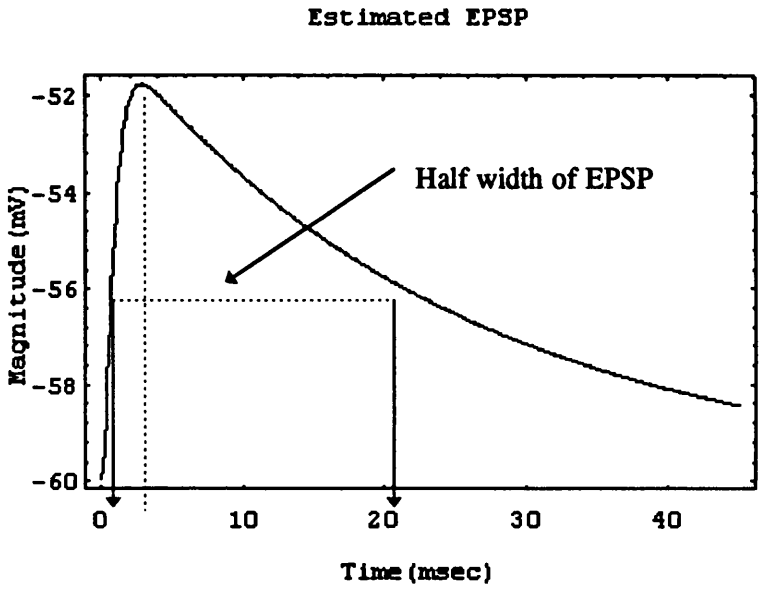


Fig.3.6.5 Illustration of the excitatory postsynaptic potential (EPSP).

polynomial recovery function needed for both models is one and because we wish always to minimise the number of unknown parameters in our model, the choice for a constant threshold in this example was assisted in this.

Fig.3.6.6b shows the recovery and threshold (constant) functions together. These are similar to those shown in Fig.3.6.3b, except that the threshold is constant in this case and the recovery function starts a little later. The probability of an output spike is very small up to about 25 msec after the previous output spike, but then increases rapidly, whereas the square root of the estimated cross-intensity function, by contrast, shown in Fig.2.5.3c, suggests a possible periodic response which was not present in the simulation. Furthermore, the square root of the cross intensity suggests that the duration of an excitatory effect is much shorter than that suggested by the estimated summation function, and again the cross-intensity approach significantly underestimates the duration of the underlying excitatory effects of a synaptic input.

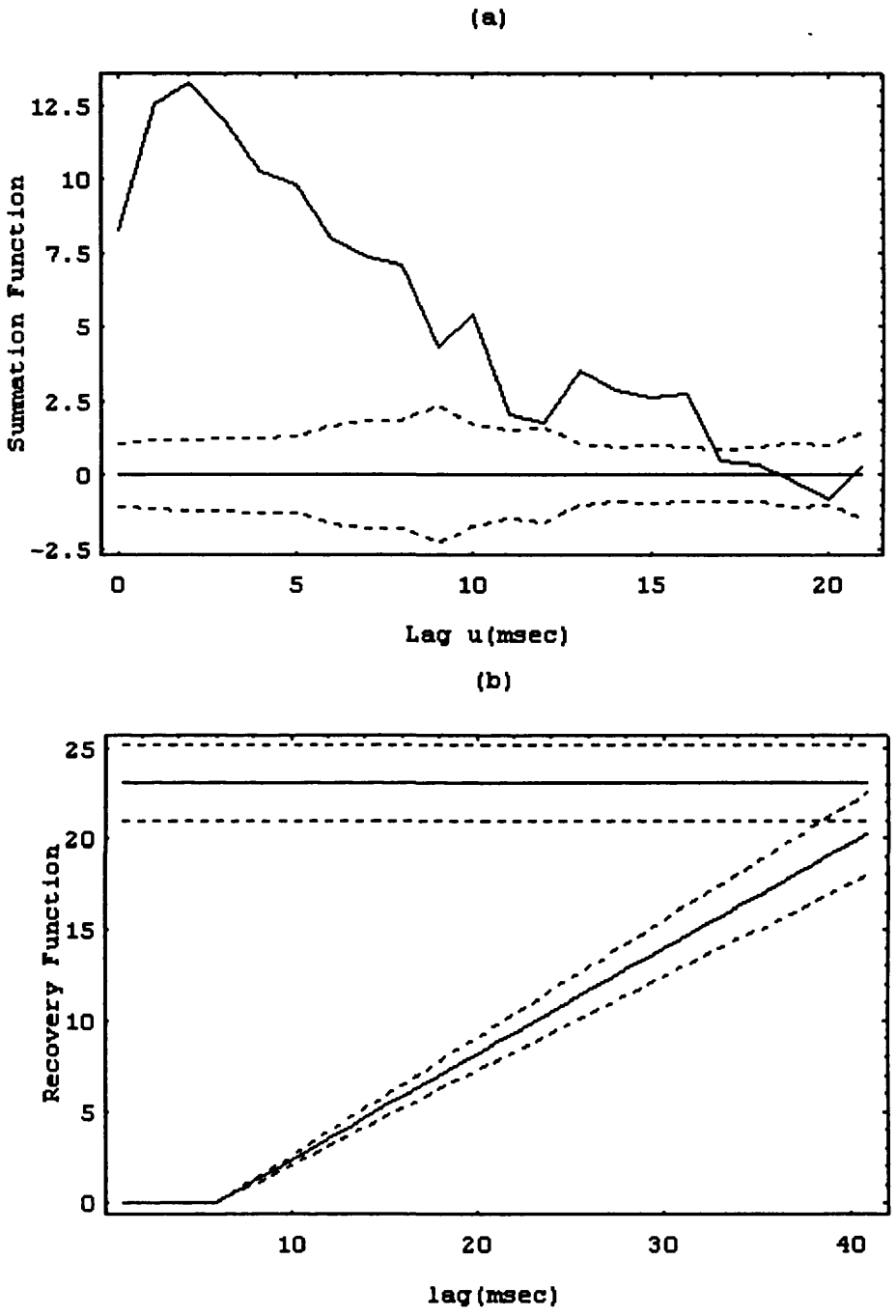


Fig.3.6.6

a) Estimated summation function, the dotted lines provide ± two standard error limits plotted around zero.

b) Estimated recovery (lower curve) and threshold (upper curve) functions, the dotted lines present the ± two standard error limits plotted around each curve.

3.7 Conclusion

The two demonstrations discussed above, both suggest that the square root of the estimated cross-intensity function is difficult to interpret and may be misleading, and that may be because it is combining information about synaptic effects and intrinsic membrane properties, and underestimating the underlying excitatory effects of a synaptic input. Consequently, it is a poor method of addressing the inter-relationships between processes. The likelihood approach, by contrast, is able to explore separately two aspects of the relationship between spike trains in a much better way than the cross-intensity function.

Assuming that the inputs are all present in the model, the recovery and threshold functions taken together appear to describe the time course of intrinsic membrane properties, whereas the estimated summation function characterises a direct relationship between an input to and output from a neurone. All the results discussed above suggest that the likelihood approach does not produce misleading results in the way that the cross-intensity function does. However, there are further results and discussions in subsequent chapters which will clarify these conclusions.

Chapter 4

4 Likelihood Applications to Spontaneous Firing and to Single Input-Output Data

4.1 Introduction

The aims of the present chapter can be summarised by two main points. First to show how best the likelihood approach can be used in analysing neuronal spike train data by applying it to many simulated data sets, each with different and known features, to ensure that the technique is capable of reflecting those features and therefore to have the confidence to apply it to real data sets. This will highlight the advantages of using the likelihood approach that are mentioned earlier in chapter 3. Further we discuss a number of issues concerning the approach, such as the assessment of the goodness of fit, the modelling of threshold and the choice between link functions. Secondly to compare some of the likelihood results with those obtained via stochastic point processes techniques reviewed in chapter 2.

We start with a discussion concerning the assessment of the goodness of fit procedure for the adequacy of the model and the relevance of the link function we introduced in chapter 3 when discussing the analysis of deviance.

4.2 The Goodness of Fit Procedure

The assessment of the goodness of fit of models based on the binomial distribution with a number of trials $m_t = 1$ is a problem which needs a great deal of care as we have seen in the analysis of deviance given in the previous chapter (section 3.5.4).

In these circumstances, we cannot rely on the deviance statistic as an absolute measure of goodness of fit (McCullagh and Nelder, 1992). It is necessary to look for alternative model checking procedures. For example, one may look at deviance tables which are based on differences between the deviances of two nested models since these differences are approximately proportional to a χ^2 random variable. As a general rule, we will choose the simplest model (i.e. the model with the smallest number of parameters) that adequately describes the data.

One way to check the validity of a model has been suggested by Brillinger (1988). This is based on a graphical comparison between an estimated probability, $\hat{P}(\eta)$, and its corresponding theoretical probability, $P(\eta)$, when both are plotted against selected values of the linear predictor of the model, η. Let $\hat{\eta}_t$ be the estimated linear predictor value of the model at any given time t, where $\hat{\eta}_t$ represents the estimated membrane potential of the cell. Then for any given linear predictor value, η, and small h, the estimated probability of firing, $\hat{P}(\eta)$, can be defined as the ratio of the number of firings at any given time t in the small interval $(\eta - h,\ \eta + h)$, and the total number of possible firings in this small interval (Brillinger, 1988), i.e.,

$$\hat{P}(\eta) = \frac{\sum_t \#\{Y_t = 1 \ \ with \ \ \eta - h < \hat{\eta}_t < \eta + h\}}{\sum_t \#\{t \ \ with \ \ \eta - h < \hat{\eta}_t < \eta + h\}} \qquad .....(4.2.1)$$

where $\#\{.\}$ indicates the number of events in the small interval $(\eta - h,\ \eta + h)$. To select a linear predictor value, η, we divide the range of the estimated linear predictor values obtained from the model into a number of small bins (usually of equal width). The centre value of each bin is considered as one selected linear predictor value, η, at which the probability of firing, $\hat{P}(\eta)$, is to be calculated. For any other selected values of η outside the above mentioned range, the denominator of expression (4.2.1) will be zero and therefore the empirical probability cannot be calculated.

Now for any given linear predictor value, η, the theoretical probability, $P(\eta)$, is defined as

$$P(\eta) = H^{-1}(\eta) \qquad .....(4.2.2)$$

where $H^{-1}(.)$ denotes the inverse of the relevant link function (see next section for the most commonly used link functions).

It is reasonable to estimate also the standard error limits for the estimated probability. This makes the graphical comparison between the estimated probabilities and their corresponding theoretical probabilities at any selected values of the linear predictor of the model, η, easier. Since neuronal spike train data are usually of a very large sample sizes, then for any given linear predictor value, η and small h, the estimated standard error for the theoretical probability denoted by $E.S.E(\hat{P}(\eta))$ can be defined as

$$E.S.E(\hat{P}(\eta)) = \sqrt{\frac{\hat{P}(\eta)\,(1 - \hat{P}(\eta))}{\sum_t \#\{t \ \ with \ \ \eta - h < \hat{\eta}_t < \eta + h\}}}. \qquad .....(4.2.3)$$

The approximate 95% standard error limits about the estimated probability, are defined as

$$\hat{P}(\eta) \pm 1.96\ E.S.E(\hat{P}(\eta)). \qquad \text{.....(4.2.4)}$$

The validity of the model and the relevance of the link function depend on both the closeness of the estimated probabilities to their corresponding theoretical probabilities and the range of the linear predictor values (i.e., the width of the range of the linear predictor values which enables larger values for predicted probabilities to be achieved, indicates how informative the model is). This method of checking the validity of the model along with the difference in deviances will be used throughout this study as a model choice criterion. It still needs to be developed however (see chapter 6).

4.3 Choice of Link Function

In the previous chapter we have introduced the idea of a link function and we have utilised three link functions which are widely used in practice. These are the logistic link function, the probit link function and the complementary log-log link function, see for example Cox (1970); McCullagh and Nelder (1992). Suppose that η is any given value of the linear predictor of the model. These three link functions are defined as

(i) the logistic link function given by

$$H(P) = \eta = log_e\left(\frac{P}{1-P}\right) \qquad \text{.....(4.3.1a)}$$

where the probability of firing at any given value of the linear predictor of the model, η, is obtained by taking the inverse of the link function as

$$P(\eta) = H^{-1}(\eta) = \frac{exp(\eta)}{1+exp(\eta)} \qquad \text{.....(4.3.1b)}$$

(ii) the probit link function given by

$$H(P) = \eta = \Phi^{-1}(P) \qquad \text{.....(4.3.2a)}$$

where $\Phi(.)$ is the standard normal cumulative distribution function.

The inverse of the link function is given by

$$P(\eta) = H^{-1}(\eta) = \Phi(\eta) \qquad \text{.....(4.3.2b)}$$

(iii) the complementary log-log link function given by

$$H(P) = \eta = log_e\left(-log_e(1-P)\right) \qquad \text{.....(4.3.3a)}$$

and the inverse of the link function is given by

$$P(\eta) = H^{-1}(\eta) = 1 - exp[-exp(\eta)] \qquad \text{.....(4.3.3b)}$$

In this section we attempt to find methods to enable us to choose between these link functions. We invariably start with the logistic link function for two primary reasons ***(i)*** it is the canonical link function for the binomial distribution as we have shown in section (3.5.1), and ***(ii)*** from the mathematical point of view, it is an extremely flexible and easily used function as we have noticed in the derivation of the log-likelihood function given in section (3.5.2).

Fig.4.3.1 provides a graphical comparison of these three link functions when plotted against selected values of the linear predictor η over the range $(-6, 6)$. This is one means of studying their differences for values of the linear predictor. The figure shows that the logistic and probit link functions have some symmetry in the sense that

$$H(P) = -H(1-P).$$

We have utilised these two link functions with over 50 different data sets, simulated as well as real, and we have noted that the two link functions usually produce similar results with similar physiological interpretations so that it is often difficult to discriminate between them. Cox (1970) and McCullagh and Nelder (1992) came to the same conclusion. The complementary log-log link function possesses no such symmetrical feature and as P approaches 1, it approaches infinity much more slowly than the other two link functions and becomes very close to the logistic link function for η less than -2. It is therefore often easier to discriminate between it and the other two. However, we have yet to find a data set for which it is the best link function.

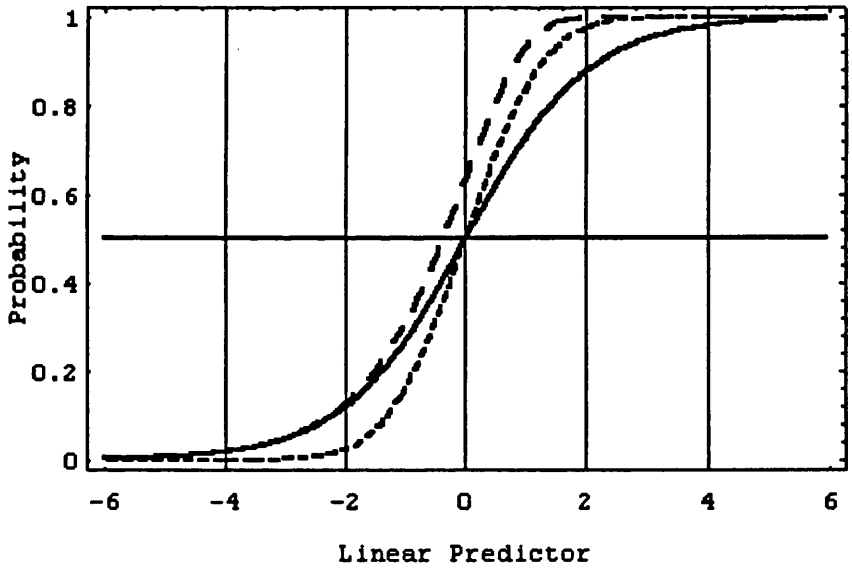


Fig.4.3.1 A graphical comparison of three link functions, the solid line represents the logistic link function, the dotted line (- - - - - - -) represents the probit link function and the complementary log-log link function is represented by the dotted line (— — — — — —).

It is reasonable to start with the logistic link function since its the canonical link for the binomial distribution (see section (3.5.1)) and if this link

function seems in some way inadequate, then to try the probit link function and then if necessary to investigate the complementary log-log link function.

In the next section we illustrate the choice of link function and the goodness of fit assessment by an example using spontaneous data.

4.4 Analysis of Spontaneous Discharge Data

The firing of a neurone is a signal-like discharge that may be considered as a common language with which neurones are able to communicate and exchange information with each other in order to carry out different tasks.

It is known that neurones fire mainly as a result of an external stimulus such as the influence of other neighbouring neurones. Another common feature that many neurones seem to have, however, is spontaneous behaviour where the neurone fires on its own without any external influence imposed by other neurones. This spontaneous behaviour occurs when an action potential rises steadily on its own irrespective of the influences of other neurones. When that potential becomes close to or exceeds an extant level called a threshold, the cell tends to fire. For more details see chapter 1.

In this section we investigate the application of the maximum likelihood approach to spontaneous neural spike train data where the nerve cell is assumed to receive no observed input and gives rise to a single observed output. In this case, we are able only to estimate the threshold and recovery functions and are unable to compare with other approaches since there are no analogous measures obtained using the traditional point process techniques.

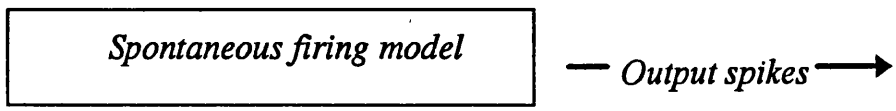


Fig.4.4.1 Spontaneous firing model with no input and a single output.

The object of this section is to apply the likelihood estimation technique described in chapter 3 to spontaneous discharge data in order to get more insight into the processes involved and to investigate some intrinsic membrane properties of the cell from which the spontaneous discharges have been generated. Further we wish to apply the goodness of fit procedure we have discussed earlier in the chapter and examine the choice of link function.

The simulated spontaneous discharge neuronal spike train data have been generated by using a conductance based neuronal model (Halliday, 1994) which has been described in chapter 1 (section 1.6) and discussed with more detail in appendix (A).

The number of spontaneous spikes observed in this illustration was 1620. The spike train was replaced by a zero-one valued series taking a sampling interval, h, of 1 msec. This led to a zero-one valued series of approximately 60000 points. The unknown parameters used to estimate the recovery and threshold functions were estimated by maximising the likelihood equation given in (3.5.7), employing the logistic, probit and complementary log-log link functions. In the actual computations it seemed simplest to first create a data file via a FORTRAN program and then to process that file via the GENSTAT statistical package. A GENSTAT program is given in appendix (B).

Fig.4.4.2 represents the inter-spike interval histogram of the spontaneous discharges, and suggests that the interval between firings is very approximately normally distributed centred around 36 msec with a minimum interval of 23 msec and a maximum interval of 59 msec between firings.

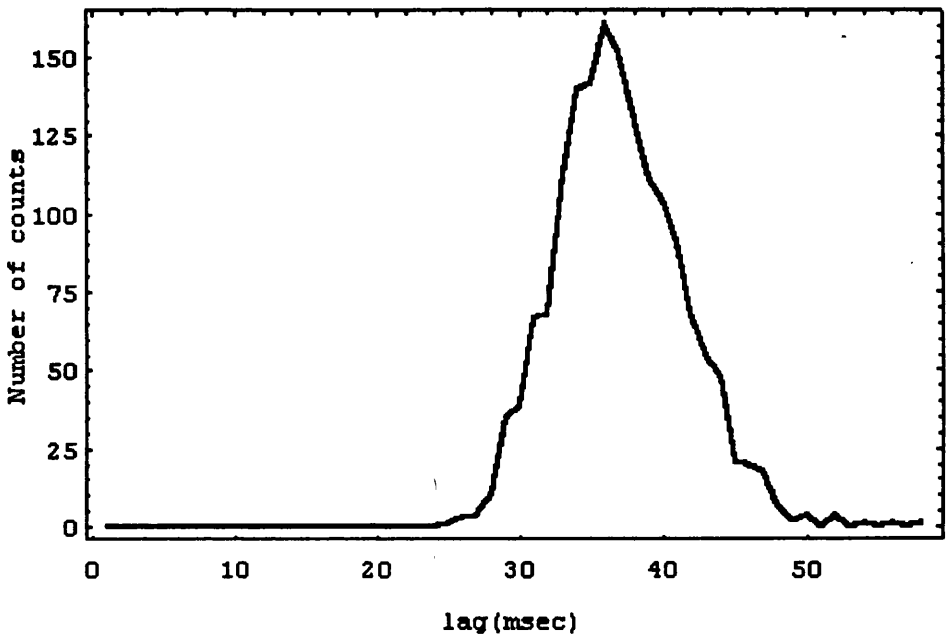


Fig.4.4.2 The inter-spike interval histogram of the spontaneous discharges.

Fig.4.4.3a represents the estimated recovery function (lower curve) and threshold (upper solid line) and their estimated ±two standard error limits (dotted lines). The estimates were obtained by employing the logistic link function given in (4.3.1b). Let V_t represent a polynomial recovery function of order k at any given time t, and given by

$$V_t = \begin{cases} \sum_{i=1}^{k} \theta_i \, (\gamma_t - \zeta_1 - 1)^i \; ; & \gamma_t \geq \zeta_1 + 1 \\ 0 \; ; & \gamma_t \leq \zeta_1 + 1 \end{cases} \qquad \text{.....(4.4.1)}$$

where γ_t denotes the time elapsed since the time of the last output spike and ζ_1 denotes the minimum of the output inter-spike intervals. In this example $\zeta_1 = 23$ msec and the order of the polynomial recovery function needed is a cubic (i.e. $k = 3$), since $\hat{\theta}_i \; ; \; i \geq 4$ are not statistically significant. We force the recovery function to be zero for $\gamma_t \leq \zeta_1 + 1$ (as given in (4.4.1) above), because there were no output data for smaller values of γ_t.

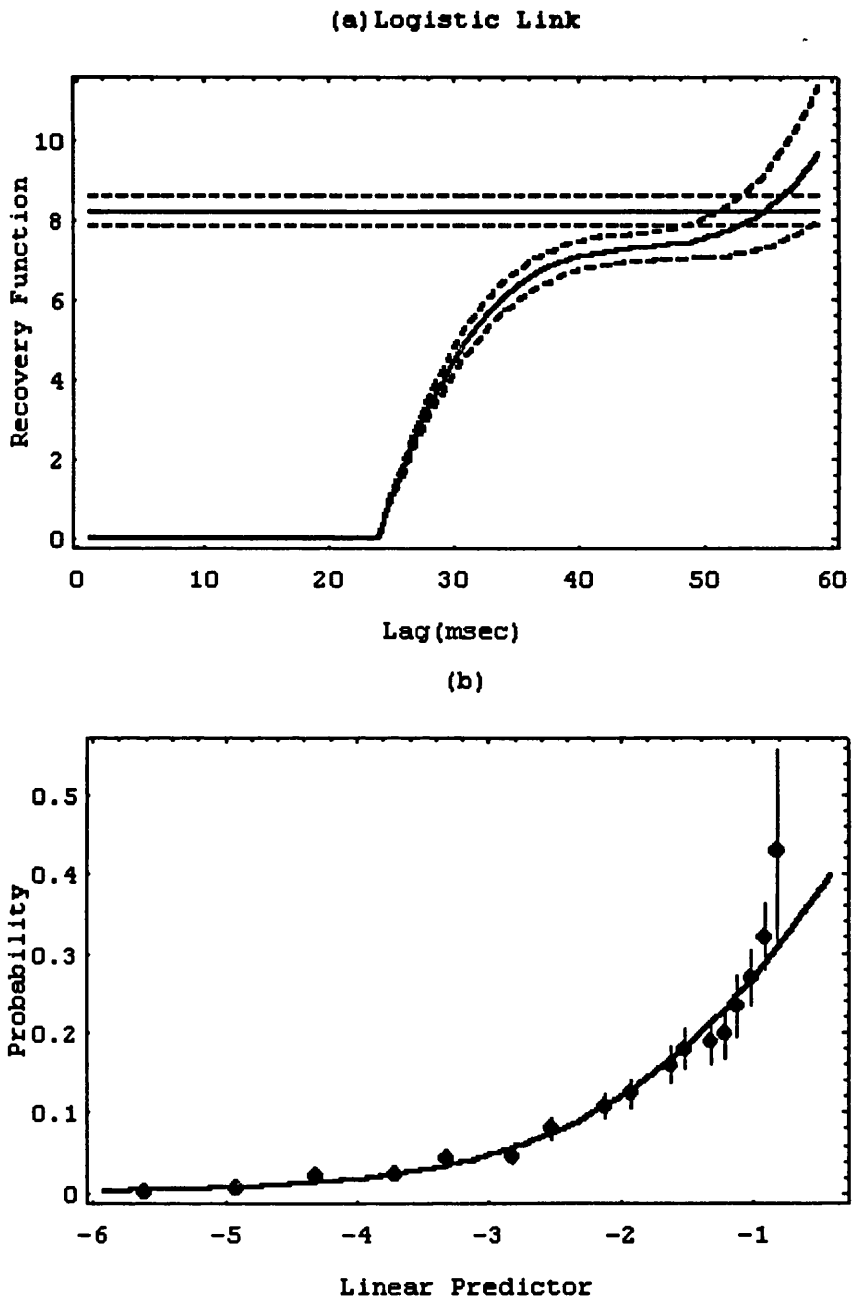


Fig.4.4.3 Illustration of the estimated recovery and threshold functions of the logistic model.

(a) represents the estimated recovery (lower solid line) and threshold (upper solid line) functions and the estimated $\pm$ two standard error limits (dotted lines) plotted around each function.

(b) represent the empirical (dots) and theoretical (smooth curve) probabilities corresponding to the model given in (a) plotted against selected linear predictor values, η. The vertical bars are the approximate 95% confidence intervals for the theoretical probabilities.

A constant threshold was used and no improvements were found with an exponential threshold neither in the simplicity of the model (i.e., the number of parameters in the model and in particular those needed in the recovery function) nor in the reduction in deviance.

The estimated recovery function which starts at 24 msec indicates that the probability of an output spike is very small up to about 30 msec after the previous output spike, but then it increases rapidly and it is approximately constant between 40 and 50 msec. Note also that the recovery function is badly estimated after about 50 msec as the standard error limits get wider due to the lack of output data as can be seen from the output inter-spike interval histogram given in Fig.4.4.2.

Fig.4.4.3b illustrates the empirical and theoretical probabilities corresponding to the logistic model given in Fig.4.4.3a and suggests a good fit for the logistic model, since the confidence intervals about each of the empirical probabilities are seen to contain the corresponding theoretical probabilities except for values of the linear predictor near -0.8.

Fig.4.4.4a and Fig.4.4.5a correspond to Fig.4.4.3a and represent the estimated recovery and threshold functions and their estimated $\pm$ 1.96 standard error limits plotted around each function, obtained by employing the probit and complementary log-log link functions given in expressions 4.3.2a and 4.3.3a, respectively. Again no improvements were found with an exponential threshold and the order of the polynomial recovery function needed in each case is a cubic (i.e. $k = 3$). The two figures reveal similar interpretations as those revealed by the logistic link function in Fig.4.4.3a where the probability of an output spike is very small up to about 30 msec after the previous output spike, but then it increases rapidly and it is approximately constant between

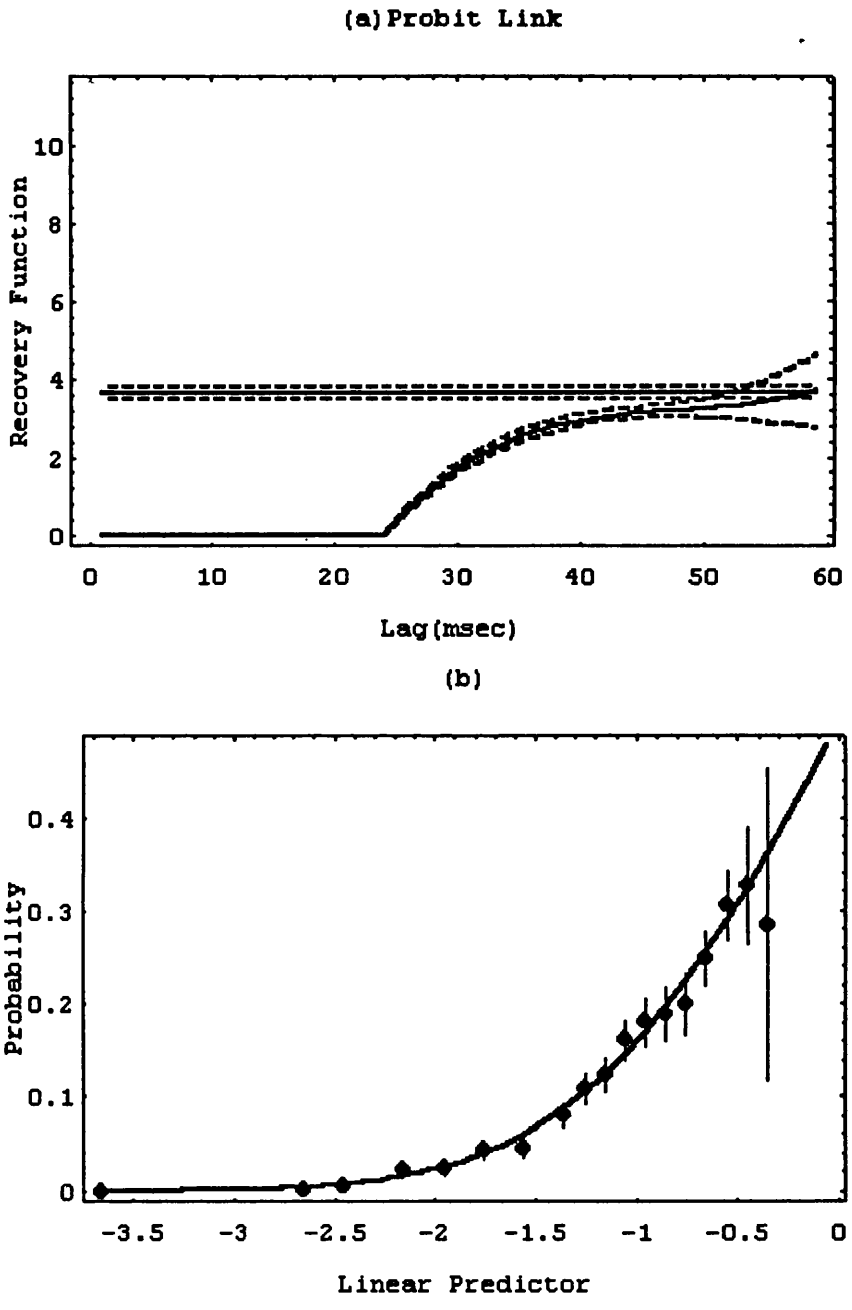


Fig.4.4.4 Illustration of the estimated recovery and threshold functions of the probit model.

(a) represents the estimated recovery (lower solid line) and threshold (upper solid line) functions and the estimated $\pm$ two standard error limits (dotted lines) plotted around each function.

(b) represent the empirical (dots) and theoretical (smooth curve) probabilities corresponding to the model given in (a) plotted against selected linear predictor values, η. The vertical bars are the approximate 95% confidence intervals for the theoretical probabilities.

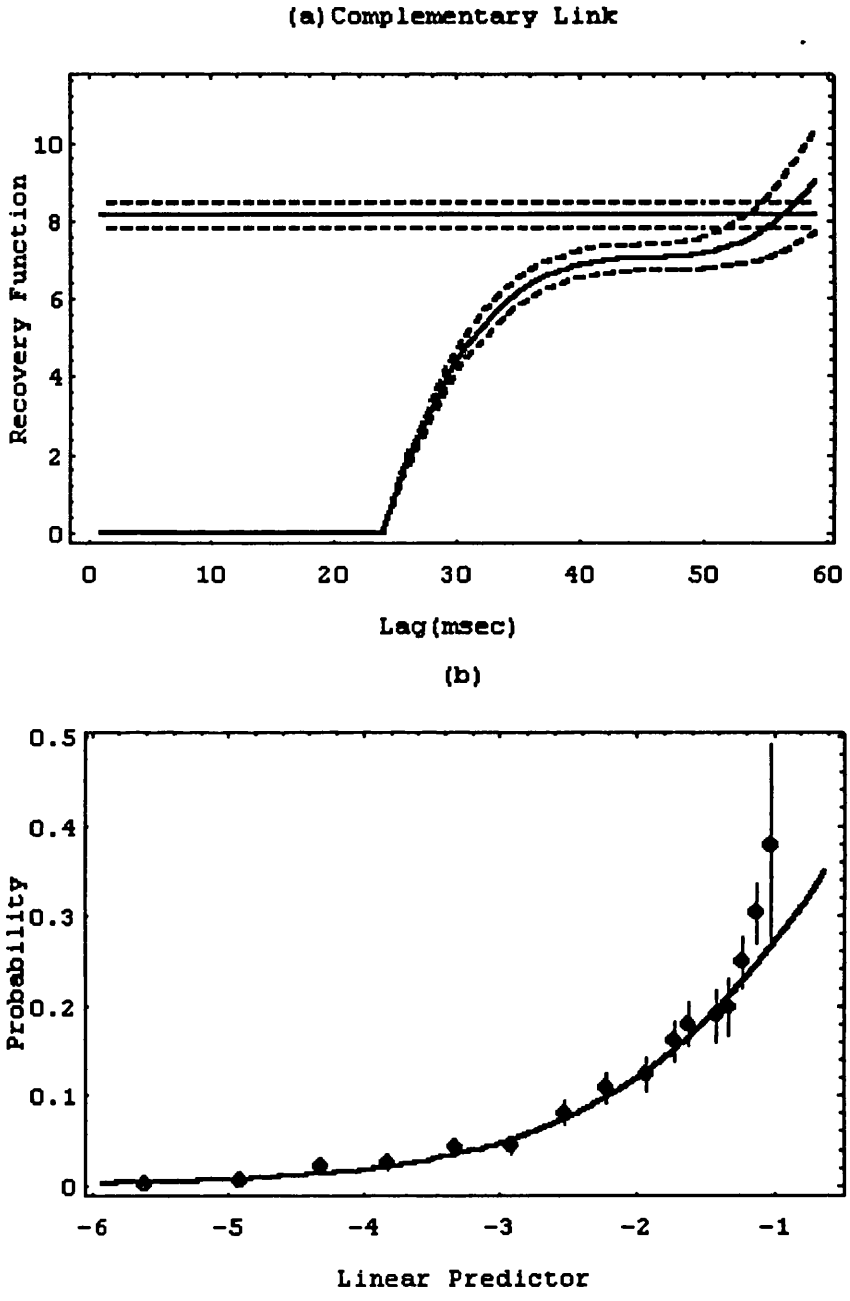


Fig.4.4.5 Illustration of the estimated recovery and threshold functions of the complementary model.

(a) represents the estimated recovery (lower solid line) and threshold (upper solid line) functions and the estimated $\pm$ two standard error limits (dotted lines) plotted around each function.

(b) represent the empirical (dots) and theoretical (smooth curve) probabilities corresponding to the model given in (a) plotted against selected linear predictor values, η. The vertical bars are the approximate 95% confidence intervals for the theoretical probabilities.

40 and 50 msec. Again the standard error limits for each of the two recovery functions get wider after about 50 msec due to the lack of output data.

There are scale differences for the threshold parameters between different link functions. However, these are not physiologically meaningful parameters and in fact what is physiologically meaningful is not the parameters in the recovery and threshold functions themselves but the difference between these two functions (i.e., the more the estimated recovery function approaches its corresponding estimated threshold, the higher is the probability of an output spike). Further explanations concerning the physiological interpretation of the recovery and threshold functions in different circumstances will be discussed in detail in chapter 5.

The goodness of fit plot given in Fig.4.4.4b corresponds to the probit model given in Fig.4.4.4a and suggests a good fit for the model, since the confidence intervals about each of the empirical probabilities (dots) are all seen to contain the corresponding theoretical probabilities whereas the goodness of fit plot for the complementary log-log model given in Fig.4.4.5b suggests that the fit is not adequate for values of the linear predictor near -1 compared to the goodness of fit plots of both the probit and logistic models. The probit model seems superior to the logistic one particularly because it takes the range of the linear predictor further than that of the logistic model. This enables larger values for predicted probabilities to be achieved. The probit model also fits better for values of the linear predictor above -1.2.

The deviance table illustrated in Fig.4.4.6 below shows that the deviances for all three models are very similar with a slightly greater reduction in deviance for the probit model. We would therefore choose the probit model, both because of the deviance table and because of our analysis of the goodness of fit plot.

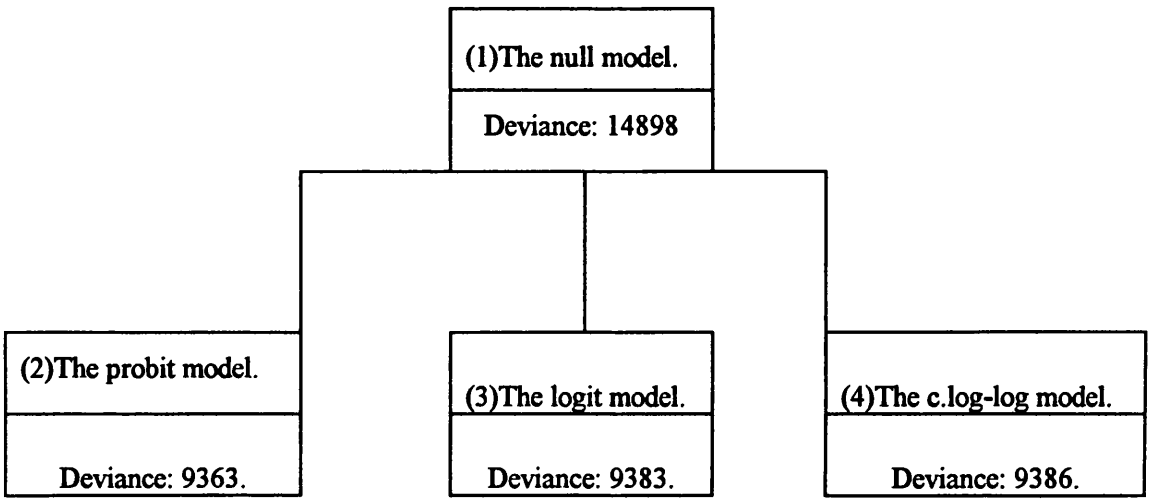


Fig.4.4.6 Diagrammatic representation of the deviance table.

(1) represents the null model. (2), (3) and (4) represent the threshold and recovery functions for the probit, logistic and complementary log-log models respectively.

4.5 Analysis of Single Input-Single Output Data

This section deals with the case where the nerve cell is assumed to receive one observed input and give rise to a single observed output. In this case, we are able to estimate threshold, recovery and summation functions and we will be able to compare results obtained via the likelihood approach with those obtained via the traditional stochastic point process techniques.

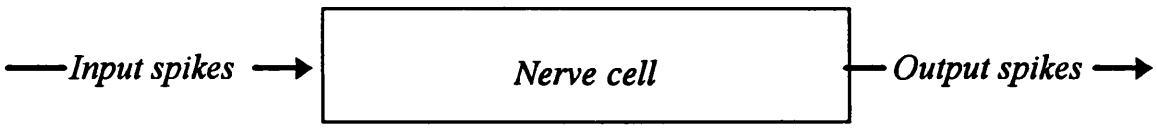


Fig.4.5 A model with observed single input and single output.

We apply the likelihood approach to estimate the summation function, the recovery function and threshold using both real and simulated data sets with excitatory and inhibitory input effects. Then we compare some of the likelihood results with those obtained via time domain and frequency domain techniques.

4.5.1 Summation Function, Recovery Function and Threshold

We start by investigating the effects of a single neurone X (as input) on the membrane potential at the trigger zone of a neighbouring neurone Y (as output), described by counting measures $X(t)$ and $Y(t)$ respectively. These are analogous to the counting measure $N(t)$ defined in section (2.3). The function which represents the effects of a neurone X as input on the firing of a neurone Y as output is called the summation function, $a(.)$ which has been described in detail in the previous chapter (section 3.4.2).

The neurone may further tend to fire spontaneously on its own when there are no external effects applied to the cell, as described earlier in this chapter and in chapter 1. The firing may be as a result of both internal and external effects combined together. For further explanation see chapter 5. The function representing these spontaneous effects at a given time t is, as previously, called the recovery function and is denoted by V_t . The membrane potential at its trigger zone at any given time t, U_t, may be given in its approximate discrete form by

$$U_t = V_t + \sum_{u=0}^{\gamma_t - 1} a_u x_{t-u} \qquad \text{.....(4.5.1)}$$

where the set of coefficients $\{a_u\}$ make up the discretised summation function

and the term V_t represents the polynomial recovery function of order k, i.e.

$$V_t = \begin{cases} \sum_{i=1}^{k} \theta_i (\gamma_t - \zeta_1 - 1)^i \; ; & \gamma_t \geq \zeta_1 + 1 \\ 0 \; ; & \gamma_t \leq \zeta_1 + 1 \end{cases} \qquad \text{.....(4.5.2)}$$

γ_t denotes the time elapsed since the time of the last output spike and ζ_1 denotes the minimum of the output inter-spike intervals. The recovery function given in (4.5.2) is forced to be zero for $\gamma_t \leq \zeta_1 + 1$, as explained in (4.4). For further details see chapter 3 (section 3.2.2). The linear predictor of the model, η_t, representing the difference between the membrane potential and the threshold may be given by

$$\eta_t = U_t - \theta_0 \qquad \text{.....(4.5.3)}$$

where θ_0 represents the constant threshold. We could also use an exponential threshold as we will see shortly in section (4.5.2).

Brillinger and Segundo (1979) and Brillinger (1988 and 1992) considered related models. However, these models were limited to a constant threshold, used only a probit link function, and only fitted recovery (and threshold) and summation functions in the absence of any other unmeasured inputs. Moreover, only one real data set was used. These limitations of their work heavily affect their interpretations as we will see shortly in the subsequent sections of this chapter and also in chapter 5.

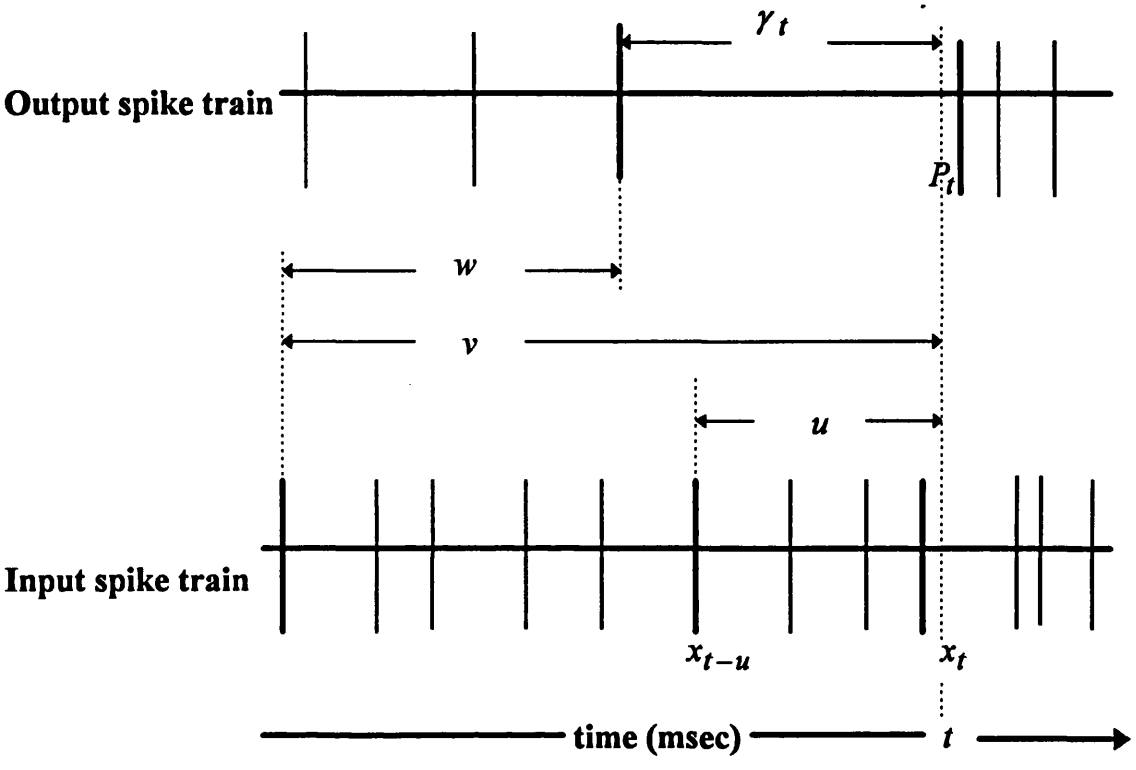


Fig.4.5.1 Diagrammatic representation of the timing convention used for the interval γ_t between an output spike and the time t at which the probability of firing P_t is to be estimated. u is the interval (starting after the last output spike) between an input spike and time t, v is the interval between an input spike and time t, and w is the interval between any input spike occurring prior to the previous output spike and the time of the previous output spike.

(a) Real Data with an Excitatory Input

In this section we analyse a real set of data obtained from a muscle spindle lying within the tenuissimus muscle in the hind limb of deeply anaesthetised cats where the fusimotor axons were stimulated with voltage pulses. The inputs and the resulting response from the spindle in the form of sequences of pulses were recorded. For further detail see chapter 1 of this thesis, Halliday et al (1988), Rosenberg et al (1989) and Amjad (1989).

The numbers of spikes observed in this real data set were 1768 for the input and 450 for the output. The spike trains were replaced for a small interval, h, of length 1 msec by a zero-one valued series of approximately 60000 points.

The inter-spike intervals between the input firings are centred around 40-45 msec with minimum and maximum intervals of about 15 and 95 msec, respectively, whereas the inter-spike intervals between the output firings are centred around 125-130 msec with minimum and maximum intervals of about 94 and 350 msec, respectively.

The unknown parameters were estimated by maximum likelihood, employing the logistic link function given in (4.3.1a). The membrane potential, threshold and linear predictor are the same as those given in (4.5.1 - 4.5.3).

Fig.4.5.2a represents the estimated summation function and shows effects from inputs are only significant in affecting the output to any great extent around 8 msec later. We may note that in this case, the cross-intensity function given in Fig.4.5.4a is almost identical to the summation function. This may be due to the reason that there was no sign of any carry-over input effects (see section 4.6) or due to the very weak input effects on the output, since the reduction in deviance when the summation function is fitted alone is a very small one as can be seen from the deviance table given in Fig.4.5.3. It may be a combination of the two. This is one of the few cases where the cross intensity function might still be a meaningful measure.

The minimum of the output inter-spike intervals, ζ_1, is 94 msec and the order of the polynomial recovery function needed is a cubic (i.e. $k = 3$), since $\{\hat{\theta}_i\}$; $i \geq 4$ are not statistically significant. We force the recovery function to be zero for $\gamma_t \leq \zeta_1 + 1$, because there were no output data for smaller values of γ_t.

The estimated recovery function given in Fig.4.5.2b which starts at 95 msec is well-estimated out to about 150 msec, and although the recovery function is rising steadily and it is approximately constant after about 135 msec, it remains far below threshold at 150 msec. These weak input and spontaneous effects on the firing of the output suggests that there must be some other unmeasured inputs present. Indeed, about 77% of the variation remained unexplained as we will see shortly when investigating the deviance table. However, the involvement of the unmeasured inputs and some other interesting results will be discussed in detail in chapter 5 when we take into account all the input information available. The dotted lines provide ± two standard error limits plotted around zero for the coefficients of the summation function and plotted around each solid curve for the recovery and threshold functions.

The goodness of fit plot given in Fig.4.5.2c indicates that only very small probabilities can be predicted which reflects the large unexplained variability. And the confidence intervals are wide.

A constant threshold was used and no improvements were found with an exponential threshold. The deviance table given in Fig.4.5.3 indicates that the recovery function is much more informative compared to the summation function, since it reduces the deviance by 1126 when fitted alone and a reduction of only 83 when fitting the summation function alone, however, a model with both functions reduces the deviance by 1199 (i.e., almost the same reduction as the sum of the reductions, 1209, when both functions are fitted separately). This suggests that the information contained in the recovery function is both greater and largely orthogonal to that contained in the summation function. The table also indicates that the data is largely dominated

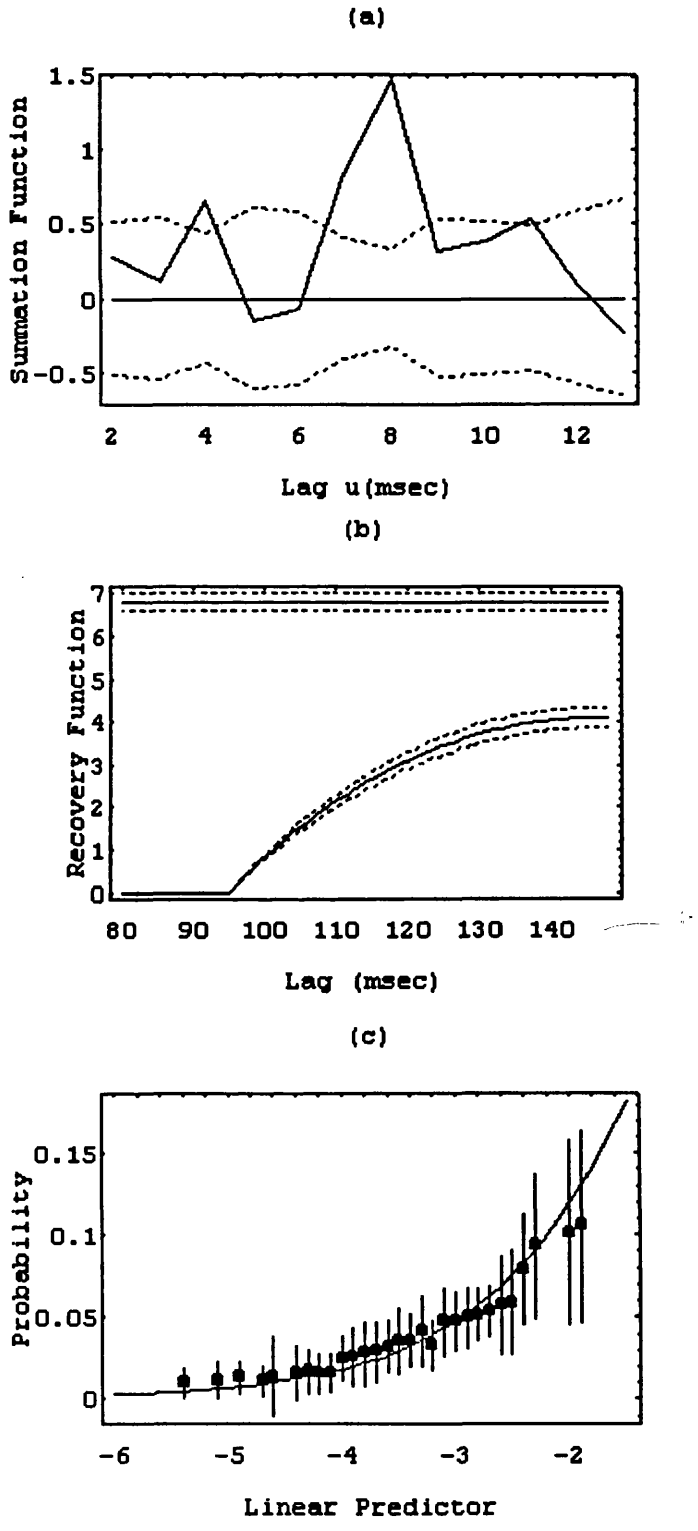


Fig.4.5.2 a) Estimated summation function. b) Estimated recovery and threshold (upper solid line) functions. The dotted lines give ± two standard error limits plotted about zero in (a) for the summation function and plotted around each function in (b) for the threshold and recovery functions. c) The goodness of fit plot with empirical (dots) and theoretical (smooth curve) probabilities plotted against selected values for the linear predictor, the vertical bars give ± two standard error limits for the theoretical probabilities.

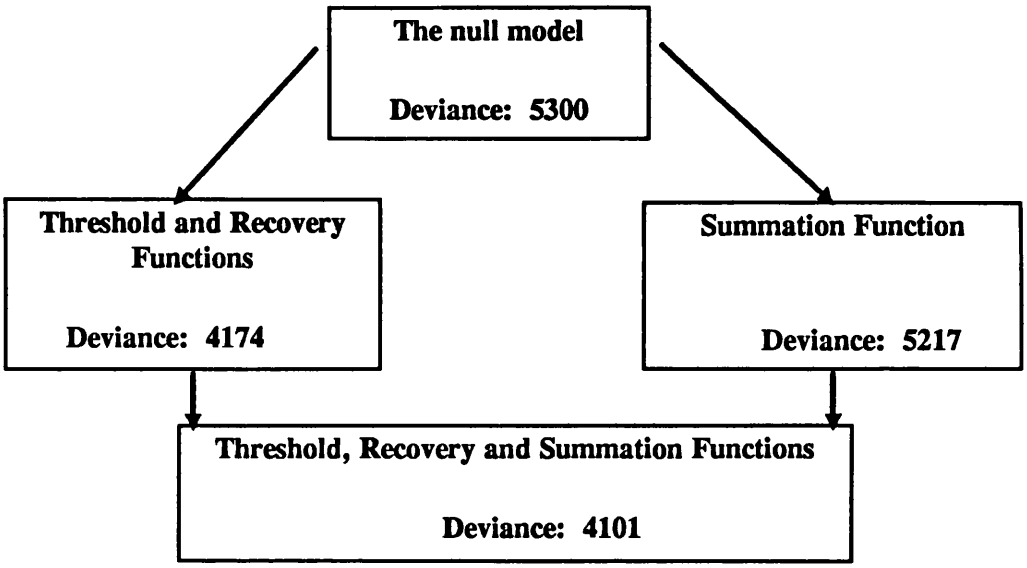


Fig.4.5.3 Diagrammatic representation of the deviance table.

by some other sources of variation (e.g., unmeasured inputs) since the amount of variation explained by the full model is only about 23%.

Fig.4.5.4a, b and c correspond to the estimates of the square root of the cross-intensity function, coherence and phase, respectively. The coherence plot suggests that the two processes are not associated with each other over any range of frequencies. The estimated phase suggests that there is no simple delay present. The likelihood approach through its threshold, recovery and summation functions and the deviance table provides more information than that provided by the time domain and frequency domain approaches.

(b) Simulated Data with an Inhibitory Postsynaptic Input

This is an example where the firing of one neurone input decreases (or inhibits) the spike activity of the second neurone output. This leads to the inhibitory post-synaptic potential (IPSP) as explained earlier in chapter 1. The data were again simulated by using a conductance based neuronal model using an inhibitory input (Halliday, 1994).

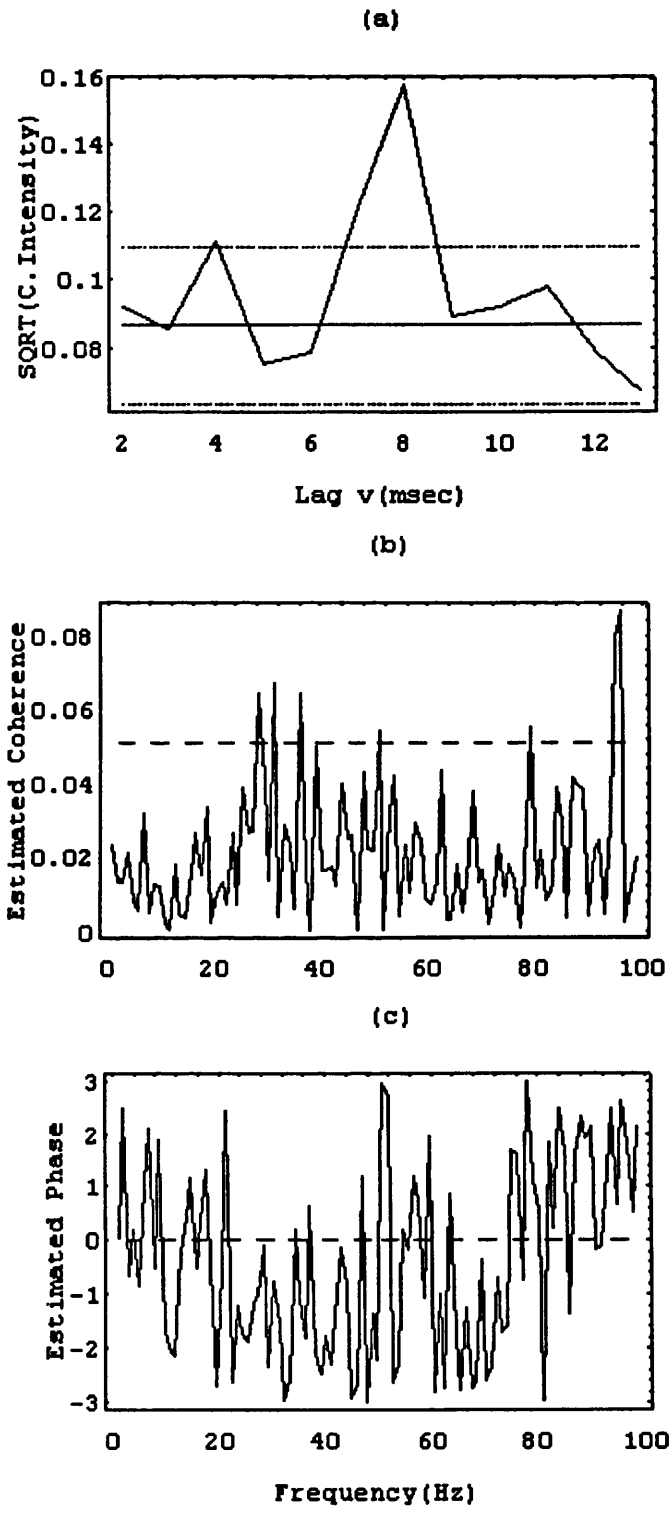


Fig.4.5.4 **a)** Estimate of the square root of the cross-intensity function. **b)** Estimate of the ordinary coherence. **c)** Estimate of the phase. The dotted line in the coherence plot represents the upper limit of the 95% confidence interval for the coherence under the hypothesis that the two processes are independent. The dotted lines in the cross-intensity plot represent approximate 95% confidence intervals under the hypothesis that the two processes are independent plotted around $\sqrt{\hat{P}_Y}$.

The numbers of spikes observed were 2997 for the input and 1400 for the output. The spike trains were replaced, for small intervals h of length 1 msec by a 0-1 valued series of approximately 60000 points (records of approximately 60 seconds). The summation, threshold and recovery functions were again estimated by maximum likelihood, employing the canonical link function, whereas the square root of the cross-intensity function, the coherence and phase were estimated by stochastic point processes techniques.

The inter-spike intervals between the input firings given in Fig.4.5.5a suggest approximately an exponential distribution (i.e. a Poisson process) with a maximum interval between firings of 120 msec whereas the inter-spike intervals between the output firings given in Fig.4.5.5b suggest very approximately a normal distribution centred around 45 msec with minimum and maximum intervals of 24 and 75 msec, respectively.

Fig.4.5.6a, b and c correspond to the estimates of the square root of the cross-intensity function, coherence and phase, respectively. The coherence plot (Fig.4.5.6b) indicates clearly that the two processes are associated with each other over the range of about 0-18 Hz. The estimated phase given in Fig.4.5.6c shows that, over the range of frequencies at which the coherence is significant, the output process is delayed, on average, by an about 1.6 msec with an approximate 95% confidence interval for the delay of (0.81, 2.39) msec. The delay suggested by the phase seems to be consistent with the location of the peak in the estimated cross intensity function. However, Fig.4.5.6c might also be considered as not demonstrating the existence of a simple delay at all.

Fig.4.5.7a, b and c correspond to the estimates of the summation function, recovery and threshold functions and the goodness of fit plot, respectively. We see that while the square root of the cross intensity function

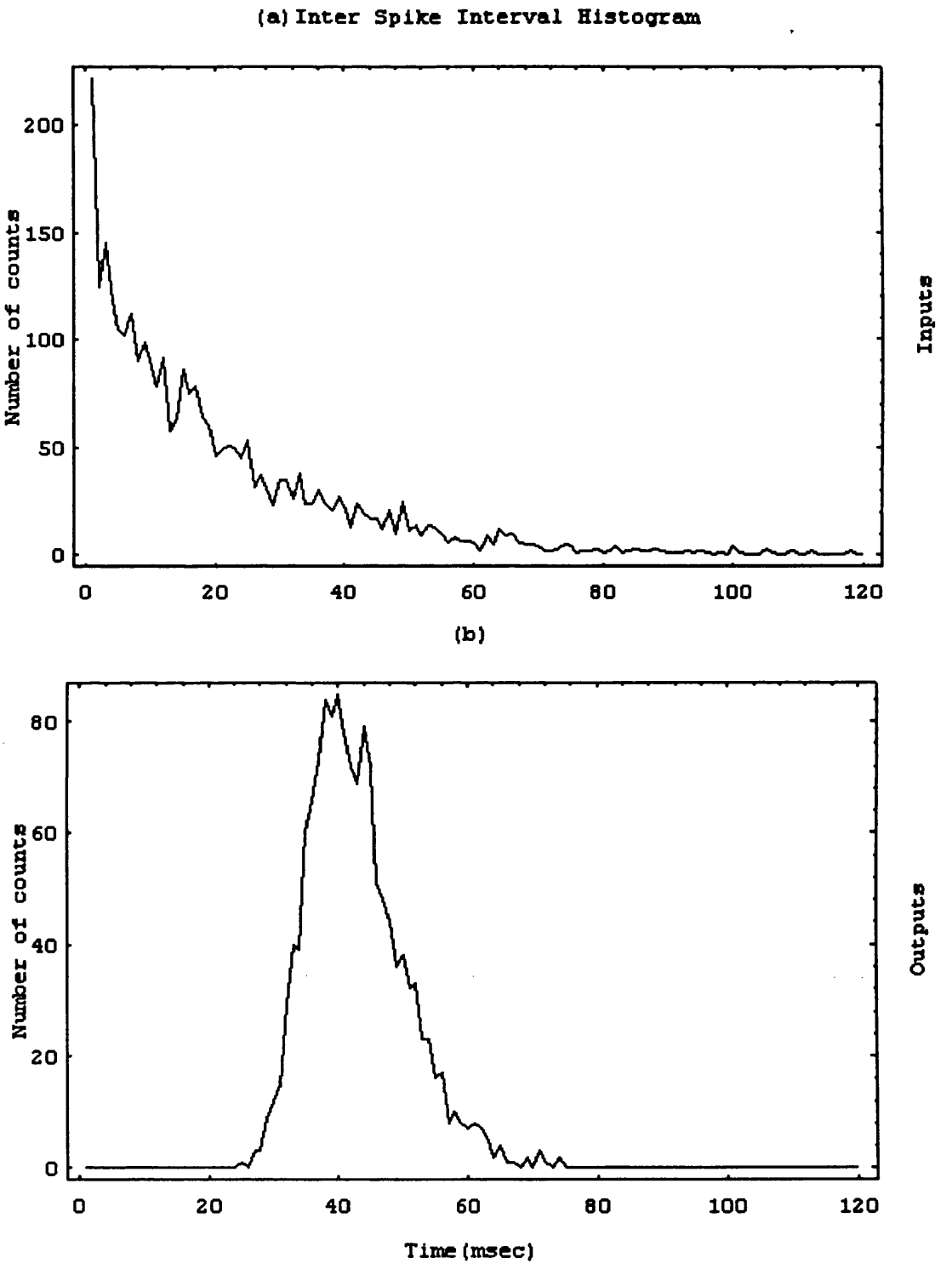


Fig.4.5.5

(a) The inter-spike interval histogram between the input firings.

(b) The inter-spike interval histogram between the output firings.

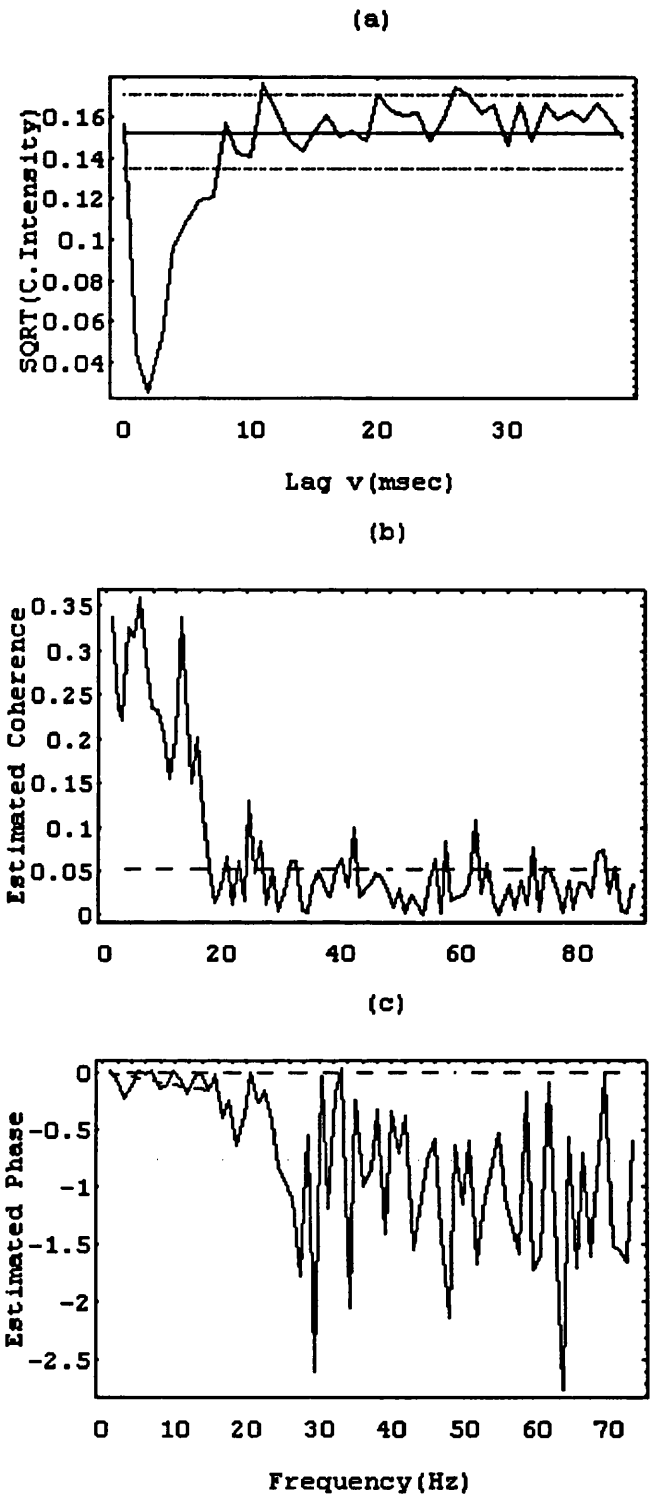


Fig.4.5.6 a) Estimated square root of the cross-intensity function. The dotted lines represent approximate 95% confidence intervals for the square root of the cross-intensity function under the hypothesis that the two processes are independent plotted around $\sqrt{\hat{P}_Y}$. b) Estimate of the ordinary coherence, the dotted line represents the upper limit of the 95% confidence interval for the coherence under the hypothesis that the two processes are independent. c) Estimate of the phase.

given in Fig.4.5.6a indicates that the inhibitory effects of an input last only about 8 msec, in contrast, the summation function given in Fig.4.5.7a reveals effects lasting 35 msec, i.e., $\{\hat{a}_u\}$; $u = 0, 1, ..., 35$, are statistically significant. We keep adding new parameters into the summation function until they stop being statistically significant. The summation function seems more consistent with the way by which the data have been simulated since its significant duration compared to the duration of the square root of the cross-intensity function corresponds closely with the 39 msec half-width of the estimated inhibitory post-synaptic potential (Fig.4.5.8) used as the synaptic input to the neuronal model used for the simulations. It appears that the square root of the cross-intensity function underestimates the underlying inhibitory effects of a synaptic input and combines information about direct synaptic effects and intrinsic membrane properties whereas the likelihood approach seems to be more capable of separating these two effects providing all the input information is presented in the model. Otherwise it appears that the recovery function along with the intrinsic membrane properties of the cell also reflects some of the unmeasured input effects. Consequently, the square root of the cross-intensity function in this case is an inadequate method for addressing the inter-relationships between the two processes.

The recovery and threshold functions (Fig.4.5.7b) are well estimated to about 60 msec (and we note very few intervals between output spikes are longer). The probability of an output spike occurring after 25 msec increases rapidly so that it is unlikely that output spikes are separated by much more than 50 msec, and this corresponds well with Fig.4.5.5b.

The goodness of fit plot given in Fig.4.5.7c indicates that only relatively small probabilities can be predicted since the model can only predict for linear predictor values less than zero and also the confidence intervals

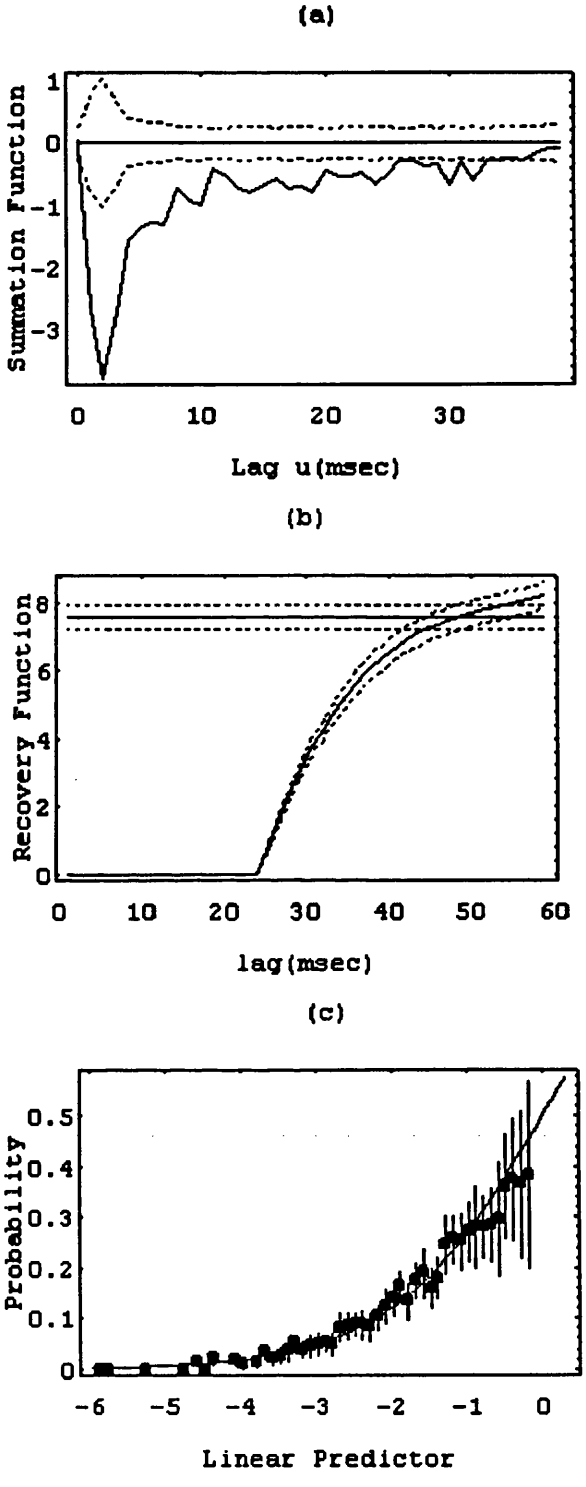


Fig.4.5.7 **a)** Estimated summation function. **b)** Estimated recovery and threshold (upper solid line) functions. The dotted lines give ± two standard error limits plotted about zero in (a) for the summation function and plotted around each function in (b) for the threshold and recovery functions. **c)** The goodness of fit plot with empirical (dots) and theoretical (smooth curve) probabilities plotted against selected values for the linear predictor, the vertical bars give ± two standard error limits for the theoretical probabilities.

near zero are wider. However, the goodness of fit plot seems satisfactory in the sense that the confidence intervals about each of the empirical probabilities are seen to contain the corresponding theoretical probabilities. A logistic link function was used and in terms of the goodness of fit test it was superior to probit or complementary models.

The distance between the two vertical arrows in Fig.4.5.8 measures the half-width of the estimated inhibitory postsynaptic potential (IPSP).

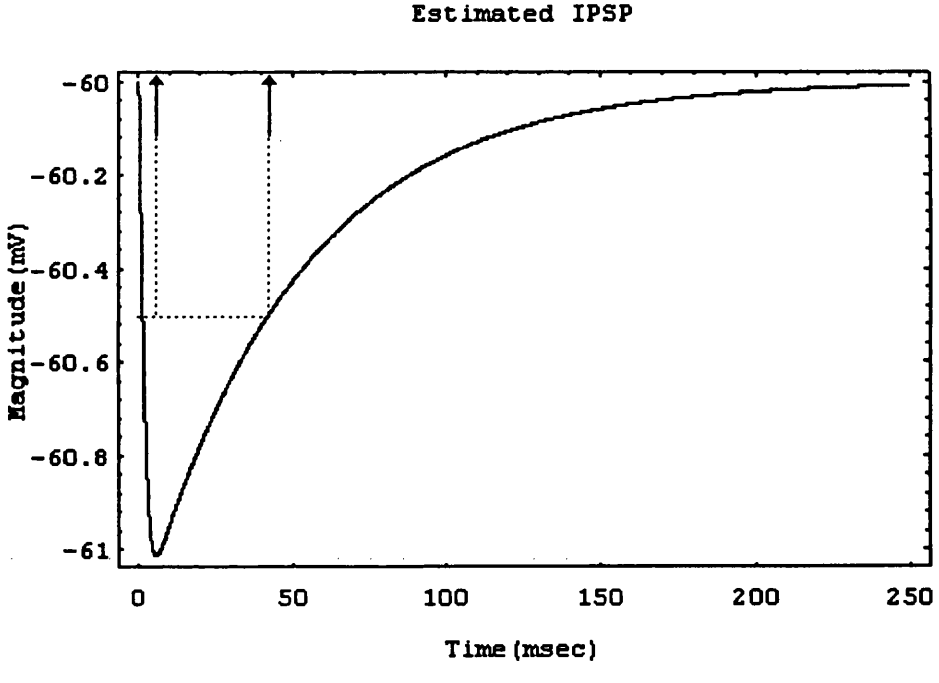


Fig.4.5.8 Illustration of the inhibitory postsynaptic potential(IPSP).

Fig.4.5.9 illustrates the table of deviances and shows that a model with only recovery and threshold functions leads to a greater reduction in deviance than a model with only a summation function. A model with recovery, threshold and summation functions leads to a further reasonable reduction in deviance. Evidently the information contained in the recovery and threshold

functions is almost orthogonal to that contained in the summation function because the sum of the two reductions in deviance when both functions are fitted separately, 5071, is quite similar to the reduction in deviance, 4870, when both functions are fitted together.

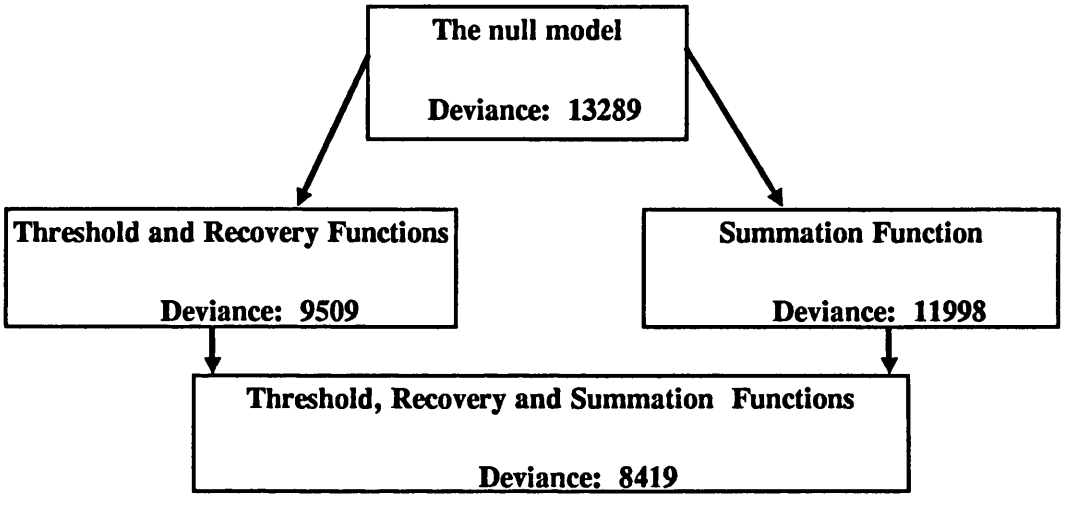


Fig.4.5.9 Diagrammatic representation of the deviance table.

Furthermore, suppose we use the likelihood approach to fit only the cross-intensity function, i.e.

$$U_t = \sum_{v=0}^{\infty} a_v^* x_{t-v}$$

where the set of coefficients $\{a_v^*\}$ represents the cross-intensity function. Here we take into account not only the time of the previous output spike (γ_t) as in the summation function, but consider all previous input spikes (i.e., because that is how the cross intensity function is traditionally calculated).

Fig.4.5.10a gives the estimated cross-intensity function (estimated via the likelihood function). The residual deviance for this model is 12847; a reduction of only 442 from the null model (i.e., a reduction of about 3.33%), providing still further evidence that the cross-intensity function in general has

very poor explanatory power. The goodness of fit plot for a model containing only the cross-intensity function (shown in Fig.4.5.10b) indicates that the fit is a very poor one compared with that for the threshold, recovery and summation functions shown in Fig.4.5.7c. This can be seen both from the very small values of the predicted probabilities, $\hat{P}(\eta) \leq 0.15$, (i.e., the short range of the linear predictor values) and from the departure of the predicted probabilities from the theoretical curve for values of the linear predictor less than -4.7 and above -3.

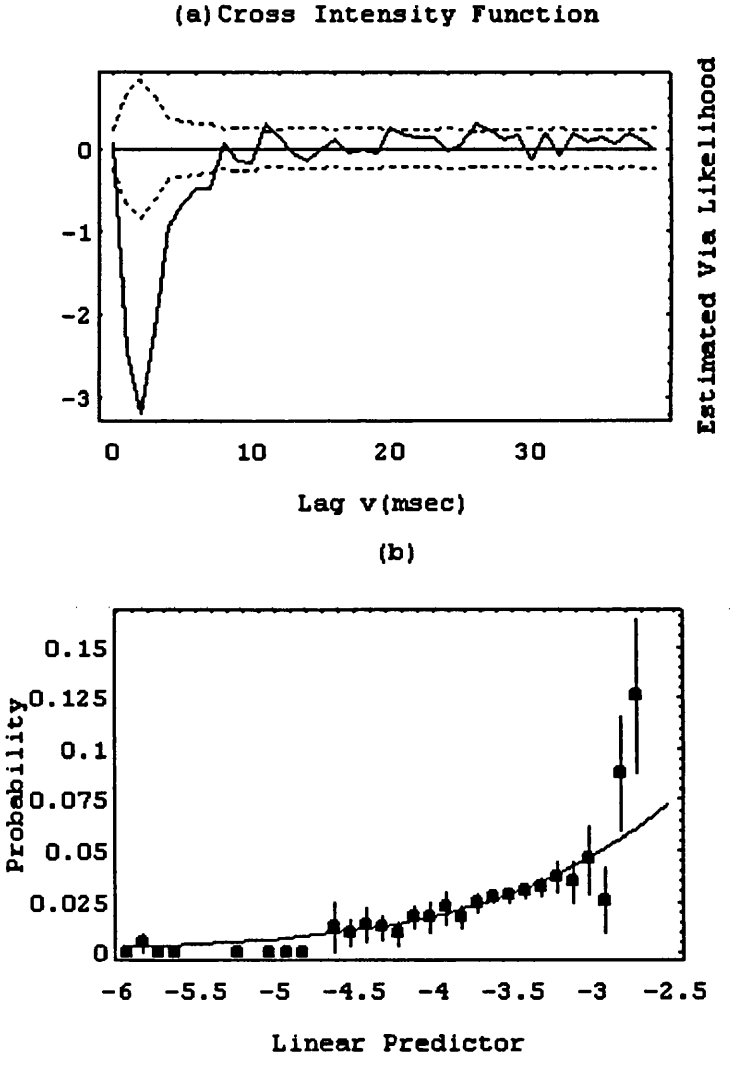


Fig.4.5.10 **a)** The cross-intensity function estimated via likelihood. The dotted lines give $\pm$ two standard error limits for the cross-intensity functions plotted about zero.

b) The goodness of fit plot corresponds to the model given in a).

We can provide a formal F-test based on the change in the deviance and the degrees of freedom under the hypothesis that there is no significant improvement made by adding new terms. The calculated and tabulated F-values are given by:

$$F_{calculated} = \frac{\{change\ in\ deviance\ /\ change\ in\ degrees\ of\ freedom\}}{deviance\ of\ the\ new\ model\ /\ its\ degrees\ of\ freedom}$$
$$= \frac{\{change\ in\ deviance\ /\ (m_1 - m_2)\}}{deviance\ of\ the\ new\ model\ /\ m_2}$$

and

$$F_{tabulated} = F_{m1-m2,\ m2;\ q}$$

where m_1 and m_2 are the degrees of freedom for the old and new models respectively, whereas q is the significance level. The F-test suggests that, the hypothesis that there is no significant effect in adding the threshold and recovery functions to the null model is rejected at any significance level since the tabulated F-value (i.e., $F_{tabulated} = F_{m1-m2,\ m2;\ q} = F_{3,\ \infty;\ 0.999} = 5.42$) is much smaller than the calculated F-value (i.e., $F_{calculated} = 7949.8$). Similarly, and for the same reasons, we reject the hypothesis that there is no significant effect in adding the summation function to the null model. The test also rejected both of the hypotheses of not adding the threshold and recovery functions to a model with only a summation function and that of not adding the summation function to a model with only threshold and recovery functions. We do not explicitly provide these F-tests subsequently in the thesis, but they can easily be derived from the deviance table.

4.5.2 Recovery Function Choice and Threshold Modelling

The recovery function is intended to describe the intrinsic membrane properties of the cell and to allow for spontaneous firing. However, we shall see in chapter 5 that it detects other features also. It is modelled as a polynomial of order k in $(\gamma_t - \zeta_1 - 1)$, where γ_t denotes the time elapsed at time t since the time of the last output spike and ζ_1 denotes the minimum of the output inter-spike intervals.

The order k of the polynomial recovery function depends on the function form that is used in modelling the threshold which may take either a constant or an exponentially decaying form. If U_t represents the membrane potential on the trigger zone of the cell at time t, we have

$$U_t = V_t + \sum_{u=0}^{\gamma_t - 1} a_u x_{t-u} \qquad \text{.....(4.5.4)}$$

where the set of coefficients $\{a_u\}$ make up the discretised summation function and the term V_t represents the polynomial recovery function of order k, i.e.

$$V_t = \begin{cases} \sum_{i=1}^{k} \theta_i (\gamma_t - \zeta_1 - 1)^i \; ; & \gamma_t \geq \zeta_1 + 1 \\ 0 \; ; & \gamma_t \leq \zeta_1 + 1 \end{cases} \qquad \text{.....(4.5.5)}$$

γ_t denotes the time elapsed since the time of the last output spike and ζ_1 denotes the minimum of the output inter-spike intervals. The recovery function given in (4.5.5) is forced to be zero for, $\gamma_t \leq \zeta_1 + 1$ as explained in (4.4). The linear predictor of the model, η_t, representing the difference between the membrane potential and the threshold may be given by

$$\eta_t = U_t - \left(\theta_0 + \mu \exp\{-\lambda \gamma_t\}\right) \qquad \text{.....(4.5.6)}$$

where the term $\left(\theta_0 + \mu \exp\{-\lambda \gamma_t\}\right)$ in (4.5.6) represents the threshold (decaying exponentially). We naturally wish to minimise the number of unknown parameters involved in η_t and particularly in V_t, and an appropriate choice for the threshold assists in this.

From a practical point of view, it is easier to start with a model of the simplest threshold form, i.e., a model with a constant threshold as given in (4.5.3), then to check the order k (i.e., the number of statistically significant parameters) required for the recovery function. If this order is high (say $k > 3$) then we may try a model with an exponentially decaying threshold as given in (4.5.6) above. The model which requires the smaller number of parameters for an adequate fit (among models with similar deviances) is the model we are usually going to choose.

To amplify the above discussion, we demonstrate the threshold modelling and recovery function choice by an example, applying maximum likelihood techniques and then comparing the results obtained for the final model with those obtained via stochastic point processes techniques.

The simulated data set demonstrated here consists of a 0-1 valued series of approximately 60000 sampling points. The unknown parameters used in the summation, threshold and recovery functions were estimated by maximising the likelihood equation given in (3.5.7). The probit link function given in (4.3.2a) was used and in terms of the reduction in deviance, the goodness of fit test and taking the range of the linear predictor a little further, was superior to a logistic model.

The number of spikes observed were 3029 and 1410 for the input and the output, respectively. The inter-spike intervals between the input firings suggest approximately an exponential distribution (i.e., a Poisson process) with minimum and maximum intervals of 1 msec and 168 msec, respectively,

whereas the inter-spike intervals between the output firings suggest a skewed distribution centred around 41 msec with minimum and maximum intervals of 5 msec and 236 msec, respectively.

(a) A Model with a Constant Threshold

The set of data has been used to fit a model with a constant threshold, recovery function and summation function, employing the probit link function, i.e.,

$$P_t = \Phi(\eta_t) = \Phi(U_t - \theta_0)$$

where $\Phi(.)$ is the standard normal cumulative distribution function and the terms η_t, U_t and θ_0, are as defined in expressions (4.5.1), (4.5.2) and (4.5.3).

Fig.4.5.11a represents the estimated summation function and estimated 95% standard error limits plotted about 0 and suggests that the effective duration of an excitatory input on the output is about 11 msec, i.e., $\{\hat{a}_u\}$; $u = 0, 1, ..., 11$, are statistically significant. The order of the polynomial recovery function needed is a cubic (i.e. $k = 3$), since $\{\hat{\theta}_i\}$; $i \geq 4$ are not statistically significant. The constant threshold (upper solid line) and (cubic) recovery (lower solid curve) functions given in Fig.4.5.11b indicate that the probability of having an output spike spontaneously is relatively small as can be seen both from the comparatively small reduction in deviance when the recovery function is added, and from the wide distance between the recovery function and the threshold over the whole range of the intervals between output spikes. Another point to notice here is that the recovery function is badly estimated at intervals greater than 75 msec because very few intervals between output spikes are this long.

The goodness of fit plot given in Fig.4.5.11c indicates that only relatively small probabilities can be predicted and also the confidence intervals

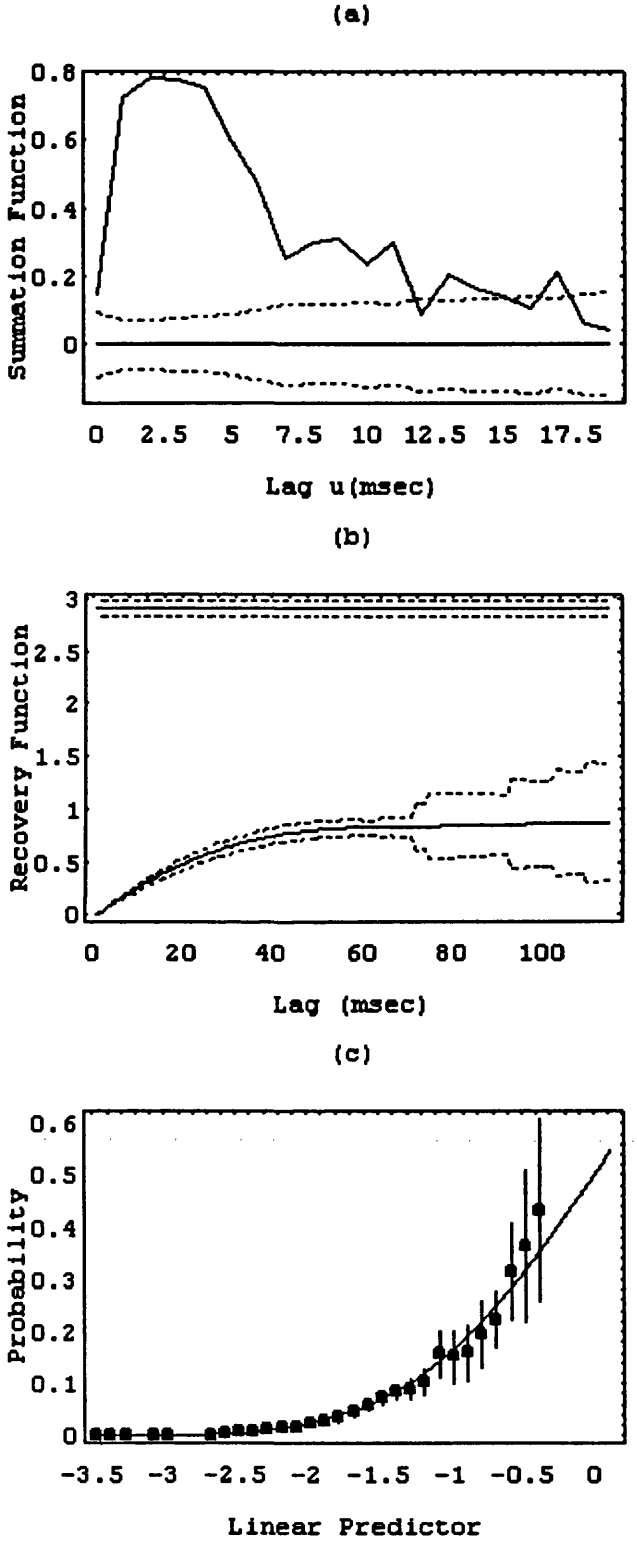


Fig.4.5.11 **a)** Estimated summation function. **b)** Estimated recovery and threshold (upper solid line) functions. The dotted lines give ± two standard error limits plotted about zero in (a) for the summation function and plotted around each function in (b) for the threshold and recovery functions. **c)** The goodness of fit plot with empirical (dots) and theoretical (smooth curve) probabilities plotted against selected values for the linear predictor, the vertical bars give ± two standard error limits for the theoretical probabilities.

near zero are wider. However, the goodness of fit plot seems satisfactory in the sense that the confidence intervals about each of the empirical probabilities are seen to contain the corresponding theoretical probabilities.

(b) A Model with an Exponentially Decaying Threshold

We may use an exponential instead of a constant threshold in the model along with a recovery function, V_t, and a summation function, i.e. setting

$$P_t = \Phi(\eta_t) = \Phi\left(U_t - \left(\theta_0 + \mu \exp\{-\lambda \gamma_t\}\right) \right)$$

where $\Phi(.)$ is the standard normal cumulative distribution function and the terms η_t, U_t and $\left(\theta_0 + \mu \exp\{-\lambda \gamma_t\}\right)$, are as defined in expressions (4.5.4), (4.5.5) and (4.5.6). This reduced the order k of the polynomial needed for the recovery function, V_t, from order 3 to order 1, since $\{\hat{\theta}_i\}$; $i \geq 2$ were not statistically significant. It also improved the fit of the model a little since it produced a slightly larger reduction in deviance than a model with a constant threshold as we will see shortly from the deviance table, although both models had the same number of parameters (i.e., both in terms of the threshold and recovery functions and because the summation function lengths were identical). Furthermore, the parameters in the threshold and recovery functions are not physiologically meaningful parameters and we cannot give them any direct interpretation. But the significant thing we are looking at is in fact the difference between these two functions (i.e., the more the recovery function approaches the threshold, the higher the probability of an output spike).

The summation function given in Fig.4.5.12a was almost identical to that given in Fig.4.5.11a above and both are with the same number of

parameters. This suggests that, in this example, the change in the threshold and recovery functions has very little effect on the summation function behaviour. This is good, as it suggests that the effect of the single input is largely contained in the summation function, and not in the recovery function.

The exponential threshold (upper solid curve) and (first order) recovery (lower solid line) functions given in Fig.4.5.12b suggest a similar interpretation as in the constant threshold model except that the recovery function in this case is well estimated (though very small in effect) over the entire range of intervals because the linear parameter is well estimated, whereas the cubic term in (4.5.5) leads to greater variability for large values of γ_t. Again we have a satisfactory goodness of fit plot (Fig.4.5.12c), very similar to Fig.4.5.11c, which makes distinguishing between the two models using this type of goodness of fit plot very difficult.

The table of deviances given in Fig.4.5.13 suggests that a model with the summation function alone leads to a modest reduction in the deviance. A model with an exponentially decaying threshold and first order polynomial recovery function alone leads to a slightly larger reduction in the deviance than that with a constant threshold and third order polynomial recovery function, but not as much as the reduction made by the summation function. The largest reduction in deviance has been achieved with a model with an exponential threshold, first order recovery function and summation function and it seems to be the best model, although only marginally better than that with the constant threshold, cubic recovery function and summation function, since the deviances are similar.

The square root of the estimated cross-intensity function as a time domain measure of the degree of association between the two processes given in Fig.4.5.14a indicates an excitatory input effect lasting about 6 msec only. This duration is shorter than that suggests by the estimated summation function

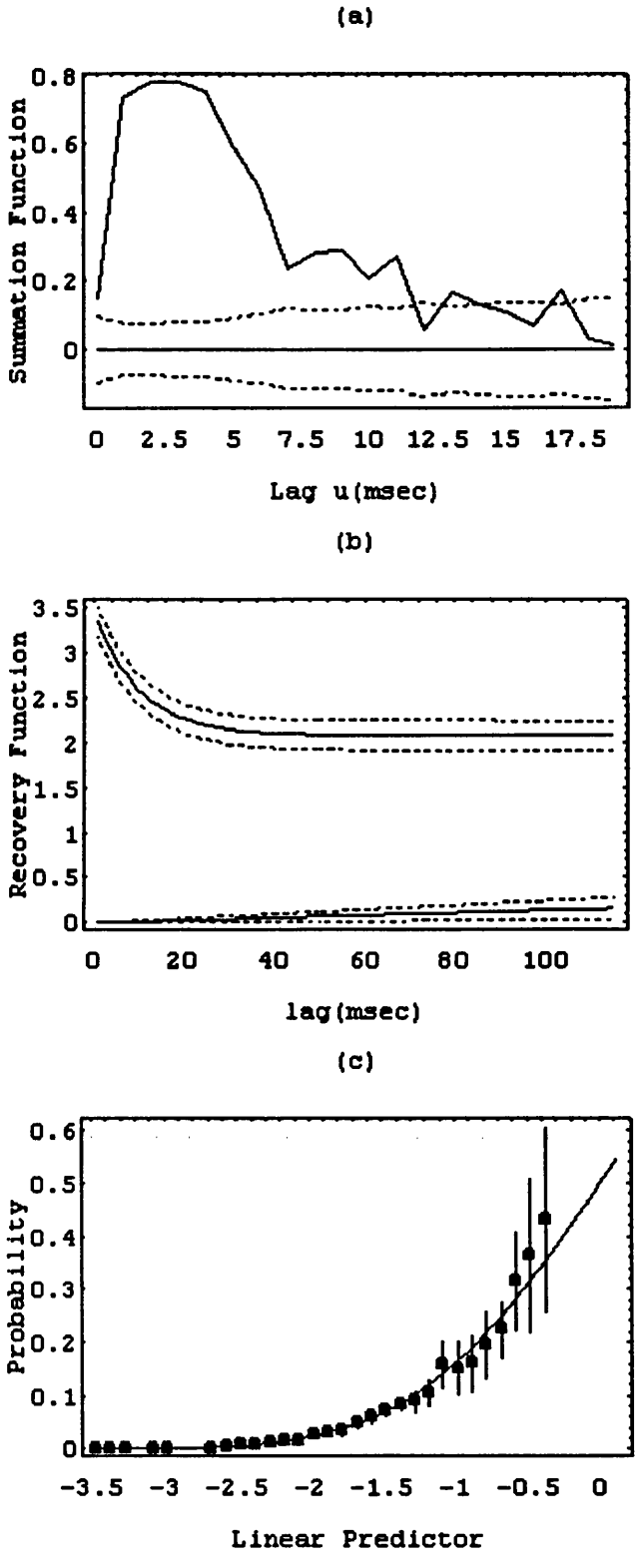


Fig.4.5.12 **a)** Estimated summation function. **b)** Estimated recovery and threshold (upper solid curve) functions. The dotted lines give ± two standard error limits plotted about zero in (a) for the summation function and plotted around each function in (b) for the threshold and recovery functions. **c)** The goodness of fit plot with empirical (dots) and theoretical (smooth curve) probabilities plotted against selected values for the linear predictor, the vertical bars give ± two standard error limits for the theoretical probabilities.

which is about 11 msec (Fig.4.5.12a). This is another example where the square root of the estimated cross-intensity function significantly underestimates the duration of the underlying excitatory effects of a synaptic input. Furthermore, the residual deviance when we use the likelihood approach to estimate the cross-intensity function (as explained in section 4.5.1-b) is 12320; a reduction of only 1044 from the null model (i.e., a reduction of about 7.8%), indicating that the cross intensity function has poor explanatory power compared to the summation function, where the reduction in deviance is 1617.

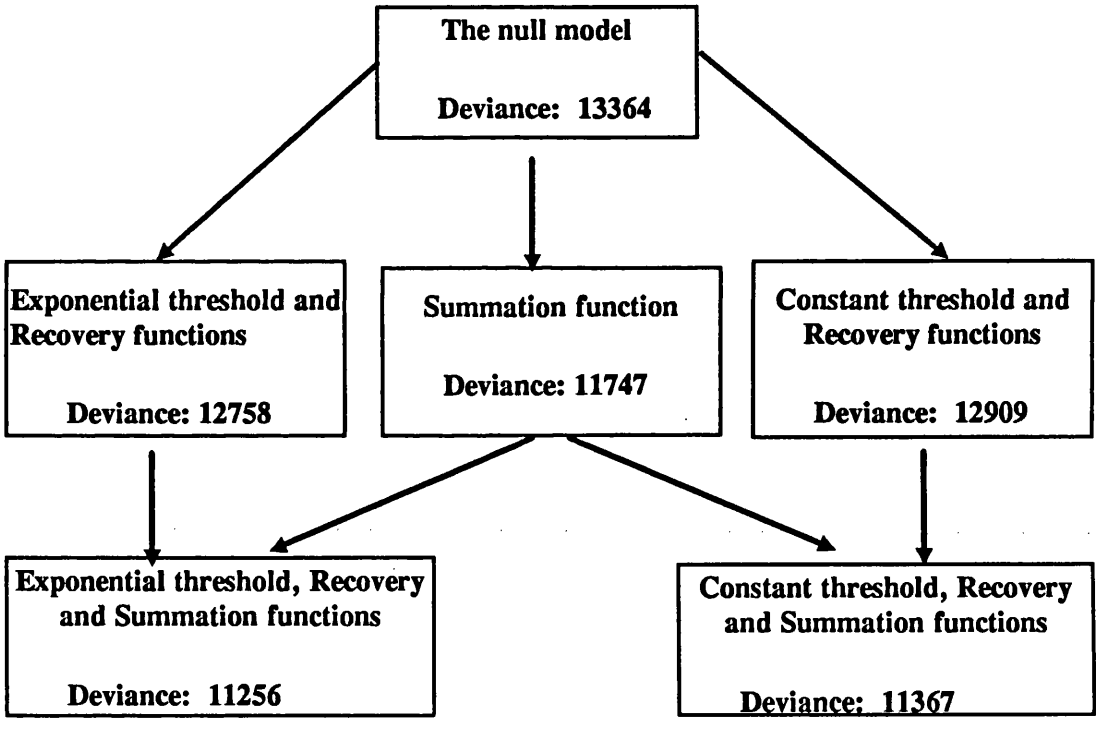


Fig.4.5.13 Diagrammatic representation of the deviance table.

The estimated coherence (Fig.4.5.14b) gives clear evidence that the two processes are significantly coupled over the range of frequencies of about (0, 90) Hz. The estimated phase given in Fig.4.5.14c shows that, over the range of frequencies at which coherence is significant, the output process is delayed,

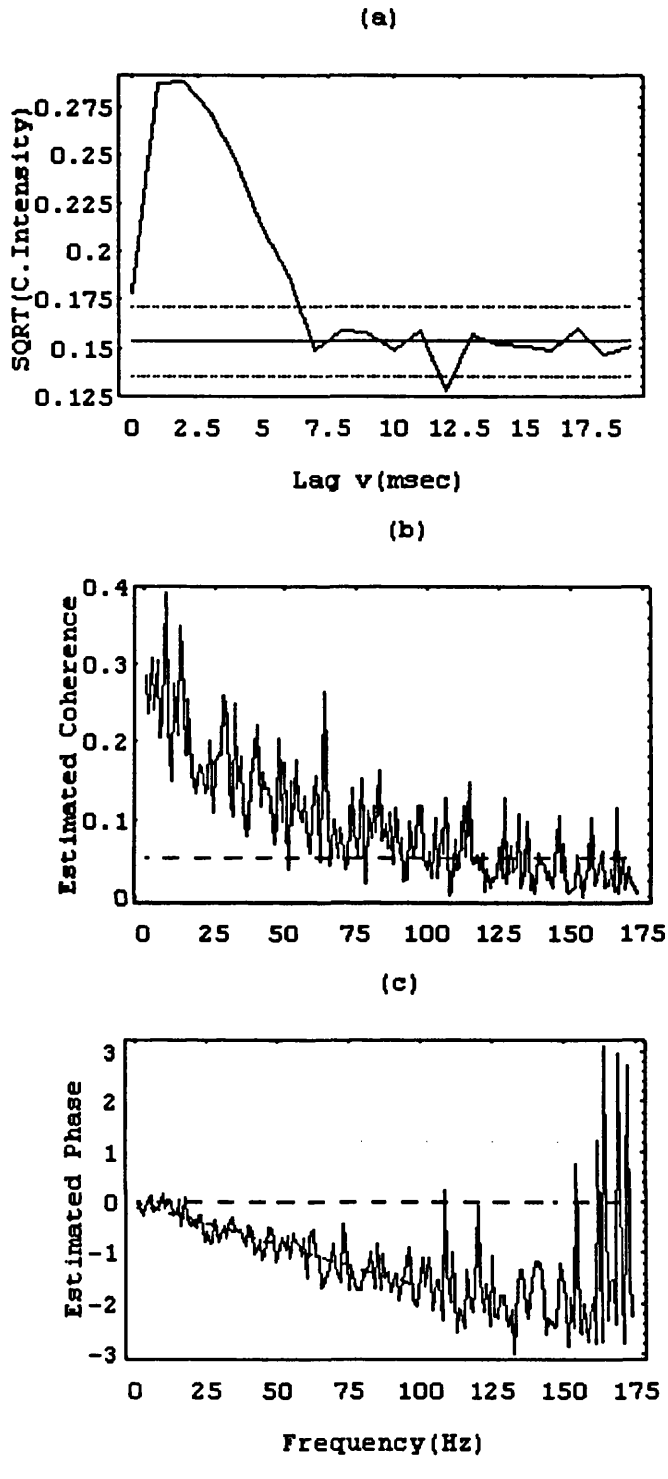


Fig.4.5.14 **a)** Estimated square root of the cross-intensity function. The dotted lines represent approximate 95% confidence intervals for the square root of the cross-intensity function under the hypothesis that the two processes are independent plotted around $\sqrt{\hat{P}_Y}$. **b)** Estimate of the ordinary coherence, the dotted line represents the upper limit of the 95% confidence interval for the coherence under the hypothesis that the two processes are independent. **c)** Estimate of the phase.

on the average, by about 2.55 msec with a 95% confidence interval, for the delay, of (2.4, 2.7) msec. The value for the delay suggested by the phase seems to be consistent with the peak at 2 msec in the square root of the estimated cross intensity function. In this particular example the peak in the summation function is centred about 2.5 msec. This is also close to the value for the delay estimated from the phase. In this example the data set is input dominated, and the delay appears as the dominant feature. This is not always the case, as we will see in subsequent examples.

4.6 Carry-over Effect Function (COE)

The effects of an input on the output have been discussed both physiologically in chapter 1, and statistically earlier in this chapter. The likelihood characterises these effects in a more informative way through the summation function, recovery and threshold functions compared to the time domain through the square root of the cross intensity function. But these effects have been considered only for those input spikes occurring after the time since the last output spike occurred, ignoring those occurring before the time of the previous output spike. This physical constraint assumes that each input spike only contributes to the firing of the next output spike, then remains without any significant effect on any subsequent output firing. These input postsynaptic effects are estimated by the summation function as we have seen before.

From the physiological point of view, it is important however to investigate these input postsynaptic effects (at lag w as shown in Fig.4.5.1) to check if they still have any significant effects after the subsequent output firing or if they just die away immediately after the next output spike has been produced.

The input postsynaptic effects therefore, can be divided into two types of effect. First is the effect of those occurring at times after the time of the last output spike (at lag u as shown in Fig.4.5.1). This type of effect has been represented by the summation function as we discussed earlier. Second is the effect at any given time t on the output of those input spikes occurring at times prior to the time of the previous output spike (at lag w as shown in Fig.4.5.1). This type of effect is termed a carry-over effect and the function quantifying these effects is called the carry-over effect function (COE) and takes the form

$$C_t = \sum_{w \geq \gamma_t} c_w x_{t-w} \qquad \text{.....(4.6.1)}$$

where γ_t is the time elapsed at time t since the time of the last output spike and the set of coefficients $\{c_w\}$ make up the carry-over effect function.

This capability of separating out the input postsynaptic effects into two components shows clearly the flexibility of the maximum likelihood approach which gives the likelihood one further advantage. The time domain and frequency domain approaches reviewed in chapter 2 seem entirely incapable of separating out the two types of input postsynaptic effects and therefore do not provide an analogous measure for the carry-over effects of a synaptic input. This gives the likelihood approach an advantage and superiority over the time domain and frequency domain approaches.

We demonstrate the above by an example where input carry-over effects are apparently present. We apply maximum likelihood to the set of data, to estimate the threshold, recovery, summation and the carry-over effect functions. Also we compare the likelihood results with those obtained via stochastic point process techniques.

The simulated data set demonstrated here consists of a 0-1 valued series of approximately 60000 sampling points. The unknown parameters used in the estimate of the threshold, recovery, summation and the carry-over effect functions were estimated by maximising the likelihood equation given in (3.5.10), employing the logistic link function given in (4.3.1a), since it was superior to the probit link function in terms of the reduction in deviance, the goodness of fit test and taking the range of the linear predictor a little further. The linear predictor of the model, η_t, representing the difference between the membrane potential and the threshold is given by

$$\eta_t = U_t - \theta_0 \qquad \text{.....(4.6.2)}$$

where θ_0 represents the constant threshold and the term U_t represents the membrane potential on the trigger zone of the cell at any given time t, and may be given in its approximate discrete form by

$$U_t = V_t + \sum_{u=0}^{\gamma_t - 1} a_u x_{t-u} + \sum_{w \geq \gamma_t} c_w x_{t-w} \qquad \text{.....(4.6.3)}$$

γ_t is the time elapsed at time t since the time of the last output spike, the two sets of coefficients $\{c_w\}$ and $\{a_u\}$ make up the carry-over effect and summation functions respectively and the term V_t, as previously defined, represents the polynomial recovery function of order k.

The number of spikes observed were 2991 and 4345 for the input and the output, respectively. The inter-spike intervals between the input firings suggest approximately an exponential distribution (i.e., a Poisson process) with minimum and maximum intervals of 1 msec and 120 msec, respectively, whereas the inter-spike intervals between the output firings suggest approximately a normal distribution centred around 18 msec with minimum and

maximum intervals of 2 msec and 41 msec, respectively. A GENSTAT printout of the results obtained for this data set is illustrated in appendix (C).

The summation function given in Fig.4.6.1a reveals an excitatory effect of an input lasting about 13 msec. Also we have an evident carry-over effect as shown in Fig.4.6.1b, lasting from about 2 msec to 20 msec. It is a relatively small effect, as can be seen both from the comparatively small reduction in deviance from a model with only summation, threshold and recovery functions to a model when a carry-over effect function is added, and from the fact that the parameters of the summation function tend to be much more statistically significant. Nonetheless it is a real effect, and the appropriate F-test for its inclusion in the model was significant. The threshold and recovery functions (Fig.4.6.1c) are well estimated to about 35 msec (and we note very few intervals between output spikes are any longer) and suggest that the probability of an output spike is small up to about 10 msec after the previous output spike, but then increases rapidly. The chance of crossing the threshold is substantial after about 20 msec.

The goodness of fit plot given in Fig.4.6.2 seems very reasonable and indicates that, in this particular example, large probabilities can be predicted by the likelihood model. The model takes the range of the linear predictor to about 1.6 which reflects the relatively large explained variability by the model compared to that in previous examples. And the confidence intervals are reasonably narrow.

Fig.4.6.3 illustrates the deviance table and shows that both summation function and recovery function when fitted separately lead to a moderate reduction in the deviance whereas a model with both functions fitted together leads to a much larger reduction in the deviance. Adding the carry-over effect function to a model with the summation function alone or to a model with both recovery and summation functions makes a relatively small difference.

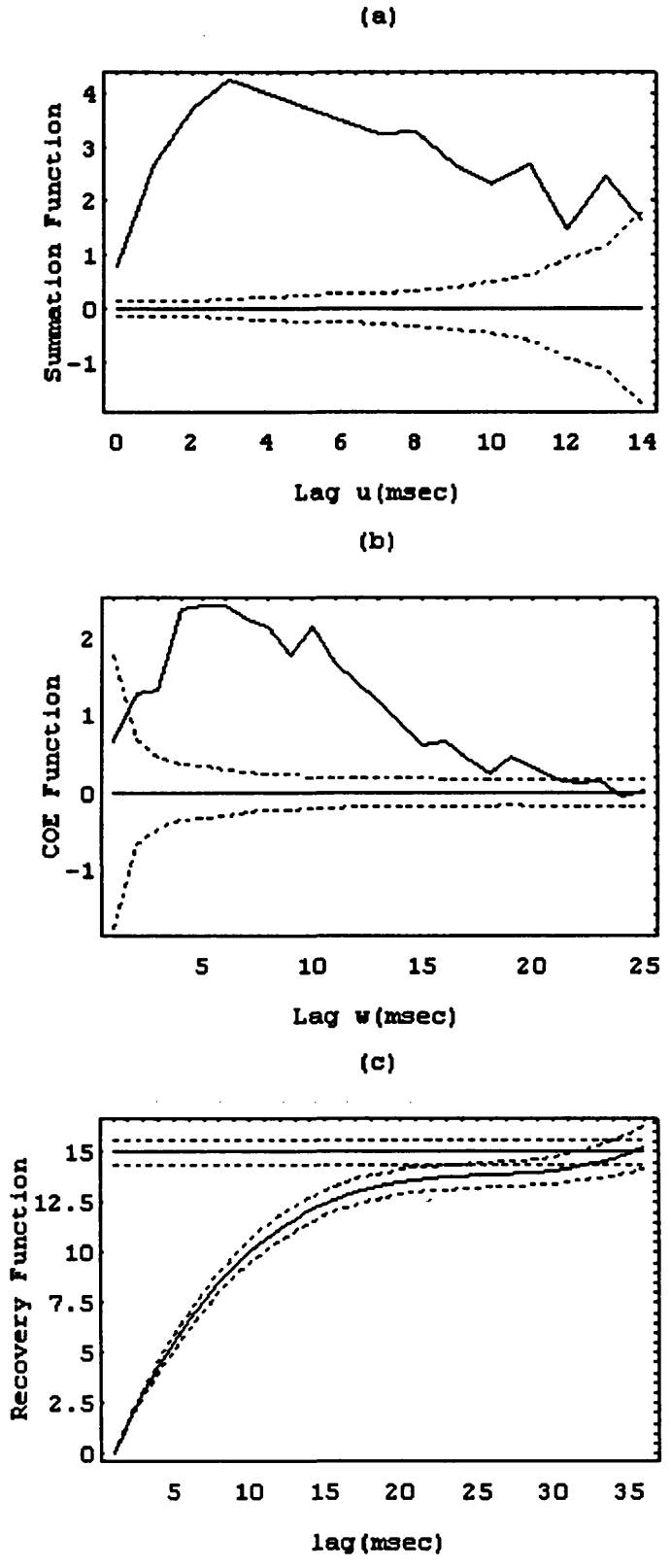


Fig.4.6.1 a) Estimated summation function. b) Estimated carry-over effect function. c) Estimated recovery (lower curve) and threshold (upper solid line) functions The dotted lines give $\pm$ two standard error limits plotted about zero in (a) and (b) for the summation and the corresponding COE functions and plotted around each function in (c) for the threshold and recovery functions.

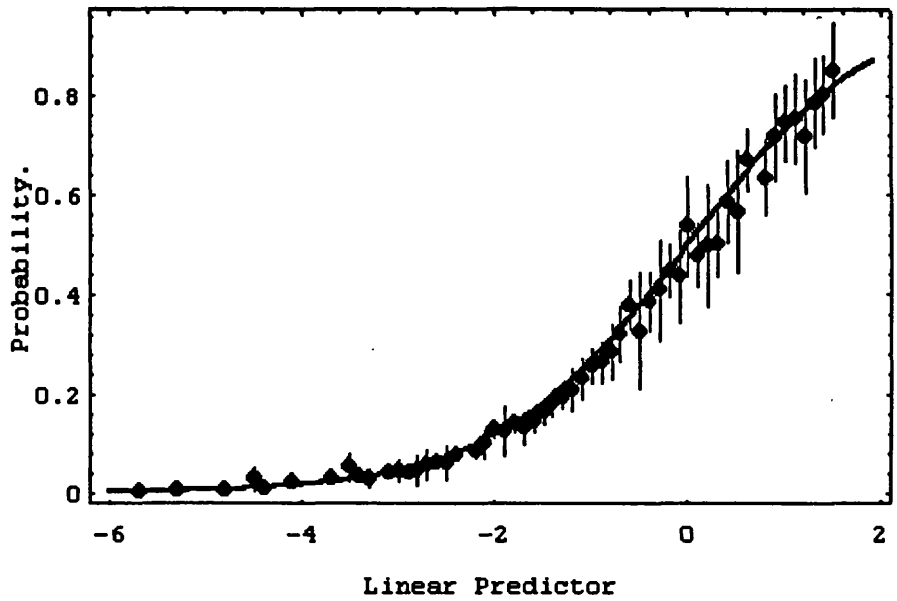


Fig.4.6.2 The goodness of fit plot with empirical (dots) and theoretical (smooth curve) probabilities plotted against selected values for the linear predictor, the vertical bars present the ± two standard error limits for the theoretical probabilities.

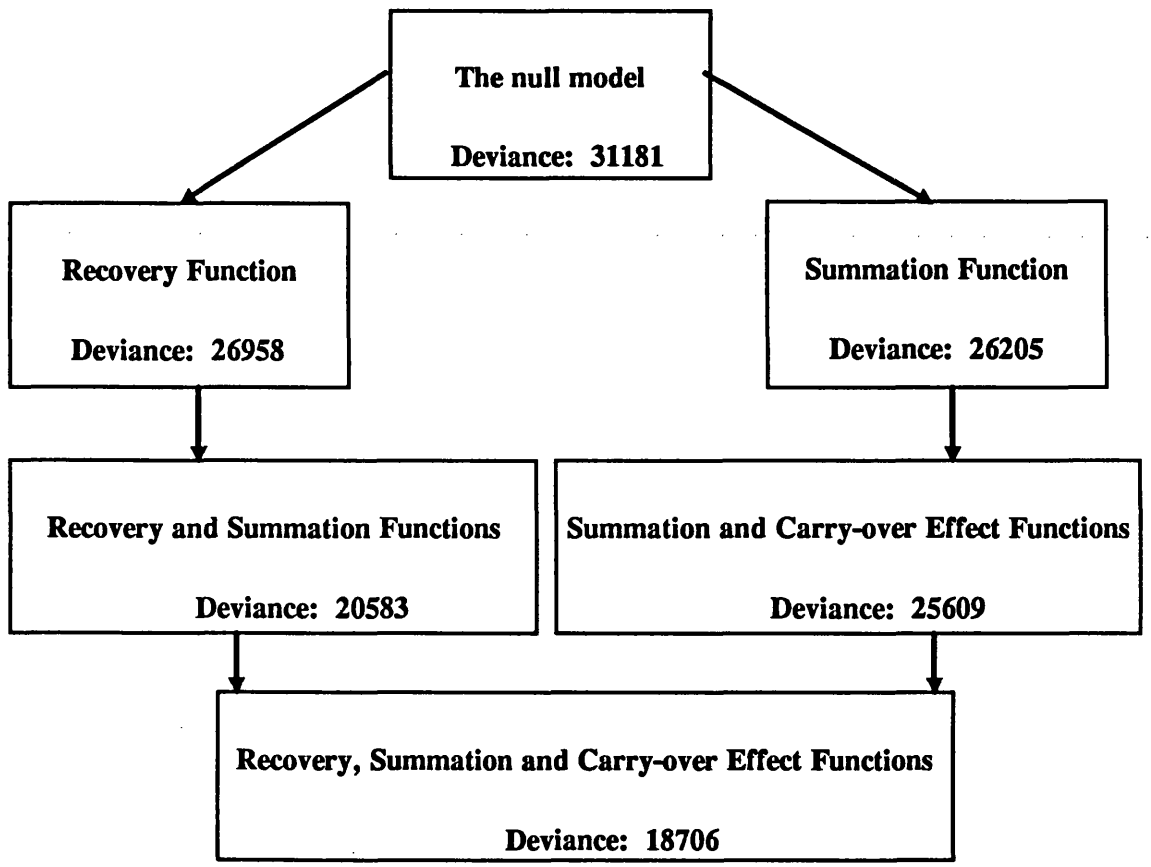


Fig.4.6.3 Diagrammatic representation of the deviance table.

The square root of the estimated cross-intensity function given in Fig.4.6.4a indicates an excitatory effect of an input lasting about 5 msec only. As in most previous examples, the square root of the estimated cross-intensity function provides little of the information available in the 4 components (threshold, recovery, summation and carry-over effect functions) of the likelihood model. The estimate of the cross-intensity function using the likelihood approach (as explained in section 4.5.1-b) providing the evidence that the cross-intensity function has very poor explanatory power as the residual deviance obtained from the likelihood is 29359; a reduction of only about 5.8 %, whereas the residual deviance obtained using the summation function model is 26205.

Fig.4.6.4b gives the estimated coherence and suggests that the two processes are strongly coupled over the range of frequencies of about (0, 75) Hz. The estimated phase given in Fig.4.6.4c shows that, over the range of frequencies at which coherence is significant, the output process is delayed, on the average, by 1.87 msec with a 95% confidence interval, for the delay, of (1.76, 1.98) msec. The value for the delay suggested by the phase is moderately close to the peak in the square root of the estimated cross intensity function. In this example the peak in the summation function is centred about 3 msec. This differs from the estimated time of the peak in the cross intensity function and from the value of the delay estimated from the phase. Unlike the previous example, this data set is not input dominated and there is a lack of agreement between the summation function and the cross intensity function. The situation seems to be more complicated than what is suggested by the phase and the cross intensity function, and it seems that perhaps summarising this in a single number and calling it a "delay" may not always be appropriate.

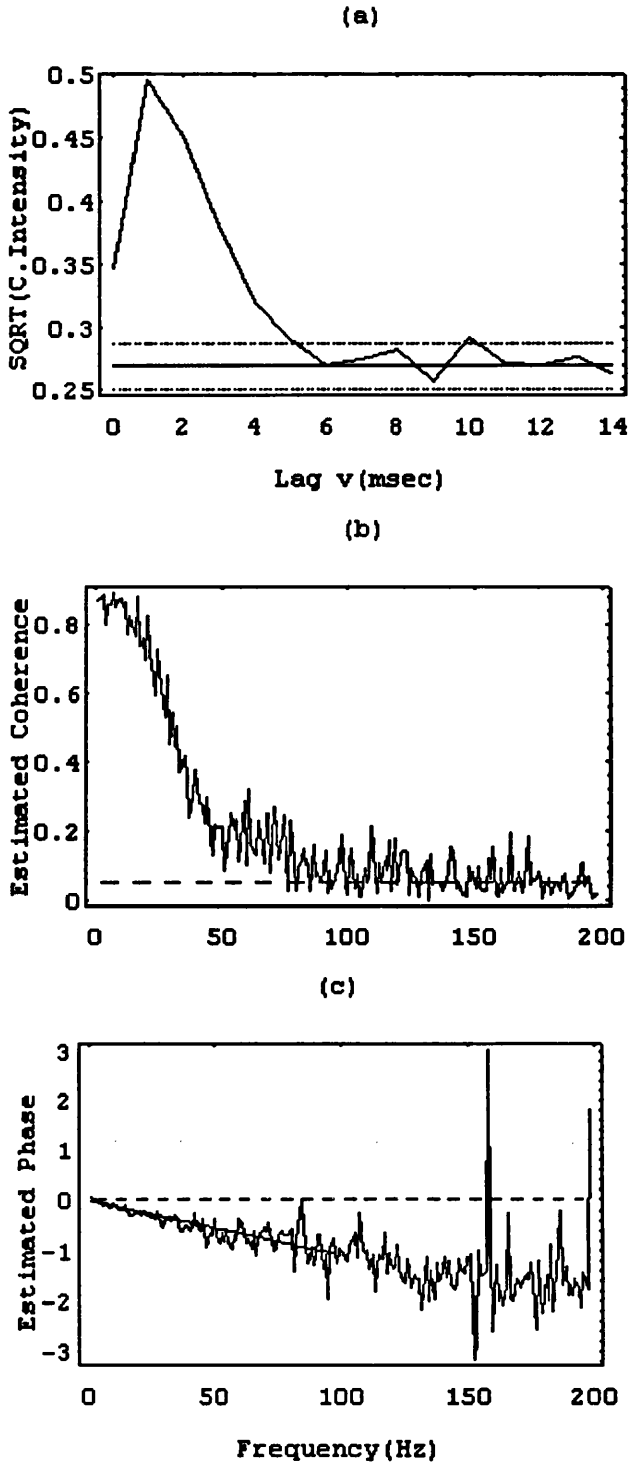


Fig.4.6.4 **a)** Estimated square root of the cross-intensity function. The dotted lines represent approximate 95% confidence intervals for the square root of the cross-intensity function under the hypothesis that the two processes are independent plotted around $\sqrt{\hat{P}_Y}$. **b)** Estimate of the ordinary coherence, the dotted line represents the upper limit of the 95% confidence interval for the coherence under the hypothesis that the two processes are independent. **c)** Estimate of the phase.

4.7 Conclusion

In this chapter the likelihood estimation procedure has been applied to various different neuronal spike train data sets, and there are many interesting features to which we draw attention. The main features can be summarised in the following points. First is the analysis of a spontaneous discharge data set (section 4.4), the case where we are only able to estimate the threshold and recovery functions. Here stochastic point process techniques do not provide an analogous measure of the spontaneous behaviour of the cell, so that the likelihood approach has a total advantage over time and frequency domain analyses. Second is the analysis of single input and single output neuronal spike train data sets, and we introduced the idea of a carry-over effect (COE) of the synaptic inputs on the firing of a neurone and we have estimated the carry-over effect function which quantifies these effects. As in the analysis of the spontaneous discharge data, the time and frequency domains do not provide an analogous measure of the carry-over effect of the synaptic inputs which again gives the likelihood approach the advantage over the stochastic point process techniques. Also these demonstrations suggest again that the square root of the cross intensity function is difficult to interpret and may be misleading and underestimating the underlying excitatory and inhibitory effects of a synaptic input. Furthermore, the cross intensity function seems to reduce the residual deviance less than the summation function. Therefore, the cross intensity approach is a poor method of investigating the association between processes. The likelihood approach is able to separate the aspects of the relationship between spike trains through the threshold, recovery, summation and carry over effect functions and such is not provided by stochastic point process techniques.

Chapter 5

5 Likelihood Applications to Multiple Input and Single Output Data

5.1 Introduction

The usefulness of the likelihood approach in analysing neural spike train data with a single input and a single output, discussed in the previous chapter, leads to a further consideration of a wide range of questions relating to more realistic situations when we extend the use of the likelihood approach to the case of neural spike train data with two inputs and a single output. We may also include a continuous input representing an "unobservable" input which can be used in the simulation as a stimulus to match comparable experimental data. Thus we use this continuous input to simulate the effects of all other unmeasured or unmeasurable inputs. This extension will give the likelihood model the ability to achieve greater insight into the processes involved, and to reflect major features of the cell. This will be seen from the very substantial reduction in the deviance as the model receives more input information.

The main aim of this chapter is to extend the application of the likelihood approach to the above case. This is to show that the approach is

sufficiently flexible, and it may further be extended in principle to the case of an arbitrary number of neurone inputs. Another target of this chapter is again to compare the results obtained using likelihood with those obtained using stochastic point process techniques.

We demonstrate the application of the maximum likelihood approach in this chapter with three simulated sets of data. For the first set of data we have one observed (spike train) input and one "unobservable" input, and one observed output whereas two observed (spike train) inputs and one "unobservable" input and one observed output are presented in the other two sets of data. The demonstration will also include a set of real spike train data recorded from the muscle spindle. The simulation was done again by using a conductance based neuronal model using an excitatory input. To increase the output firing rate in order to match comparable experimental data, a continuous input representing a population of "unobservable" inputs has been used to stimulate the cell (Halliday, 1994). The real data were obtained from a muscle spindle lying within the tenuissimus muscle in the hind limb of a deeply anaesthetised cat. More details of the recording procedure will be illustrated in the chapter.

The general application of the likelihood methods to the case of more than one input is the same as that in the case of single input discussed in chapter 4. But it has some computational implications (i.e., it requires more computational space and computational time than the case of a single observed input). For this computational reason the data sets used in this chapter have been either regrouped with binwidth two (i.e., $h = 2$ msec) in the cases where the data then remains as a 0-1 valued series (i.e., maintains the 0-1 property required for the binomial distribution) or have been split up into two or more disjoint segments if the 0-1 property cannot be maintained. In this latter case,

the individual likelihood estimates for each segment were averaged to obtain the final estimates.

We start the demonstration with simulated data with one observed (spike train) input and one "unobservable" input, and one observed output. Then we discuss the case of two observed spike train inputs and one "unobservable" input, and one observed output.

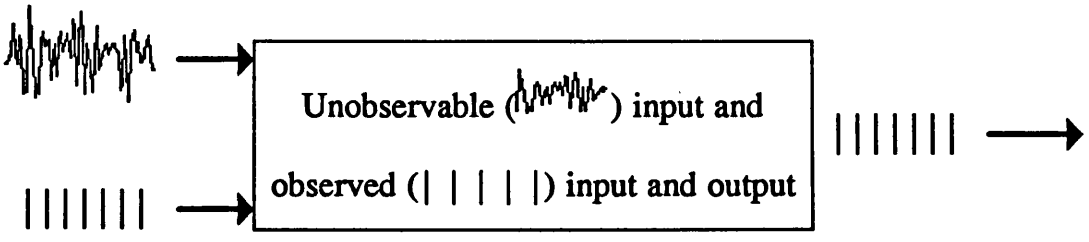
5.2 Analysis of Simulated Neuronal Spike Train Data

5.2.1 "Unobservable" Effect Function (UOE)

It is known that nerve cells are not isolated but rather interconnected with each other. The firing of nerve cells mostly depends on the influence imposed by a large number of neighbouring neurones (as many as 2000 neurones).

From the computational point of view, it seems possible only to investigate the behaviour of a very small number of neurones within a moderate sized network of neurones and therefore the effects of a large number of neurones on the firing of a particular neurone will not be available. We refer to them as "unobservable" (or unmeasured) inputs. This is more obvious in real experimental situations where the signals of neurones can only be recorded simultaneously from a small number of cells. In simulations, however, this is a less difficult problem, since the model used to simulate an artificial observed neuronal spike train data can also provide simulated synaptic data from a population of "unobservable" inputs by using a non-zero mean normal distribution. These "unobservable" inputs are used as a stimulus to the cell to increase its firing rate as well as to mimic the behaviour of real cells. For more details see Lüscher (1990) and Halliday (1994), also see section (1.6) and appendix (A) of this thesis.

Thus we can use the likelihood approach to analyse a set of simulated data which contains along with the observed input and output spike train data, a continuous input z_t which is discretised over small intervals of length 1 msec. This has been scaled such that z_t takes on values between 0 and 1 and represents the "unobservable" inputs (Halliday, 1994). A model with a single input and output together with "unobservable" inputs may be represented diagramatically as follows



The simulated data set demonstrated here consisted of a 0-1 valued series of approximately 60000 points for each of the observed input and output, along with 60000 points for the continuous input which represents the "unobservable" inputs. For computational purposes the data were split up into three disjoint segments each of 19960 points and the individual maximum likelihood estimates for each segment were averaged to obtain the final estimates. The linear predictor of the model, η_t, will take the same form as in (4.6.2) and the membrane potential on the trigger zone of the cell at any given time t may be given in its approximate discrete form by

$$U_t = V_t + \sum_{u=0}^{\gamma_t - 1} a_u x_{t-u} + \sum_{w \geq \gamma_t} c_w x_{t-w} + \sum_{u=0}^{\gamma_t - 1} b_u z_{t-u} + \sum_{w \geq \gamma_t} d_w z_{t-w} \qquad \text{.....(5.2.1)}$$

where γ_t is the time elapsed at time t since the time of the last output spike, the two sets of coefficients $\{a_u\}$, $\{c_w\}$ represent the summation and carry-over effect functions for the observed input, respectively, whereas the two sets of coefficients $\{b_u\}$, $\{d_w\}$ represent the summation and carry-over effect functions for the "unobservable" input, respectively. The term V_t, as defined earlier, represents the polynomial recovery function of order k.

The numbers of spikes observed were 2398 and 2991 for the observed input and output, respectively. The inter-spike intervals between the input firings suggest approximately an exponential distribution (a Poisson process) with minimum and maximum intervals of 1 msec and 135 msec respectively, whereas the inter-spike intervals between the output firings suggest a skewed distribution with a mode around 24 msec and with minimum and maximum intervals of 9 msec (i.e. ζ_1=9 msec) and 51 msec, respectively.

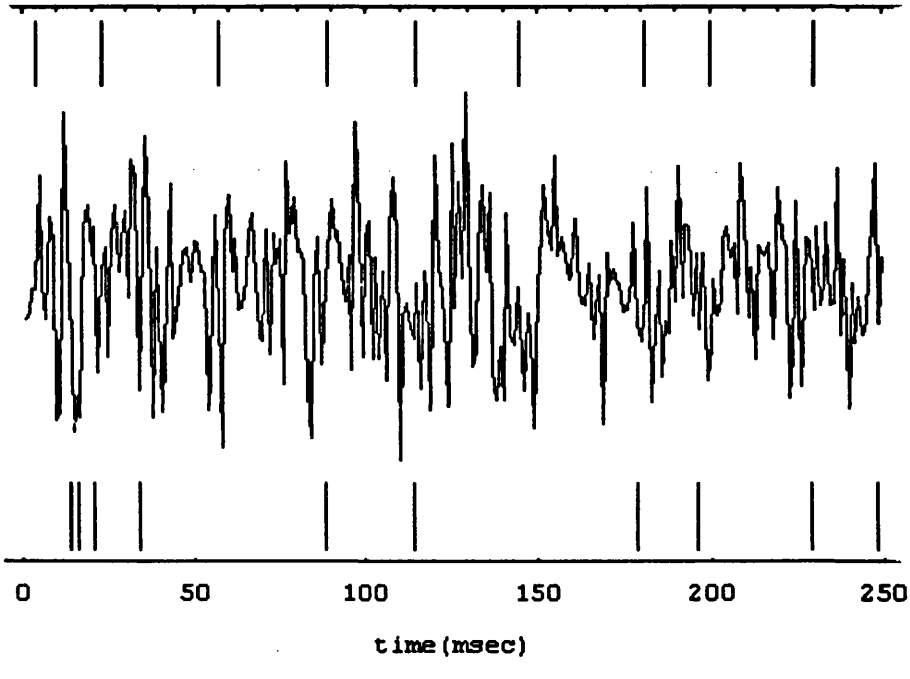


Fig.5.2.1 The lower set of spikes represents the times of the observed input to the neurone, the middle trace provides a stretch of continuous input to the neurone and the upper set of spikes gives the corresponding times at which the neurone fired (the observed output times).

The fitting was done for both probit and logistic link functions, and was found to be better for the logistic model. Both constant and exponentially-decaying thresholds were tried. Although the constant threshold model increased the number of parameters needed for the recovery function (i.e., a cubic recovery function with constant threshold and linear recovery function with exponentially-decaying threshold), the number of parameters needed for each model was the same. Both models reduced the deviance by a very similar amount (i.e., a model with constant threshold reduced the deviance by 4581 and a model with exponentially-decaying threshold reduced the deviance by 4578). The constant threshold model was chosen, arbitrarily.

Fig.5.2.2a and Fig.5.2.3a represent the two estimated summation functions, $\{\hat{a}_u\}$ and $\{\hat{b}_u\}$ for the observed and "unobservable" inputs respectively, and suggest that, while the summation function for the observed inputs $(\text{Summation Function})_1$ reveals an excitatory effect lasting about 26 msec, the summation function for the "unobservable" inputs $(\text{Summation Function})_2$ shows longer excitatory effects lasting about 34 msec. The "unobservable" inputs also seem to have larger effects than the observed inputs as can be seen both from the reduction in deviance (Fig.5.2.4) when the two summation functions are fitted separately, and from the fact that the estimated coefficients, $\{\hat{b}_u\}$ for the "unobservable" summation function are much more statistically significant at any given lag than the estimated coefficients, $\{\hat{a}_u\}$ for the observed summation function.

Carry-over effects for the observed as well as for the "unobservable" inputs are present. Fig.5.2.2b represents the estimated carry-over effect function for the observed inputs $(\text{COE})_1$ and suggests excitatory effects lasting from about 16 to 27 msec. The estimated carry-over effect function for the "unobservable" inputs $(\text{COE})_2$ as given in Fig.5.2.3b suggests excitatory effects

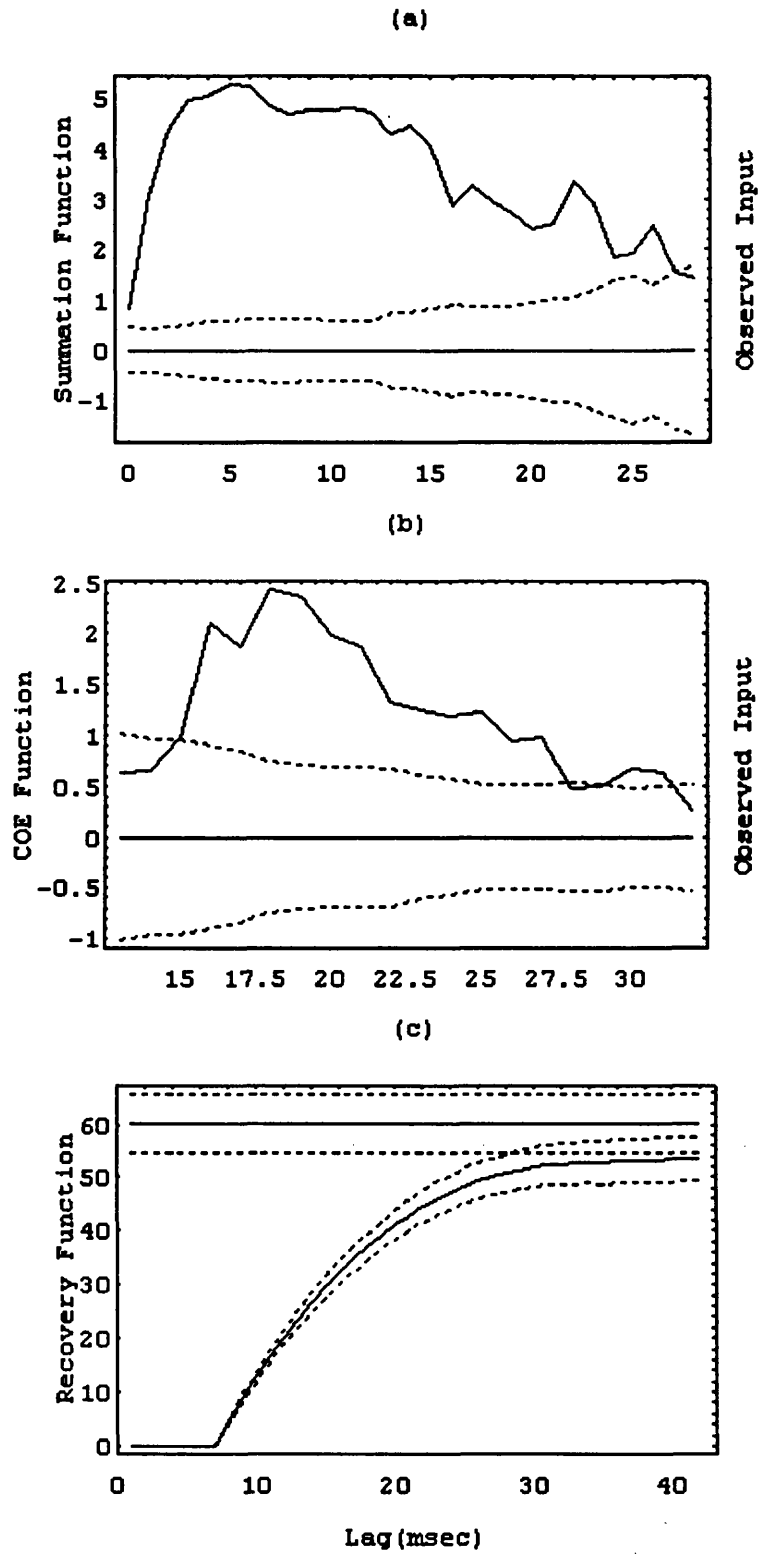


Fig.5.2.2 **a)** Estimated summation function for the observable input. **b)** Estimated carry-over effect function for the observable input. **c)** Estimated recovery (lower curve) and threshold (upper solid line) functions. The dotted lines give ± two standard error limits plotted about zero in (a) and (b) for the summation and the corresponding COE functions and plotted around each function in (c) for the threshold and recovery functions.

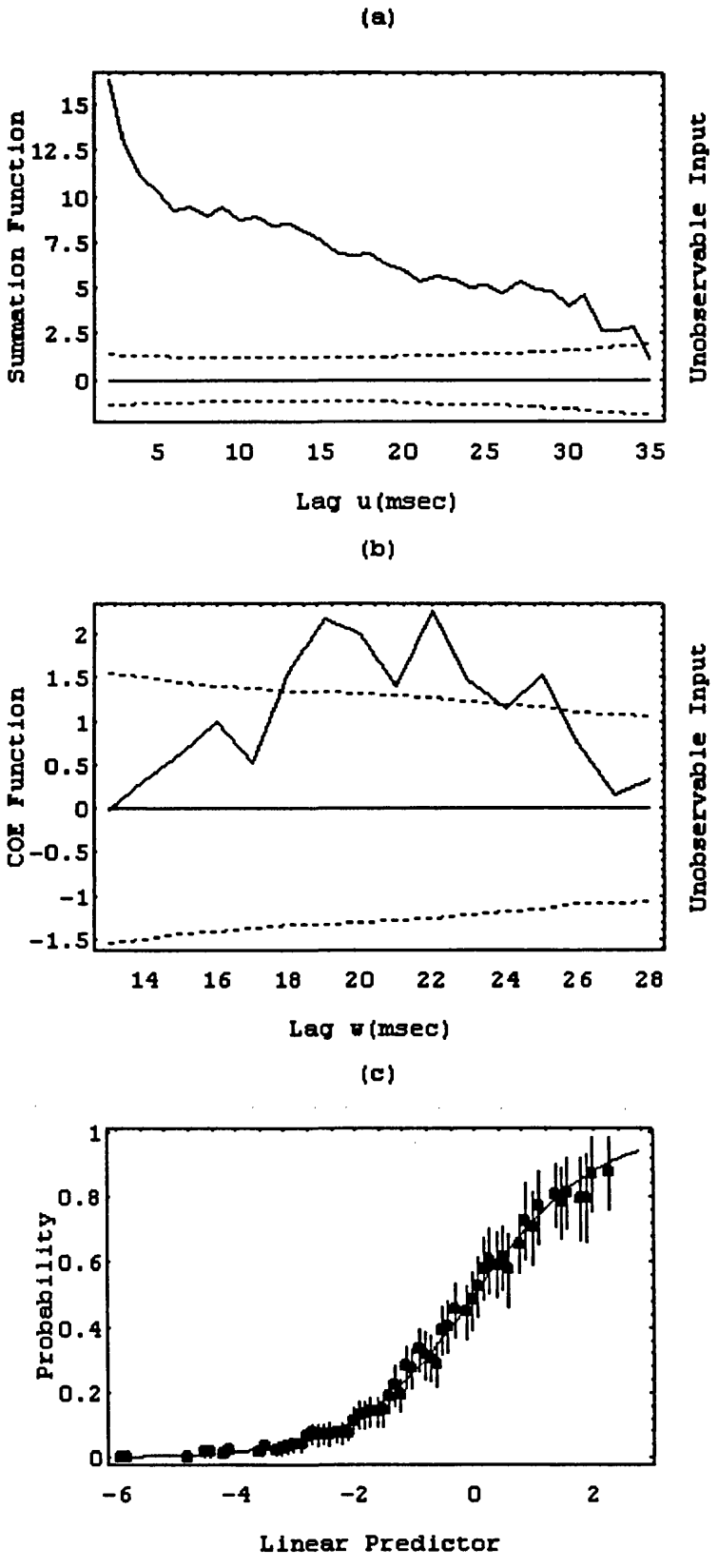


Fig.5.2.3 a) Estimated summation function for the "unobservable" input. b) Estimated carry-over effect function for the "unobservable" input. The horizontal dotted lines in (a) and (b) give ± two standard error limits plotted about zero for the summation and the corresponding COE functions for the "unobservable" input. c) The goodness of fit plot for the full model.

lasting from about 18 to 25 msec. But each of the two carry-over effect functions has a relatively small effect compared to their corresponding summation functions, as can be seen both from the comparatively small reduction in deviance from models with only summation functions to models with both summation and carry-over effect functions, and from the fact that the parameters of the two summation functions tend to be much more statistically significant than those of the corresponding carry-over effect functions. However, these can clearly be considered as real effects, and this was confirmed by F-tests on the deviances.

The threshold and recovery functions are well-estimated up to about 42 msec as shown in Fig.5.2.2c and suggest that the probability of an output spike is small up to about 20 msec, but it then increases rapidly and the chance of an output spike becomes quite large after about 30 msec.

The deviance table given in Fig.5.2.4 illustrates the sequential fitting of a set of successively more complex models in the most complete situation available to us; namely one where "unobservable" inputs can be taken into consideration. There are many interesting features to which we draw attention. The first is that the unobservable inputs explain more of the variability than either the summation function or the recovery function when fitted alone; although each of them is sufficiently informative to be worth fitting. This is a feature of the way in which the data have been simulated. The second is that fitting all five components (as given in 5.2.1) reduces the deviance from 20111 (the initial model) to 6708; a reduction of about 67% and the best we have so far been able to achieve. The goodness of fit test (Fig.5.2.3c) corresponding to the final model reveals that the fit of the model is very satisfactory. It is possible to use values of the linear predictor much larger than we have been able to in previous data sets because more input information was available to us in this case. This leads to circumstances where the probability of an output

spike is very large. The third is to note some of the various effects of adding the recovery function to a previous model. Adding it to the null model reduces the deviance by 4581. Adding it to the "unobservable" inputs reduces the deviance by only 176. Adding it to the summation function reduces the deviance by 4103. This requires careful interpretation. Evidently the information contained in the recovery function is largely orthogonal to that contained in the summation function because the two reductions in deviance, 4581 and 4103, are quite similar. However the recovery function contains almost no extra information to that contained in the "unobservable" inputs, as the additional reduction in deviance is very small. Evidently the recovery function "explains" part of the effects of the "unobservable" inputs if these latter are not (or cannot be) modelled. We shall need to be very careful therefore not to give the recovery function a physiological interpretation which may not be meaningful. It seems therefore in general that, unless all inputs are modelled, the recovery function will contain some input information.

The square root of the estimated cross-intensity function (estimated by the stochastic point process techniques) given in Fig.5.2.5b indicates an excitatory effect of an input lasting about 5 msec only. This duration is very short compared with the 26 msec duration of an excitatory effect suggested by the summation function for the observed inputs $(\text{Summation Function})_1$ given in Fig.5.2.2a. As in previous examples, the square root of the cross-intensity function seems to underestimate the underlying excitatory effects of a synaptic input and provides little or none of the information available in the likelihood model with all five components (as given in 5.2.1).

Fig.5.2.5a gives the estimated cross-intensity function (estimated via the likelihood function as described in section 4.5.1-b). The residual deviance for this model is 19720; a reduction of only 391 from the null model, providing

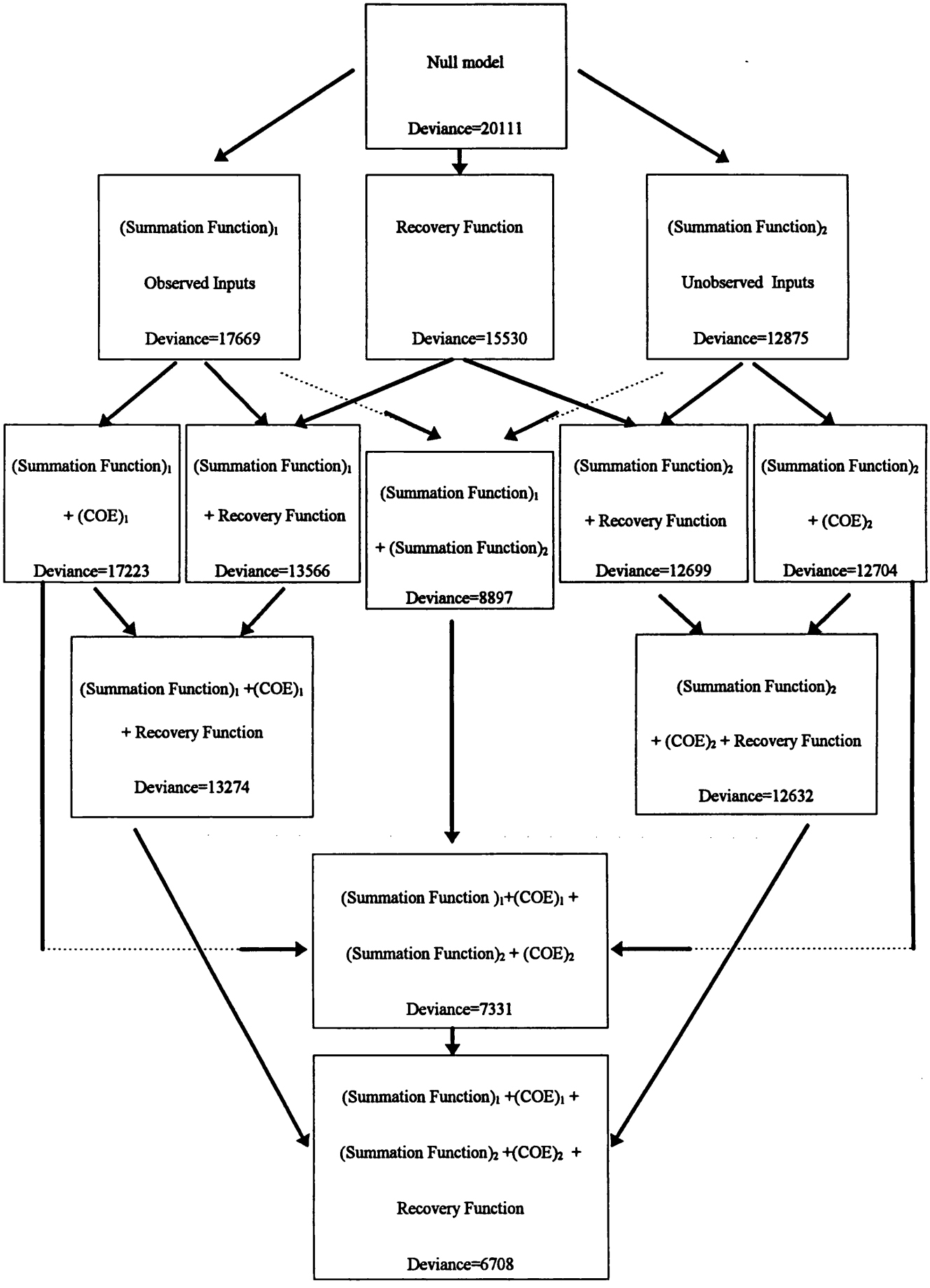


Fig.5.2.4 Diagrammatic representation of the deviance table.

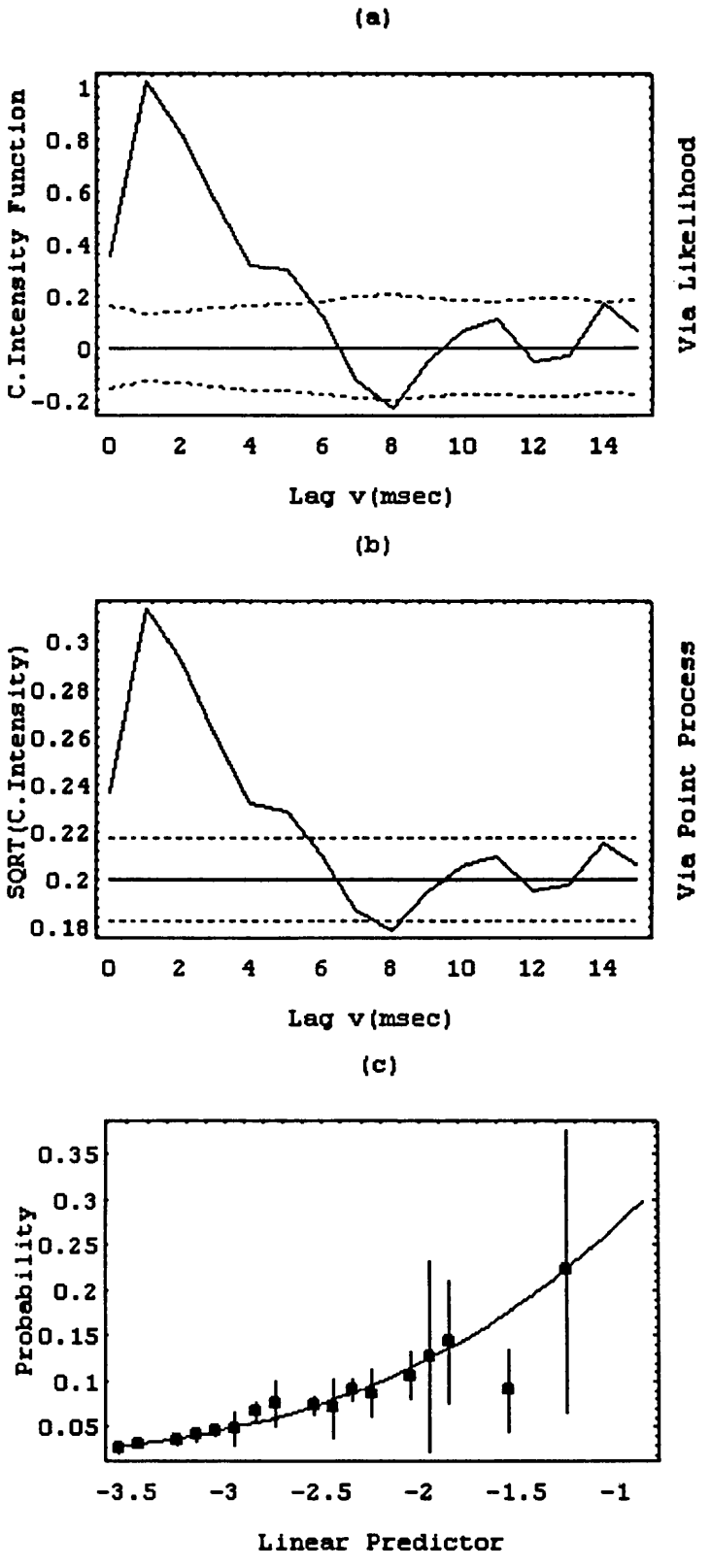


Fig.5.2.5 **a)** The cross-intensity function estimated via likelihood. **b)** The square root of the cross-intensity function estimated via the point process approach. The dotted lines in a) and b) give $\pm$ two standard error limits for the cross-intensity functions plotted about zero in (a) and about the square root of estimated output mean rate in (b). **c)** The goodness of fit plot corresponds to the model given in a).

still further evidence that the cross-intensity function in general has very poor explanatory power.

The goodness of fit plot for a model containing only the cross-intensity function (shown in Fig.5.2.5c) indicates that the fit is a very poor one compared with that for the five components model shown in Fig.5.2.3c. This can be seen both from the relatively small values of the predicted probabilities ($\hat{P}(\eta) \leq 0.2$), (i.e., the short range of the linear predictor values) and from the departure of the predicted probabilities from the theoretical curve for values of the linear predictor near -1.5.

Fig.5.2.6a and b correspond to the estimates of the coherence and phase, respectively. The coherence plot (Fig.5.2.6a) indicates clearly that the two processes are associated with each other over the range of about 0-28 Hz. The estimated phase given in Fig.5.2.6b shows that, over the range of frequencies at which coherence is significant, the output process is delayed, on the average, by an amount 1.20 msec with a 95% confidence interval, for the delay, of (0.90, 1.50) msec. The average value of the delay suggested by the phase seems to be consistent with the 1 msec peak in the square root of the estimated cross intensity function. In this example the peak in the summation function is centred about 5 msec which is considerably different from both the peak in the cross intensity function and from the average value of the delay estimated from the phase. This lack of agreement of detecting similar estimates for the delay between the likelihood and point process techniques may be due to the fact that this data set is not input dominated and in such a case the delay may not be the dominant feature, or even particularly meaningful as a single parameter.

Next we demonstrate with two simulated sets of data each with two observed spike train inputs and one "unobservable" input, and one observed

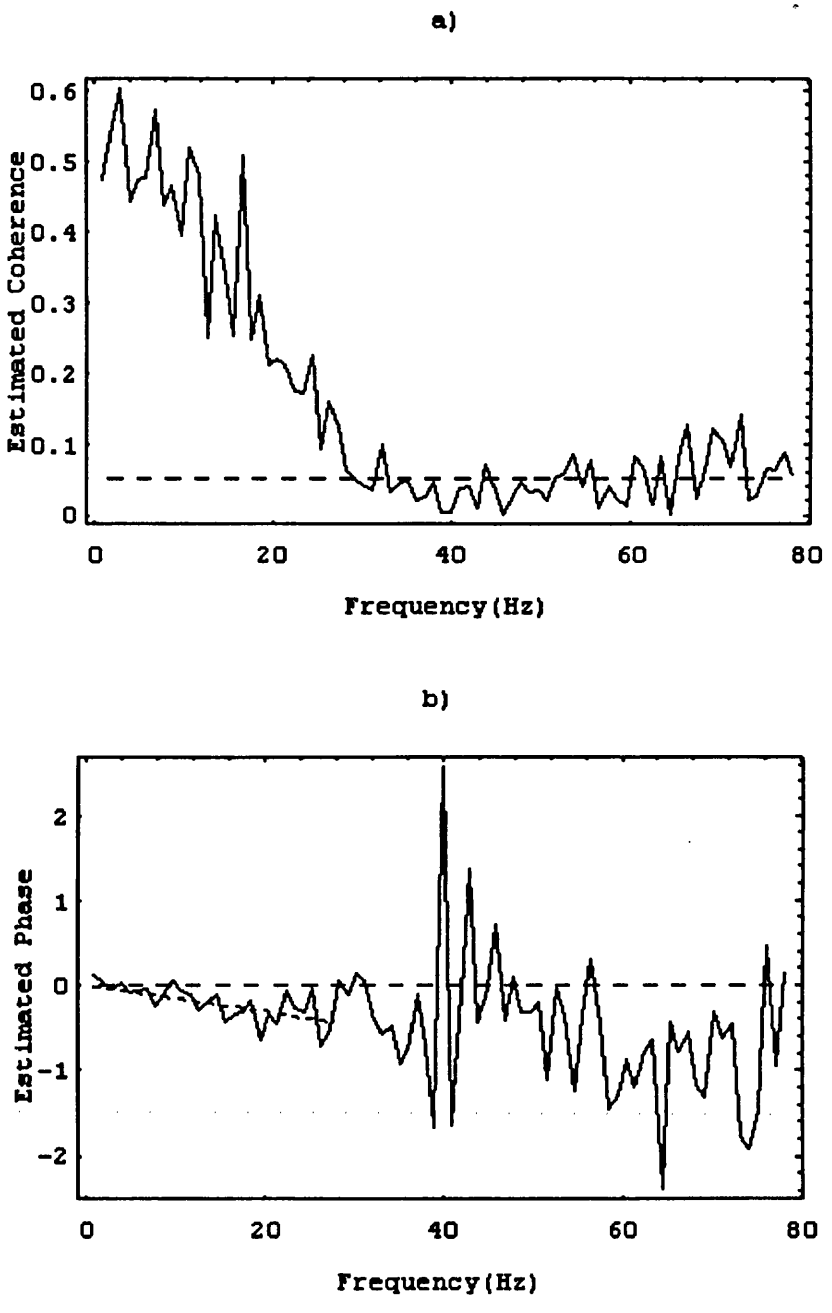


Fig.5.2.6 Illustration of the ordinary coherence and phase.

a) Estimate of the ordinary coherence. The dotted line represents the upper limit of the 95 % confidence interval for the coherence under the hypothesis that the two processes are independent.

b) Estimate of the phase. The diagonal dotted line represents the weighted least squares line over the range of frequencies at which the corresponding coherence is significant.

output. The two observed spike train inputs behave independently in the first example and are correlated in the second. In both cases, they excite the occurrence of the output spikes, whether in the absence or presence of "unobservable" input. The aim of this demonstration is to see how the likelihood techniques treat these two types of data set compared to the stochastic point process techniques.

5.2.2 Two Uncorrelated Spike Train Inputs

The neuronal network demonstrated in this section is illustrated in Fig.5.2.7 where the two spike train inputs to the neurone are denoted by ${}_1X$ and ${}_2X$ and the spike train output is denoted by Y whereas Z denotes the "unobservable" input that can be added to simulate the effects due to all other unmeasured synaptic inputs. A diagrammatic representation of the first 500 msec segment of the two spike train inputs and the spike train output and also a stretch of the corresponding "unobservable" input are shown in Fig.5.2.8.

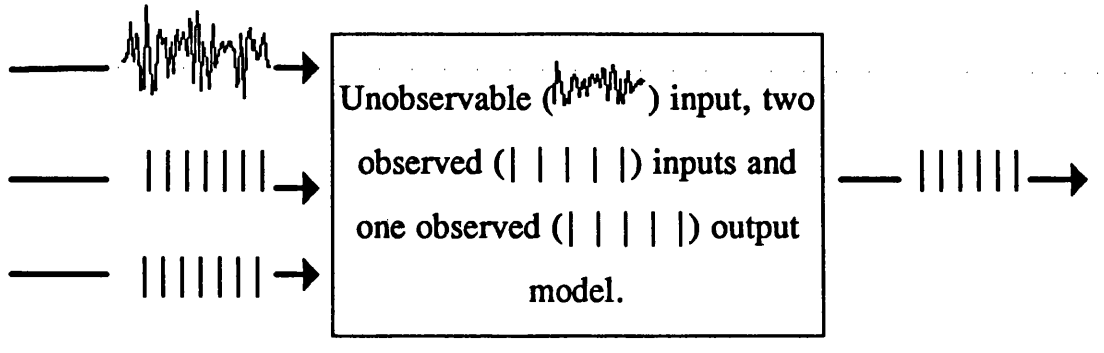


Fig.5.2.7 Diagrammatic representation of a neuronal network in which two independent observed inputs excite the occurrence of an output in the present of a continuous input.

The simulated data set demonstrated here consists of a 0-1 valued series of approximately 60000 points for each of the two spike train inputs and the

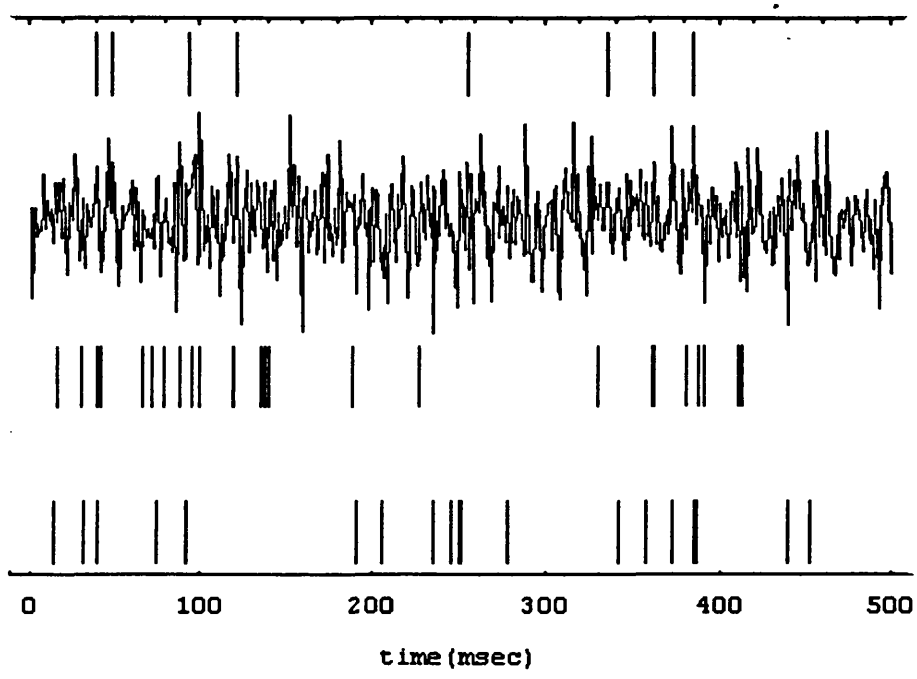


Fig.5.2.8 A segment of the first 500 msec of the data. The lower two sets of spikes represent the times for each of the two observed inputs to the neurone, the solid curve provides a stretch of a continuous input to the neurone and the upper set of spikes represents the corresponding times at which the neurone fired (i.e., the times of the observed output).

output with the "unobservable" input present to represent all other unmeasured inputs. The numbers of spikes observed were 3111 and 3039 for the first ($_1X$) and second ($_2X$) inputs respectively, whereas the number of spikes observed for the output (Y) was 1605. The two spike train inputs were driven by independent Poisson processes.

For computational purposes (i.e., to get more computational space and time) the data were regrouped with binwidth two (i.e. $h = 2$ msec). The data remained as a 0-1 valued series (i.e. maintained the 0-1 property required for the binomial distribution). All the illustrated figures in this demonstration are given in the original data scale of 1 msec.

We start by using the stochastic point process techniques, where the square root of the cross-intensity function, the ordinary and partial coherences

and the phase were all estimated. Then we apply the likelihood techniques to our data to see what extra information it can produce.

Fig.5.2.9a and Fig.5.2.9b represent the square roots of the estimated cross-intensity functions between the output Y and each of the inputs $_1X$ and $_2X$, respectively. The horizontal dotted lines in each figure represent approximate 95 % confidence intervals for the square root of the cross-intensity functions, under the hypothesis of inputs $_1X$ and $_2X$ each being independent of the output Y, plotted around the square root of the estimated output mean rate (horizontal solid line). Each of the two functions suggest excitatory synaptic input effects lasting about 6 msec.

Fig.5.2.10b represents the estimated ordinary coherence between the first input $_1X$ and the output Y and suggests a relatively weak coupling. The estimated ordinary coherence between the second input $_2X$ and the output Y (Fig.5.2.10c) also indicates a relatively weak coupling similar to that shown in Fig.5.2.10b. These weak couplings might give the impression that the output is not mainly driven by the effects of the two observed inputs but is also driven by the effects of all other "unobservable" inputs or perhaps as a combination of the two types of input as we will investigate shortly. The estimated ordinary coherence between the two inputs $_1X$ and $_2X$ as shown in Fig.5.2.10a indicates that the two inputs are not associated with each other over the whole range of frequencies. This feature also can be seen from a direct comparison between the two figures (Fig.5.2.11a and b) illustrating the estimated partial coherences between the output and each input after removing the effect of the other input and those (Fig.5.2.10b and c) of the ordinary coherences between the output and each input. The two figures Fig.5.2.10b and Fig.5.2.11a and the two figures Fig.5.2.10c and Fig.5.2.11b of the ordinary and partial coherences

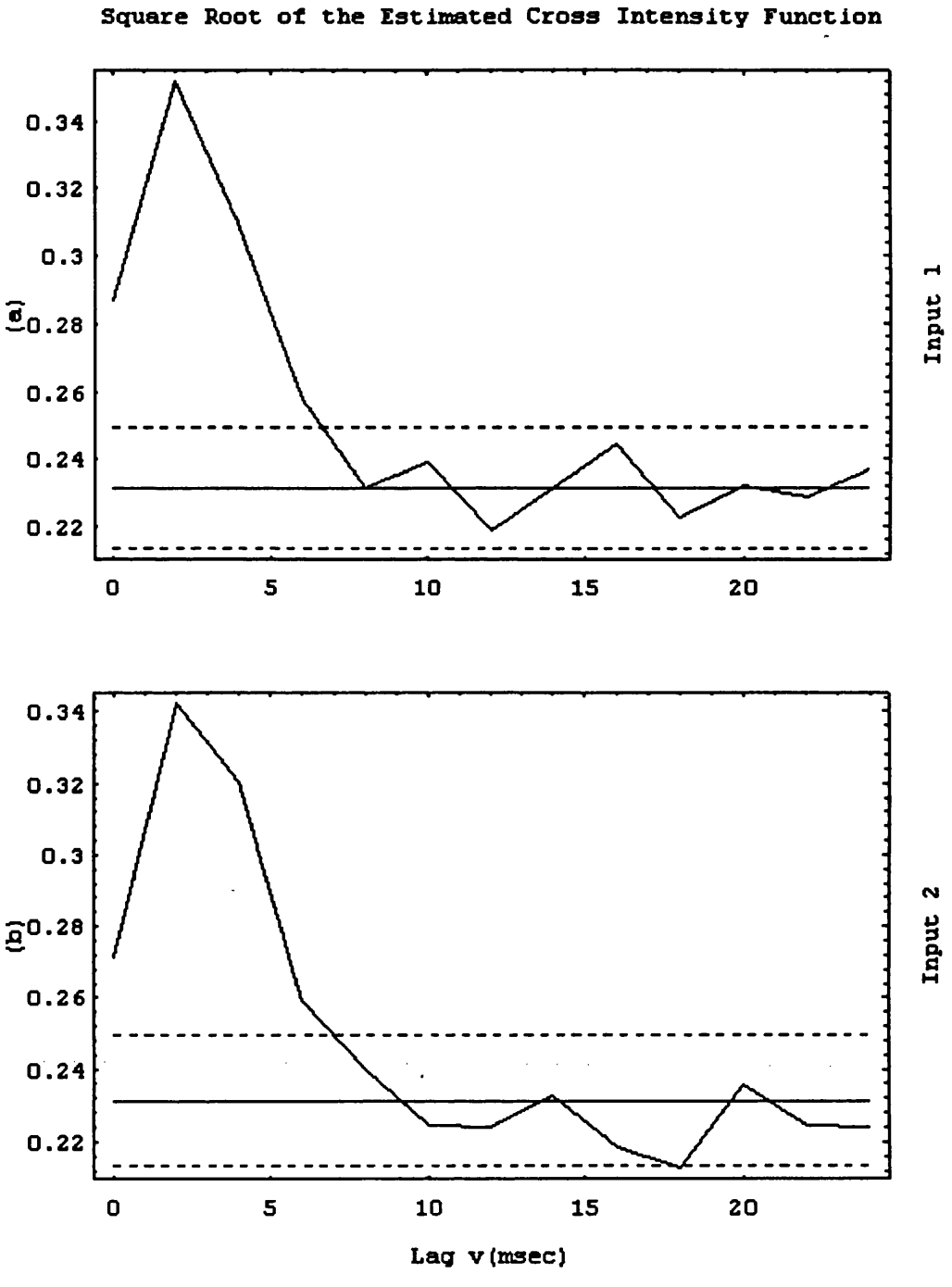


Fig.5.2.9 Estimated square root of the cross-intensity function between the output Y and a) the first input ${}_1X$ b) the second input ${}_2X$. The horizontal dotted lines in (a) and (b) represent approximate 95 % confidence intervals for the estimated square roots of the cross-intensity functions under the hypothesis that the two processes are independent, plotted around the square root of the estimated output mean rate (horizontal solid line).

are almost identical to each other which suggests that the two observed inputs act independently to affect the occurrence of the output spikes. The horizontal dotted line in the coherence figures represents the upper limit of the approximate 95% confidence intervals for the coherence under the hypothesis that the two processes have zero coherence.

The estimated phase between the first observed input ${}_1X$ and the output Y given in Fig.5.2.12a suggests that there is no simple delay present. Similarly, there is no simple delay present between the second observed input ${}_2X$ and the output Y as suggested by the estimated phase given in Fig.5.2.12b. Each phase figure cannot be represented by a straight line.

Fig.5.2.13a represents the estimated partial phase between the first observed input ${}_1X$ and the output Y after removing the linear effects of the second observed input ${}_2X$ whereas Fig.5.2.13b represents the estimated partial phase between the second observed input ${}_2X$ and the output Y after removing the linear effects of the first observed input ${}_1X$. As in the estimated phase figures given in Fig.5.2.12a and b, the estimated partial phase figures suggest that there is no simple delay present between the output and each input after removing the linear effects of the other input.

The figures of the estimated phase (Fig.5.2.12a and Fig.5.2.12b) between the output and each input are almost identical to the figures of their corresponding estimated partial phase (Fig.5.2.13a and Fig.5.2.13b) between the output and each input after removing the linear effects of the other input. This suggests that the two observed inputs act almost independently to affect the occurrence of the output spikes. This seems to be in agreement with the ordinary and partial coherences discussed earlier.

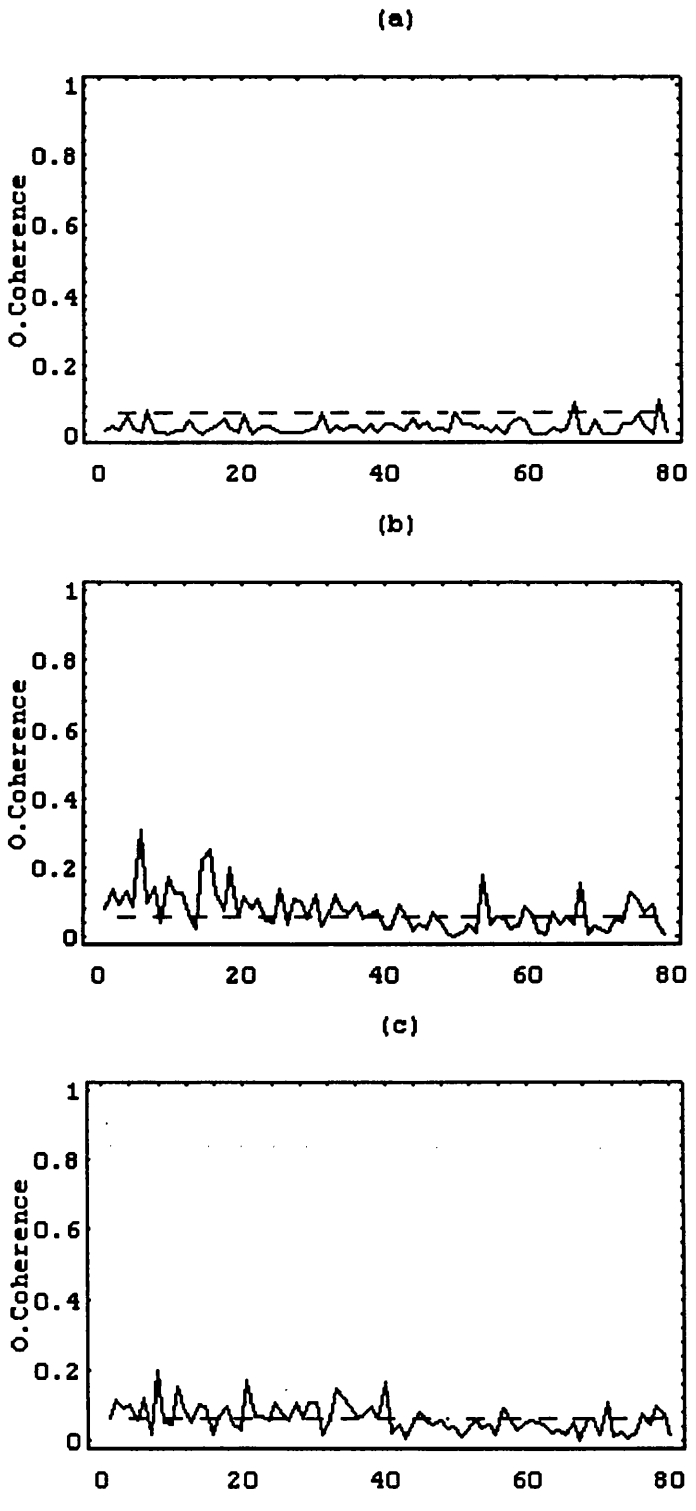


Fig.5.2.10 Estimated ordinary coherences between

a) the two observed inputs ${}_1X$ and ${}_2X$ **b)** the output Y and the first input ${}_1X$, and **c)** the output Y and the second input ${}_2X$. The dotted lines correspond to the upper limit of the 95% confidence intervals for the coherence under the hypothesis of zero coherence.

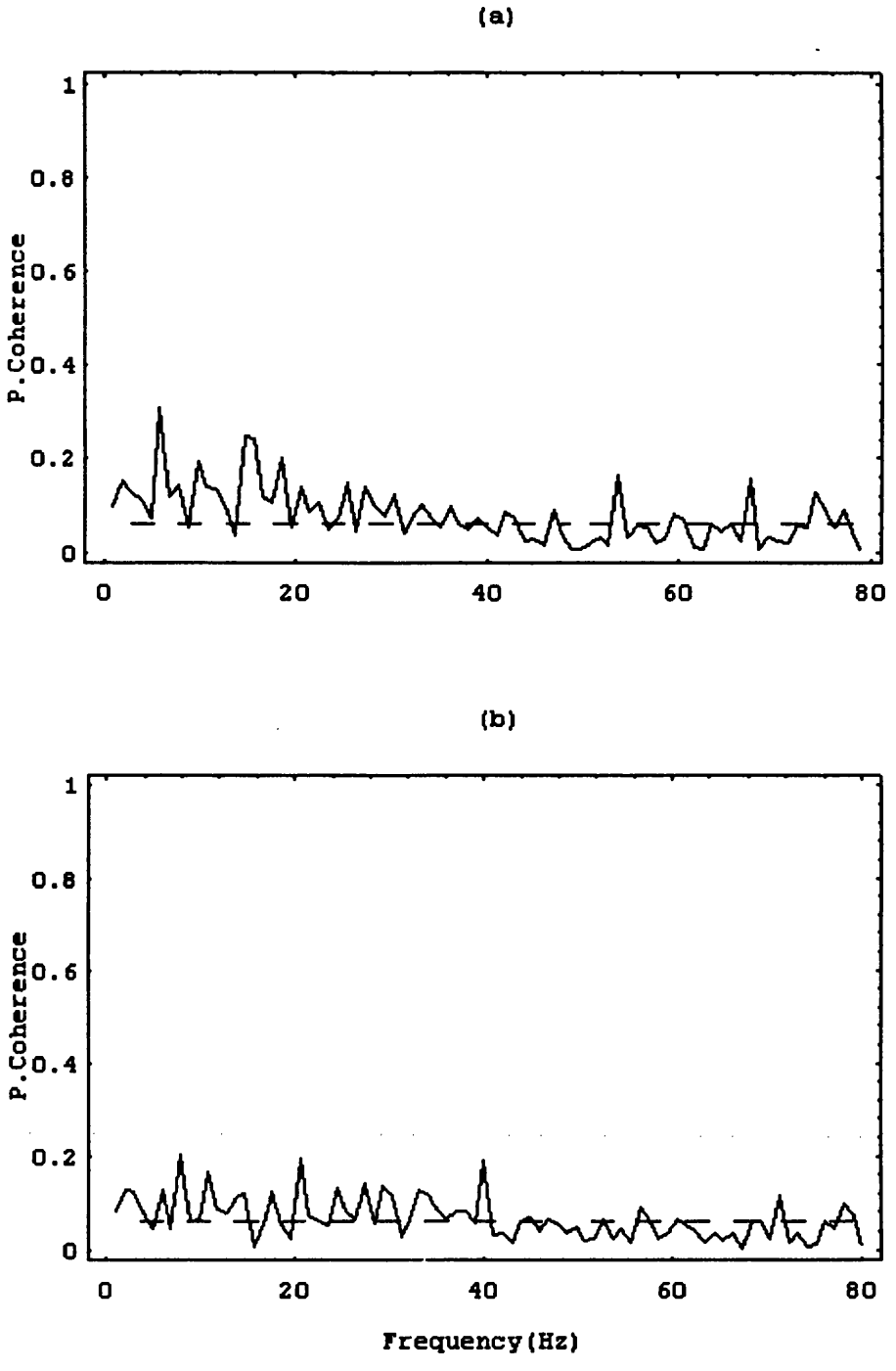


Fig.5.2.11 Estimated partial coherences between the output Y and

a) the first input ${}_1X$ after removing the effect of the second input ${}_2X$ and **b)** the second input ${}_2X$ after removing the effect of the first input ${}_1X$. The dotted lines correspond to the upper limit of the 95% confidence intervals for the coherence under the hypothesis of zero partial coherence.

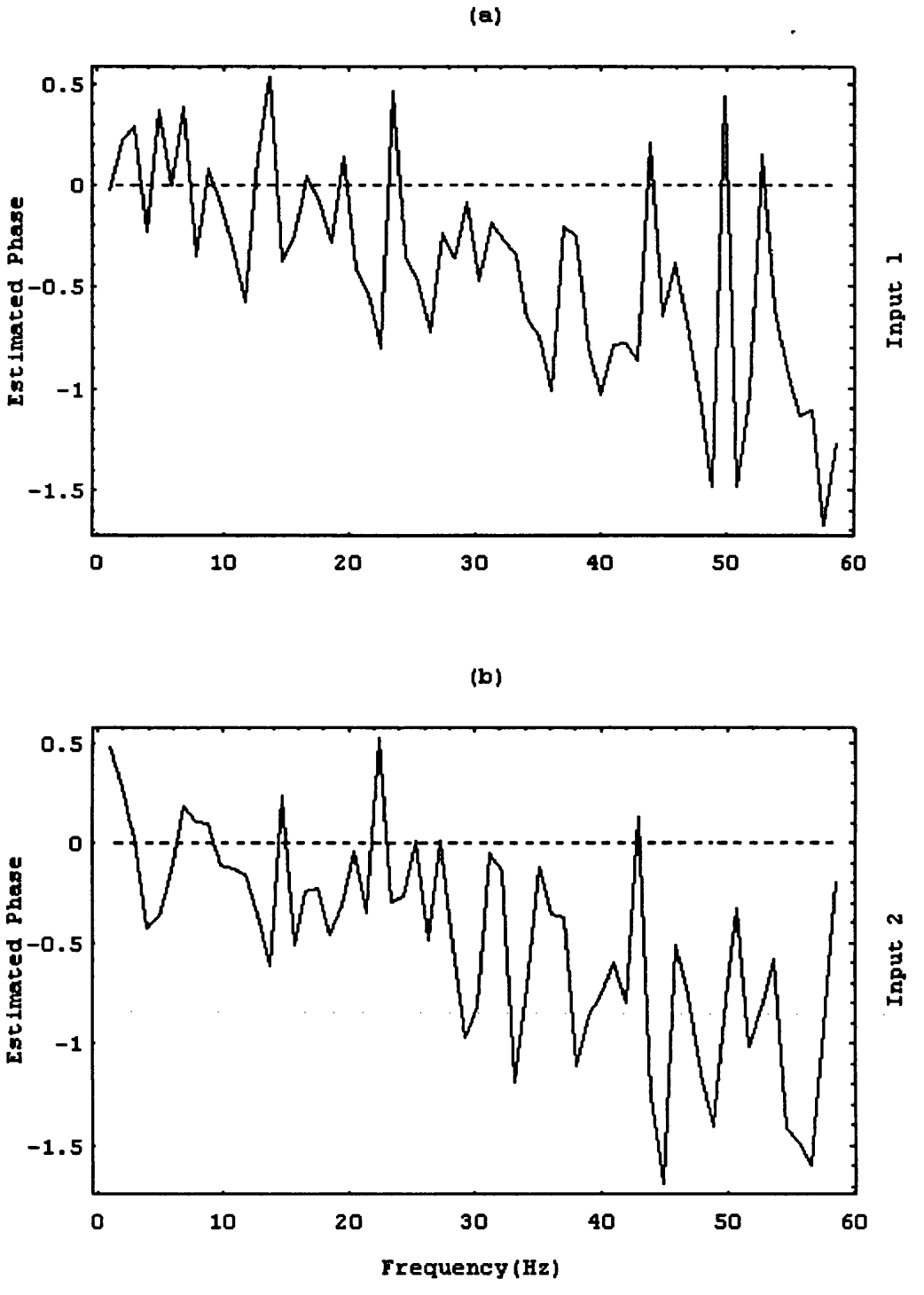


Fig.5.2.12 illustration of the phase functions.

a) Estimated phase between the output Y and the first observed input ${}_1X$.

b) Estimated phase between the output Y and the second observed input ${}_2X$.

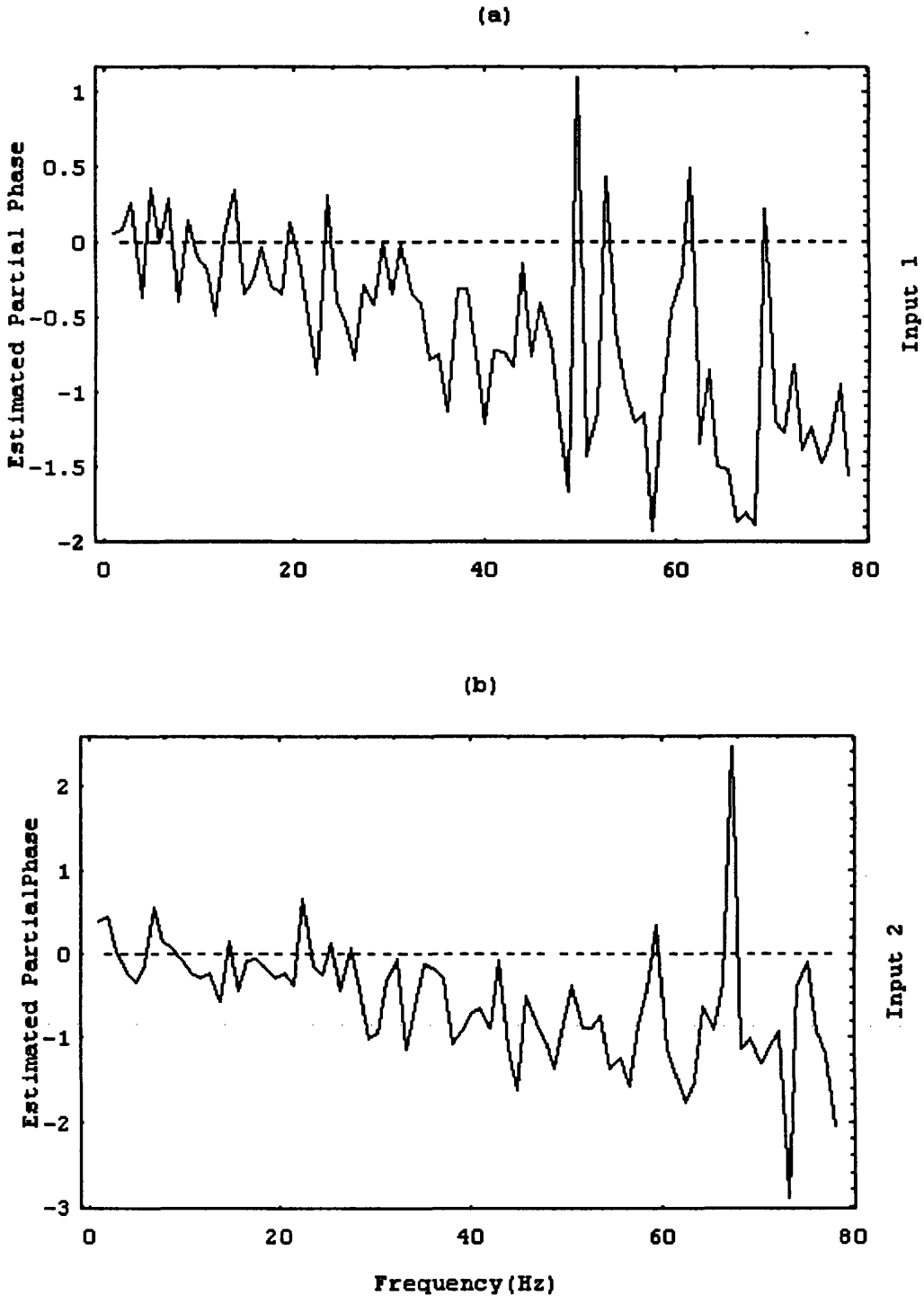


Fig.5.2.13 illustration of the partial phase functions.

a) Estimated partial phase between the output Y and the first observed input after removing the linear effects of the second input. **b)** Estimated partial phase between the output Y and the second observed input after removing the linear effects of the first input.

We next apply the likelihood procedure to the same set of data and our aim is to compare some of its results such as the significant durations of the summation functions with those of their corresponding cross intensity functions. In order to find how much explanatory power the cross intensity functions have, we again use the likelihood procedure to estimate them. Also we see if the likelihood procedure (through the table of deviances) can detect features similar to those revealed by the ordinary and partial coherences and phase.

The maximum likelihood estimation technique has been applied to the same set of simulated data, employing the canonical link function (as given in 4.3.1a-b). The membrane potential, U_t, on the trigger zone at time t is of the form

$$\begin{aligned} U_t = V_t &+ \sum_{u=0}^{\gamma_t - 1} {}_1a_u \; {}_1x_{t-u} + \sum_{w \geq \gamma_t} {}_1c_w \; {}_1x_{t-w} \\ &+ \sum_{u=0}^{\gamma_t - 1} {}_2a_u \; {}_2x_{t-u} + \sum_{w \geq \gamma_t} {}_2c_w \; {}_2x_{t-w} \\ &+ \sum_{u=0}^{\gamma_t - 1} b_u \, z_{t-u} + \sum_{w \geq \gamma_t} d_w \; z_{t-w} \end{aligned} \qquad \text{.....(5.2.3)}$$

where γ_t is the time elapsed since the time of the last output spike, the sets of coefficients $\{{}_i a_u\}$ and $\{{}_i c_w\}$; $i = 1,2$ represent the summation and carry-over effect functions for the two observed inputs respectively, the summation and carry-over effect functions for the "unobservable" input are represented by the two sets of coefficients $\{b_u\}$ and $\{d_w\}$ respectively and the term V_t represents

the polynomial recovery function of order k, i.e.

$$V_t = \begin{cases} \sum_{i=1}^{k} \theta_i (\gamma_t - \zeta_1 - 1)^i & ; \ \gamma_t \geq \zeta_1 + 1 \\ 0 & ; \ \gamma_t \leq \zeta_1 + 1 . \end{cases} \quad .....(5.2.4)$$

The recovery function given in (5.2.4) is forced to be zero for $\gamma_t \leq \zeta_1 + 1$, as explained earlier in chapter 4. The linear predictor of the model, η_t, representing the difference between the membrane potential and the threshold may be given by

$$\eta_t = U_t - \theta_0 \quad .....(5.2.5)$$

where the term θ_0 in (5.2.5) represents the constant threshold, which the deviance table showed was superior to the exponential threshold model.

Fig.5.2.14a, b and c represent the estimated (Summation Function)$_1$, $\{_1\hat{a}_u\}$; estimated (Summation Function)$_2$, $\{_2\hat{a}_u\}$; for the first and second observed inputs respectively, and the estimated (Summation Function)$_3$, $\{\hat{b}_u\}$, for the "unobservable" input. The figures suggest that the summation functions for the two observed inputs reveal excitatory effects lasting about 20 msec for the first input and about 22 msec for the second input and the (Summation Function)$_3$ for the "unobservable" input shows longer effects lasting about 28 msec. The "unobservable" input is also seen to have much larger effects than the two observed inputs as can be seen both from the reduction in deviance (Fig.5.2.18) when the three summation functions are fitted separately, and from the fact that the coefficients, $\{\hat{b}_u\}$ for the "unobservable" input seem to be statistically more significant at any given lag than either of the two sets of coefficients, $\{_i\hat{a}_u\}$; $i = 1,2$ for the observed inputs.

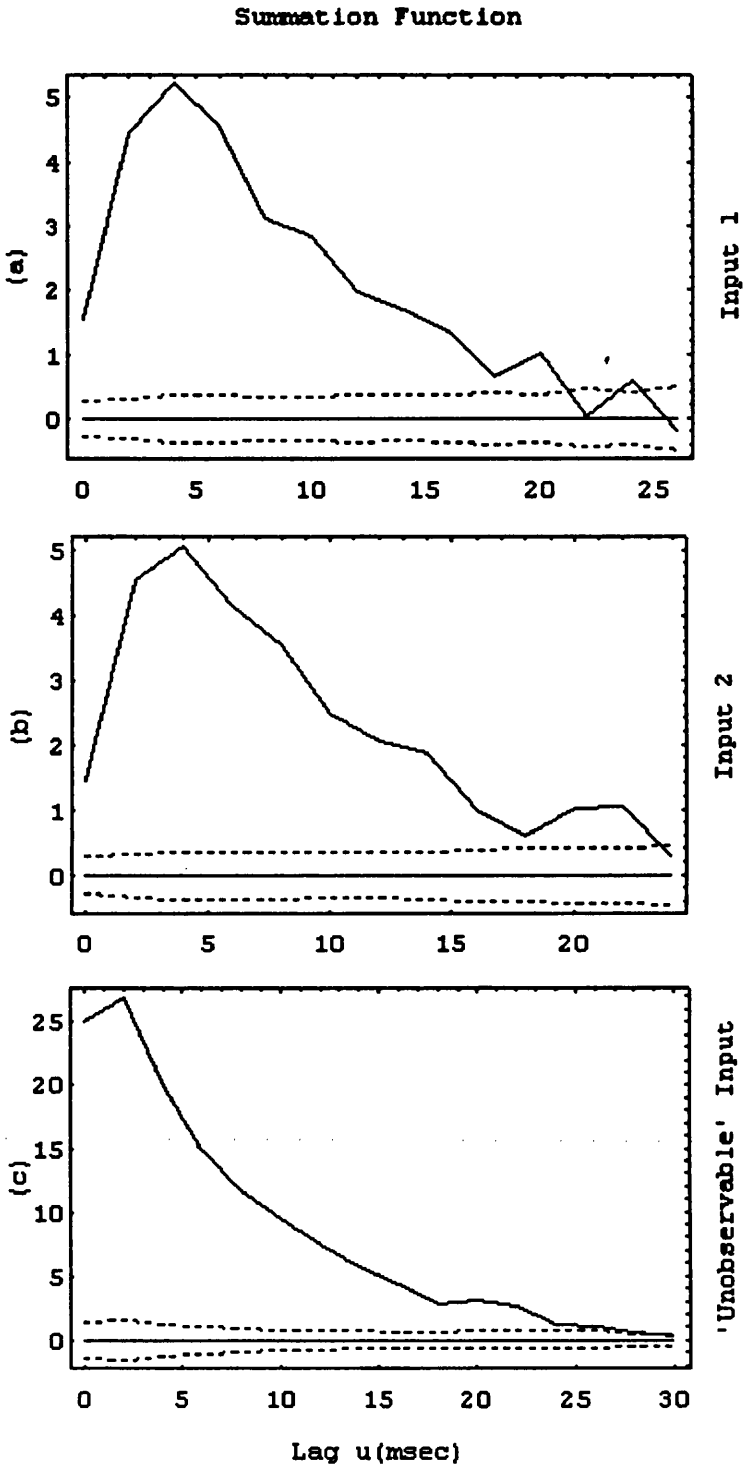


Fig.5.2.14 **a)** Estimated summation function $\{{}_1\hat{a}_u\}$ for the first observed input ${}_1X$. **b)** Estimated summation function $\{{}_2\hat{a}_u\}$ for the second observed input ${}_2X$. **c)** Estimated summation function $\{\hat{b}_u\}$ for the "unobservable" input Z. The horizontal dotted lines give $\pm$ two standard error limits plotted about zero for the summation functions.

The significant durations of the summation functions for the two observed inputs given in Fig.5.2.14a and b are much longer than the 6 msec significant durations of each of the two corresponding cross-intensity functions as shown in Fig.5.2.9a and b. The two cross-intensity functions seem to underestimate the underlying excitatory effects of a synaptic input and provide little or none of the information available in the likelihood model with all seven components (as given in 5.2.3). Furthermore, suppose we use the likelihood approach to fit only the cross-intensity functions for the two observed inputs, as explained in (5.2.2). The residual deviance for this model is 11711; a reduction of only 811 from the null model (i.e., the variation explained by the two cross-intensity functions when fitted together is only 6.48 %), providing still further evidence that the cross-intensity function in general has very poor explanatory power, whereas the residual deviance for a model with the two summation functions is 11070.

Evident carry-over effects for each of the two observed inputs and for the "unobservable" input are shown in Fig.5.2.15a, b and c respectively, and suggest that while the estimated carry-over effect function $(COE)_3$ for the "unobservable" input (Fig.5.2.15c) suggests an excitatory effect lasting from about 1 to 23 msec, the estimated carry-over effect functions $(COE)_1$ and $(COE)_2$ suggest excitatory effects lasting from about 6 to 14 msec for the first input and about 4 to 18 msec for the second input as shown in Fig.5.2.15a and b respectively. These latter two excitatory input effects are smaller in magnitude than the estimated carry-over effect function $(COE)_3$ for the "unobservable" input as can be seen from the smaller reduction in deviance they made compared to that for the "unobservable" input. However, they can clearly be considered as real effects, and this was confirmed by F-tests.

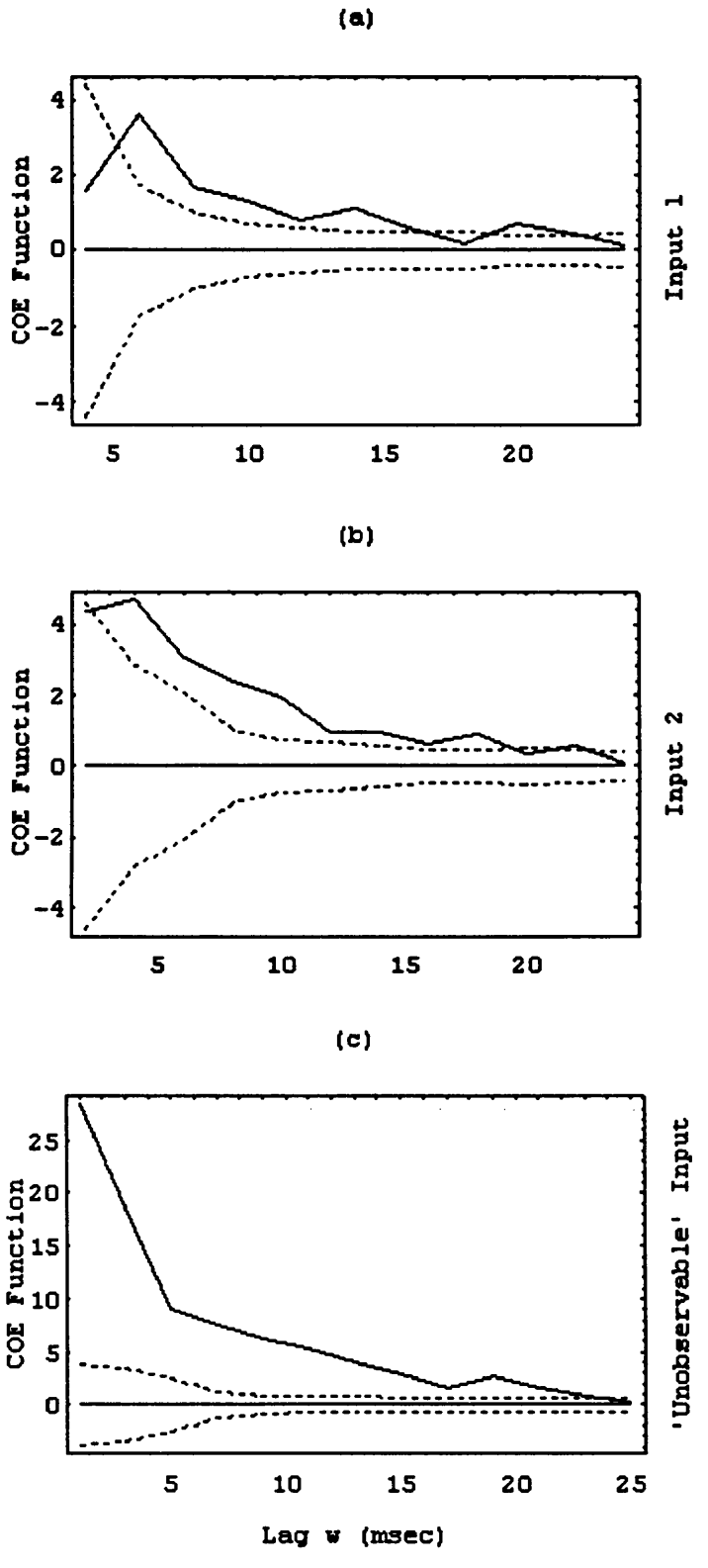


Fig.5.2.15 **a)** Estimated carry-over effect function $\{{}_1\hat{c}_w\}$ for the first observed input ${}_1X$. **b)** Estimated carry-over effect function $\{{}_2\hat{c}_w\}$ for the second observed input ${}_2X$. **c)** Estimated carry-over effect function $\{\hat{d}_w\}$ for the "unobservable" input Z. The horizontal dotted lines give $\pm$ two standard error limits plotted about zero for the carry-over effect functions.

The threshold and recovery functions given in Fig.5.2.16a suggest that the probability of an output spike occurring spontaneously is very small over the whole range of intervals (as the recovery function remains far below threshold over the whole range of intervals). This also suggests that the recovery function has almost no effect on the production of output spikes in the presence of the "unobservable" inputs. The order of the recovery function is 3 (cubic recovery function) when fitted either alone or with the summation functions (and carry over effect functions) for the observed inputs. But in the presence of the "unobservable" inputs the order is reduced from 3 to 1 (linear recovery function), i.e., the numbers of statistically significant parameters involved in the recovery function in the absence and presence of the "unobserved" inputs are 3 and 1 respectively. This change in the order of the recovery function, in the absence and presence of the "unobserved" inputs, also affects its role in accelerating the output firing. This can also be seen clearly from the deviance table (as we will discuss shortly), which suggests again that the recovery function is acting as a proxy for unmeasured inputs when these latter are not (or cannot be) modelled. The recovery function reflects the intrinsic properties of the cell only in the case where all inputs are present in the model. If any unmeasured inputs are not (or cannot be) present in the model, we cannot give the recovery function this intrinsic physiological interpretation.

The goodness of fit plot given in Fig.5.2.16b seems very reasonable and indicates that, in this particular example, high probabilities can be predicted by the likelihood model. this is because all the available input information is included. The model takes the range of the linear predictor to above 1.7 which reflects the large explained variability by the model. And the confidence intervals are reasonably narrow.

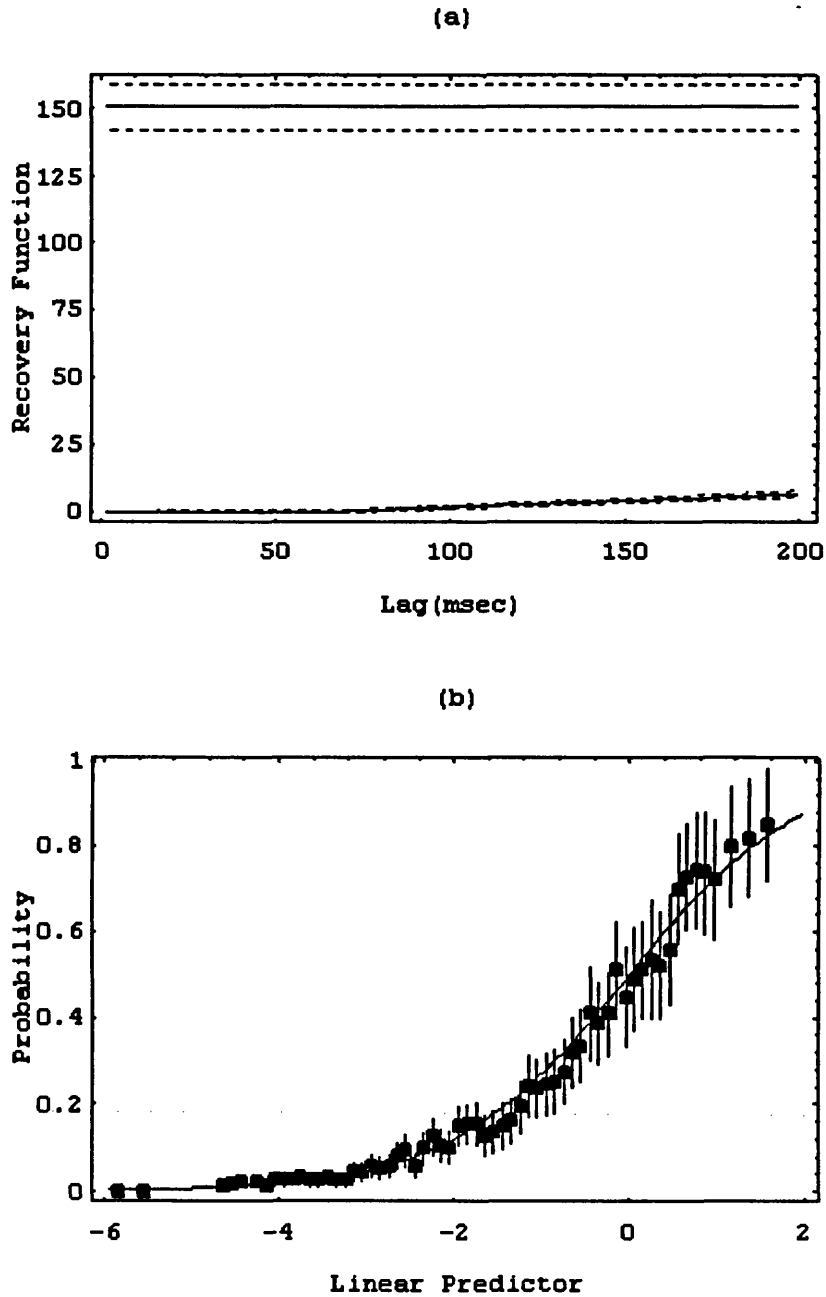


Fig.5.2.16 illustration of the threshold and recovery functions and the goodness of fit plot.

a) Estimated recovery (lower curve) and threshold (upper solid line) functions. The dotted lines give ± two standard error limits plotted around each function for the threshold and recovery functions.

b) The goodness of fit plot with empirical (dots) and theoretical (smooth curve) probabilities plotted against selected values for the linear predictor, the vertical bars present the ± two standard error limits for the theoretical probabilities.

Fig.5.2.17a and b represent the estimated (Summation Function)$_1$ and the estimated (Summation Function)$_2$ for the first and second observed inputs respectively, when each function is estimated in the presence of its corresponding carry-over effect function and the recovery function only (i.e., in the absence of the other observed input and the "unobservable" input). Fig.5.2.17c represents the estimated (Summation Function)$_3$ for the "unobservable" input when estimated in the presence of the carry-over effect function (COE)$_3$ and the recovery function only (i.e., in the absence of the two observed inputs). The figures suggest that these summation functions reveal excitatory effects lasting about 16 msec for the first input, about 14 msec for the second input and about 21 msec for the "unobservable" input. Comparing these summation functions with those given in Fig.5.2.14a, b and c, when the summation functions (and their corresponding carry-over effect functions) for the two observed and the "unobservable" inputs and the recovery function are estimated together (i.e., the full model as given in 5.2.3 - 5.2.5) reveals that every summation function seems to extend its significant duration (i.e., from 16 to 20 msec for the input $_1X$; 14 to 22 for the input $_2X$ and from 21 to 26 for the "unobservable" input Z). They also seem to shift the position of their peaks from 2 msec to 4 msec for each of the observed inputs and from 0 msec to 2 msec for the "unobservable" input. Also in the case of the full model (i.e., when all functions are present) there is a much larger reduction in deviance.

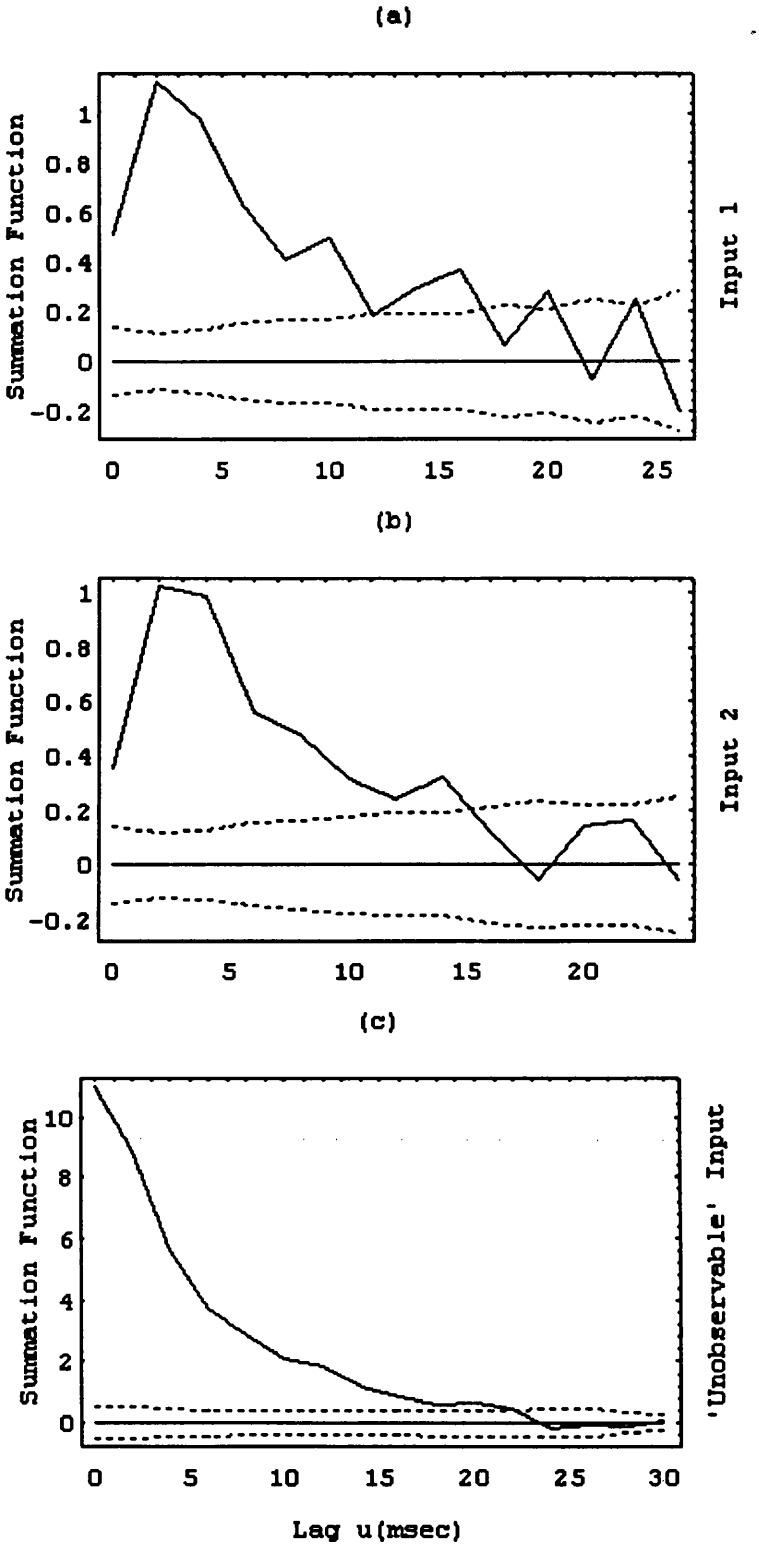


Fig.5.2.17 **a)** Estimated summation function for the input ${}_1X$ in the absence of the input ${}_2X$ and the "unobservable" input Z. **b)** Estimated summation function for the input ${}_2X$ in the absence of the input ${}_1X$ and the "unobservable" input Z. **c)** Estimated summation function for the "unobservable" input Z in the absence of both observed inputs ${}_1X$ and ${}_2X$. The horizontal dotted lines give $\pm$ two standard error limits plotted about zero for the summation functions.

The changes in the shape, the position of the peak and the significant duration are due to the fact that the estimates of each summation function coefficient (when estimated in the presence of its corresponding carry-over effect function and the recovery function only) are statistically biased estimates because all the available variability has not been taken into account and they become statistically unbiased estimates only as we consider all the available variability (i.e., as represented in the final model) where the reduction in deviance is much larger as we reach the final model compared to all previous models.

The deviance table given in Fig.5.2.18 illustrates the sequential fitting of a set of successively more complex models in the situation of two observed inputs and one observed output and also where "unobservable" inputs can be taken into consideration. The table reveals the following interesting features,

(1) fitting all seven components reduces the deviance from 12522 (the initial model) to 2860; a reduction of about 77.2 % which is very substantial compared to previous examples.

(2) fitting a model with only the summation functions (and their corresponding carry-over effect functions) for the two observed inputs and a recovery function reduces the deviance from 12522 (the initial model) to 10759; a reduction of 1763. But a larger reduction in deviance is achieved as we move from a model with only the summation function (and the corresponding carry-over effect function) for the "unobservable" input to the full model which reduces the deviance from 6986 to 2860; a reduction of 4126. Despite the fact that these two reductions in deviance (i.e., 1763 and 4126) represent the effects of the two observed inputs and the recovery process in the absence and the presence of the "unobservable" input

effects respectively, they are quite different, i.e., the effects of the two observed inputs and the recovery process seem to explain a larger amount of the variability in the presence of the "unobservable" input than they do in its absence. One reason for this dramatic change in deviance might be due to the fact that the estimates of the summation function coefficients are statistically biased estimates in the first case where all the available variability has not been taken into account and they become statistically unbiased estimates as we reach the final model where all the available information is utilised.

(3) fitting a model with only a recovery function makes very little reduction in deviance. This gives the impression that the set of data is input dominated. The recovery function contains almost no extra information to that contained in each of the summation functions and their corresponding carry-over effect functions for observed and "unobservable" inputs as can be seen when the recovery function is added to other models.

(4) evidently the information contained in the summation function and its corresponding carry-over effect function for the first observed input is largely orthogonal to that contained in the summation function and its corresponding carry-over effect function for the second observed input because these two summation functions and their corresponding carry-over effect functions when added separately to the null model reduce the deviance by 1002 and 941 respectively. But when added together to the null model, they reduce the deviance by 1750 which is quite similar to the sum of the reductions when each summation function and its corresponding carry-over effect function is fitted alone. This orthogonal feature between the two observed inputs gives the impression that the two inputs are acting almost independently to excite the firing of the neurone, and this seems to be consistent with the coherence figures shown in Fig.5.2.10a, b and c which also reveal the same sorts of orthogonal

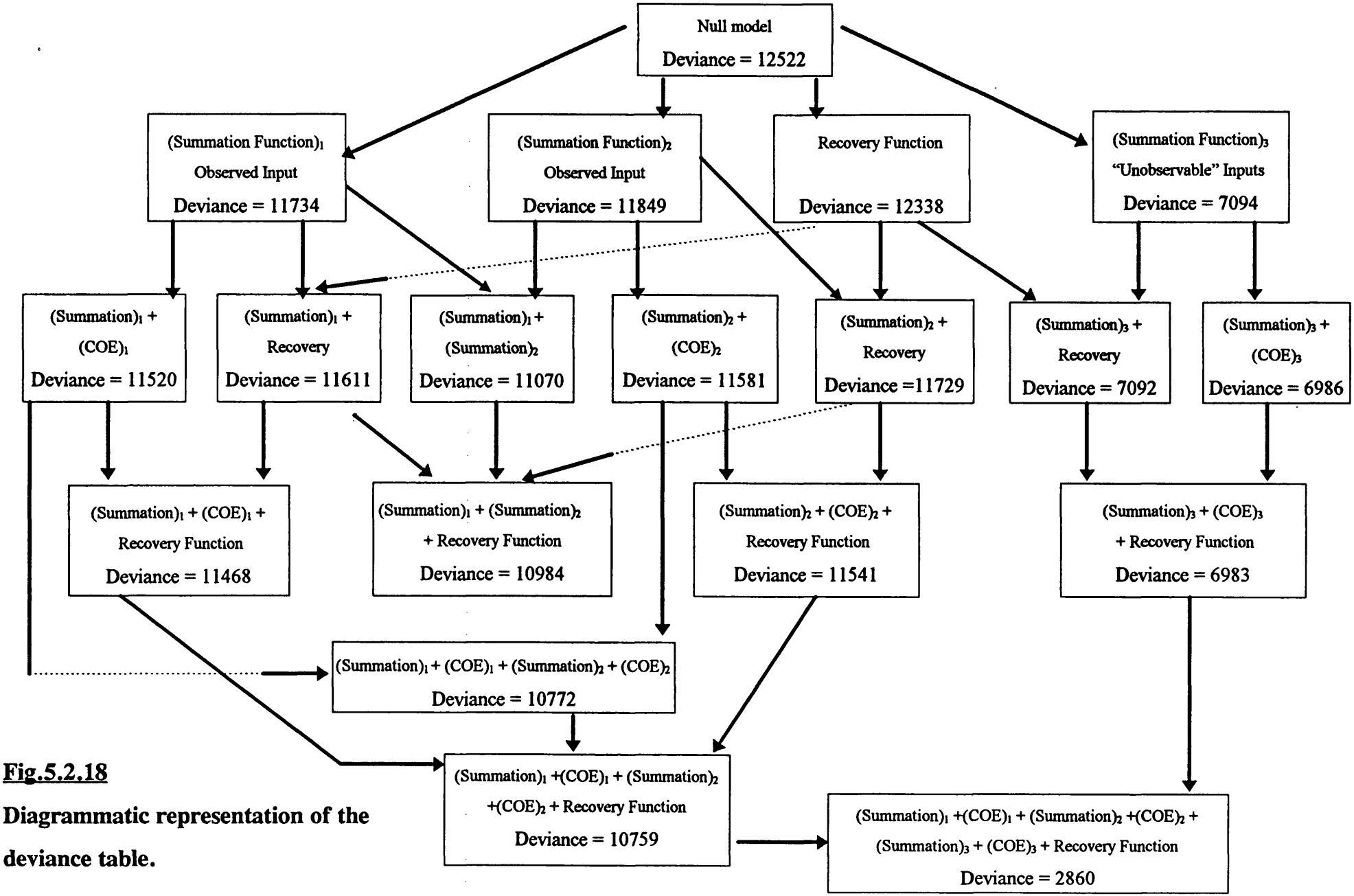


Fig.5.2.18

Diagrammatic representation of the deviance table.

feature (i.e., the two observed input processes are uncorrelated with each other) as we discussed earlier.

(5) the "unobservable" inputs in this set of data seem to explain more of the variability than the recovery, threshold and the summation functions for the two observed inputs when all those functions are fitted together; although each of them is sufficiently informative to be worth fitting. This is a feature of the way in which the data have been simulated, however, this is to be expected since the "unobservable" inputs contain all the other unmeasured inputs.

(6) the information contained in the carry-over effect functions is very small compared to their corresponding summation functions as can be seen from the small reduction in deviance in every case where a carry-over effect function is added to a previous model. This is again a feature of the way in which the data in this example have been simulated. However, F-tests demonstrated that each carry-over effect function was statistically significant, and worth adding to the model.

5.2.3 Two Correlated Spike Train Inputs

In the previous section we discussed the case of two spike train observed inputs and one "unobservable" input where these two observed inputs are uncorrelated (i.e. they act independently to excite the neurone to produce an output spike). The present section deals with the case of two spike train observed inputs and one "unobservable" input where the two observed inputs are correlated with each other. The aim is again to apply the likelihood technique and to compare it with time domain and frequency domain techniques, and to see if the known features of the simulation are reflected in the likelihood analysis.

The schematic diagram of the neuronal network demonstrated in this section is illustrated in Fig.5.2.19 where the two spike train inputs to the neurone are denoted by $_1X$ and $_2X$, the spike train output is denoted by Y and Z denotes the "unobservable" input that is added to represent the effects due to all other unmeasured synaptic inputs.

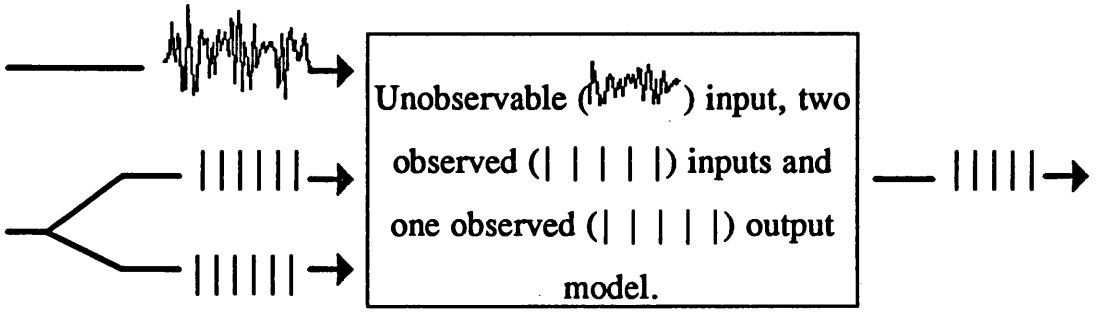


Fig.5.2.19 Diagrammatic representation of a neuronal network in which two dependent observed inputs excite the occurrence of an output in the presence of "unobservable" inputs.

The simulated data set demonstrated in this section is again a 0-1 valued series of approximately 60000 points for each of the two spike train inputs and the output. It also includes a continuous input representing all other "unobservable" inputs. The numbers of spikes observed were 3431 and 3410 for the first ($_1X$) and second ($_2X$) inputs respectively, whereas the number of spikes observed for the output (Y) was 1731.

As in the previous section, the data set was regrouped with binwidth two (i.e., $h = 2$ msec) where it remained as a 0-1 valued series (i.e. maintained the 0-1 property required for the binomial distribution). For comparison purposes, the square root of the cross-intensity function, the ordinary and partial coherences and phases were also estimated. All the illustrated figures in this demonstration are given in the original data scale of 1 msec.

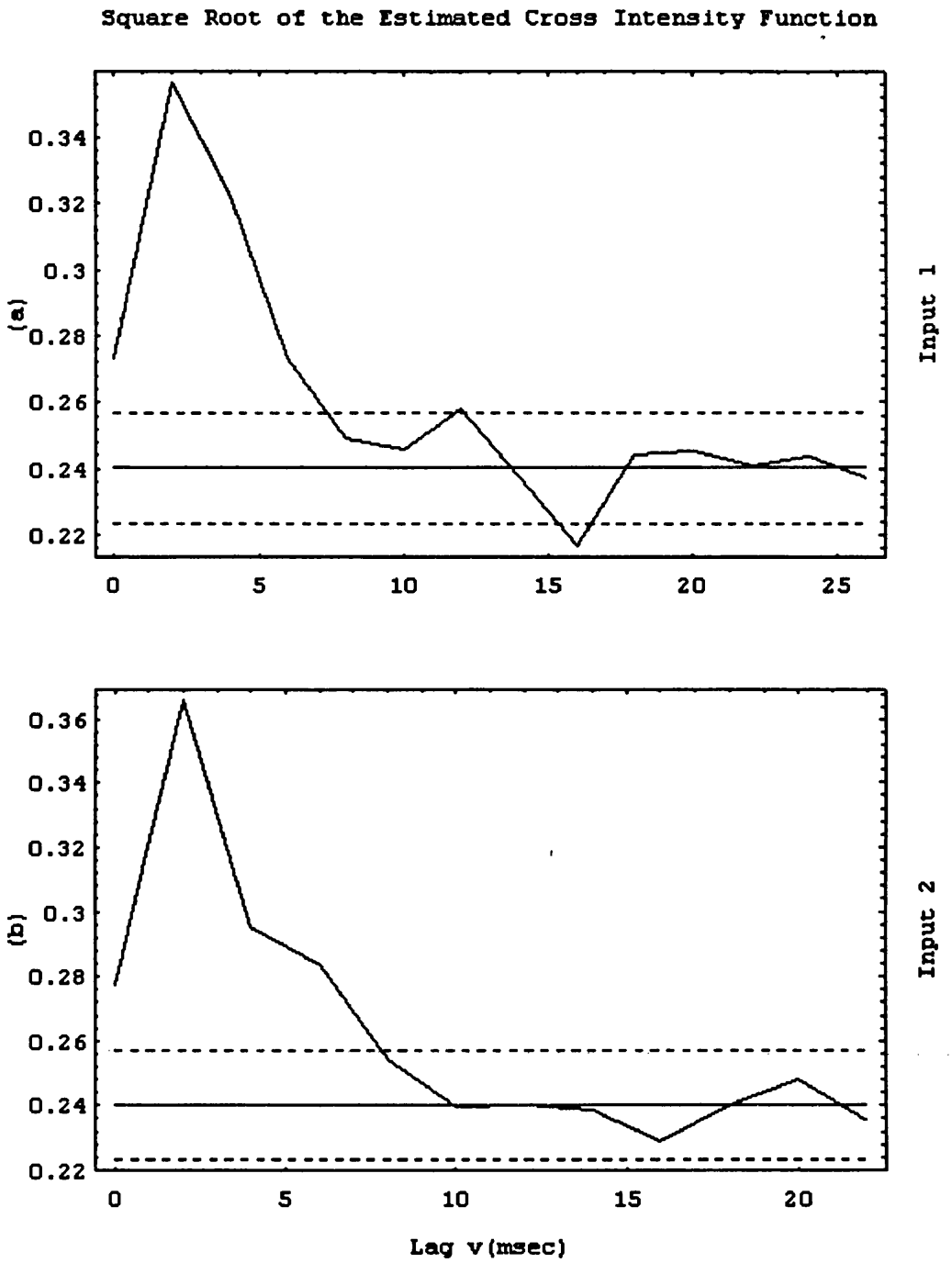


Fig.5.2.20 Estimated square root of the cross-intensity function between the output Y and **a)** the first input ${}_1X$ and **b)** the second input ${}_2X$. The horizontal dotted lines in (a) and (b) represent approximate 95 % confidence intervals for the estimated square roots of the cross-intensity functions under the hypothesis that the two processes are independent, plotted around the square root of the estimated output mean rate (horizontal solid line).

Fig.5.2.20a and Fig.5.2.20b represent the square roots of the estimated cross-intensity functions between the output Y and each of the inputs $_1X$ and $_2X$, respectively. Both cross intensity functions suggest excitatory synaptic input effects lasting about 6 msec.

The estimated ordinary coherence between the first input $_1X$ and the output Y illustrated in Fig.5.2.21b suggests that the first input and the output are weakly coupled over the range of frequencies of about (0 to 75) Hz and the estimated ordinary coherence between the second input $_2X$ and the output Y illustrated in Fig.5.2.21c also indicates a weak coupling over the range of frequencies of about (0 to 75) Hz similar to that shown in Fig.5.2.21b.

The estimated ordinary coherence between the two inputs $_1X$ and $_2X$ as shown in Fig.5.2.21a indicates that the two inputs are well-coupled with each other over the range of frequencies of about (0 to 85) Hz. This dependence can also be seen from a direct comparison between the estimated partial coherence between the output and the first input $_1X$ after removing the effect of the second input $_2X$ (Fig.5.2.22a) and the estimated ordinary coherence between the output and the first input $_1X$ (Fig.5.2.21b) and also a comparison between the estimated partial coherence between the output and the second input $_2X$ after removing the effect of the first input $_1X$ (Fig.5.2.22b) and the estimated ordinary coherence between the output and the second input $_2X$ (Fig.5.2.21c). Here, removing the effect of one input significantly effects the coupling between the output and the other input which suggests that the two observed inputs act dependently to affect the occurrence of the output spikes. Indeed, both partial coherences are virtually statistically non significant, and that suggests that given one input, the other adds little to our ability to predict the output.

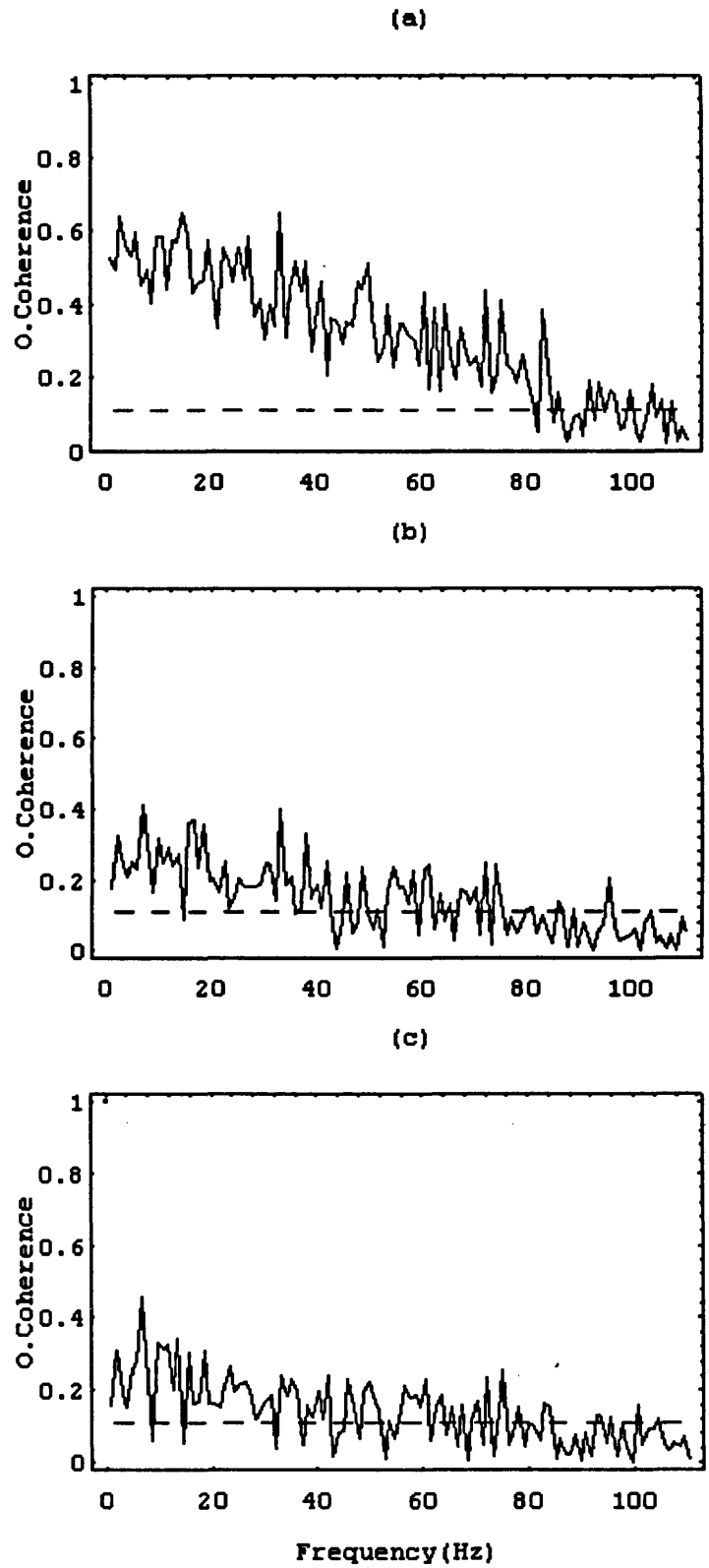


Fig.5.2.21 Estimated ordinary coherence between

a) the two observed inputs ${}_1X$ and ${}_2X$, **b)** the output Y and the first input ${}_1X$ and **c)** the output Y and the second input ${}_2X$. The dotted lines correspond to the upper limit of the 95% confidence intervals for the coherence under the hypothesis of zero coherence.

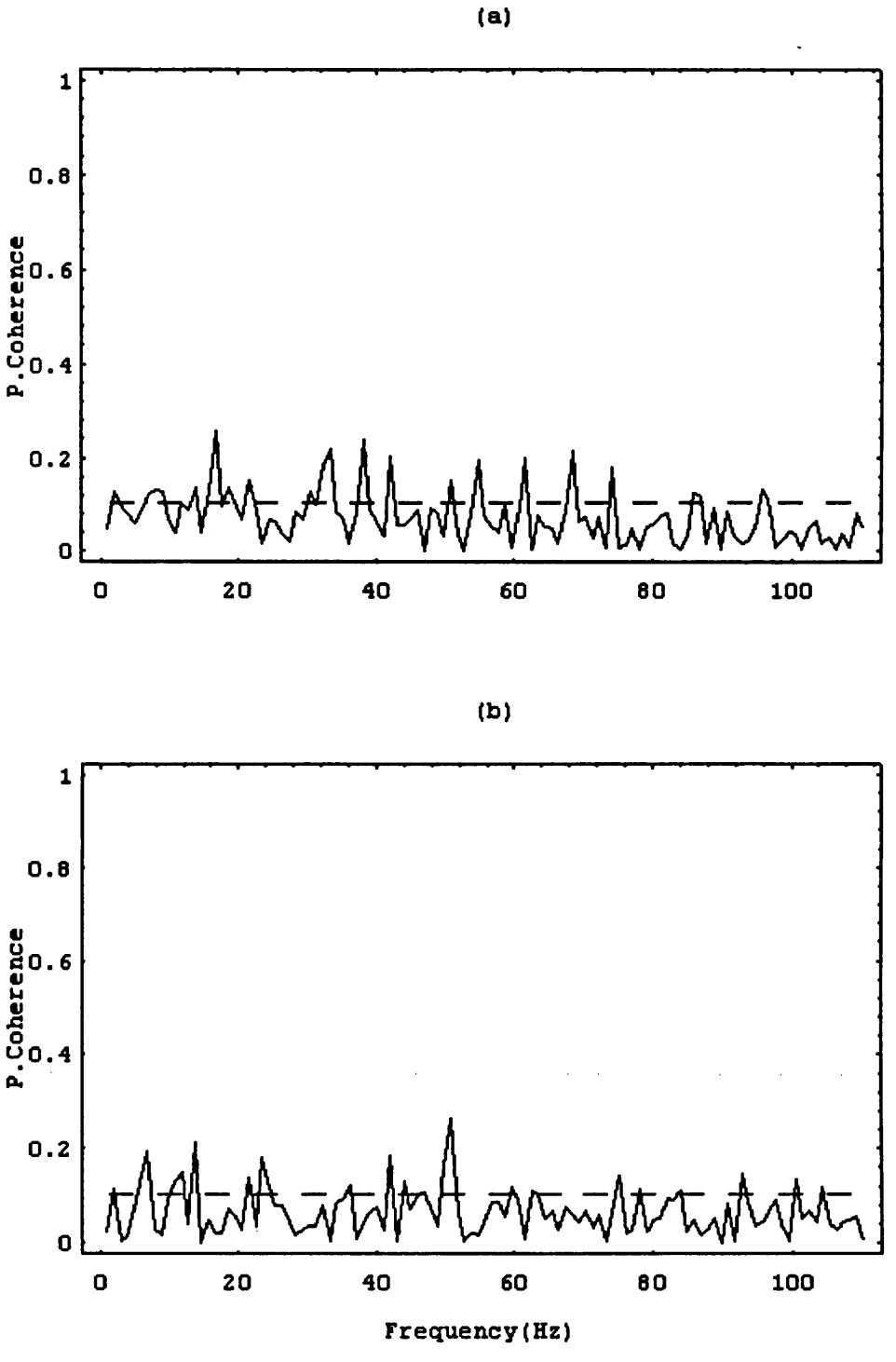


Fig.5.2.22 Estimated partial coherence between the output Y and

a) the first input ${}_1X$ after removing the effect of the second input ${}_2X$ and **b)** the second input ${}_2X$ after removing the effect of the first input ${}_1X$. The dotted lines correspond to the upper limit of the 95% confidence intervals for the coherence under the hypothesis of zero partial coherence.

The estimated phase between the first observed input $_1X$ and the output Y given in Fig.5.2.23a suggests that there is no simple delay present in the data. Similarly, there is no simple delay present between the second observed input $_2X$ and the output Y as given in Fig.5.2.23b (i.e., each phase figure cannot be represented by a straight line).

Fig.5.2.24a and b represent the estimated partial phase between each observed input and the output Y after removing the linear effects of the other observed input. As in the estimated phase plots given in Fig.5.2.23a and b, the estimated partial phase figures suggest also that there is no feature such as a simple delay present.

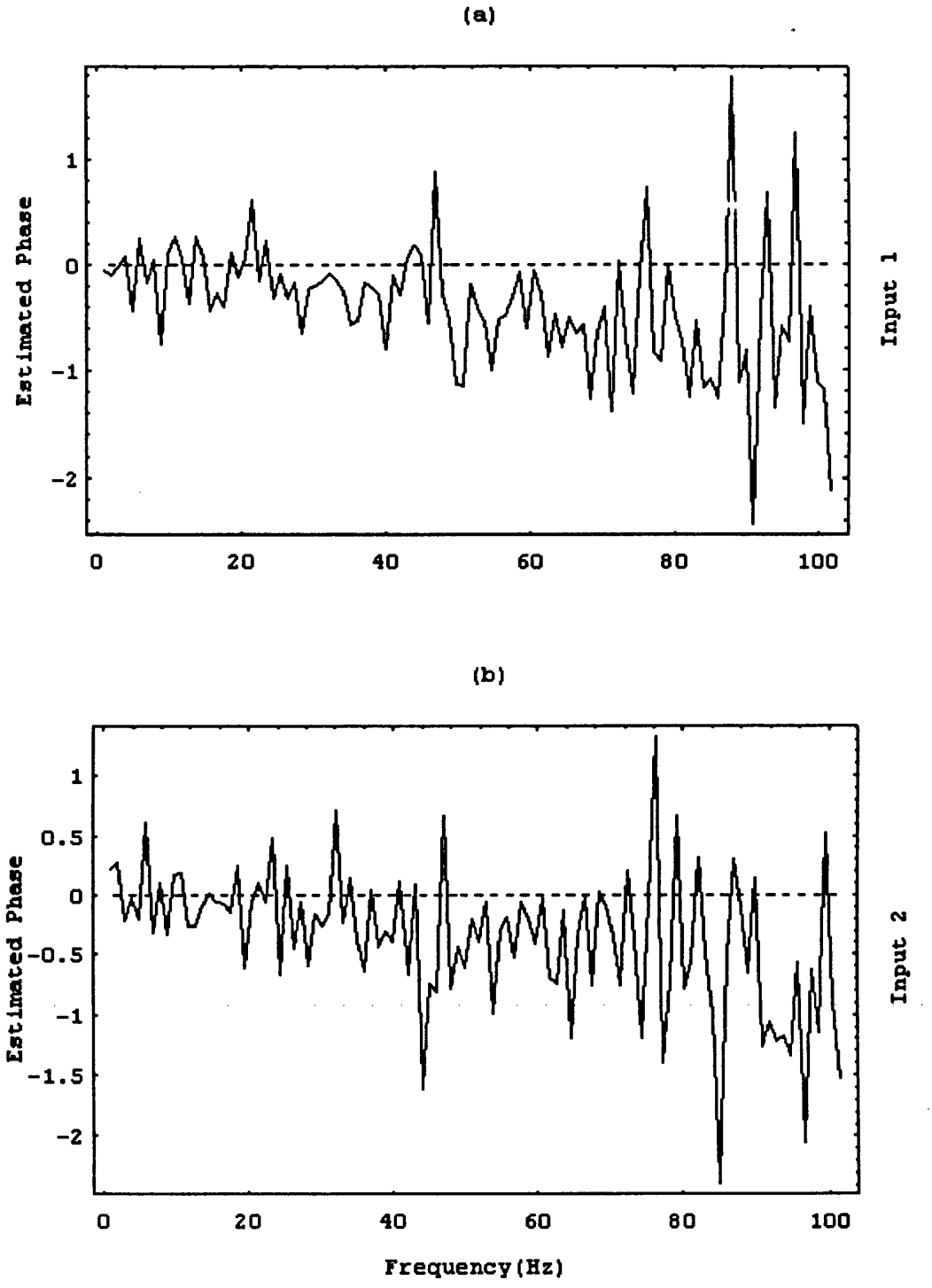


Fig.5.2.23 illustration of the phase functions.

a) Estimated phase between the output Y and the first observed input ${}_1X$.

b) Estimated phase between the output Y and the second observed input ${}_2X$.

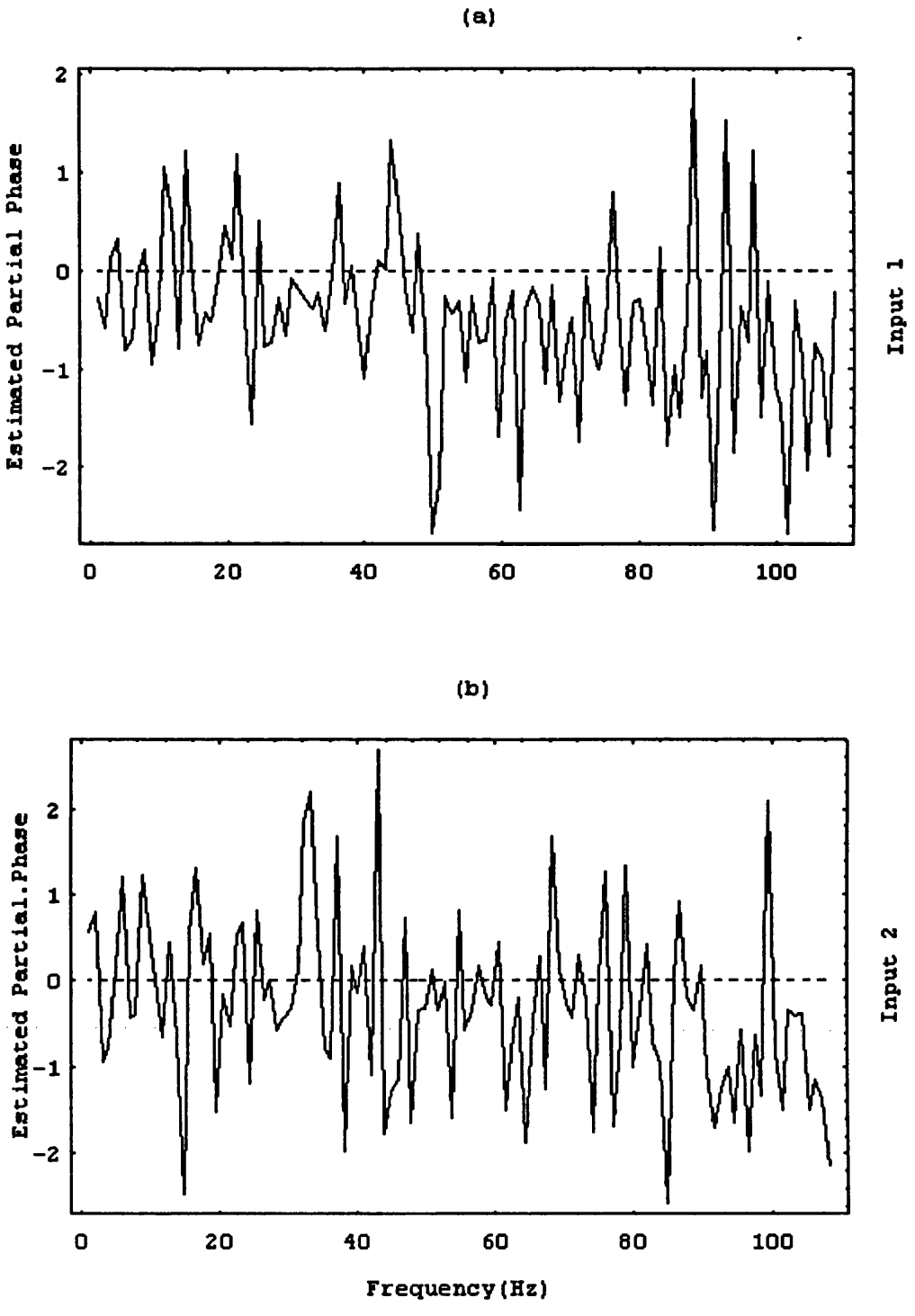


Fig.5.2.24 illustration of the partial phase functions.

Estimated partial phase between **a)** the output Y and the first observed input after removing the effects of the second input and **b)** the output Y and the second observed input after removing the effects of the first input.

The maximum likelihood estimation technique has been applied to the same set of simulated data. The logistic link function was used and found to be superior to the probit link both in terms of the goodness of fit and also the reduction in deviance. The membrane potential U_t, the polynomial recovery function of order k, V_t, and the linear predictor of the model η_t are the same as (5.2.3 - 5.2.5) given in the previous section, respectively. As in the previous section and for the same reasons, the threshold is modelled as a constant.

Fig.5.2.25a, b and c represent the estimated (Summation Function)$_1$, $\{{}_1\hat{a}_u\}$, the estimated (Summation Function)$_2$, $\{{}_2\hat{a}_u\}$, for the first and second observed inputs respectively, and the estimated (Summation Function)$_3$, $\{\hat{b}_u\}$, for the "unobservable" input. The figures suggest that the summation functions for the two observed inputs reveal excitatory effects lasting about 20 msec for the first input and about 16 msec for the second input and the estimated (Summation Function)$_3$ for the "unobservable" input suggests excitatory effects lasting about 22 msec. As in the previous section and for the same reasons, the "unobservable" inputs suggest larger effects than either of the two observed inputs as can be seen both from the reduction in deviance (Fig.5.2.28) when the three summation functions are fitted separately, and from the fact that the coefficients, $\{\hat{b}_u\}$ for the "unobservable" input are statistically more significant at a given lag than either of the two sets of coefficients, $\{{}_1\hat{a}_u\}$ and $\{{}_2\hat{a}_u\}$ for the first and second observed inputs respectively.

The significant durations for each of the two summation functions for the observed inputs given in Fig.5.2.25a and b above are much longer than the 6 msec significant durations of the corresponding cross-intensity functions as shown in Fig.5.2.20a and b. This suggests the same conclusion as in the previous example (section 5.2.2), that the square root of the cross-intensity function is underestimating the synaptic input effects. The cross intensity

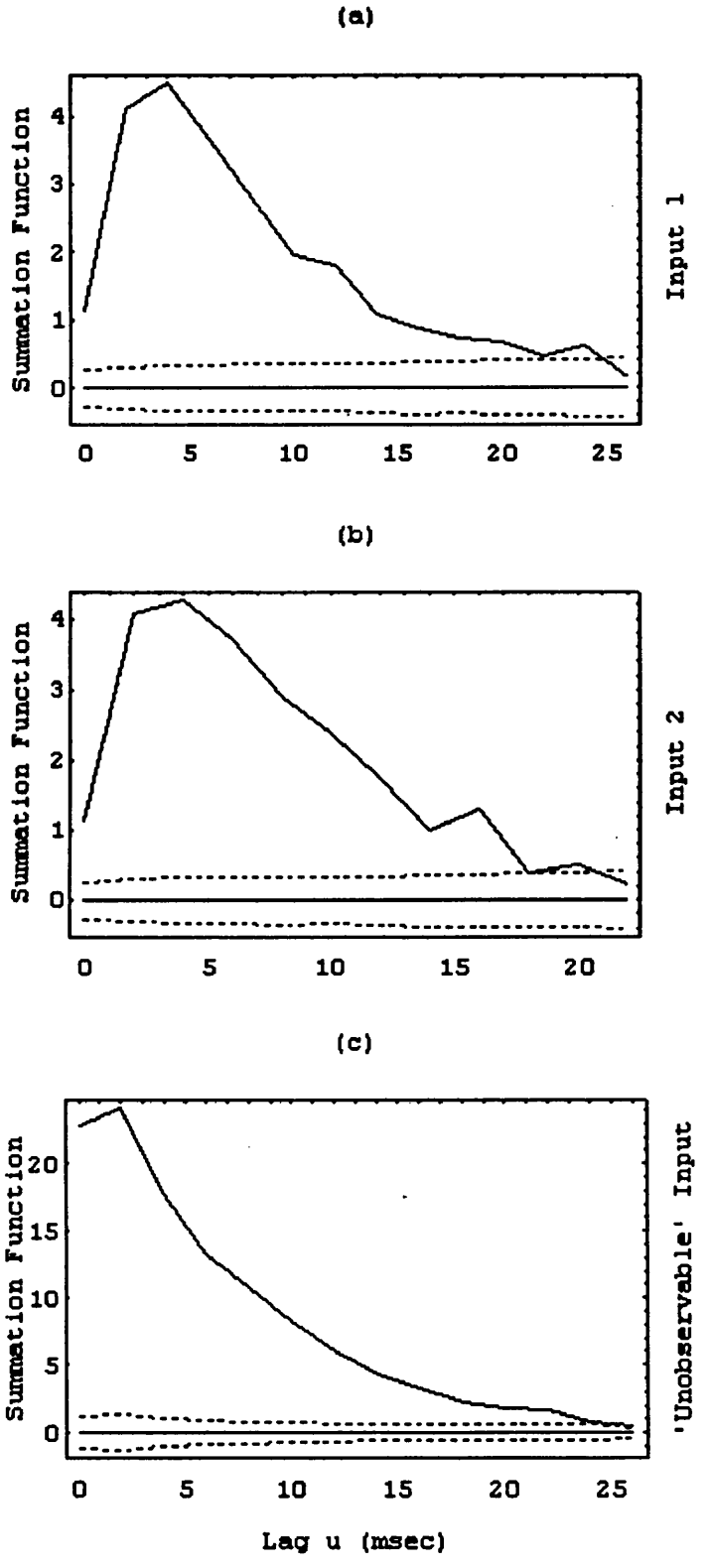


Fig.5.2.25 a) Estimated summation function $\{{}_1\hat{a}_u\}$ for the first observed input ${}_1X$. b) Estimated summation function $\{{}_2\hat{a}_u\}$ for the second observed input ${}_2X$. c) Estimated summation function $\{\hat{b}_u\}$ for the "unobservable" input Z. The horizontal dotted lines give $\pm$ two standard error limits plotted about zero for the summation functions.

functions for the two observed inputs were estimated via the likelihood approach, as in the previous example (see also section 5.2.1). The residual deviance for this model is 12369; a reduction of only 866 from the null model (i.e., the variation explained by the two cross-intensity functions when fitted together is only 6.5 %), providing still further evidence that the cross-intensity function in general has very poor explanatory power, whereas the reduction for a model involving the two summation functions was 1696.

Fig.5.2.26a, b and c represent the carry-over effect functions for the first and second observed inputs and the "unobservable" input respectively, and suggest that while the estimated carry-over effect function $(COE)_3$ for the "unobservable" input (Fig.5.2.26c) suggests excitatory effects lasting from about 1 to 17 msec, the estimated carry-over effect functions $(COE)_1$ and $(COE)_2$ suggest excitatory effects which seem to be much smaller in statistical significance. This can be seen also from the reduction in deviance as we will see shortly.

The threshold and recovery functions for the full model given in Fig.5.2.27a suggest that the probability of an output spike occurring spontaneously is very small over the whole range of intervals (as the recovery function remains far below threshold over the whole range of intervals). Adding the recovery function to other models seems to have very little effect as we will see shortly from the table of deviances which indicates clearly that the recovery function is almost negligible.

The goodness of fit plot shown in Fig.5.2.27b suggests that the fit of the seven component model is satisfactory.

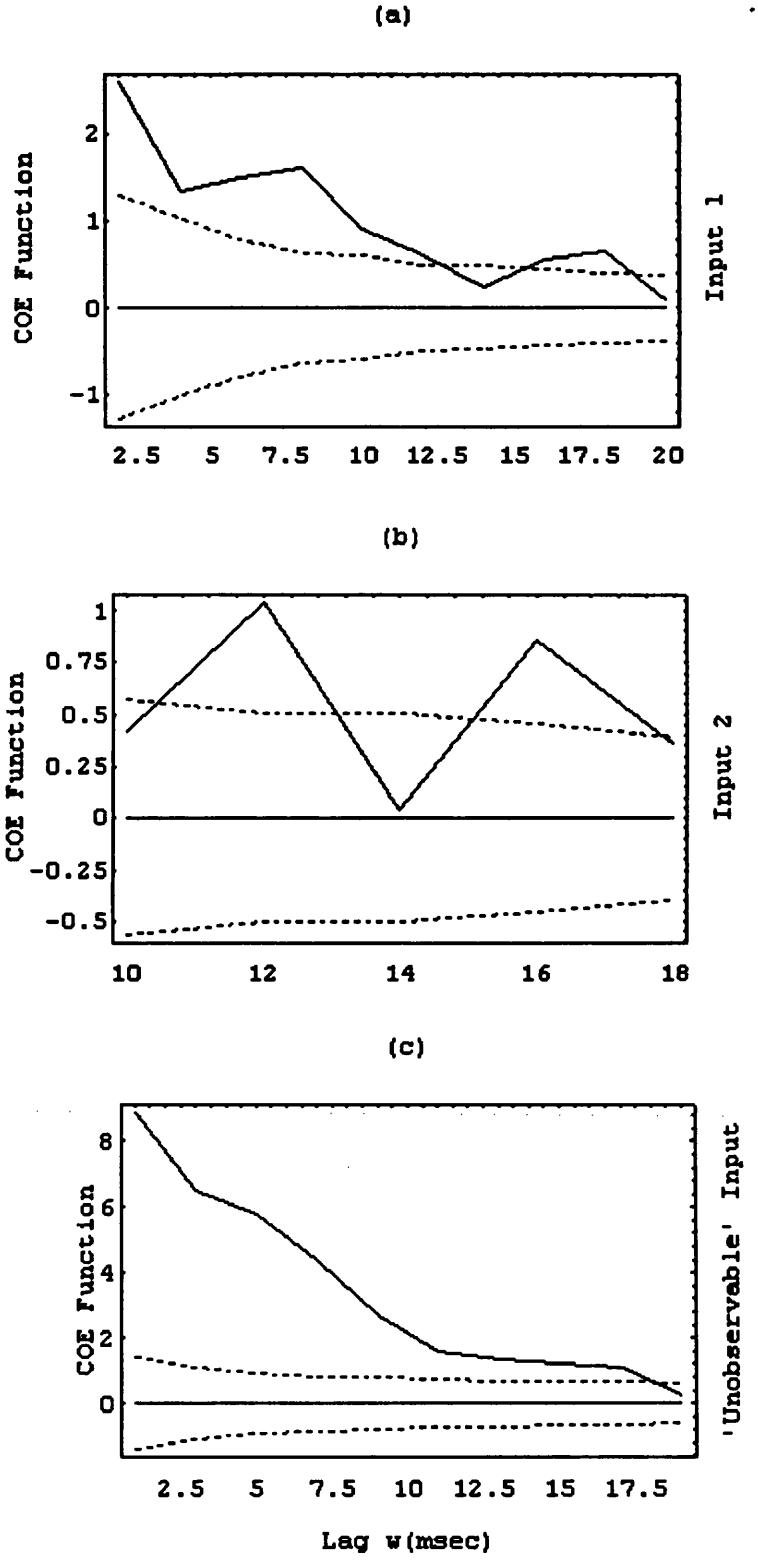


Fig.5.2.26 a) Estimated carry-over effect function $\{{}_1\hat{c}_w\}$ for the first observed input ${}_1X$. b) Estimated carry-over effect function $\{{}_2\hat{c}_w\}$ for the second observed input ${}_2X$. c) Estimated carry-over effect function $\{\hat{d}_w\}$ for the "unobservable" input Z. The horizontal dotted lines give $\pm$ two standard error limits plotted about zero for the carry-over effect functions.

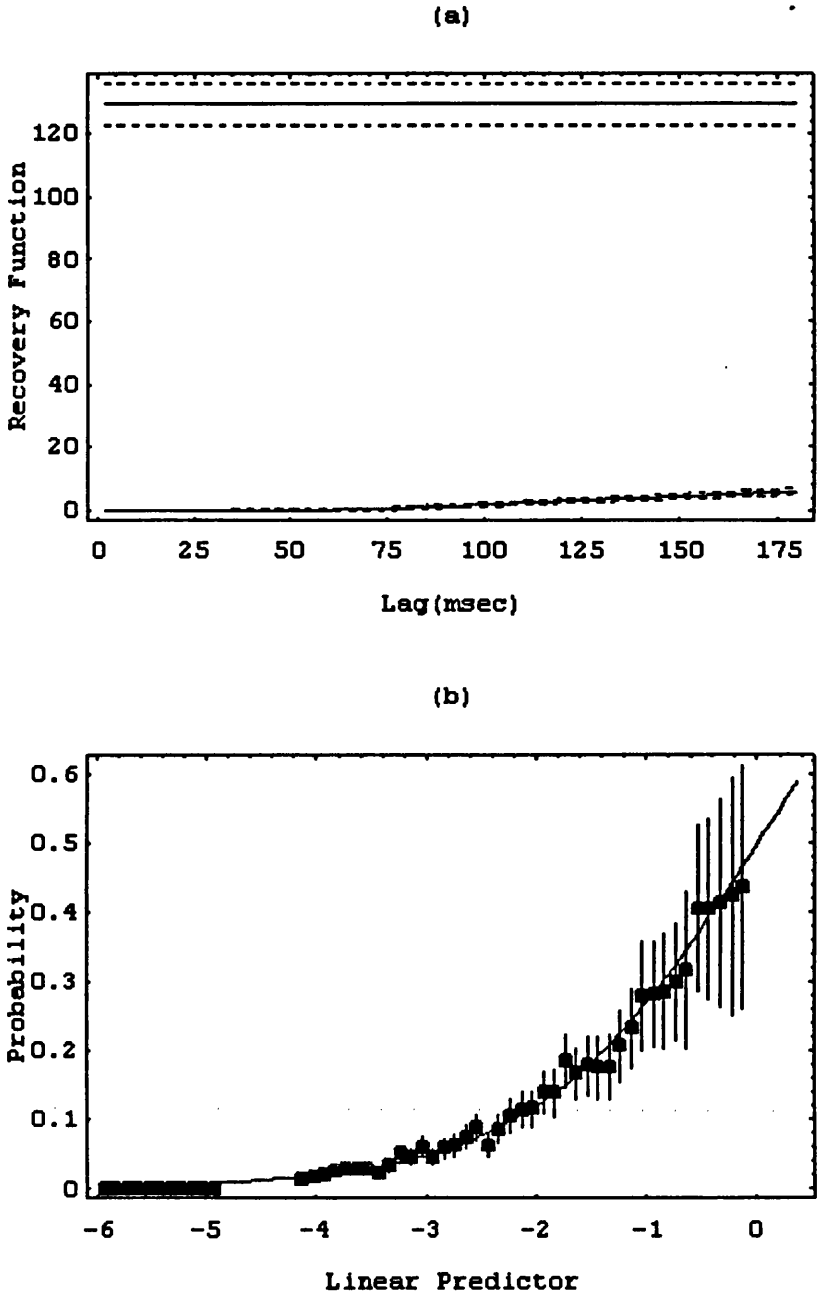


Fig.5.2.27 illustration of the threshold and recovery functions and the goodness of fit plot.

a) Estimated recovery (lower curve) and threshold (upper solid line) functions. The dotted lines give ± two standard error limits plotted around each function for the threshold and recovery functions.

b) The goodness of fit plot with empirical (dots) and theoretical (smooth curve) probabilities plotted against selected values for the linear predictor, the vertical bars present the ± two standard error limits for the theoretical probabilities.

The deviance table given in Fig.5.2.28 illustrates the sequential fitting of a set of successively more complex models and reveals the following interesting features,

(1) fitting all seven components reduces the deviance from 13235 (the initial model) to 3335; a reduction of about 75 % which is a substantial reduction.

(2) as we have seen in the previous example, fitting a model with only the summation functions (and their corresponding carry-over effect functions) for the two observed inputs and a recovery function reduces the deviance from 13235 (the initial model) to 11284; a reduction of 1951. But a larger reduction in deviance is achieved as we move from a model with only the summation function (and the corresponding carry-over effect function) for the "unobservable" input to the full model which reduces the deviance from 7541 to 3335; a reduction of 4206. Both of these reductions in deviance (i.e., 1951 and 4206) represent the effects of the two observed inputs and the recovery process in the absence and the presence of the "unobservable" input effects respectively. They are quite different, i.e., the effects of the two observed inputs and the recovery process seem to explain a larger amount of the variability in the presence of the "unobservable" input than they do in the absence of the "unobservable" input. The reason for this change in deviance, as we mentioned earlier in the previous analysis, might be due to the fact that the estimates of the summation function coefficients are statistically biased estimates in the first case where all the available variability has not been taken into account and they become statistically unbiased estimates as we reach the final model where all the available inputs are included.

(3) fitting a model with only a recovery function makes very little reduction in deviance. Adding a recovery function to other models seems to

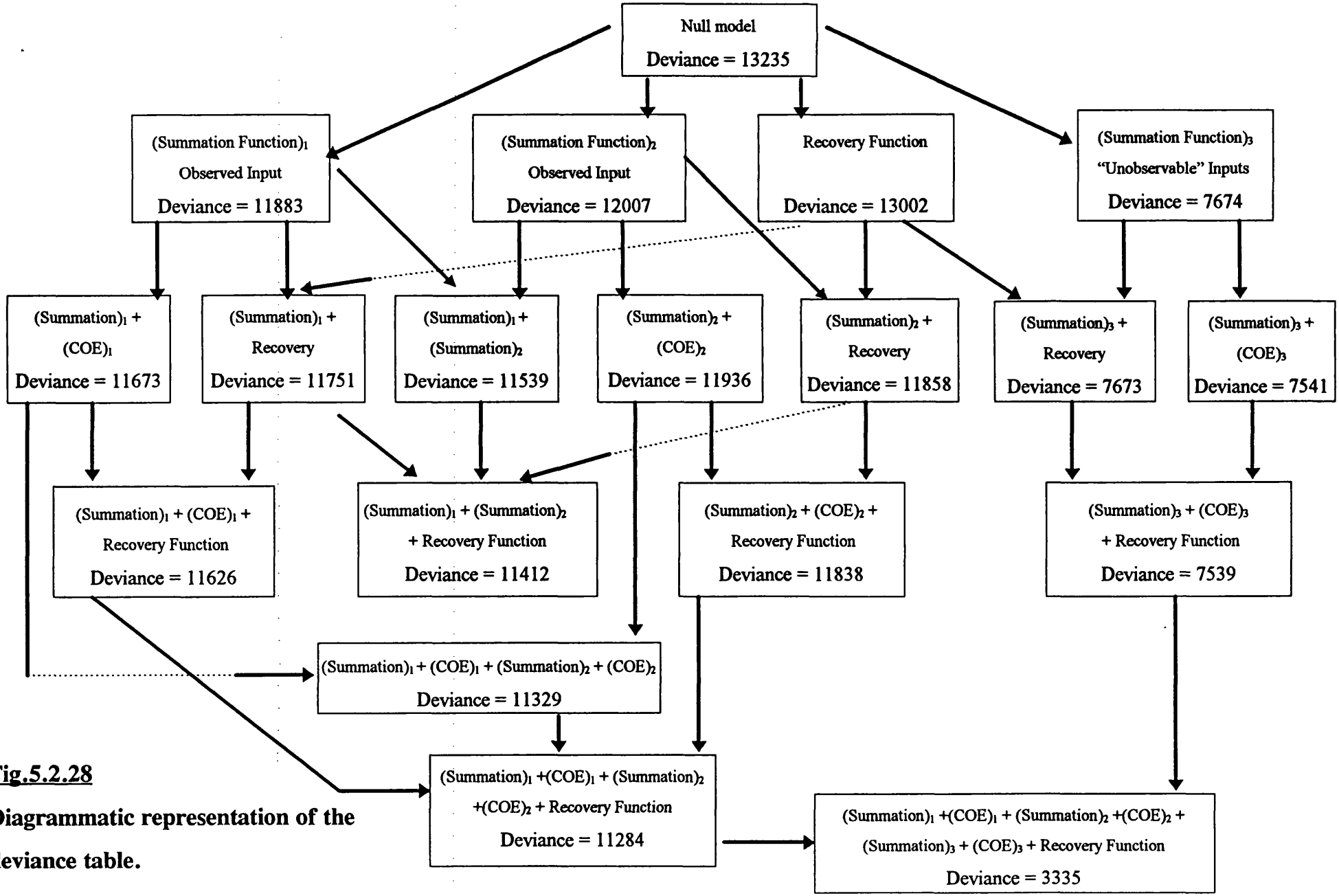


Fig.5.2.28

Diagrammatic representation of the deviance table.

provide almost no extra information to that contained in each of the summation functions and their corresponding carry-over effect functions for both observed and "unobservable" inputs. This gives the impression that the set of data is an input dominated one (i.e., the cell fires as a result of the excitatory synaptic input effects rather than spontaneously).

(4) in the previous section we noticed from the deviance table that the amounts of information contained in each of the summation functions (and their corresponding carry-over effect functions) for the two observed inputs were largely orthogonal to each other. But in the case we are discussing now, the two summation functions (and their corresponding carry-over effect functions) for the two observed inputs seem to share information (i.e., the amounts of information contained in each of the two summation functions overlap), since the two summation functions and their corresponding carry-over effect functions when added separately to the null model reduce the deviance by 1562 and 1299 respectively. But when added together to the null model, they reduce the deviance by only 1906 which is much smaller than the sum of the reductions (2861) when they are fitted separately. This feature, unlike the previous example, gives the impression that the two inputs act dependently to excite the firing of the neurone, and this seems to be consistent with the coherence shown in Fig.5.2.21a which also reveals that the two observed input processes are well-correlated with each other. Also it seems in agreement with the conclusions made when we compared the figures of the estimated coherence and phase with the figures of their corresponding partial coherence and partial phase which suggested also that the two observed inputs seem to act dependently to affect the occurrence of the output spikes.

(5) in this example and as in the previous section, the "unobservable" inputs again seem to explain much more of the variability (about 43 %), more than the recovery, threshold functions and the summation functions for the two observed inputs when all fitted together (about 15 %), although each of them is sufficiently informative to be worth fitting. This is again a feature of the simulation procedure used to simulate the data set in this example and is anyway to be expected since the "unobservable" inputs contain all the other unmeasured inputs.

(6) all the three carry-over effect functions seem to contain very little information compared to their corresponding summation functions as can be seen from a negligible reduction in deviance every time a carry-over effect function is added to a previous model. But each carry-over effect function was statistically significant and worth adding to the model, and this was verified using F-tests. The conclusions given in (5) and (6) above seem to be consistent with the way in which the data have been simulated.

5.3 Real Data Set Obtained From Mammalian Muscle Spindle

In the previous two sections we discussed the application of likelihood techniques to simulated data in the case of two spike train inputs and one output in the presence of an "unobservable" input representing all other unmeasured inputs. In this section we turn our attention to the application of likelihood methods to a real set of data in the case of two inputs and one output (see Fig.5.3.1 below) obtained from a muscle spindle (see section 1.4 of chapter 1) lying within the tenuissimus muscle in the hind limb of a deeply anaesthetised cat. The fusimotor axons were stimulated with voltage pulses, and the resulting primary (Ia) and secondary (II) responses in the form of sequences of pulses from the spindle were recorded. For further detail concerning the recording technique used, see chapter 1 of this thesis, Halliday et al (1988), Rosenberg et al (1989) and Amjad (1989).

The data set demonstrated here consists of a 0-1 valued series of approximately 50000 points for each of the two inputs and the output. The numbers of spikes observed were 3180 and 1524 for the fusimotor axons γ_0 and γ_b inputs respectively, whereas the number of spikes observed for the primary (Ia) and secondary (II) outputs were 1164 and 1728 respectively.

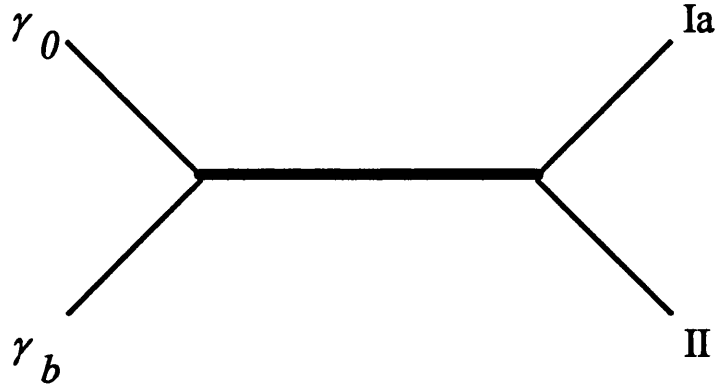


Fig.5.3.1 Diagrammatic representation of a muscle spindle with two observed fusimotor axon inputs, γ_0 and γ_b, and the two primary (Ia) and secondary (II) outputs.

Using the maximum likelihood estimation technique the threshold, recovery, summation and carry-over effect functions for each of the outputs separately were estimated. For computational purposes the data have been split up into two disjoint segments of 24950 points and the individual maximum likelihood estimates for each segment were averaged to obtain the final estimates. Using the stochastic point process techniques reviewed in chapter 2 the square root of the cross-intensity function, the ordinary and partial coherences and phase were also estimated. The analysis was done in two stages where we consider both inputs and only one output in each stage. The possibility of considering both outputs together (i.e., two input and two output likelihood model) where very useful results might be obtained remains as a possible area in which this work may be extended (see chapter 6). We start with the two observed fusimotor axon inputs, γ_a and γ_b, and the secondary (II) output.

5.3.1 The Two Fusimotor Inputs and the Secondary (II) Output

We start with the stochastic point process techniques. Fig.5.3.2a and b represent the two square roots of the estimated cross-intensity functions between the secondary (II) output and each of the inputs γ_a and γ_b, respectively. The square root of the estimated cross-intensity function corresponding to the first input γ_a suggests an excitatory synaptic input effect lasting from about 13 to 21 msec whereas the square root of the estimated cross-intensity function between the output and the input γ_b suggests very small excitatory effects of synaptic input but these effects are almost negligible.

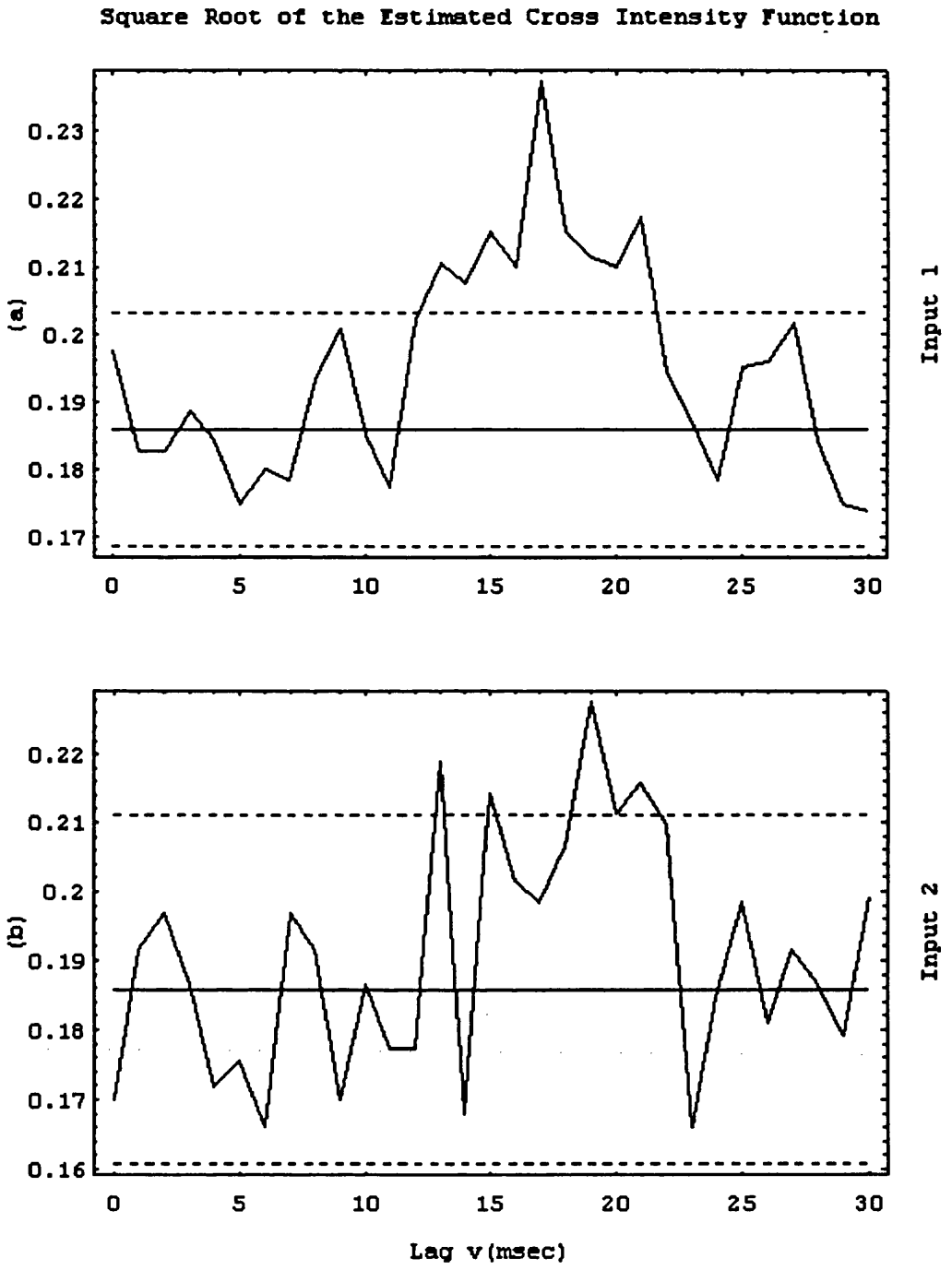


Fig.5.3.2 a) Estimated square root of the cross-intensity function between the secondary (II) output and the input γ_0. b) Estimated square root of the cross-intensity function between the secondary (II) output and the input γ_b. The horizontal dotted lines in (a) and (b) represent approximate 95 % confidence intervals for the estimated square roots of the cross-intensity functions under the hypothesis that the two processes are independent, plotted around the square root of the estimated output mean rate (horizontal solid line).

Fig.5.3.3b represents the estimated ordinary coherence between the first input γ_0 and the secondary (II) output and suggests that the two processes are well coupled over the range of frequencies of 0 to about 25 Hz. The estimated ordinary coherence between the second input γ_b and the output (Fig.5.3.3c) indicates a weaker coupling over the range of frequencies of 0 to about 12 Hz. The estimated ordinary coherence between the first input γ_0 and the second input γ_b is shown in Fig.5.3.3a and indicates that the two inputs are not correlated with each other over the whole range of frequencies. This feature can also be seen from a direct comparison between the two figures (Fig.5.3.4a-b) of the estimated partial coherences between the output and each input after removing the effect of the other input and those of the ordinary coherences between the output and each input given in Fig.5.3.3b and Fig.5.3.3c. In each case, the ordinary and partial coherences are quite similar to each other. This suggests that the two inputs act almost independently to affect the occurrence of the output spikes.

Fig.5.3.5a represents the estimated partial phase between the first observed fusimotor axon input, γ_0, and the secondary (II) output after removing the linear effects of the second observed fusimotor axon input, γ_b. The weighted least squares line (dotted) fitted to the partial phase curve over the range of frequencies where the corresponding partial coherence (Fig.5.3.4a) is significantly different from zero shows that, the output process is delayed, on the average, by an amount 18.15 msec with a 95% confidence interval, for the delay, of (16.46, 19.84) msec. The value of the delay suggested by the partial phase seems to be consistent with the 17 msec peak in the corresponding square root of the estimated cross intensity function (Fig.5.3.2a).

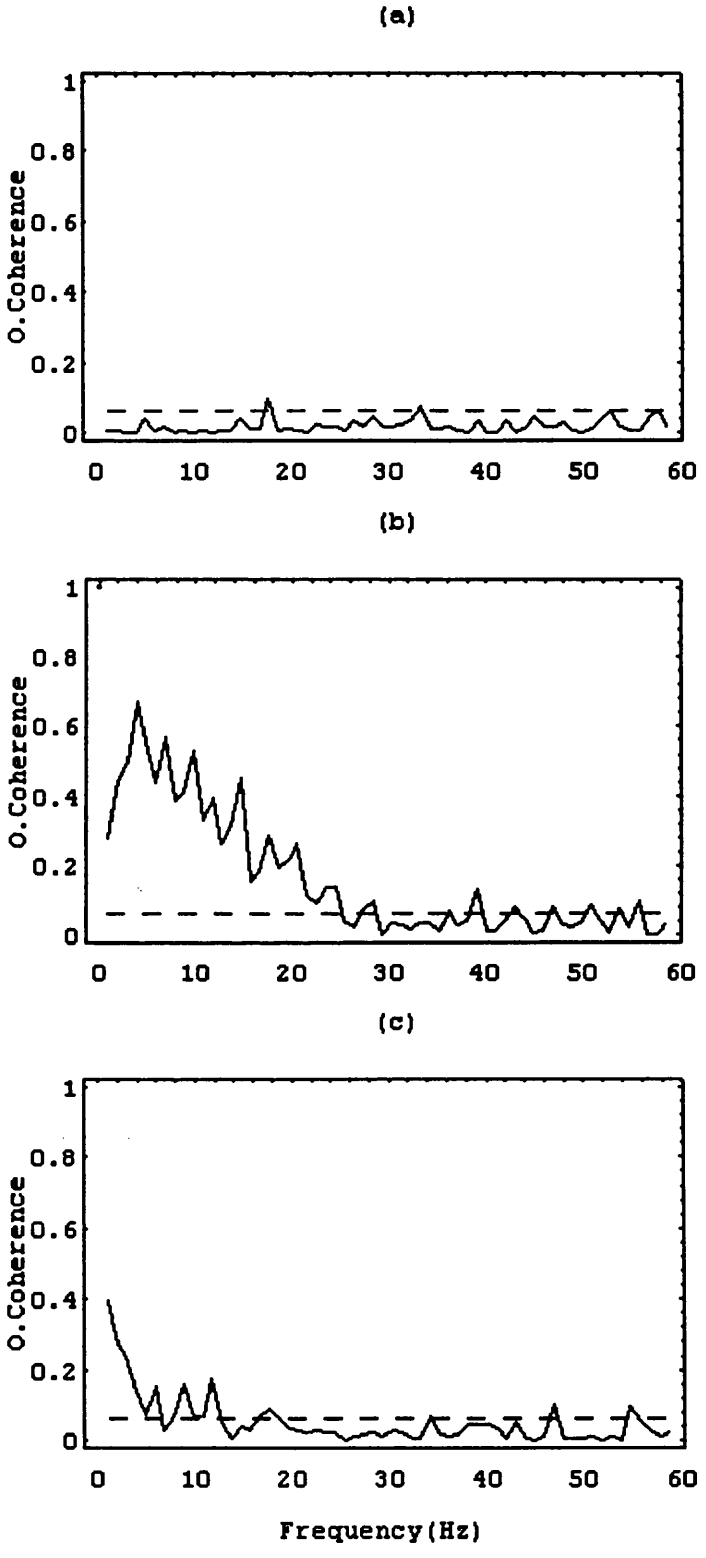


Fig.5.3.3 Estimated ordinary coherence between

a) the two inputs γ_0 and γ_b, **b)** the secondary (II) output and the input γ_0 and **c)** the secondary (II) output and the input γ_b. The dotted lines correspond to the upper limit of the 95% confidence intervals for the coherence under the hypothesis of zero coherence.

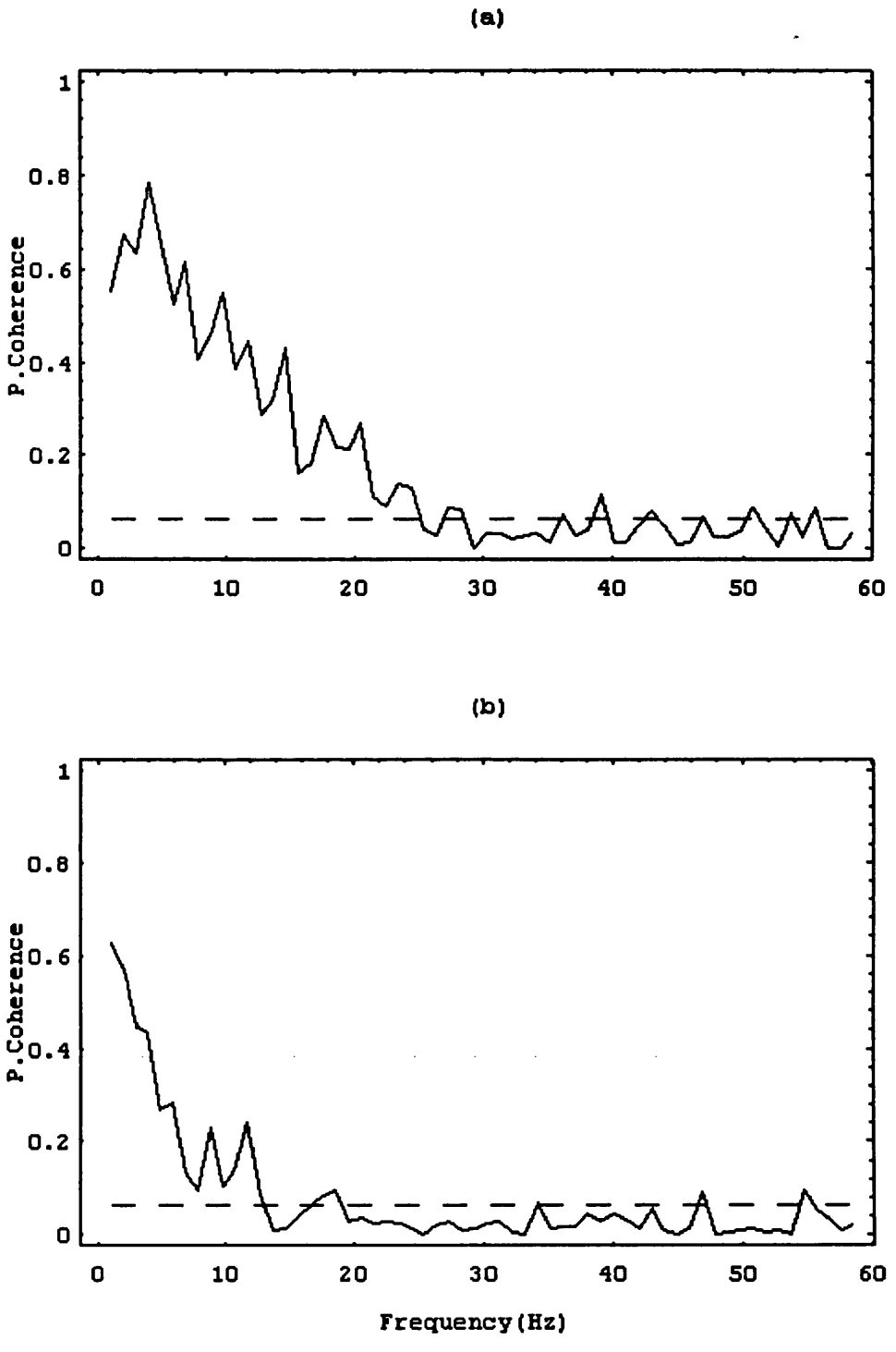


Fig.5.3.4 Estimated partial coherence between the secondary (II) output and

a) the first input γ_0 after removing the effect of the second input γ_b and **b)** the second input γ_b after removing the effect of the first input γ_0. The dotted lines correspond to the upper limit of the 95% confidence intervals for the coherence under the hypothesis of zero partial coherence.

The estimated partial phase between the second observed fusimotor axon input, γ_b, and the secondary (II) output after removing the linear effects of the first input is presented in Fig.5.3.5b. The estimated slope of the weighted least squares line (dotted) over the range of frequencies at which the corresponding partial coherence (Fig.5.3.4b) is significantly different from zero shows that, the output process is delayed, on the average, by an amount 20.16 msec with a 95% confidence interval, for the delay, of (13.06, 27.26) msec. The value of the delay suggested by the partial phase again seems to be consistent with the 19 msec peak in the corresponding square root of the estimated cross intensity function (Fig.5.3.2b).

In this real data the peaks in the two summation functions are centred about 24 msec and 22 msec for the first (γ_0) and second (γ_b) inputs respectively. These differ from the estimated times of the peaks in the corresponding cross intensity functions and from the two values of the delay estimated from the partial phase. As can be seen from the deviance table (Fig.5.3.8), this data set is not input dominated. This lack of agreement suggests that the delay may not be the most appropriate description and the situation is more complicated than what is suggested by the partial phase and the cross intensity function. Again it seems that perhaps summarising this in a single number and calling it a "delay" may not always be sensible.

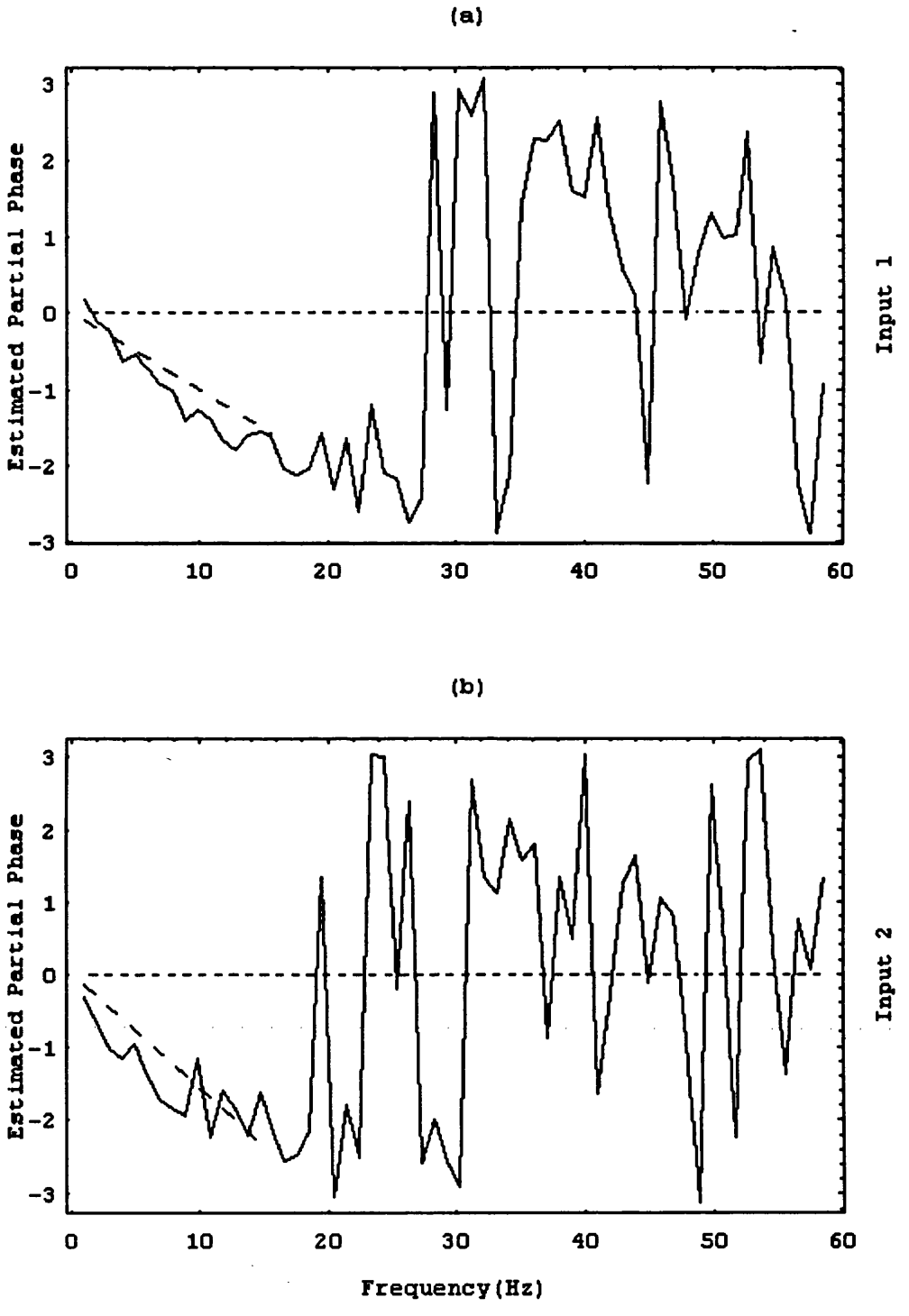


Fig.5.3.5 Estimated partial phase between the secondary (II) output and

a) the first input γ_0 after removing the effect of the second input γ_b and **b)** the second input γ_b after removing the effect of the first input γ_0. The slopes of the fitted weighted least squares lines (dotted) correspond to the estimated time delays.

The maximum likelihood estimation technique has been applied to the same set of muscle spindle real data. The logistic link function was used and found to be superior to the probit link both in terms of the goodness of fit and also the reduction in deviance. The membrane potential, U_t, the polynomial recovery function of order k, V_t, and the linear predictor of the model η_t are the same as (5.2.3 - 5.2.5) given in section (5.2.2) respectively, except in this real set of data the "unobservable" inputs are not available to us.

Fig.5.3.6a and b represent the estimated (Summation Function)$_1$, $\{{}_1\hat{a}_u\}$, and the estimated (Summation Function)$_2$, $\{{}_2\hat{a}_u\}$, for the first (γ_0) and second (γ_b) inputs respectively. The figures suggest that, while the summation function for the first input reveals an excitatory effect lasting from about 13 to 29 msec, the summation function for the second input suggests an excitatory effect lasting from about 18 to 22 msec, but these effects are relatively weaker and also shorter compared to those for the first input γ_0.

The significant duration of the summation function for the first input (Fig.5.3.6a) is also longer than that for the corresponding cross intensity function as illustrated in Fig.5.3.2a. This suggests again that the square root of the cross-intensity function underestimates the synaptic input effects. The significant duration of the summation function for the second input (Fig.5.3.6b) however is quite similar to that for the corresponding cross intensity function as shown in Fig.5.3.2b.

The threshold and recovery functions are well estimated over the whole range of intervals (as given in Fig.5.3.6c) and indicate that the probability of an output spike is very small up to about 20 msec after the previous output spike, but it then increases rapidly and the chance of an output spike becomes quite large after about 40 msec. As we have seen in the figure the recovery function

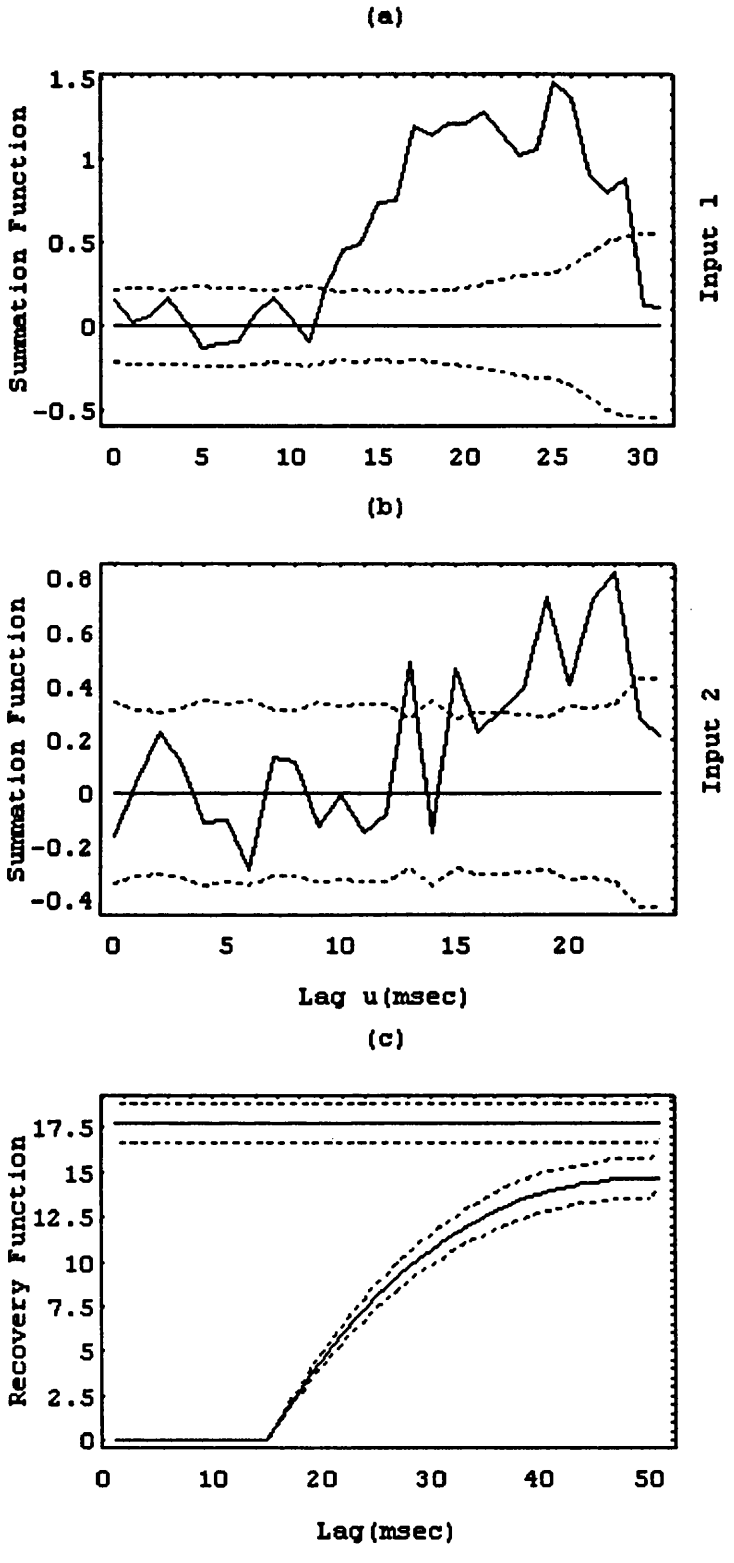


Fig.5.3.6 a) Estimated summation function $\{_1\hat{a}_u\}$ for the first input γ_0. b) Estimated summation function $\{_2\hat{a}_u\}$ for the second observed input γ_b. c) Estimated threshold and recovery functions. The dotted lines give $\pm$ two standard error limits plotted about zero in (a) and (b) for the summation functions and plotted around each function in (c) for the threshold and recovery functions.

is forced to be zero for intervals less than the minimum of the output inter-spike intervals ζ_1, because there will be no output data available at those corresponding intervals.

Fig.5.3.7a and b represent the carry-over effect functions for the first and second inputs. The estimated carry-over effect function $(COE)_1$ for the first input suggests excitatory effects lasting from about 21 to 56 msec whereas the estimated carry-over effect function $(COE)_2$ for the second input suggest excitatory effects lasting from about 26 to 54 msec.

The goodness of fit plot given in Fig.5.3.7c seems reascnable in the sense that the confidence interval about each of the empirical probabilities (dots) are all seen to contain the corresponding theoretical probabilities and the confidence intervals are reasonably narrow. As the unexplained variability by the model is relatively large compared to that in previous examples where the "unobservable" inputs were also considered, the model takes the range of the linear predictor only to values near zero. This affects the ability of the model to predict higher probabilities.

The deviance table given in Fig.5.3.8 reveals the following interesting features,

(1) the recovery and threshold functions taken together explain more of the variability (about 24.84 %) than the summation functions for the two observed inputs when both are fitted together (about 15.04 % of the variability). This suggests that the output spikes are produced mainly by the effects of other unmeasured inputs or by spontaneous activity or a combination of the two, since the recovery function can demonstrate the intrinsic properties of the cell as well as the effects due to any unmeasured inputs if these latter are not modelled (as we have discussed earlier in sections 4.7 and 5.2).

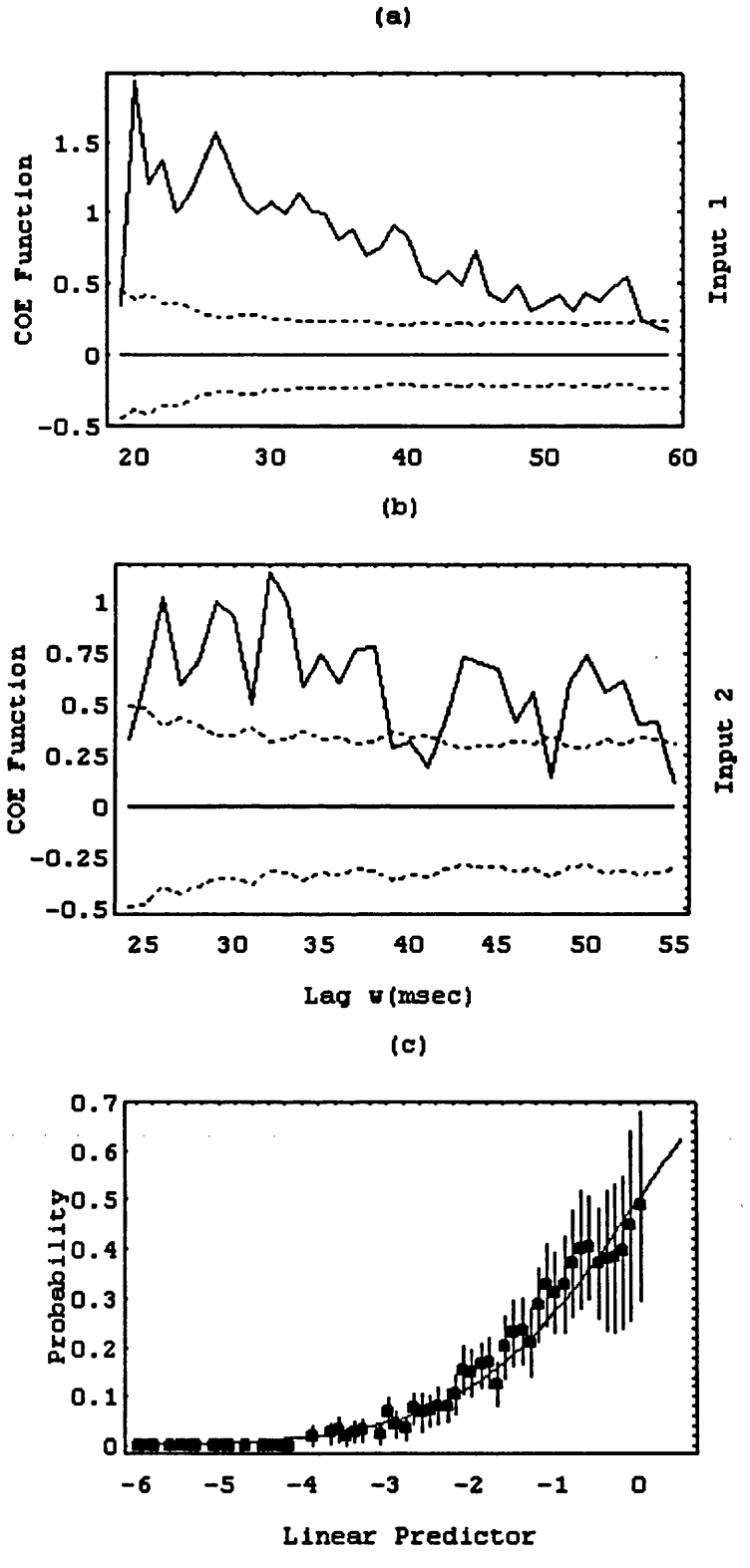


Fig.5.3.7 illustration of the two carry-over effect functions and goodness of fit plot.

a) Estimated carry-over effect function $\{{}_1\hat{c}_w\}$ for the fusimotor input γ_0. **b)** Estimated carry-over effect function $\{{}_2\hat{c}_w\}$ for the fusimotor input γ_b. The dotted lines in a) and b) give $\pm$ two standard error limits plotted about zero for the summation functions. **c)** The goodness of fit plot.

(2) fitting all five components reduces the deviance from 14992 (the initial model) to 9303; a reduction of about 38 % and the reason for this relatively small reduction in deviance when using the full model compared to the previous three simulated examples given in sections 5.2.1-5.2.3 is the fact that in this real data set the "unobservable" inputs are not available to us.

(3) fitting a model with only (Summation Function)$_2$ for the second "γ_b" input makes only a small reduction in deviance (a reduction of 467) whereas a model with only (Summation Function)$_1$ for the first "γ_0" input makes a larger reduction in deviance (a reduction of 1984). But a model with both functions leads to a total reduction in deviance of 2255. This suggests that the amounts of information contained in each of the summation functions are largely orthogonal to each other, in other words the two inputs are acting largely independently to excite the firing of the neurone. This seems consistent with the coherence figure shown in Fig.5.3.3a which indicates that the two observed input processes are uncorrelated with each other.

Now, suppose we use the likelihood approach to fit only the two cross-intensity functions, i.e.

$$U_t = \sum_{v=0}^{\infty} {}_1a_v^* \, {}_1x_{t-v} + \sum_{v=0}^{\infty} {}_2a_v^* \, {}_2x_{t-v} \qquad \text{.....(5.2.6)}$$

where the two sets of coefficients $\left\{{}_i a_v^*\right\}$; $i = 1, 2$ represent the cross-intensity functions between the secondary (II) output and each of the fusimotor axon inputs, γ_0 and γ_b respectively. As we explained earlier in section (5.2.1), we take into our account not only the time of the previous output spike (γ_t) as in the summation function, but we consider all previous input spikes, i.e., consider all the input postsynaptic effects at lag v as shown in Fig.4.5.1.

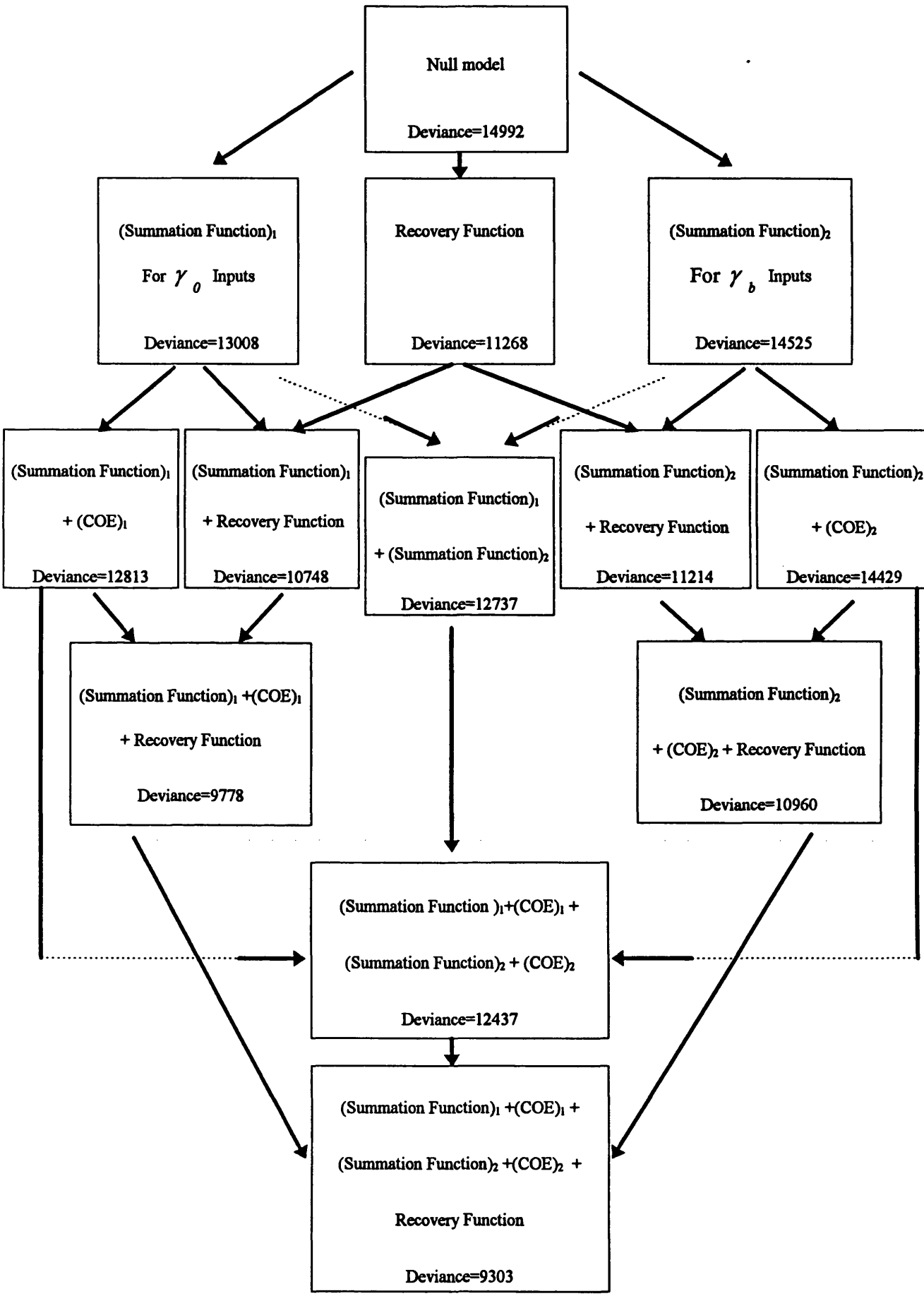


Fig.5.3.8 Diagrammatic representation of the deviance table.

Fig.5.3.9a and Fig.5.3.9b represent the estimated cross-intensity functions (estimated via the likelihood function) between the secondary (II) output and each of the fusimotor axon inputs, γ_0 and γ_b respectively. The residual deviance for this model is 14804; a reduction of only 188 (i.e., a reduction of only about 1.25 %) from the null model, providing still further evidence that the cross-intensity function in general has very poor explanatory power. The two summation functions and their corresponding carry-over effect functions seem to be much more better than the two cross intensity functions they reduce the deviance by 2255, and so they are more informative and reliable.

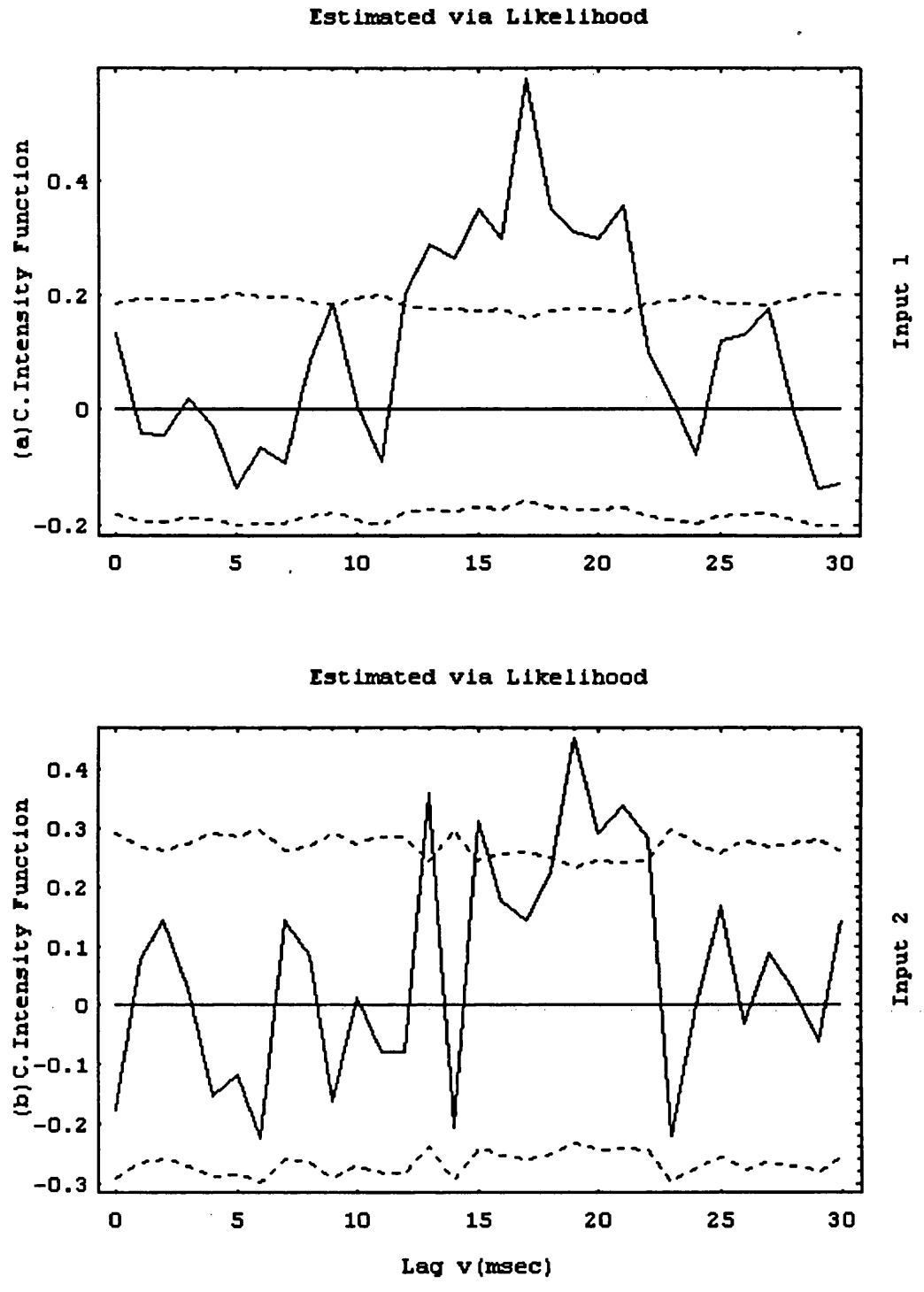


Fig.5.3.9 The two cross intensity functions estimated via likelihood.

a) Estimated cross intensity function between the secondary (II) output and the input γ_0. **b)** Estimated cross-intensity function between the secondary (II) output and the input γ_b. The horizontal dotted lines in (a) and (b) give $\pm$ two standard error limits about zero for the cross-intensity functions.

5.3.2 The Two Fusimotor Inputs and the Primary (Ia) Output

In this section we discuss the effects of the two fusimotor Inputs γ_0 and γ_b on the primary (Ia) output. As in the previous section we start first with the stochastic point process techniques, then use the likelihood technique and as in the previous examples we compare some results of the two approaches.

Fig.5.3.10a and b represent the square roots of the two estimated cross-intensity functions between the primary (Ia) output and each of the inputs γ_0 and γ_b respectively. The cross-intensity function for the fusimotor input γ_0 suggests excitatory synaptic input effects lasting from about 11 to 19 msec and the cross-intensity function between the primary (Ia) output and the fusimotor input γ_b suggests excitatory synaptic input effects lasting from about 9 to 23 msec.

Fig.5.3.11a represents the estimated ordinary coherence between the fusimotor input γ_0 and the primary (Ia) output and suggests that the two processes are weakly coupled over the range of frequencies of 0 to about 28 Hz and the estimated ordinary coherence between the other fusimotor input γ_b and the primary (Ia) output (Fig.5.3.11b) indicates also a weak coupling over the range of frequencies of 0 to about 42 Hz.

The similarity of the two estimated partial coherences between the primary (Ia) output and each fusimotor input after removing the effect of the other input (Fig.5.3.12a and Fig.5.3.12b) and those of the ordinary coherences between the primary (Ia) output and each fusimotor input (Fig.5.3.11a and Fig.5.3.11b) suggests that the two inputs are again acting almost independently in affecting the occurrence of the primary (Ia) output spikes.

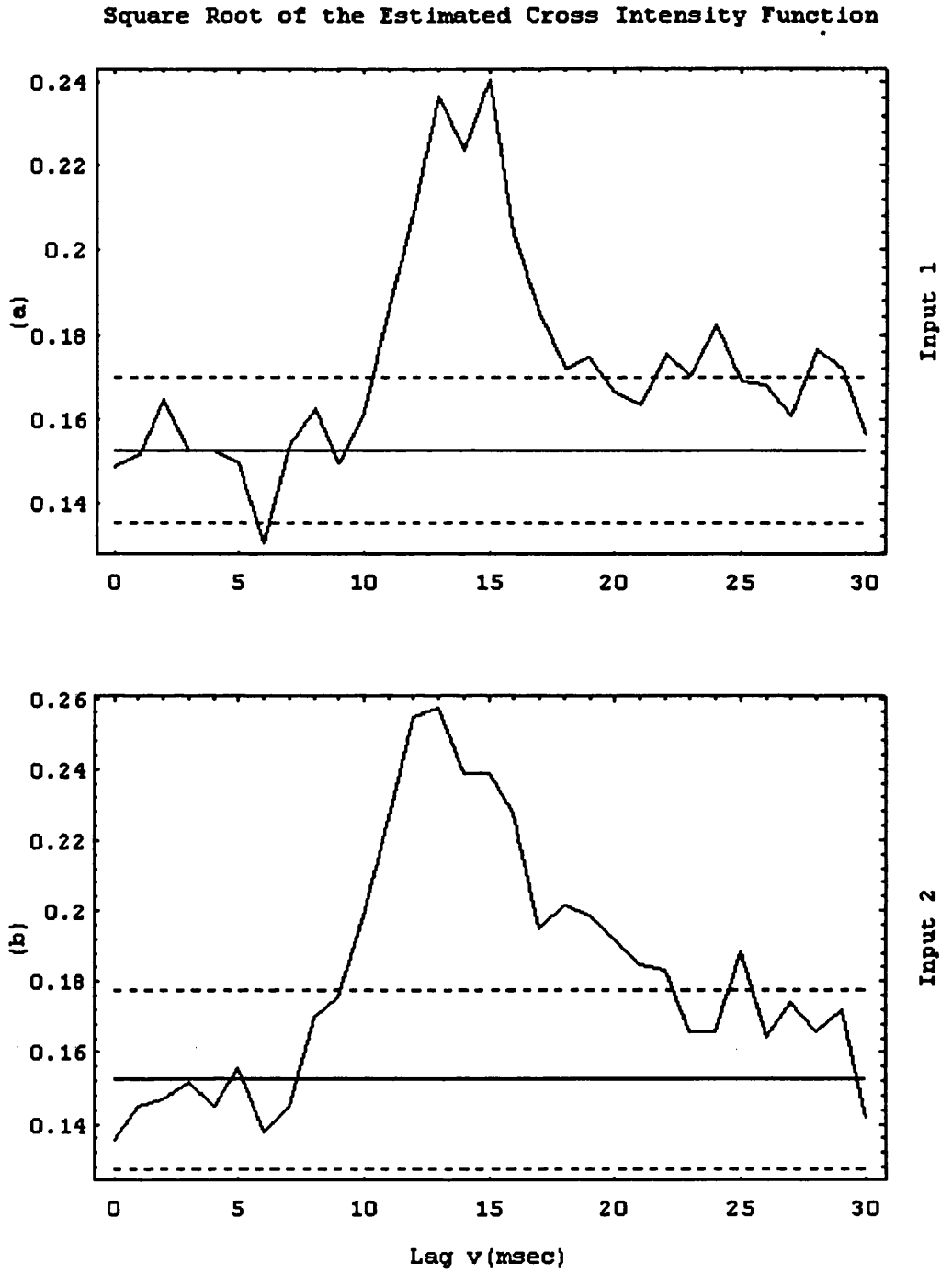


Fig.5.3.10 a) Estimated square root of the cross-intensity function between the primary (Ia) output and the input γ_0. b) Estimated square root of the cross-intensity function between the primary (Ia) output and the input γ_b. The horizontal dotted lines in (a) and (b) represent approximate 95 % confidence intervals for the estimated square roots of the cross-intensity functions under the hypothesis that the two processes are independent, plotted around the square root of the estimated output mean rate (horizontal solid line).

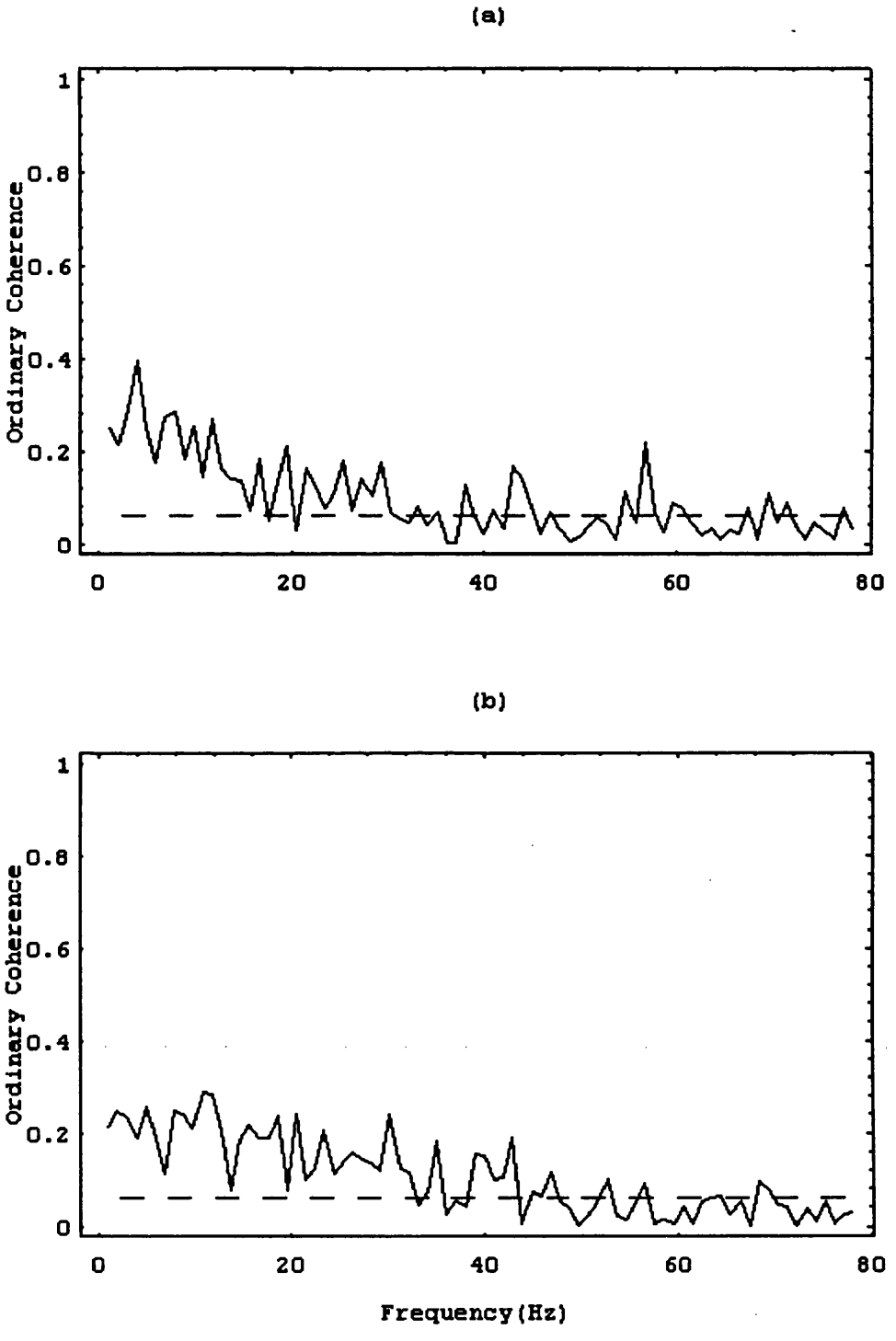


Fig.5.3.11 Estimated ordinary coherence between

a) the primary (Ia) output and the first fusimotor input γ_0 and **b)** the primary (Ia) output and the second fusimotor input γ_b. The dotted lines correspond to the upper limit of the 95% confidence intervals for the coherence under the hypothesis of zero coherence.

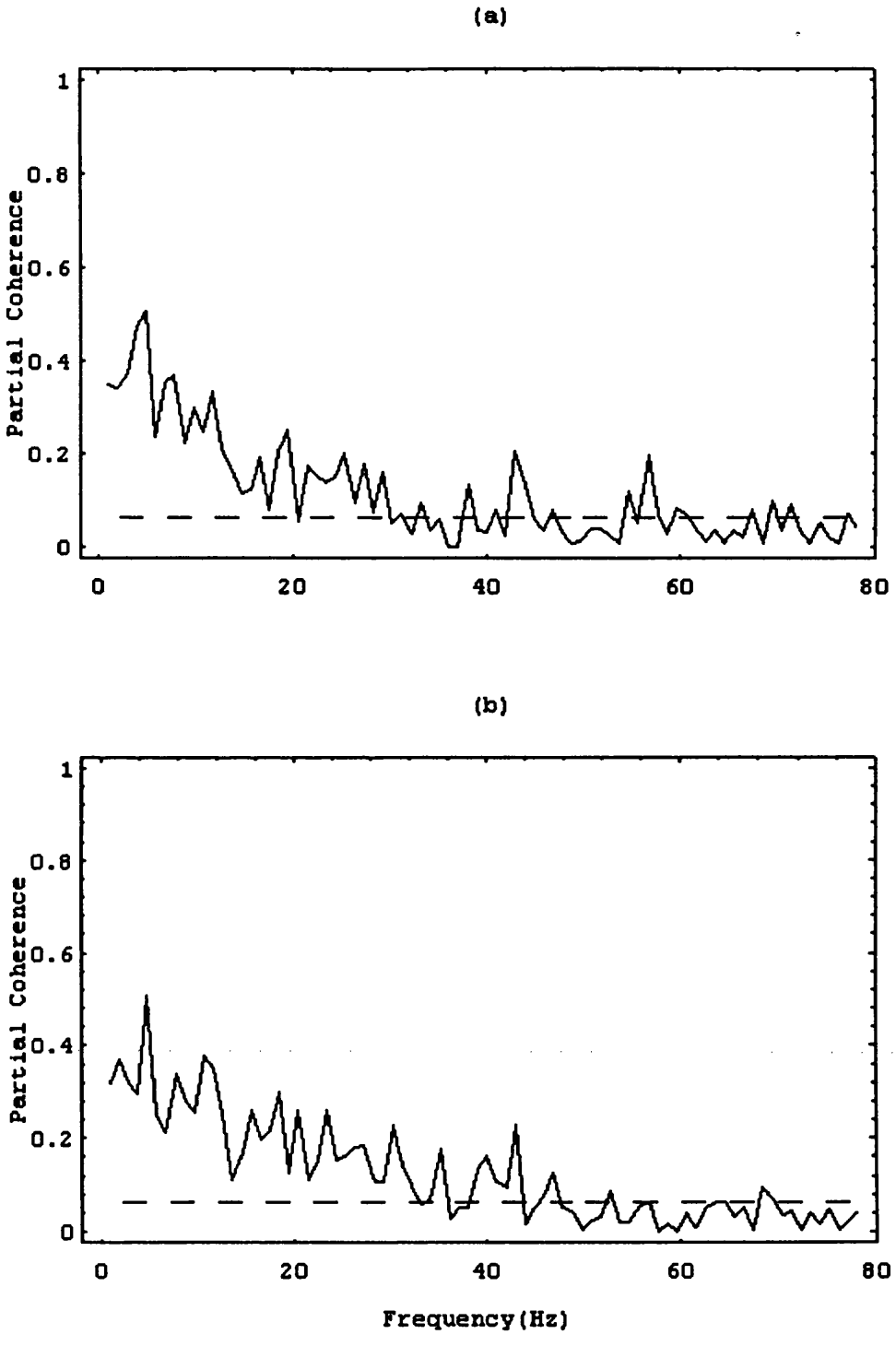


Fig.5.3.12 **a)** Estimated partial coherence between the primary (Ia) output and the first fusimotor input γ_0 after removing the effect of the second input γ_b. **b)** Estimated partial coherence between the primary (Ia) output and the second fusimotor input γ_b after removing the effect of the first input γ_0. The dotted lines correspond to the upper limit of the 95% confidence intervals for the partial coherence under the hypothesis of zero partial coherence.

Fig.5.3.13a represents the estimated partial phase between the first observed fusimotor axon input, γ_0, and the primary (Ia) output after removing the linear effects of the second observed fusimotor axon input, γ_b. The weighted least squares line (dotted) fitted to the partial phase curve over the range of frequencies at which the corresponding partial coherence (Fig.5.3.12a) is significantly different from zero shows that, the output process is delayed, on the average, by an amount 14.70 msec with a 95% confidence interval, for the delay, of (14.09, 15.30) msec. The value of the delay suggested by the partial phase again seems to be consistent with the 15 msec peak in the corresponding square root of the estimated cross intensity function (Fig.5.3.10a).

Fig.5.3.13b represents the estimated partial phase between the second observed fusimotor axon input, γ_b, and the primary (Ia) output after removing the linear effects of the first input. The estimated slope of the weighted least squares line (dotted) over the range of frequencies at which the corresponding partial coherence (Fig.5.3.12b) is significantly different from zero shows that, the output process is delayed, on the average, by an amount 14.45 msec with a 95% confidence interval, for the delay, of (13.89, 15.00) msec. The value of the delay suggested by the partial phase again seems to be consistent with the 13 msec peak in the corresponding square root of the estimated cross intensity function (Fig.5.3.10b).

As in the previous section, the peaks in the two summation functions are centred about 16 msec and 17 msec for the first (γ_0) and second (γ_b) inputs respectively. These differ from the estimated times of the peaks in the corresponding cross intensity functions and from the two values of the delay estimated from the partial phase. This is again might be due to the fact that the data set is not input dominated (as can be seen from the deviance table

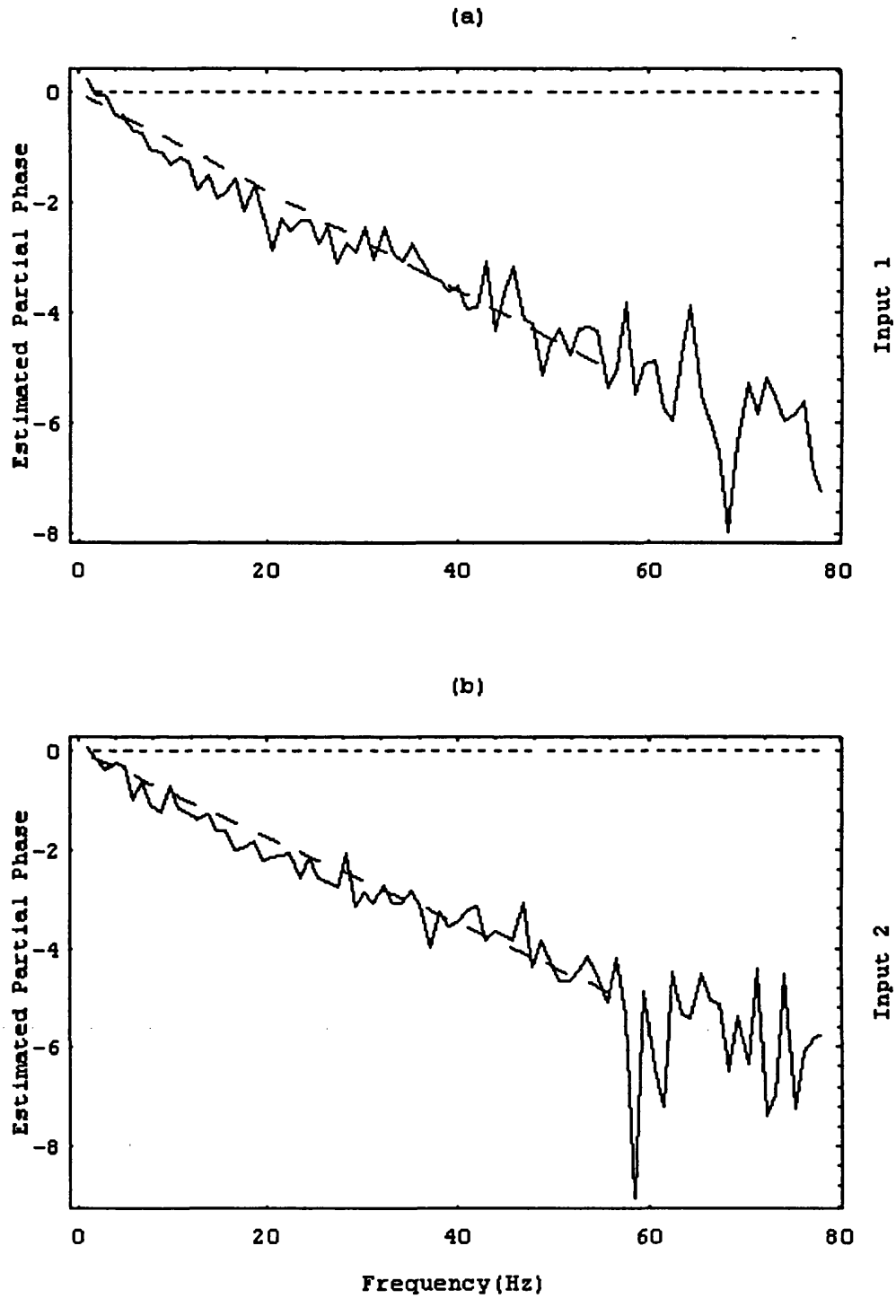


Fig.5.3.13 Estimated partial phase between the primary (Ia) output and

a) the first input γ_0 after removing the effect of the second input γ_b and **b)** the second input γ_b after removing the effect of the first input γ_0. The slopes of the fitted weighted least squares lines (dotted) correspond to the estimated time delays.

given in Fig.5.3.16). The lack of agreement between the summation function and the point process techniques of detecting similar values for the delay suggests again that the delay may not be the most relevant description. The situation seems to be more complicated than what is suggested by the partial phase and the cross intensity function.

We now turn our attention to the applications of the likelihood method to the same set of muscle spindle real data where the structure of the likelihood model is the similar as previously described in section (5.3.1). The logistic link function was again used since it found to be superior to the probit link in terms of the goodness of fit, in taking the range of the linear predictor a little further and also from the greater reduction in deviance.

Fig.5.3.14a represents the estimated (Summation Function)$_1$, $\{{}_1\hat{a}_u\}$, for the first fusimotor input γ_a and suggests an excitatory effect lasting from about 11 to 29 msec whereas the estimated (Summation Function)$_2$, $\{{}_2\hat{a}_u\}$, for the second fusimotor input γ_b (Fig.5.3.14b) suggests an excitatory effect lasting from about 9 to 25 msec.

The two summation functions for the first and second fusimotor inputs (Fig.5.3.14a and b) have a longer significant duration than their corresponding cross intensity functions (Fig.5.3.10a and b). As in the previous example, the difference in duration suggests again that the square root of the cross-intensity function underestimates the synaptic input effects for each of the fusimotor inputs. (as we will see shortly from the residual deviance for the cross intensity functions when estimated using the likelihood approach).

The threshold and recovery functions given in Fig.5.3.14c suggest that the probability of an output spike occurring spontaneously is very small over the whole range of intervals after the previous output spike as the recovery function remains far below threshold over the whole range of intervals.

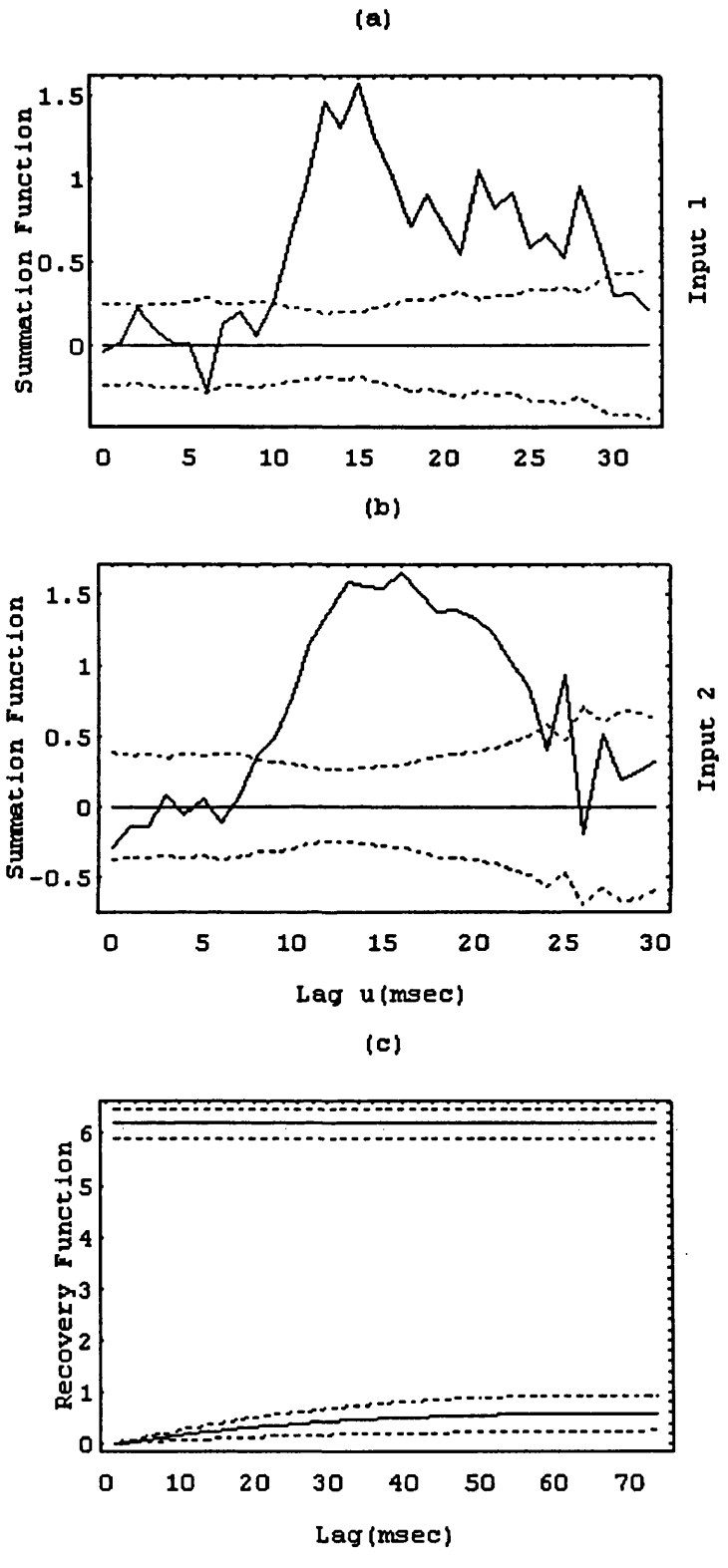


Fig.5.3.14 **a)** Estimated summation function ${}_1\hat{a}_u$ for the first input γ_0. **b)** Estimated summation function ${}_2\hat{a}_u$ for the second observed input γ_b. **c)** Estimated threshold and recovery functions. The horizontal dotted lines give $\pm$ two standard error limits plotted about zero for the summation functions in a) and b) and around each function in c).

Fig.5.3.15a and b represent the carry-over effect functions for the first and second fusimotor inputs respectively. The estimated carry-over effect function $(COE)_1$ for the first fusimotor input suggests excitatory effects lasting from about 12 to 30 msec whereas the estimated carry-over effect function $(COE)_2$ for the second fusimotor input suggests excitatory effects lasting from about 11 to 33 msec. But both carry-over effects are relatively small compared to their corresponding summation functions.

The goodness of fit plot given in Fig.5.3.15c seems reasonable but not as good as that given in Fig.5.3.7c when analysing the effects of the two fusimotor inputs on the secondary (II) output, in the sense that the range of the linear predictor is shorter and the confidence intervals are wider and also from the smaller reduction in deviance. The model, in this example, takes the range of the linear predictor up to around -1.1. The model cannot predict high probabilities of an output which reflects the large unexplained variability in the model.

The deviance table given in Fig.5.3.16 illustrates the sequential fitting of a set of successively more complex models and reveals the following interesting features,

(1) the recovery and threshold functions taken together explain very little of the variability (about 5.1 %). The two summation functions fitted together explain about 11.12 % of the variability whereas fitting them separately explains about 6.64 % for the first input and about 5.2 % for the second input. These small reductions in the deviance give the impression that the output spikes are not produced only by the influence of the two observed inputs or by an intrinsic cell property, but produced mainly by some other unmeasured inputs.

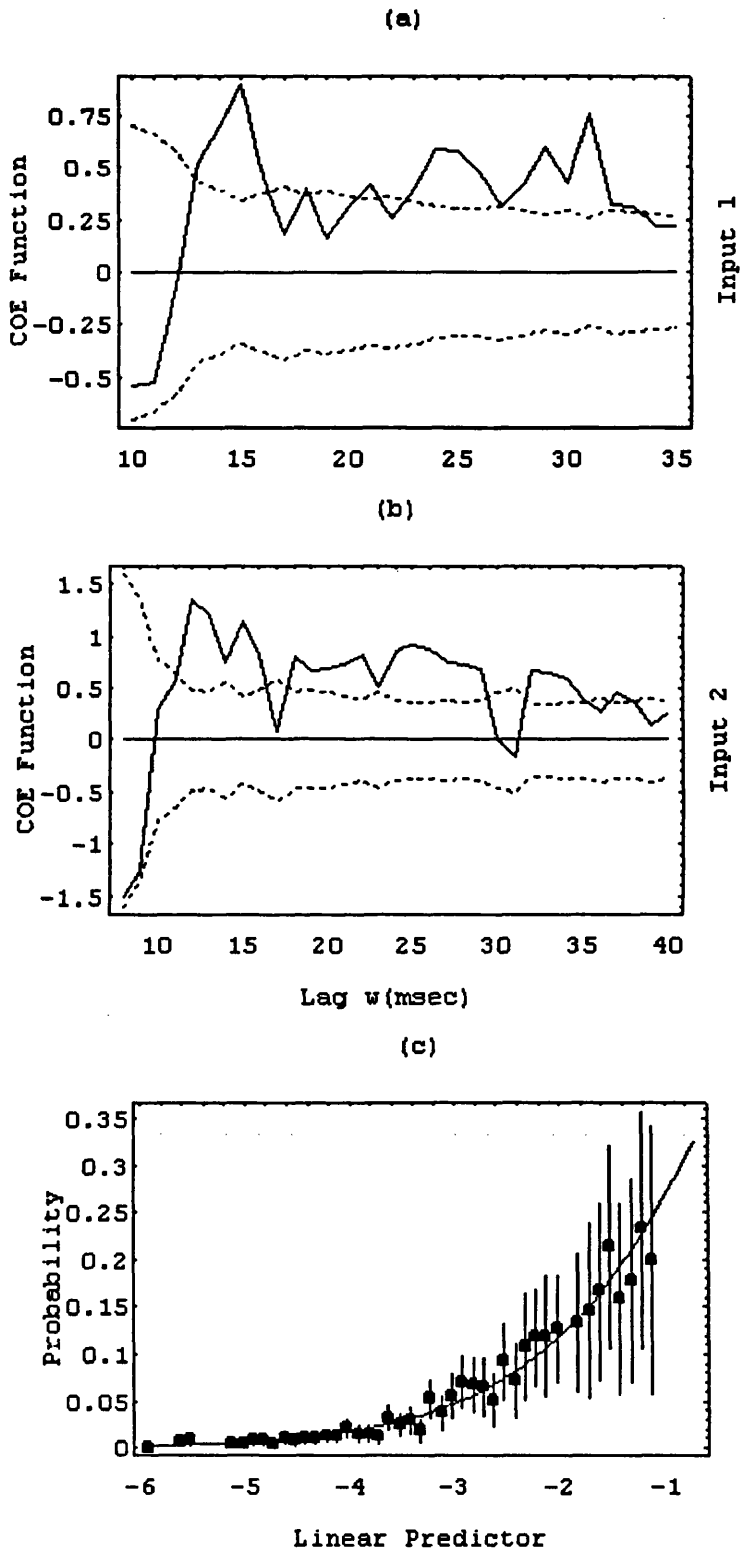


Fig.5.3.15 illustration of the two carry-over effect functions and goodness of fit plot.

a) Estimated carry-over effect function ${}_1\hat{c}_w$ for the first fusimotor input γ_0. **b)** Estimated carry-over effect function ${}_2\hat{c}_w$ for the second fusimotor input γ_b. The horizontal dotted lines in (a) and (b) give $\pm$ two standard error limits plotted about zero for the summation functions **c)** The goodness of fit plot.

(2) fitting all five components reduces the deviance from 11026 (the initial model) to 9249; a reduction of only about 16.12 % and the reason for this small reduction in deviance when using the full model, as in the previous section, is the fact that the "unobservable" inputs are not included since they are not available to us in this real experimental situation.

(3) as we have seen in (1) above, fitting a model with both summation functions leads to a reduction in deviance of 1226 which is very similar to the sum of the reductions in deviance when each summation function is fitted separately (a reduction of 732 and 570 for the first and second input summation functions respectively). This suggests that the amounts of information contained in each of the summation functions are largely orthogonal to each other, i.e., the two inputs act almost independently to excite the firing of the neurone. This seems consistent with the coherence and phase figures we have seen earlier.

(4) in this example both carry-over effect functions seem to contain very little information compared to their corresponding summation functions as can be seen from a small reduction in deviance every time a carry-over effect function is added to a previous model. But the adding of the carry-over effect functions was verified using F-tests.

Suppose we use the likelihood approach, as explained in the previous section, to fit a model with only the two cross-intensity functions.

Fig.5.3.17a and Fig.5.3.17b represent the estimated cross-intensity functions (estimated via the likelihood function) between the primary (Ia) output and each of the fusimotor axon inputs, γ_0 and γ_b respectively. The residual deviance for this model is 9944; a reduction of 1082 (i.e., a reduction of about 9.81 %) from the null model, providing still further evidence that the cross-intensity function in general has poor explanatory power, although in this

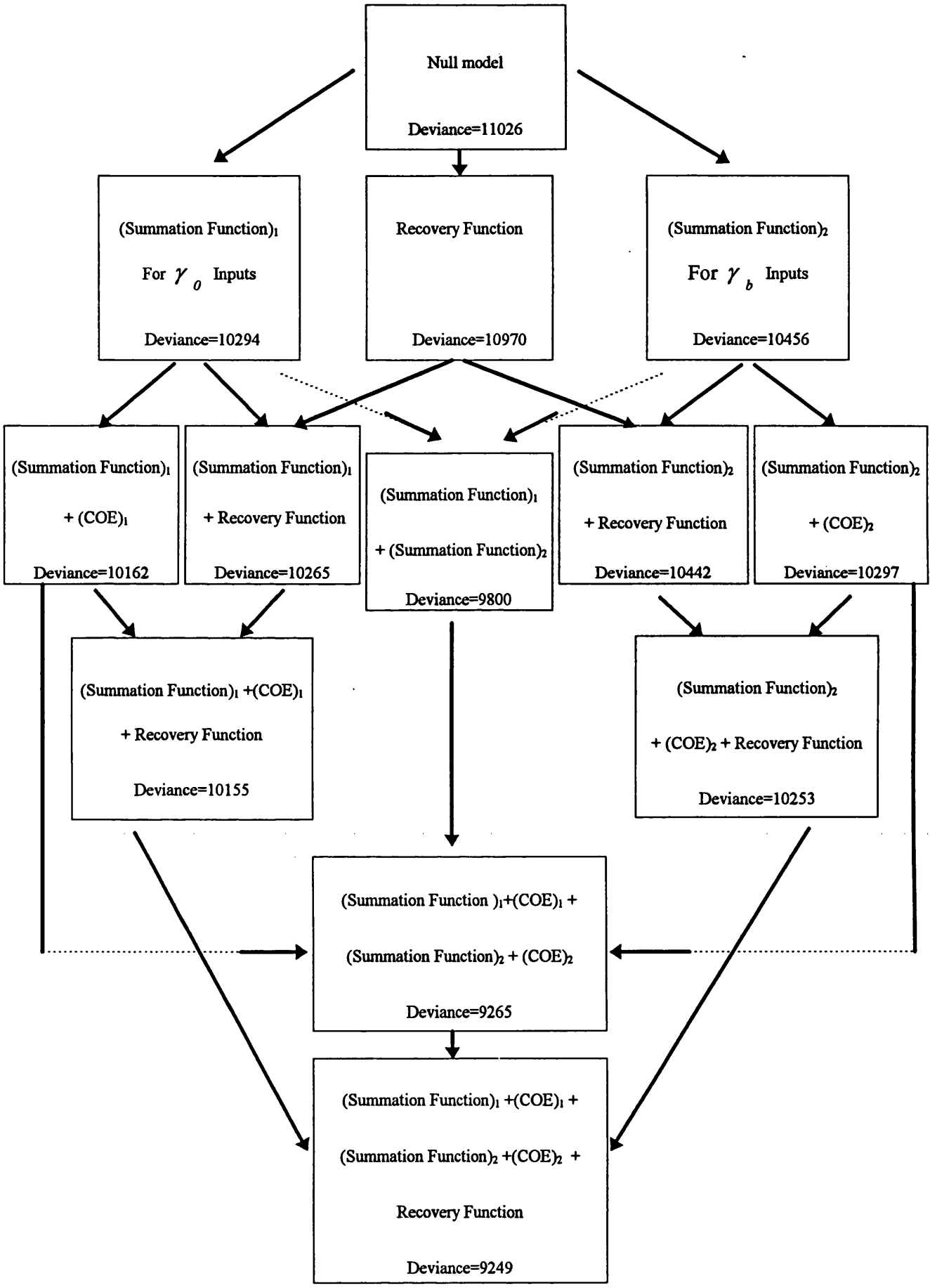


Fig.5.3.16 Diagrammatic representation of the deviance table.

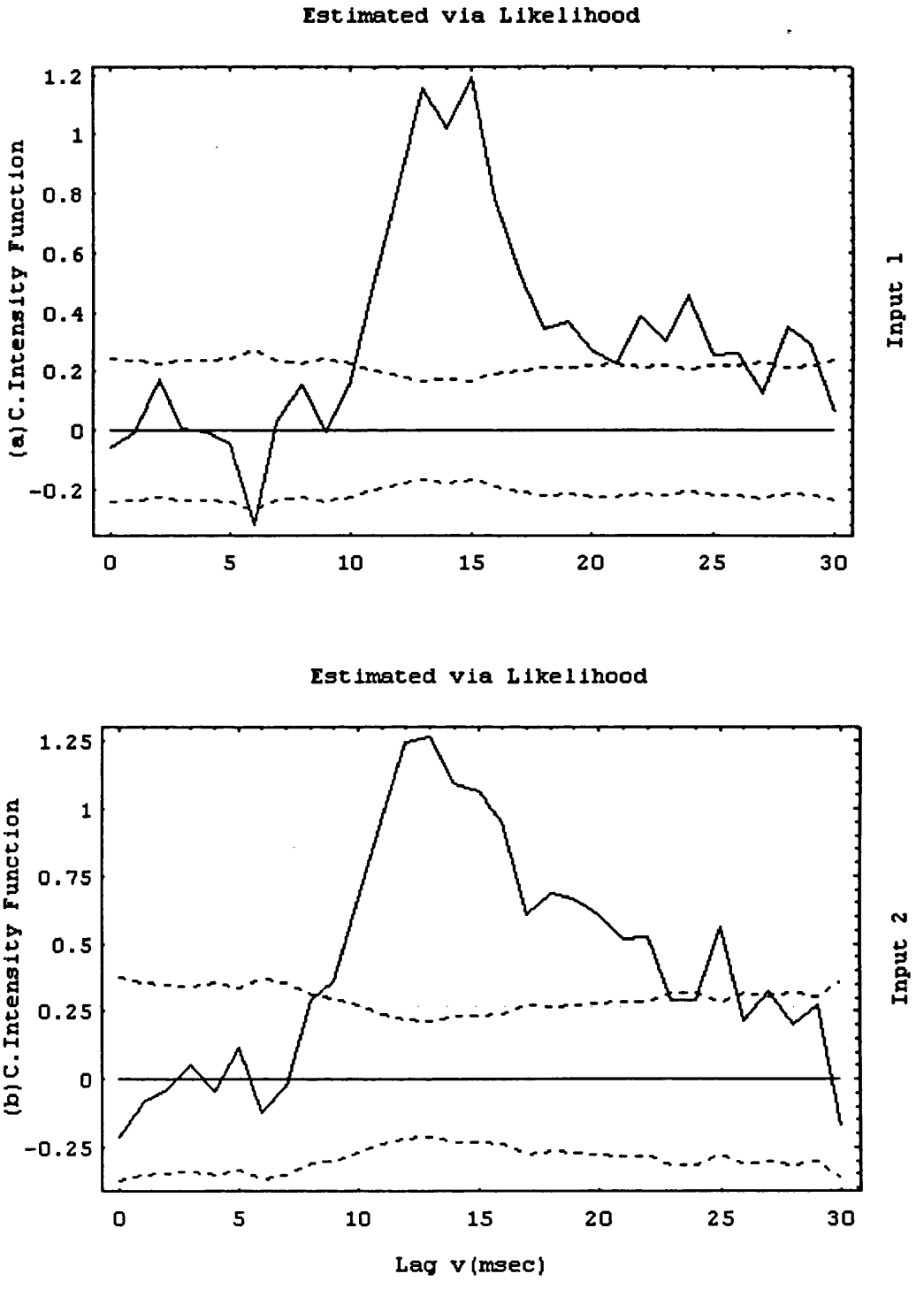


Fig.5.3.17 The two cross intensity functions estimated via likelihood.

a) Estimated cross intensity function between the primary (Ia) output and the input γ_0. **b)** Estimated cross-intensity function between the primary (Ia) output and the input γ_b. The horizontal dotted lines in (a) and (b) give $\pm$ two standard error limits about zero for the cross-intensity functions.

example, the two cross-intensity functions are only slightly less informative than the two summation functions (which reduce the deviance by 1226).

We now use the frequency domain technique to discuss the effects of the two fusimotor inputs on the coupling between the two outputs. The reason of presenting this analysis at the end of this chapter is due to the fact that the likelihood method has not been yet developed to the case of multiple output (see chapter 6).

Fig.5.3.18a represents the estimated ordinary coherence between the primary (Ia) and secondary (II) outputs and suggests that the two outputs are coupled over the range of frequencies of about 0 to 18 Hz. Fig.5.3.18b and c represent the estimated partial coherences between the primary (Ia) and secondary (II) outputs after removing the effect of each fusimotor input. They suggest that removing the effect of the first fusimotor input γ_0 (Fig.5.3.18b) significantly weakens the coupling between the two outputs whereas the partial coherence between the two outputs after removing the effect of the second fusimotor input γ_b (Fig.5.3.18b) remains almost identical to the ordinary coherence between the two outputs (Fig.5.3.18a). This weakening in coupling between the two outputs when removing the effect of the first fusimotor input γ_0 may be due to the fact that the coupling between each of the two outputs and the first fusimotor input γ_0 is stronger (as given in Fig.5.3.3b and Fig.5.3.11a) than that between each of the two outputs and the second fusimotor input γ_b (as given in Fig.5.3.3c and Fig.5.3.11b). A similar conclusion can also be drawn from the two deviance tables.

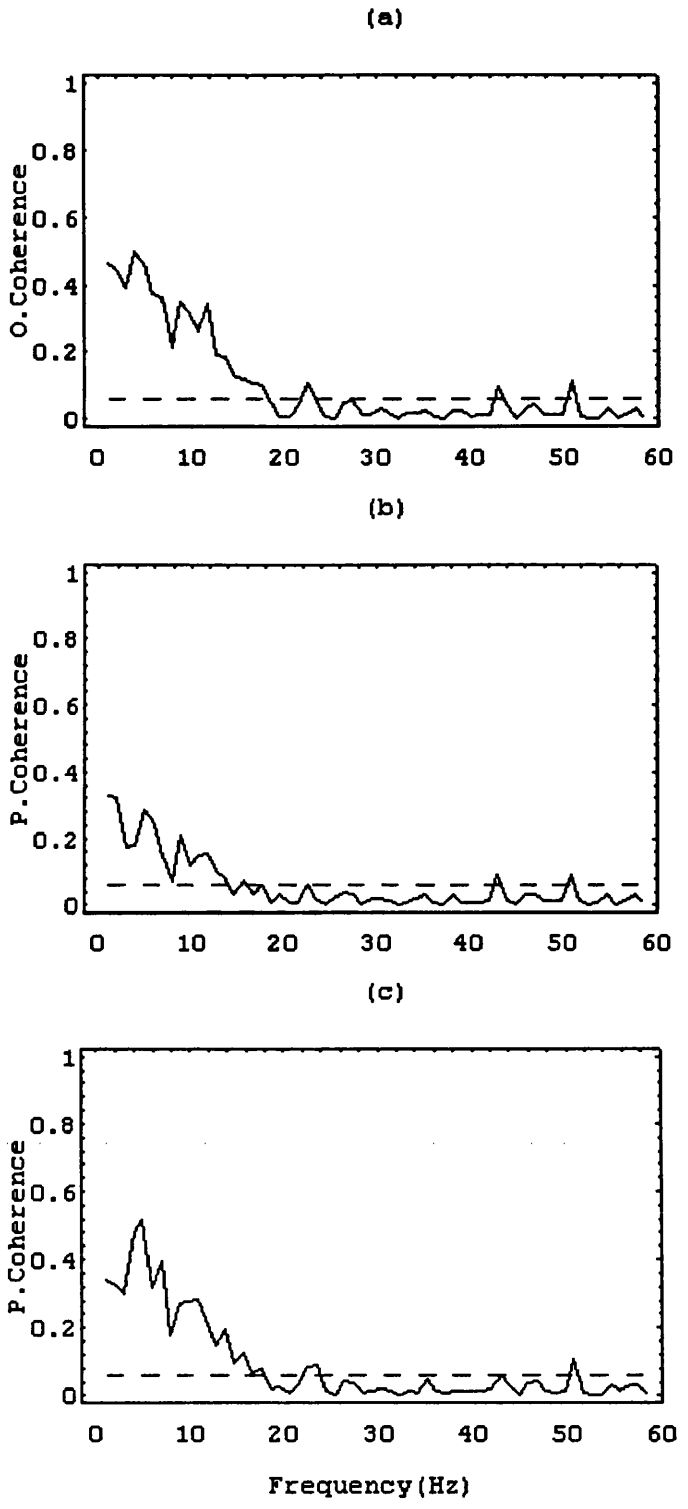


Fig.5.3.18 a) Estimated ordinary coherence between the primary (Ia) and the secondary (II) outputs.

Estimated partial coherence between the primary (Ia) and the secondary (II) outputs, b) after removing the effect of the input γ_0 and c) after removing the effect of the input γ_b. The dotted lines correspond to the upper limit of the 95% confidence intervals for the coherence under the hypothesis of zero coherence.

5.4 Conclusion

In this chapter the maximum likelihood approach has been applied to three simulated sets of data and also to a real neuronal spike train data set. There are many interesting features to which we draw attention. First is the analysis of the first set of data with one observed (spike train) input and one "unobservable" input, and one observed output where summation and carry-over effect functions for the "unobservable" inputs can be estimated. In this case we have noticed that the "unobservable" inputs explained more of the variability than either the summation function or the recovery function (for the observed inputs) when fitted alone. The deviance table in the latter case shows that the recovery function contains almost no extra information to that contained in the "unobservable" inputs which gives the impression that the recovery function "explains" part of the effects of the unmeasured inputs. There is an evident warning not to try to give the recovery function a physiological interpretation in any circumstance where there are unmeasured inputs. The summation function by contrast seem to have clear physiological interpretations in every case. Second is the analysis of two simulated sets of data each with two observed (spike train) inputs and one "unobservable" input, and one observed output where the two observed inputs are uncorrelated with each other in the first set of data and correlated in the second set of data. The demonstrations shows again that the likelihood is a very flexible approach and has the ability to separate the aspects of the relationship between spike trains through the threshold, recovery, summation (and carry over effect) functions for each of the two observed inputs and for the "unobservable" input and no such ability is provided by the traditional stochastic point process techniques. Also the analyses suggest that the square root of the cross intensity function is again a poor method of investigating the association between processes since it is regularly underestimating the underlying excitatory effects of a synaptic

input. The reduction in deviance for a model with only the cross intensity function (when estimated via likelihood) from the null model is very small, providing still further evidence that the cross-intensity function in general has very poor explanatory power. The third point to notice here is that the likelihood approach suggests similar features, i.e., whether the inputs act dependently or independently to affect the output spikes, to those revealed by the ordinary and partial coherences as we have seen in sections 5.2 and 5.3. Finally we have the case of real muscle spindle data where we investigated the relationships between the two fusimotor γ_0 and γ_b inputs and each of the sensory Ia and II outputs. Again similar features to those mentioned in the two simulated cases have been noticed here except in this real experiment the "unobservable" inputs cannot be taken into consideration. The likelihood approach still needs to be developed to include the case of two or more outputs.

Chapter 6

6. General Conclusions and Further Work

6.1 General Conclusions

In general terms, the work described in this thesis shows that the maximum likelihood approach is a very powerful estimation procedure which can be used effectively to analyse neuronal spike train data. Throughout the course of this thesis, the likelihood approach has been applied to many simulated as well as real data sets, and very useful results have been obtained. The main interesting features of these results to which we draw attention can be summarised as follows:

- The approach shows great flexibility. The recovery (and threshold) functions, when all inputs can be measured, represent intrinsic properties of the neurone and no analogous measure is available using the traditional stochastic point process techniques. However, the parameters in the recovery and threshold functions are not physiologically meaningful parameters and we cannot give them any direct interpretation. The meaningful feature in fact is the difference between

the estimated values of these two functions (i.e., the rate at which the recovery function approaches threshold is the significant and physiologically meaningful thing we are looking at).

- The linear summation of the effects of the input spike train on the membrane potential have been further separated into the effects of input spikes occurring at times after the time of the last output spike, i.e., at lag u as shown in Fig.4.5.1 and the effects of input spikes occurring at times prior to the time of the previous output spike, i.e., at lag w as shown in Fig.4.5.1. These two types of input effect are measured by the summation function and its corresponding carry-over effect function respectively. Again the time and frequency domain analyses do not provide analogous measures of the carry-over effect of the synaptic inputs. Unlike the parameters in the threshold and recovery functions, the estimated durations of the summation function (or the carry-over effect) are interpretable and physiologically meaningful.
- The analyses of most data sets we have looked at during the preparation of this thesis, suggest that the square root of the cross intensity function as a time domain measure of the degree of associations between two processes usually underestimates the underlying excitatory and inhibitory effects of a synaptic input. In some cases it may further produce results that contradict the way in which the data have been simulated, e.g., it shows an inhibition or periodic behaviour where there are no such features present in the simulation. The summation function, by contrast, provides an alternative measure which seems to be more informative and reliable in terms of reduction in deviance. Also it seems to be more consistent with the way in which the data are simulated, i.e., it reflects features which are present in the simulation. The only case where the cross intensity function still could

be a useful measure is the case where the effects on the neurone cell are dominated by a single input with no significant carry-over effect and also if there is a negligible recovery function as in chapter 5.

- Unlike the cross-intensity approach, the likelihood procedure also allows for continuous "unobservable" inputs to be involved in the analyses, and both summation and carry-over effect functions for the "unobservable" inputs can be estimated.
- In the case where "unobservable" inputs are taken into consideration, the deviance table sometimes shows that the recovery function contains almost no extra information to that contained in the "unobservable" inputs. This gives the impression that the recovery function may explain both intrinsic properties of the neurone as well as part of the effects of unmeasured inputs if these latter cannot be modelled. This will invariably be the case for real data.
- The likelihood model can be extended to the case of two observed inputs and one observed output and where "unobservable" inputs can also be taken into consideration. This gives the approach some more explanatory power.
- The likelihood approach through the table of deviances, in most data sets we have looked at in this thesis, suggests similar features, i.e., whether the inputs act dependently or independently to affect the output spikes, to those revealed by the ordinary and partial coherences.
- The simulations have provided useful insights into the interpretation of experimental results, and advances in the physiological field emphasise the development of the simulation to match experimental data as much as possible.
- Although it is a very computationally intensive and time consuming procedure compared to both the cross intensity and Fourier approaches

(which are equally rapid in terms of computational time), the likelihood approach shows numerous advantages which make it a very powerful tool for analysing neuronal spike train data where a considerable amount of extra information about the processes involved can be obtained.

- The value of the delay suggested by the phase seems usually consistent with the peak in the square root of the estimated cross intensity function. This is expected as the two approaches are mathematically equivalent and therefore reveal similar results. The location of the peak in the summation function is usually larger than the location of the peak in the cross intensity function and the value of the delay estimated from the phase (unless the summation and cross intensity functions are not very different in those circumstances mentioned earlier). This lack of agreement between the likelihood and the point process techniques may be due to the fact that some data sets are not input dominated and the delay may not be the dominant feature or even a relevant one. In this case the situation is more complicated than what is suggested by the phase and the cross intensity function, and it seems that summarising this in a single number and calling it a "delay" may not always be appropriate.
- The recently proposed maximum likelihood approach to the analysis of neuronal spike train data by Brillinger, see Brillinger and Segundo (1979) and Brillinger (1988 and 1992), is limited to a constant threshold, used only a probit link function, and only fitted recovery (and threshold) and summation functions in the absence of unmeasured inputs. Also the procedure was applied only to a single real data set. However, in this thesis we not only extended this previous work but also developed a variety of functional forms for the summation and recovery processes, considering different link functions. We fit a model with both functions

together in the presence and absence of a continuous "unobservable" input representing any unmeasured input. This enables us, as mentioned earlier, to recognise the cases in which the recovery function is purely reflecting the intrinsic properties of the neurone. Our main contributions in this present work are not only the application of this approach to as many as 50 different simulated as well as real data sets and comparing results obtained via likelihood with those obtained via stochastic point process techniques and the extension and the development of existing work, but also we introduced a variety of new functional forms of the likelihood model, such as:

a) The carry-over effect function which is seen to provide a measure of any carry-over effects that occur after an output spike.

b) Constructing a table of deviances through which the improvement of the hierarchy of forms of the model can be assessed at each stage of complexity. A significant reduction in deviance in proceeding to higher levels of complexity reflects an improvement in the model. Also formal F-tests based on the change in the deviance and the degrees of freedom, to justify the adding of new terms to our model, can easily be derived from the deviance tables. Again no such deviance tables can be obtained using the traditional stochastic point process techniques.

c) Considering two types of threshold, the constant and the exponentially decaying threshold forms. The choice between the two types of threshold forms can be assessed by their reduction of deviance as well as the reduction of the number of parameters used in the corresponding recovery function.

d) The estimate of the cross intensity function via the likelihood approach where its adequacy as a representation of the underlying

processes is assessed in terms of a reduction of deviance. The very small reduction in deviance for a model with only the cross intensity function from the null model, compared to that for the summation function, provides still further evidence that the cross-intensity function in general has very poor explanatory power.

6.2 Further Work

There are many situations and related questions which have arisen during the course of the present work which are not fully understood and still require further investigation. The development and application of likelihood estimation procedures to more realistic situations has led to areas of possible future work. The following list sets out some possible areas in which the work of this thesis may be extended:

◊ The application of the likelihood model should not be limited to a single input - single output and two inputs - single output cases but it should be extended further to a multiple input - multiple output cases.

◊ The likelihood model should be extended to include, along with linear input effects, quadratic effects of synaptic inputs which will represent interactive effects within and between individual input processes. The interactive effects could be estimated within each individual input by a likelihood model of the form (6.2.1) and that between different inputs by a model of the form (6.2.2). Assuming that the membrane potential at the trigger zone of a neurone at any given time t is denoted by U_t, then the two suggested likelihood models are of the form

$$U_t = V_t + \sum_{u=0}^{\gamma_t - 1} a_u x_{t-u} + \sum_{u=0}^{\gamma_t - 1} \sum_{u'=0}^{\gamma_t - 1} a^*_{u,u'} x_{t-u} x_{t-u'} \quad ; u \neq u' \qquad \text{.....(6.2.1)}$$

and

$$U_t = V_t + \sum_{u=0}^{\gamma_t - 1} {}_1a_u \, {}_1x_{t-u} + \sum_{u=0}^{\gamma_t - 1} {}_2a_u \, {}_2x_{t-u} + \sum_{u=0}^{\gamma_t - 1} \sum_{u'=0}^{\gamma_t - 1} a^*_{u,u'} \, {}_1x_{t-u} \, {}_2x_{t-u'} \quad ; u \neq u' \qquad \text{.....(6.2.2)}$$

where γ_t is the time elapsed at time t since the time of the last output spike, the term V_t represents the polynomial recovery function of order k, the two sets of coefficients $\{a_u\}$, $\{a^*_{u,u'}\}$ given in (6.2.1) represent the summation function and the quadratic effects within the input, respectively, whereas the sets of coefficients $\{_1a_u\}$, $\{_2a_u\}$ given in (6.2.2) represent the summation functions for the first and second inputs respectively and the set of coefficients $\{a^*_{u,u'}\}$ given in (6.2.2) represents the quadratic effects between the two inputs.

◊ In chapter 4 we have introduced a goodness of fit assessment procedure for testing the validity of the model and the adequacy of the link function. But this procedure is partially based on visual comparisons between empirical and corresponding theoretical probabilities and is to some extent an informal procedure. A formal test would be very useful to have.

◊ The application of the likelihood approach to the mammalian muscle spindle data discussed in chapter 5 (section 5.3) discussed the relationships between the two fusimotor γ_0 and γ_b inputs and each of the sensory Ia and II outputs. This investigation could be extended to include the case of considering both outputs together or it could be extended to include the effect of a length change of the muscle spindle under different conditions of other stimuli, on the sensory discharges from the same muscle spindle. This length change of the muscle spindle, as a continuous input to the cell, could be included in our model in the same way the unmeasured inputs have been included.

Appendices

Appendix (A)

Simulation Description

The reconstruction of a single neurone is represented by conductance based neurone simulations (Getting, 1989 and recently Halliday, 1994) by assuming an ionic transmembrane current to flow through channels with a linear instantaneous current voltage obeying Ohm's law (Hille, 1984). Fig.1 shows the equivalent electrical circuit used to represent a single neurone.

For each cell, the intracellular membrane potential is described by

$$C_m \frac{dV_m}{dt} = -I_{leak}(V_m) - \sum_{j=1}^{n} I^j_{syn}(V_m, t) - \sum_{i=1}^{k} I^i_{ahp}(V_m, t) - I_{ext}(t) \quad \ldots\ldots(1)$$

where V_m represents the membrane potential at time t, C_m is the cell capacitance and R_m the cell input resistance. $I_{leak}(V_m)$ is the passive leakage current, $I^j_{syn}(V_m,t)$ is

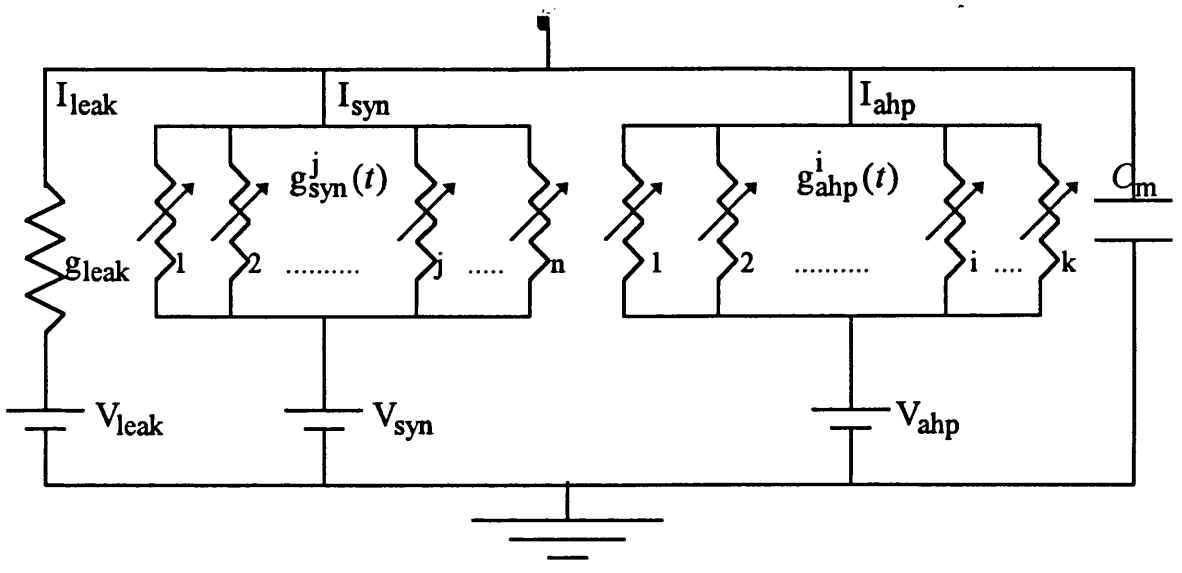


Fig.1 Equivalent electrical circuit for a single neurone, including the time-dependent conductance changes $g^{j}_{syn}(t)$ and $g^{i}_{ahp}(t)$; j=1,2,...,n and i=1,2,...,k and a single leakage conductance, g_{leak}, in series with a constant battery (V_{leak}) through which the passive ionic current flows. The cell capacitance is represented by C_m.

current due to the j^{th} pre-synaptic spike, with the summation over the total number of pre-synaptic spikes, *n*. The afterhyperpolarization (AHP) current due to the i^{th} post-synaptic spike is $I^{i}_{ahp}(V_m,t)$, with the summation over the total number of post-synaptic spikes, *k*. $I_{ext}(t)$ is a time dependent external current applied to the cell which is used to simulate a population of unobserved inputs responsible for spontaneous background firing. In practice this is achieved by using a non-zero mean normal distribution to simulate synaptic noise (Lüscher, 1990).

The cell leakage current is estimated as $I_{leak}(V_m) = (V_m - V_r)/R_m$, where V_r is the resting value of the membrane threshold. The synaptic current at time t due to a single pre-synaptic spike at time 0 is estimated as $I_{syn}(V_m,t) = g_{syn}(t)\ (V_m - V_{syn})$, where $g_{syn}(t)$ is the time dependent conductance change associated with the opening of the ionic channels and V_{syn} is the equilibrium potential for the ionic current. The AHP current due to a single post-synaptic spike at time 0 is estimated as

$I_{ahp}(V_m,t) = g_{ahp}(t)\ (V_m - V_{ahp})$, where $g_{ahp}(t)$ is the time dependent conductance change associated with the opening of the ionic channels and V_{ahp} is the equilibrium potential for the ionic current. Each pre-synaptic input spike activates one extra term in the synaptic summation in equation (1), which lasts for the duration of the particular synaptic time dependent conductance change $g_{syn}(t)$ for that input. Similarly, each post-synaptic spike activates one extra term in the afterhyperpolarization summation in equation (1), which lasts for the duration of the particular afterhyperpolarization time dependent conductance change $g_{ahp}(t)$ for that cell.

The voltage, V_m, is compared with the threshold voltage, V_{th}, at each time step to determine if an action potential has occurred. A time varying threshold is incorporated into the simulation; this allows point neurone simulations to duplicate a wide range of repetitive firing characteristics (Getting, 1989). The threshold is specified by three variables, the asymptotic level, θ_∞, the level to which the threshold is elevated after each output spike, θ_0, and the decay time constant, τ_θ, through which the threshold decays to the asymptotic level.

A further simplification used in some simulations is the representation of a neurone as an integrate-and-fire device, where the membrane threshold is reset to the resting value, V_r, after each output spike. In this case terms involving $I_{ahp}(V_m,t)$ are omitted from equation (1).

The selection of simulation parameters is done in sequence, and at each stage parameters are selected so that the behaviour of the simulation matches experimental observations for the type of cell being simulated. Firstly passive parameters are selected, cell membrane (input resistance), R_m, and time constant, τ_m, are chosen, where $\tau_m = R_m\ C_m$. This determines the cell capacitance, C_m. Secondly the cell resting potential, V_r, and threshold parameters, θ_∞, θ_0 and τ_θ, are chosen. These determine the rheobase current required for repetitive firing of the cell. If AHP currents are included, the time course of the AHP can be adjusted

under constant current stimulus by altering the conductance function $g_{ahp}(t)$. The characteristics of a single Excitatory Post Synaptic Potential (EPSP), or a single Inhibitory Post Synaptic Potential (IPSP) from rest can be adjusted by altering the conductance $g_{syn}(t)$, and the equilibrium potential V_{syn}. The resulting EPSP or IPSP can be characterised by rise time, half width and magnitude. EPSP, IPSP and AHP conductances are modelled by the conductance function:

$$g_{syn}(t) = g_{ahp}(t) = A\ \{(t/\tau_a)\ \exp(-t/\tau_a)\},$$

(Rall, 1967) requiring the choice of a scaling factor, A, and a time constant, τ_a. Once these have been determined the firing rates for pre-synaptic inputs have to be chosen. Selecting an appropriate mean firing rate for the input, along with any applied external current, $I_{ext}(t)$, determines the mean output firing rate of the simulation, and can be adjusted to give the desired output rate for each cell.

Appendix (B)

A GENSTAT Program

There follows a GENSTAT program for analysing a set of data in the case of two inputs and their corresponding carry-over effects, and a single output. Let the linear predictor of the model, η_t, representing the difference between the membrane potential on the trigger zone (U_t) and the constant threshold (θ_0), be of the form

$$\eta_t = V_t + \sum_{u=0}^{\gamma_t - 1} {}_1a_u \; {}_1x_{t-u} + \sum_{w \geq \gamma_t} {}_1c_w \; {}_1x_{t-w} + \sum_{u=0}^{\gamma_t - 1} {}_2a_u \; {}_2x_{t-u} + \sum_{w \geq \gamma_t} {}_2c_w \; {}_2x_{t-w} - \theta_0 \qquad \text{.....(2)}$$

where γ_t is the time elapsed since the time of the last output spike, the sets of coefficients $\{{}_ia_u\}$ and $\{{}_ic_w\}$; $i = 1,2$ represent the summation and carry-over effect functions for the two observed inputs respectively, and the term V_t represents the polynomial recovery function of order k, i.e.

$$V_t = \begin{cases} \sum_{i=1}^{k} \theta_i \, (\gamma_t - \zeta_1 - 1)^i & ; \; \gamma_t \geq \zeta_1 + 1 \\ 0 & ; \; \gamma_t \leq \zeta_1 + 1 \end{cases} \qquad \text{.....(3)}$$

ζ_1 is the minimum of the output inter-spike intervals (see chapter 3 and 4).

Assume that there exist three data files called "NERVE1.DAT", "NERVE2.DAT" and "NERVE3.DAT". The data file "NERVE1.DAT" represents the first set of inputs A[1...10] and their corresponding carry-over effect CA[1...15] which contain a matrix made up of 25 parallel vectors (A[1...10] to represent the first input at lag u, ${}_1x_{t-u}$; $u = 1, 2, \ldots, 10$ and CA[1...15] to represent the carry-over effect for the first input at lag w, ${}_1x_{t-w}$; $w = 1, 2, \ldots, 15$ as in the model form given in (2) above).

Similarly the second set of inputs B[1...8] and their corresponding carry-over effect CB[1...20]) are represented by the data file "NERVE2.DAT" (i.e., B[1...8] to represent the second input at lag u, ${}_2x_{t-u}$; $u = 1, 2, ..., 8$ and CB[1...20] to represent the carry-over effect for the second input at lag w, ${}_2x_{t-w}$; $w = 1, 2, ..., 20$ as given in (2) above). The third data file "NERVE3.DAT" represents the two vectors Y and Z[1] where Y represents the output discharges (i.e. $Y = y_t$) and $Z[1] = (\gamma_t - \zeta_1 - 1)$ which is used in the recovery function as given in (3).

VARIATE [NVAL=30000] Y, Z[1], V1

VARIATE [NVAL=30000] A[1...10], CA[10...15]

VARIATE [NVAL=30000] B[1...8], CB[1...20]

"Remark 1: NVAL is the total number of points or sample size of the data set and V1 is the total number of binomial trials which is always 1 since the process will take only the values 0 or 1, i.e., y_t=1 if there is an event and y_t=0 otherwise."

CALCULATE V1=1

"Remark 2: NERVE1.DAT is a binary data file of 25 parallel columns and 30000 rows for the 10 inputs A[1...10] and their corresponding carry-over effect CA[1...15] respectively."

OPEN NAME='NERVE1.DAT'; CHANNEL=2; FILETYPE= INPUT

READ [CHANNEL=2;END=*] A[ ], CA[ ]

"Remark 3: NERVE2.DAT is a binary data file of 28 parallel columns and 30000 rows for the 8 inputs B[1...8] and their corresponding carry-over effect CB[1...20] respectively."

OPEN NAME='NERVE2.DAT'; CHANNEL=3; FILETYPE= INPUT

READ [CHANNEL=3;END=*] B[ ],CB[ ]

OPEN NAME='NERVE3.DAT'; CHANNEL=4; FILETYPE= INPUT

READ [CHANNEL=4;END=*] Y,Z[1]

CLOSE CHANNEL=2, 3, 4

CALCULATE Z[2]=Z[1]2**

CALCULATE Z[3]=Z[1]3**

"Remark 4: the distribution is binomial with V1 trials. We use the logistic link function."

MODEL [DIST=BINOMIAL;LINK=LOGIT] Y; NBINOMIAL=V1

FIT [PRINT=E] Z[ ], A[ ], CA[ ], B[ ], CB[ ]

"Remark 5: VC1 denotes the variance covariance matrix needed for calculating the standard errors for the recovery function, D1 denotes the deviance and LP1 denotes the linear predictor."

RKEEP VCOVAR=VC1; DEVIANCE=D1; LINEARP=LP1

PRINT STRUCTURE=D1

PRINT STRUCTURE=VC1; DECIMALS=12

OPEN NAME='LINPR.DAT'; CHANNEL=4; FILETYPE=OUTPUT

PRINT [CHANNEL=4;IPRINT=*] STRUCTURE=Y, LP1; /

DECIMALS=5

STOP

END

Appendix (C)

A GENSTAT Printout

In this section we give a list of a GENSTAT printout of the results when analysing the set of simulated data that we discussed earlier in chapter four (section 4.6) when introducing the idea of the carry-over effect function (COE). A GENSTAT printout of the results as given by the package, without making any changes to it, is of the form

Genstat 5 Release 2.2 (80386 based DOS PCs) 21-Mar-1995 09:49:45
Copyright 1990, Lawes Agricultural Trust (Rothamsted Experimental Station)

```
 1  VARIATE [NVAL=60000] V, Z[1]
 2  OPEN NAME='Z502.DAT'; CHANNEL=3; FILETYPE=INPUT
 3  READ [CHANNEL=3;END=*] Z[1]
 Identifier  Minimum     Mean  Maximum  Values  Missing
      Z[1]    0.000    7.503   38.000   60000        0  Skew
 4  CLOSE CHANNEL=3
 5  CALCULATE V=1
 6  VARIATE [NVAL=4345] A1[1...6]
 7  OPEN NAME='A1.DAT'; CHANNEL=2; FILETYPE=INPUT
 8  READ [CHANNEL=2;END=*] A1[1...6]
 Identifier  Minimum     Mean  Maximum  Values  Missing
     A1[1]       16    30239    59995    4345        0
     A1[2]       16    39712    59994    4345        0
     A1[3]       28    42114    59995    4345        0
     A1[4]       29    46992    59905    4345        0  Skew
     A1[5]       30    51057    59906    4345        0  Skew
     A1[6]       31    54024    59907    4345        0  Skew
 9  CALCULATE Y = EXPAND(A1[1]; 60000)
10  CALCULATE A[0] = EXPAND(A1[2]; 60000)
11  CALCULATE A[1] = EXPAND(A1[3]; 60000)
12  CALCULATE A[2] = EXPAND(A1[4]; 60000)
13  CALCULATE A[3] = EXPAND(A1[5]; 60000)
14  CALCULATE A[4] = EXPAND(A1[6]; 60000)
```

```
15  VARIATE [NVAL=611] A1[1...5]
16  OPEN NAME='A2.DAT'; CHANNEL=3; FILETYPE=INPUT
17  READ [CHANNEL=3;END=*] A1[1...5]
 Identifier  Minimum    Mean  Maximum  Values  Missing
    A1[1]       32   30333    59908     611        0
    A1[2]       33   37886    59909     611        0
    A1[3]      536   44133    59910     611        0
    A1[4]     1421   49113    59911     611        0  Skew
    A1[5]     1422   53536    59912     611        0  Skew
18  CALCULATE A[5] = EXPAND(A1[1]; 60000)
19  CALCULATE A[6] = EXPAND(A1[2]; 60000)
20  CALCULATE A[7] = EXPAND(A1[3]; 60000)
21  CALCULATE A[8] = EXPAND(A1[4]; 60000)
22  CALCULATE A[9] = EXPAND(A1[5]; 60000)
23  VARIATE [NVAL=81] A1[1...5]
24  OPEN NAME='A3.DAT'; CHANNEL=4; FILETYPE=INPUT
25  READ [CHANNEL=4;END=*] A1[1...5]
 Identifier  Minimum    Mean  Maximum  Values  Missing
    A1[1]     1434   29568    59913      81        0
    A1[2]     5121   45345    59891      81        0
    A1[3]     6815   50614    56248      81        0  Skew
    A1[4]     6816   53059    56249      81        0  Skew
    A1[5]    36925   54817    55393      81        0  Skew
26  CALCULATE A[10] = EXPAND(A1[1]; 60000)
27  CALCULATE A[11] = EXPAND(A1[2]; 60000)
28  CALCULATE A[12] = EXPAND(A1[3]; 60000)
29  CALCULATE A[13] = EXPAND(A1[4]; 60000)
30  CALCULATE A[14] = EXPAND(A1[5]; 60000)
31  CLOSE CHANNEL=2,3,4
32  VARIATE [NVAL=2522] A1[1...5]
33  OPEN NAME='C1.DAT'; CHANNEL=2; FILETYPE=INPUT
34  READ [CHANNEL=2;END=*] A1[1...5]
 Identifier  Minimum    Mean  Maximum  Values  Missing
    A1[1]       17   53310    59881    2522        0  Skew
    A1[2]       18   44914    59996    2522        0
    A1[3]       19   38008    59997    2522        0
```

```
    A1[4]     20   33160   59998    2522      0
    A1[5]     21   30542   59999    2522      0
35  CALCULATE C[1] = EXPAND(A1[1]; 60000)
36  CALCULATE C[2] = EXPAND(A1[2]; 60000)
37  CALCULATE C[3] = EXPAND(A1[3]; 60000)
38  CALCULATE C[4] = EXPAND(A1[4]; 60000)
39  CALCULATE C[5] = EXPAND(A1[5]; 60000)
40  VARIATE [NVAL=2949] A1[1...5]
41  OPEN NAME='C2.DAT'; CHANNEL=3; FILETYPE=INPUT
42  READ [CHANNEL=3;END=*] A1[1...5]
 Identifier  Minimum     Mean  Maximum  Values  Missing
    A1[1]     22   33425   60000    2949      0
    A1[2]     23   32353   59936    2949      0
    A1[3]     24   31452   59937    2949      0
    A1[4]     25   30849   59938    2949      0
    A1[5]     26   30548   59939    2949      0
43  CALCULATE C[6] = EXPAND(A1[1]; 60000)
44  CALCULATE C[7] = EXPAND(A1[2]; 60000)
45  CALCULATE C[8] = EXPAND(A1[3]; 60000)
46  CALCULATE C[9] = EXPAND(A1[4]; 60000)
47  CALCULATE C[10] = EXPAND(A1[5]; 60000)
48  VARIATE [NVAL=2987] A1[1...5]
49  OPEN NAME='C3.DAT'; CHANNEL=4; FILETYPE=INPUT
50  READ [CHANNEL=4;END=*] A1[1...5]
 Identifier  Minimum     Mean  Maximum  Values  Missing
    A1[1]     27   30675   59940    2987      0
    A1[2]     28   30598   59941    2987      0
    A1[3]     29   30556   59942    2987      0
    A1[4]     30   30524   59943    2987      0
    A1[5]     31   30518   59944    2987      0
51  CALCULATE C[11] = EXPAND(A1[1]; 60000)
52  CALCULATE C[12] = EXPAND(A1[2]; 60000)
53  CALCULATE C[13] = EXPAND(A1[3]; 60000)
54  CALCULATE C[14] = EXPAND(A1[4]; 60000)
55  CALCULATE C[15] = EXPAND(A1[5]; 60000)
```

```
56  CLOSE CHANNEL=2,3,4
57  VARIATE [NVAL=2990] A1[1...5]
58  OPEN NAME='C4.DAT'; CHANNEL=2; FILETYPE=INPUT
59  READ [CHANNEL=2;END=*] A1[1...5]
 Identifier  Minimum    Mean  Maximum  Values  Missing
    A1[1]       32   30541    59945     2990        0
    A1[2]       33   30540    59946     2990        0
    A1[3]       34   30541    59947     2990        0
    A1[4]       35   30540    59948     2990        0
    A1[5]       36   30541    59949     2990        0
60  CALCULATE C[16] = EXPAND(A1[1]; 60000)
61  CALCULATE C[17] = EXPAND(A1[2]; 60000)
62  CALCULATE C[18] = EXPAND(A1[3]; 60000)
63  CALCULATE C[19] = EXPAND(A1[4]; 60000)
64  CALCULATE C[20] = EXPAND(A1[5]; 60000)
65  VARIATE [NVAL=2990] A1[1...5]
66  OPEN NAME='C5.DAT'; CHANNEL=3; FILETYPE=INPUT
67  READ [CHANNEL=3;END=*] A1[1...5]
 Identifier  Minimum    Mean  Maximum  Values  Missing
    A1[1]       37   30542    59950     2990        0
    A1[2]       38   30543    59951     2990        0
    A1[3]       39   30544    59952     2990        0
    A1[4]       40   30545    59953     2990        0
    A1[5]       41   30546    59954     2990        0
68  CALCULATE C[21] = EXPAND(A1[1]; 60000)
69  CALCULATE C[22] = EXPAND(A1[2]; 60000)
70  CALCULATE C[23] = EXPAND(A1[3]; 60000)
71  CALCULATE C[24] = EXPAND(A1[4]; 60000)
72  CALCULATE C[25] = EXPAND(A1[5]; 60000)
73  CLOSE CHANNEL=2,3
74  CALCULATE Z[2]= Z[1]**2
75  CALCULATE Z[3]= Z[1]**3
76  MODEL [DIST=BIN;LINK=LOGIT] Y;NBIN=V
```

77 FIT [PRINT=E] Z[ ], A[ ], C[ ]

	estimate	s.error	t
Constant	-15.035	0.321	-46.87
Z[1]	1.6259	0.0575	28.28
Z[2]	-0.06487	0.00343	-18.94
Z[3]	0.00088	0.000065	13.56
A[0]	0.7879	0.0730	10.79
A[1]	2.6667	0.0654	40.75
A[2]	3.7277	0.0801	46.51
A[3]	4.2869	0.0964	44.46
A[4]	4.002	0.111	36.09
A[5]	3.731	0.125	29.80
A[6]	3.439	0.139	24.82
A[7]	3.161	0.151	20.94
A[8]	3.161	0.165	19.19
A[9]	2.548	0.208	12.28
A[10]	2.196	0.249	8.83
A[11]	2.555	0.315	8.11
A[12]	1.396	0.484	2.89
A[13]	2.368	0.593	3.99
A[14]	1.581	0.922	1.71
C[1]	1.052	0.890	1.18
C[2]	1.617	0.359	4.50
C[3]	1.636	0.250	6.56
C[4]	2.639	0.191	13.79
C[5]	2.668	0.175	15.25
C[6]	2.631	0.155	17.00
C[7]	2.438	0.136	17.87
C[8]	2.280	0.124	18.33
C[9]	1.887	0.119	15.89
C[10]	2.224	0.102	21.79
C[11]	1.714	0.101	17.05
C[12]	1.4456	0.0958	15.09
C[13]	1.1690	0.0925	12.64
C[14]	0.8562	0.0935	9.16

C[15]	0.5807	0.0928	6.26
C[16]	0.6310	0.0872	7.23
C[17]	0.4159	0.0887	4.69
C[18]	0.2055	0.0904	2.27
C[19]	0.4368	0.0844	5.17
C[20]	0.3059	0.0871	3.51
C[21]	0.1483	0.0895	1.66
C[22]	0.1085	0.0894	1.21
C[23]	0.1425	0.0878	1.62
C[24]	-0.0553	0.0914	-0.60
C[25]	0.0195	0.0881	0.22

```
78    RKEEP VCOVAR=VC1;DEVIANCE=D;LINEARP=LP1
79    PRINT STRUCTURE=Deviance

                Deviance
                 18706
80    PRINT STRUCTURE=Vcov; DECIMALS=12

                    Vcov
          Constant 0.1029E+00
             Z[1]  -0.1753E-01  0.3306E-02
             Z[2]   0.9457E-03 -0.1917E-03  0.1173E-04
             Z[3]  -0.1601E-04  0.3422E-05 -0.2186E-06  0.4231E-08
81    STOP
82    END.
```

References

Amjad, A.M. **(1989).**

Identification of Point Process Systems with Application to Complex Neuronal Networks.

Ph.D. Thesis, University of Glasgow, Glasgow.

Amjad, A.M., Breeze, P., Conway, B.A., Halliday, D.M. and Rosenberg, J.R. **(1989).**

A Framework for the Analysis of Neuronal Networks.

Progress in Brain Research, Vol 80, 243-255.

Amjad, A.M., Breeze, P., Gladden, M.H., Halliday, D.M. and Rosenberg, J.R. **(1987).**

The Application of Partial Coherence to the Analysis of the Association between Spike Trains.

J. Physiol., London, 392: 9P.

Anderson, T.W. **(1983).**

Statistical Analysis of Time Series.

Wiley & Sons, New York.

Baldissera, F. and Gustafsson, B. **(1974).**

Afterhyperpolarization Conductance Time Course in Lumbar Motoneurones.

Acta. Physiol. Scand., 91, 512-527.

Bessou, P. and Pages, B. **(1975).**

Cinematographic Analysis of Contractile Events produced in Intrafusal Fibres by Stimulation of Static and Dynamic Fusimotor Axons.

Journal of Physiology, 252, 397-427.

Bhat, U.N. **(1969).**

Sixty Years of Queuing Theory.

Management, Sci., 15, 280-294.

Bloomfield, P. **(1976).**

Fourier Analysis of Time Series: An Introduction. 2nd edition.

Wiley & Sons, New York.

Borisyuk, G.N. , Borisyuk, R. M. , Kirillov, A.B. , Kovalenko, E.I. and Kryukov, V.I. **(1985).**

A new Statistical Method for Identifying Interactions between Neuronal Network Elements.

Biological Cybernetics, 52, 301-306.

Box, G.E.P. **(1954).**

Some Theorems on Quadratic Forms applied in the Study of Analysis of Variance Problems.

Ann. Math. Statist., 25, 290-302.

Box, G.E.P. and Jenkins, G.M. **(1970).**

Time Series Analysis: Forecasting and Control.

Holden-Day, San Francisco, California.

Box, G.E.P. and Tidwell, P.W. **(1962).**

Transformation of the Independent Variables.

Technometrics, 4, 531-550.

Boyd, I.A. **(1962).**

The Structure and Innervation of the Nuclear Bag Muscle Fibre System and the Nuclear Chain Muscle Fibre System in Mammalian Muscle Spindles.

Phil. Trans. R. Soc. B., 245, 83-136.

Boyd, I.A. **(1980).**

The Isolated Mammalian Muscle Spindle.

Trends in Neuroscience , 3, 258-265.

Brillinger, D.R. **(1975b).**

Estimation of Product Densities.

Comp. Sci. Statist. , Ann. Symp. Interface 8th, 431-438, UCLA, Los Angeles.

Brillinger, D.R. **(1976a).**

Estimation of Second Order Intensities of a Bivariate Stationary Point Process.

J. R. Statist. Soc. , B 38, 60-66.

Brillinger, D.R. **(1978a).**

Comparative Aspects of the Study of Ordinary Time Series and of Point Processes.

Developments in Statistics, volume 1, 33-134, ed. Krishnaiah, P.R. Academic Press, New York.

Brillinger, D.R. **(1978b).**

A Note on the Estimation of Evoked Response.

Biol. Cybernetics, 31, 141-144.

Brillinger, D.R. **(1980).**

Some Aspects of the Analysis of Evoked Response Experiments.

Proceedings of the International Symposium of Statistics and Related Topics, Ottawa, Canada.

Brillinger, D.R. **(1981).**

The General Linear Model in the Design and Analysis of Evoked Response Experiments.

J. Theoret. Neurobiol, 1, 105-119.

Brillinger, D.R. **(1983).**

The Finite Fourier Transform of a Stationary Process.

Handbook of Statistics, eds. Brillinger, D.R. and Krishnaiah, P.R. , Elsevier, Amesterdam.

Brillinger, D.R. **(1986).**

Some Statistical Methods for Random Process Data from Seismology and Neurophysiology.

Technical Report no. 84, Department of Statistics, University of California, Berkeley.

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. **(1992).**

Nerve Cell Spike Train Data Analysis: A Progression of Technique.
Journal of the American Statistical Association, Vol.87, No. 418, 260-271.

Brillinger, D.R. , Bryant, H.L, JR., and Segundo, J.P. **(1976).**

Identification of Synaptic Interaction.
Biological Cybernetics, 22, 213-228.

Brillinger, D.R. and Rosenblatt, M. **(1967a).**

Asymptotic Theory of k-th Order Spectra.
Spectral Analysis of Time Series, ed. Harris, B., 153-188, Jhon **Wiley & Sons**, New York.

Brillinger, D.R. and Rosenblatt, M. **(1967b).**

Computation and Interpretation of k-th Order Spectra.
Spectral Analysis of Time Series, ed. Harris, B. , 189-232, **Wiley & Sons**, New York.

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. and Tukey, J. W. **(1984).**

Spectral Analysis in the Presence of Noise: Some Issues and Examples.
The Collected Works of J. W.Tukey, 1002-1141.
Monterey, Wadsworth, California.

Bryant, H.L. ,JR. , Ruiz Marcos, A. and Segundo, J. P. **(1973).**

Correlation of Neuronal Spike Discharges produced by Monosynaptic Connections and Common Inputs.
Journal of Neurophysiology , 36, 205-225.

Burke, R.M. and Rudomin,P. **(1977).**

Spindal Neurones and Synapses.

Handbook of Physiology, Section 1: The Nervous System, Vol. 1, Cellular Biology of Neurones, Part 2, Eds Brookhart, J.M. and Mountcastle, V.B. , American Physiological Society, Bethesda, USA, 877-944.

Cajal, S.R. **(1955).**

Histologie du Système Nerveux: De L'Homme and Des Vertébrés.

Consejo Superior de Investigaciones Cientificas, Madrid.

Chatfield, C. **(1980).**

The Analysis of Time Series: An Introduction, 2nd edition.

Chapman and Hall, London.

Conway, B.A. , Halliday, D.M. and Rosenberg, J.R. **(1993).**

Detection of Weak Synaptic Interactions between Single Ia Afferent and Motor-Unit Spike Trains in the Decerebrate Cat.

Journal of Physiology. (London), 471, 379-409.

Cox, D.R. and Isham, V. **(1980).**

Point Processes.

Chapman and Hall, London.

Cox, D.R. and Lewis, P.A.W. **(1966).**

The Statistical Analysis of Series of Events.

Chapman and Hall, London.

Cox, D.R. and Lewis, P.A.W. **(1972).**

Multivariate Point Processes.

Proc. 6th Berkeley Symp. , 2, 401-448.

Daley.D. J. and Vere-Jones, D. **(1988).**

An Introduction to the Theory of Point Processes .

Springer , New York.

Draper, N.R. and Smith, H. **(1981).**

Applied Regression Analysis. (2nd edition).

Wiley & Sons, New York.

Emonet-Denand, F. , Laporte, Y., Matthews, P.B.C. and Petit, J. **(1977).**

On the Subdivision of Static and Dynamic Fusimotor Axons on the Primary Endings of the Cat Muscle Spindle.

Journal of Physiology., 268, 827-860.

Feinberg, S.E. **(1974).**

Stochastic Models for Single Neurone Firing Trains: A Survey.

Biometrics., 30, 399-427.

Getting, P.A. **(1989).**

Reconstruction of Small Networks. In Methods in Neuronal Modelling: From Synapses to Networks, eds Kock. C. Segev, I.

MIT Press, 171-194.

Graybill, F.A. **(1976).**

Theory and Application of the Linear Model.

Wadsworth Publishing Company, Inc. , Belmont. California.

Haberman, S.J. **(1977).**

Maximum Likelihood Estimates in Exponential Response Models.

Ann. Statist., 5, 815-841.

Halliday, D.M. **(1986).**

Application of Point Process System Identification Techniques to Complex Physiological Systems.

Ph.D. Thesis, University of Glasgow, Glasgow.

Halliday, D.M. **(1994).**

Effects of Electrotonic Spread of EPSP's on Synaptic Transmission in Motoneurones: A Simulation Study. In: Alpha and Gamma Motor Systems, Eds. Taylor, Gladden and Durbaba.

Plenum (in press).

Halliday, D.M. **(1994).**

Effects of EPSP duration on the Correlation Between Pre- and Post-Synaptic Motoneurone Discharges: A Simulation Study. Technical Report No. 019 - Computational and Experimental Neuroscience Group.

Department of Physiology, University of Glasgow.

Height, F.A. **(1967).**

Handbook of the Poisson Distribution.

Wiley & Sons, New York.

Hille, B. **(1984).**

Ionic Channels of Excitable Membranes.

Sinauer.

Holden, A.V. **(1976).**

Models for the Stochastic Activity of Neurones.

Wiley & Sons, New York.

Hosmer, D.W. and Lemeshow, S. **(1989).**

Applied Logistic Regression.

Springer, Berlin.

Johannesma, P.I.M. **(1968).**

Diffusion Models for the Stochastic Activity of Neurones.

Neural Networks. , 116-144, ed. Caraniello, E.K. , Berlin, Springer-Verlag.

Johansson, H. **(1981).**

Reflex Control of γ -motoneurones.

Umea University, Medical Dissertation, 55-57, Dept. of Physiology, University of Umea, Sweden.

Kandel, E.R. , Schwartz, J. H. and Jessell, T.M. **(1991).**

Principles of Neural Science. 3rd edition.

Appleton and Lange. Simon, Schuster and Professional Group. Prentice Hall, Englewood Cliffs, New Jersey.

Kendall, M.G. and Stuart, A. **(1977).**

The Advanced Theory of Statistics, Vol. **1.**

Griffen, London.

Knight, B.W. **(1972).**

Dynamics of Encoding in a Population of Neurons.

J. Gen. Physiol. , 59, 734-766.

Knox, C.K. **(1974).**

Cross-Correlation Functions for a Neuronal Model.

J. Biophys., 14, 567-582.

Knox, C.K. , Kubota. S. and Poppele, R.E. **(1977).**

A Determination of Excitability Change in Dorsal Spinocerebellar Tract Neurones from Spike Train Analysis.

J. Neurophysiol. , 40, 626-646.

Koopmans, L.H. **(1974).**

The Spectral Analysis of Time Series.

Academic Press, New York.

Koopmans, L.H. **(1983).**

A Spectral Analysis Primer.

Handbook of Statistics, Vol. 3. eds. Brillinger, D.R. and Krishnaiah, P.R. , 169-184, Elsevier, Amesterdam.

Kuffler, S.W. and Nicholls, J.G. **(1976).**

From Neuron to Brain: A Cellular Approach to the Function of the Nervous System.

Sinauer Associates, Inc. , Sunderland, Massachusetts.

Lau, W.N. **(1991).**

Point Process Analysis Techniques: Theory and Applications to Complex Neurophysiological Systems.

Ph.D. Thesis, University of Glasgow, Glasgow.

Lewis, P.A.W. **(1970).**

Remarks on the Theory, Computation and Application of the Spectral Analysis of Series of Events.

J. Sound Vib. , 12, 353-375.

Lewis, P.A.W. **(1972).**

Stochastic Point Processes.

Wiley & Sons, New York.

Lusch, H.R. **(1990).**

Transmission Failure and its Relief in the Spinal Monosynaptic Reflex Arc. In The Segmental Motor System, eds Binder, M.D. and Mendell, L.M.

Oxford University Press. , 328-348.

Matthews, P.B.C. **(1962).**

The Differentiation of Two Types of Fusimotor Fibre by their Effects on the Dynamic Response of Muscle Spindle Primary Endings.

Q. Jl. Exp. Physiol., 47, 324-333.

Matthews, P.B.C. **(1981).**

Review Lecture: Evolving views on the Internal Operation and Functional Role Of the Muscle Spindle.

J. Physiol., 320, 1-30.

Matthews, P.B.C. and Stein, R.B. **(1969).**

The Regularity of the Primary and Secondary Muscle Spindle Afferent Discharges.

J. Physiol., 202, 59-82.

McCullagh, P. **(1985).**

On the Asymptotic Distribution of Pearson's Statistic in Linear Exponential-Family Models.

Int. Statist. Rev., 53, 61-67.

McCullagh, P. and Nelder, J.A.. **(1992).**

Generalised Linear Models (2nd edition.).

London, Chapman and Hall.

Mood, A.M., Graybill, F.A. and Boes, D.C. **(1985).**

Introduction to the Theory of Statistics.

McGraw-Hill, London.

Rall, W. **(1967).**

Distinguishing Theoretical Synaptic Potentials computed for different Soma-Dendritic Distributions of Synaptic Inputs.

J. Neurophysiol., 30, 1138-1168.

Rall, W. **(1977).**

Core Conductor Theory and Cable Properties of Neurones. In Handbook of Physiology: The Nervous System, vol. 1, part 1, eds Kandel, E.R., Brookhart, J.M. and Mountcastle V.B.

William and Wilkins, Maryland., 39-97.

Rao, C.R. **(1984).**

Linear Statistical Inference and its Applications.

Wiley & Sons, New York.

Rigas, A.G. **(1983).**

Point Processes and Time Series Analysis: Theory and Application to Complex Physiological Problems.

Ph.D. Thesis, University of Glasgow, Glasgow.

Rosenberg, J.R. , Amjad, A.M. , Breeze, P. , Brillinger, D.R. and Halliday, D.M. **(1989).**

The Fourier Approach to the Identification of Functional Coupling Between Neuronal Spike Trains.

Prog. Biophysics. Mol. Biol., 53: 1-31.

Rosenberg, J.R. , Murray-Smith, D.J. and Rigas, A. **(1983).**

An Introduction to the Application of System Identification Techniques to Elements of the Neuromuscular System.

Trans. Inst. M.C. , 4: 187-201.

Rosenberg, J.R. and Rigas, A. **(1985).**

Spectral Composition of Muscle Spindle Ia Responses to Combined Length and Fusimotor Inputs.

I.A. Boyd and M.H. Gladden (Eds.), The Muscle Spindle, Macmillan, London, 397-402.

Sampath, G. and Srinivasan, S.K. **(1977).**

Stochastic Spike Trains of Single Neurones.

Lecture Notes in Biomathematics, Vol. 16, ed. Levin, S. 1-188. Springer-Verlag, Berlin.

Shepherd, G.M. **(1974).**

The Synaptic Organisation of the Brain, 1st edition, pp79-110.

Oxford University Press. London.

Snyder, D.L. **(1975).**

Random Point Processes.

Wiley & Sons, New York.

Stegler, S.M. **(1986).**

The History of Statistics.

Belknap Press, Cambridge, Mass.

Stevens, C.F. **(1968).**

Synaptic Physiology.

Proc. IEEE. , 56, 916-930.

Tukey, J.W. **(1959).**

The Estimation of Power Spectra and Related Quantities.

On Numerical Approximation. 389-411. Univ. of Wisconsin Press, Madison.

Tukey, J.W. **(1977).**

Exploratory Data Analysis.

Addison-Wesley, USA.

Tukey, J.W. **(1978).**

Can We Predict where "Time Series" should go next ?.

Directions in Time Series. eds. Brillinger, D.R. and Tiao, G.C. Hayward: Inst. Math. Statist.

Wedderburn, R.W.M. **(1974).**

Quasilikelihood Function, Generalized Linear Models and the Gauss-Newton Method.

Biometrika, 61, 439-447.

Weisberg, S. **(1985).**

Applied Linear Regression (2nd edition).

Wiley & Sons, New York.

Westergaard, H. **(1968).**

Contributions to the History of Statistics.

Agathon, New York.

GLASGOW UNIVERSITY