2011

New statistical methods to derive functional connectivity from multiple spike trains

Masud, Mohammad Shahed

NEW STATISTICAL METHODS TO DERIVE FUNCTIONAL CONNECTIVITY FROM MULTIPLE SPIKE TRAINS

M. S. MASUD

PhD

2011

Copyright Statement

New statistical methods to derive functional connectivity from multiple spike trains

Mohammad Shahed Masud

School of Computing and Mathematics

University of Plymouth

A thesis submitted to the University of Plymouth in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

April 2011

Mohammad Shahed Masud

New statistical methods to derive functional connectivity from multiple spike trains

Abstract

Analysis of functional connectivity of simultaneously recorded multiple spike trains is one of the major issues in the neuroscience. The progress of the statistical methods to the analysis of functional connectivity of multiple spike trains is relatively slow. In this thesis two statistical techniques are presented to the analysis of functional connectivity of multiple spike trains. The first method is known as the modified correlation grid (MCG). This method is based on the calculation of cross-correlation function of all possible pair-wise spike trains. The second technique is known as the Cox method. This method is based on the modulated renewal process (MRP). The original paper on the application of the Cox method (Borisyuk et al., 1985) to neuroscience data was used to analyse only pairs and triplets of spike trains. This method is further developed in this thesis to support simultaneously recorded of any possible set of multiple spike trains.

A probabilistic model is developed to test the Cox method. This probabilistic model is based on the MRP. Due to the common probabilistic basis of the probabilistic model and the Cox method, the probabilistic model is a convenient technique to test the Cox method. A new technique based on a pair-wise analysis of Cox method known as the Cox metric is presented to find the groups of coupled spike trains. Another new technique known as motif analysis is introduced which is useful in identifying interconnections among the spike trains. This technique is based on the triplet-wise analysis of the Cox method.

All these methods are applied to several sets of spike trains generated by the Enhanced Leaky and Integrate Fire (ELIF) model. The results suggest that these methods are successful for analysing functional connectivity of simultaneously recorded multiple spike trains. These methods are also applied to an experimental data recorded from cat's visual cortex. The connection matrix derived from the experimental data by the Cox method is further applied to the graph theoretical methods.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to take this opportunity to thank a number of people, whose support has been crucial to me throughout my work on this thesis.

Primarily I would like to acknowledge my supervisor Professor Roman Borisyuk, for first giving me the chance to pursue this interesting line of research, and subsequently for his wise and patient guidance over the past three years. I also owe thanks to my second supervisor Dr. Liz Stuart, always generous with her time, her aid and advice was a treat.

My examiners, Professor Leslie Smith and Dr. Thomas Wennekers suggested corrections and changes that significantly improved the thesis.

I also thank Dr. Abul Kalam Al Azad for taking the time to read this thesis, and whose helpful comments were welcome. I am grateful for the support of the entire department, all of whom are great colleagues, many of whom are friends. In particular I thank Jonathan Waddington for his support.

Finally, I am grateful to my family members especially to my parents for their continuous support and encouragement throughout my entire study period.

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Graduate Committee.

This study was financed with the aid of a studentship from the Engineering and Physical Sciences Research Council (EPSRC).

Relevant scientific seminars and conferences were regularly attended at which work was presented on several occasions. One research paper has been published in refereed journals. One abstract has also been published.

Referred journal article:

Masud M.S. and Borisyuk R. Statistical technique for analysing functional connectivity of multiple spike trains. *Journal of Neuroscience Methods*, 196 (2011); 201-219.
(Associated with chapter 4 and 5)

Abstract:

Masud M.S. and Borisyuk R. Modulated renewal process and statistical analysis of multiple spike trains. *Frontiers in Neuroinformatics 2009. Conference Abstract: 2nd INCF Congress of Neuroinformatics*, pp 49.

Poster presentations:

2009: Workshop on Mathematical Neuroscience, Edinburgh. A new method to study functional connectivity of spike trains.

2009: The 8th International Workshop on Neural Coding, Tainan, Taiwan. Modulated renewal process and statistical analysis of multiple spike trains.

2010: Workshop on Spike Train Measures and Their Applications to Neural Coding, Plymouth. Spike train metric based on Cox method.

Other conferences attended:

2008: Workshop on Mathematical Neuroscience, Edinburgh.

2008: Workshop on Nonlinear Neurodynamics, Exeter.

2009: Workshop on Spike Train Data Analysis, Newcastle.

2009: Workshop on Robotics and Neural Systems, Plymouth.

Word count for the main body of this thesis: 48528

Signed: Mohammad Shahed Masud

Date: 12.08.2011

Chapter 1

Introduction

The brain consists of billions of cells of two types – glia cells and neurons. The neurons are considered as the main units dealing with information processing. One way that information is transmitted between neurons is through changes in their electrical activity known as action potential or spike. Typically action potentials or spikes have duration of 1-2 milliseconds. A chain of action potentials emitted by a single neuron is called a spike train.

Usually, generation of action potential happens as a response to the incoming signal either from other neurons or external input. Thus, it is important to study both activity of a single neuron and a group of interactive neurons recorded simultaneously. The main mechanism of neural interaction is called synaptic transmission. Synapse is a special part of the neuron which provides a possibility to transmit electrical pulse from one neuron to another. Although single unit activity is irregular and complex enough, a small system of interconnected neurons can exhibit complex behaviour and information-processing capabilities not present in a single neuron. The understanding of networks of coupled neurons is one of the major issues in the neuroscience.

For decades a common experimental method in neuroscience was based on the activity of a single neuron. It was especially useful for studies of the effects of sensory inputs. A substantial part of recent research in neuroscience involves the study of the activity of neurons to behaviour and cognition. Such studies require the simultaneous recording of many neurons. Recent advances in multi electrode neural recording systems have made it possible to record activity from a large number of neurons simultaneously.

One of the important problems of the simultaneously recorded spike trains is the study of the functional connectivity. The term functional connectivity is used to identify statistical dependencies and influences between spike trains. Despite the developments in recording technology, the progress in the methods to study functional connectivity has been rather slow. Most of the existing techniques that are used to identify functional connectivity are based on pair-wise analysis. These pair-wise techniques are usually focused on a pair of spike trains but they fail to consider all possible influences from other simultaneously recorded spike trains. For this reason, these pair-wise estimates of the functional connectivity sometimes can lead to inaccuracies. Therefore, new techniques are required which can capture all possible influences from other simultaneously recorded spike trains and can estimate accurate functional connectivity.

1.1 Objective of the study

This section provides a brief summary of the objectives of the study. The questions addressed to infer functional connectivity are important in computational neuroscience and they need the appropriate statistical techniques to answer them. There are some recent developments of statistical techniques that are used to answer the questions related to functional connectivity of multiple spike trains. Although there are a number of statistical techniques used for the analysis of functional connectivity of multiple spike trains, a great deal of work is still necessary in this area in order to provide researchers with appropriate tools for the analysis of multiple spike train data. The objective of this study is to develop new statistical techniques for analysing multiple spike trains which can provide some insights on the understanding of complex neural activity and neural interactions.

1.2 Tour of the thesis

This section provides an overview of the structure and the contents of the thesis. The thesis contains eight chapters.

Chapter 1 is an introduction where the objectives are formulated.

Chapter 2 is devoted to the review of literature and the papers relevant to study of functional connectivity of spike trains are discussed. The techniques that are used to infer functional connectivity of spike trains are reviewed. Studies of the graph theoretical methods that are used to analyse a connectivity pattern are also discussed.

Chapter 3 presents a new statistical method known as modified correlation grid (MCG) to analyse functional connectivity of multiple spike trains. This MCG method is related to the correlation grid (Stuart et al., 2005). The MCG method is able to distinguish the direct connections from the spurious (common source and indirect) connections and thus reveals the functional connectivity of multiple spike trains. This method is based on the calculation of cross-correlation function of all possible pair-wise spike trains. A clustering algorithm is applied to the significant peak and corresponding time shift to distinguish the connections. Application of this method is shown for two sets of spike trains generated by the Enhanced Leaky Integrate and Fire (ELIF) model (Borisyuk, 2002). The description, dynamics and parameters of the ELIF model are presented in appendix A.

Chapter 4 describes a probabilistic model for generation of dependent spike train. This probabilistic model is based on the theory of modulated renewal process (MRP). To generate a dependent spike train from an influence of independent spike train, an influence function is described which is used in neuroscience to describe synaptic connectivity between neurons. To assess how the generated dependent spike train agrees

with the theoretical probability distribution two 'goodness of fit' tests are used. An optimization procedure is described by which the parameters of the probabilistic model can be adjusted in such a way that these parameter values can be used to generate spike trains similar to the 'integrate and fire' neuron model. This probabilistic model is used to test the Cox method.

Chapter 5 presents a statistical method known as the Cox method to analyse functional connectivity of multiple spike trains. This technique is based on the theory of modulated renewal processes (MRP) and is the generalization of the Cox method developed by Borisyuk et al. (1985). Application of this method is shown for several sets of spike trains generated by the ELIF model. A comparison of the Cox method with the cross-correlation function (CCF) is presented using pairs of spike trains generated by the probabilistic model. Due to the common probabilistic basis of the probabilistic model and the Cox method, the probabilistic model is a convenient technique to test the Cox method. The analysis of a set of three spike trains is used to demonstrate that the Cox method can find a scheme of connections in a case of 'common source' connection architectures. Similarly, the method is successful when it is used to analyse sets of three spike trains with 'indirect connection' architecture. A new technique based on the pair-wise analysis of the Cox method known as the Cox metric is demonstrated to find the groups of coupled spike trains. An application of the Cox metric is shown for a set of twenty spike trains generated by the ELIF model. Another new technique known as motif analysis is introduced which is useful in identifying interconnections among the spike trains. This technique is based on the triplet-wise analysis of the Cox method. An application of this technique is demonstrated by a set of twenty spike trains generated by the ELIF model.

Chapter 6 presents the application of the MCG method and the Cox method to a set of experimental data recorded from the cat's visual cortex (Nikolic, 2007; Schneider et al., 2006). The experimental conditions include six different stimuli corresponding to different orientation of the stimulation grid. For each stimulus, functional connectivity of 29 spike trains are analysed by the MCG method and by the Cox method. The connections that are common to the MCG method and the Cox method are presented for each stimulus. Also the connections that are common to all stimuli identified separately by the MCG method and the Cox method are shown. The Cox metric is applied to the 29 spike trains to identify the groups of similar spike trains for each stimulus. Similarly, to identify the interconnection among the 29 spike trains the motif analysis is conducted for each stimulus. In this chapter, the results of analysing functional connectivity of stimulus 1 are presented. The results of analysing functional connectivity of another 5 stimuli are presented in appendix B.

Chapter 7 describes some graph theoretical methods (Rubinov and Sporns, 2010) that are used for the comprehensive analysis of the connectivity derived from statistical analysis of multiple spike trains. These graph theoretical methods are applied to the connectivity matrix of each stimulus identified by the Cox method. A statistical model known as P1 model (Holland and Leinhardt, 1981) which is used to identify the attractive and influential people in the social science network, is also applied to the connectivity matrix. Application of this model to the connectivity matrix is useful to identify the influential and attractive spike trains. In this chapter, the results of the graph theoretical methods of stimulus 1 are presented. The results of graph theoretical methods of another 5 stimuli are presented in appendix C.

Chapter 8 draws contribution and conclusion regarding the techniques proposed to analyse functional connectivity and their applications to the experimental data.

Chapter 2

Review of spike train analysis

In this chapter a review of studies related to the interaction between stimulus and spike train is discussed. In addition to these studies, methods of analysing functional connectivity of simultaneously recorded multiple spike trains are discussed. Studies of the analysis of functional connectivity by graph theoretical methods are also discussed.

2.1 Introduction

The brain receives, processes, and transmits information about a particular stimulus through stereotyped electrical discharges called action potentials or spikes. The signals from the stimulus are transformed into sequences of these spikes at an early stage of processing within the central nervous system. Spike trains are the starting point for virtually all of the processing performed by the brain (Kandel, 2000; Dayan and Abbott, 2001). Characterizing the relationship between the stimulus and the spike trains is an important issue in neuroscience to understand how the brain works in response to the stimulus. Many studies have been done to understand the relationship between stimulus and spike trains (Espinosa and Gerstein, 1988; Gochin et al., 1990, 1991; Eggermont, 1991; Lindsey et al., 1992c; Vaadia et al., 1995; Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996; Li et al., 1999; Shannon et al., 2000; Louie and Wilson, 2001).

Besides the relationship between stimulus and spike trains, it is also important to study the functional connectivity between spike trains in response to a particular stimulus. This is a challenging problem in neuroscience which needs statistical methods to analyse multiple spike trains (Brown et al., 2004). To study the functional connectivity

of the spike trains it is obligatory to observe the spiking activity of multiple single neurons recorded simultaneously.

Recent advances in multi electrode (micro electrode) neural recording systems have made it possible to record spiking activity from a large number of neurons simultaneously (Boven et al., 2006). Despite the pioneering work by Thomas et al. (1972), Wise et al. (1970), and Gross (1979), a remarkable step forward in Multi Electrode Array (MEA) applications has been achieved only over the last ten years. With the advent of affordable computing power (Boven et al., 2006) and commercial MEA hardware and software (Potter, 2001), multi electrode recording technology is now common in neuroscience study. This technology has been applied to record from the hippocampus (Wilson and McNaughton, 1993; Harris et al., 2003), retina and primary visual cortex (Pillow et al., 2008; Jermakowicz et al., 2009), cortical sensorimotor areas in behaving nonhuman primates (Hatsopoulos et al., 1998; Nicolelis et al., 2003; Riehle et al., 1997), and humans (Hochberg et al., 2006; Truccolo et al., 2008a, 2008b; Truccolo et al., 2010).

The detection and identification of neural spiking activity from multi electrode technology is an important problem that is a prerequisite for studying multiple spike trains (Lewicki, 1998). There are three stages between the multi electrode recording and the identification of spikes: (i) detection of spikes, (ii) determination of the number of neurons being recorded and (iii) assignation to the neurons that produced the spikes (Brown et al., 2004). These three steps together are termed as spike sorting. There are many algorithms for spike sorting (Fee et al., 1996; Lewicki, 1998; Harris et al., 2000; Quiroga et al., 2004; Shahid and Smith, 2008; Shahid et al., 2010). Different algorithms produce different results due to non-stationary background noise, electrode drift and

proper spike alignment. The decision about which one is appropriate depends on the requirements of the experiment (Lewicki, 1998).

2.2 Spike train analysis

After assigning the spike train to the corresponding neuron, the next step is to analyse the functional connectivity of these spike trains. The analysis of functional connectivity can be divided into two groups, namely: (i) pair-wise spike train analysis and (ii) multiple spike train analysis.

2.2.1 Pair-wise spike train analysis

In neuroscience, the Cross-Correlation Function (CCF) is a widely used measure of functional connectivity between spike trains (Perkel et al., 1967). CCF has been applied to many neural systems to make powerful inferences about functional connectivity. This statistical technique is used for testing the independence of two spike trains using the theory of stochastic point processes. This technique is also applied to assess oscillation, propagation delay, effective connection strength, synchronization, and spatiotemporal structure of a network (Konig et al., 1995; Brown et al., 2004). In order to make statistically significant judgements of the CCF, Brillinger (1976c) introduced a normalization technique of the CCF with the confidence interval. Peaks exceeding the confidence interval of the CCF are considered as significant.

Although the CCF has been widely used in neuroscience study over the last few decades this technique has several limitations. This technique assumes that the two spike trains are stationary which can be hard to justify in many cases (Brown et al., 2004). The use of CCF also assumes that spike trains are sufficiently large in sample size (Aertsen and Gerstein, 1985). The result of cross-correlation function is inaccurate if the sample size of the recorded spike trains is small (Shao and Tsau, 1996). The reason for the

inaccuracy is that the confidence interval is too wide to find the significant connection. Peaks in the CCF usually indicate the interaction of spike timing between neurons but there are two biological plausible ways that can draw peaks similar to the CCF and interpreting the interaction between spike trains may be a problem (Brody, 1999).

A related measure of functional interaction between spike trains is the cross intensity function (Cox and Lewis, 1972; Brillinger, 1976b, 1992). This function estimates the spike rate of one neuron at different lags relative to the spiking activity of a second neuron. There are other methods to characterize the relationship between spike trains. These include product densities, cumulant densities, cumulant spectra, methods of moments (Bartlett, 1966; Brillinger, 1975a, 1975b) and coherence (Brillinger, 1976a, 1992). These techniques are usually applied to characterize the dependencies between pairs of neurons at a time, ignoring possible effects from other neurons (Okatan et al., 2005), which lead to erroneous functional connectivity in many cases.

Since behaviour provides a key window into neural responses, it is of great interest to know how relationships between neurons vary as a function of stimuli or external events. A statistical display joint peri-stimulus-time histogram (JPSTH) (Gerstein and Perkel, 1969), which reflects the dynamics of the correlation, is used for this analysis. This technique is a logical extension of the peri-stimulus-time histogram (PSTH). It is a two-dimensional histogram of the joint occurrence of the pairs of spike trains. The main diagonal of the JPSTH displays the observed rate at which both neurons fire simultaneously. A modification of the JPSTH, the normalized JPSTH (Aertsen et al., 1989) is also used. The JPSTH has been used to investigate relationships between neurons (Vadia et al., 1988; Eggermont, 1994).

Like the CCF, the JPSTH has some drawbacks. The major drawback is applying the statistical calculation on nonstationary data. However, recent advances in the JPSTH

have identified some promising approaches to mitigating the nonstationarity problem. Another problem is the difficulty in interpretation of small quantities of data. Like CCF, this technique also applies to characterizing the dependencies between pairs of spike trains at a time, ignoring possible effects from other spike trains.

Besides the CCF and JPSTH, there are some techniques which are used to find the interaction of two spike trains such as partial correlation analysis (Eichler et al., 2003) and partial cross-correlation matrices (PCCM) (Stark et al., 2006). As with the CCF, these methods requires stationary spike trains and sufficiently large number of spikes in the spike train.

Pair-wise correlation methods discussed above measure the statistical association of a single spike train to each member of the ensemble separately but for simultaneously recorded spike trains one should be able to make inference about the collective properties of all spike trains. Gravitational clustering is one approach to an efficient search for evidence for interactions among simultaneously recorded spike trains. Although the representation is still fundamentally pair-wise, it allows evaluation of all observed spike trains at the same time. This method can be used for nonstationary data.

Gravitational clustering is a direct visualization method for examination of dynamic interactions among a group of simultaneously recorded neurons. In the gravitational clustering algorithm (Gerstein & Aertsen, 1985; Gerstein et al, 1985), the activity of neurons is mapped into motions of particles in Euclidean space of the appropriate dimension. The forces exerted on particles by others are due to 'charges' that represent interactions between the corresponding neurons. The particles are allowed to move about until they begin to cluster as a result of the charges and the resulting aggregation of the particles into smaller subgroups then presumably represents the functionally related, or cooperative, subgroups that are sought. Baker and Gerstein (2000) presented

several modifications to the gravitational clustering approach to improve its sensitivity in detecting neural synchronization.

An implementation of the original Gravity Transform algorithm is presented by Stuart et al. (2002) along with simulation results. The gravitational clustering approach has been successfully applied to a range of experimental data (Aertsen et al., 1987, 1991; Gochin et al., 1990; Lindsey et al., 1989, 1992a, 1992b, 1997; Lindsey 2001; Maldonado and Gerstein 1996; Gerstein et al., 1998; Arata et al., 2000; Martin 2001; Morris et al., 2003). Lindsey and Gerstein (2006) presented a new three dimensional display method that allows possible relationships between the observed spike trains.

Spike train similarity or dissimilarity measures are important tools in quantifying the relationship between pairs of spike trains. The use of distance between pairs of spike trains known as spike train metric can be used to find similarity or dissimilarity among pairs of spike trains. If the distance between a pair of spike trains is small enough, one can assume that the pairs of spike trains are identical. Several measures of spike train similarity based on distance have been proposed for the analysis of spike trains. Among these measures Victor and Purpura (1996, 1997) introduced a distance measure based on a cost function. This function evaluates the cost needed to transform one spike train into the other, using certain elementary steps. They define the distance between two spike trains in terms of the minimum cost of transforming one spike train into the other by using three basic operations: spike insertion, spike deletion and spike movement. In their work they proposed several spike train distances: spike time distances, spike interval distances and spike count distances. Spike time distances are further extended to multineuronal data by Aronov et al. (2003). There are a number of applications of spike time distances on the electrosensory system (Kreiman et al., 2000), vision (Mechler et al., 1998; Keat et al., 2001; Reich et al., 2001; Reinagel and Reid, 2002;

Grewe et al., 2003; Samonds et al., 2003; Samonds and Bonds, 2004; Chichilnisky and Rieke, 2005), the auditory system (Machens et al., 2001), chemical senses (MacLeod et al., 1998; Di Lorenzo and Victor 2003), and the motor system (Vaknin et al., 2005).

Another metric proposed by van Rossum (2001), measures the Euclidean distance between the two spike trains after convolution of the spikes with an exponential function. A generalization of van Rossum (2001) measures to multiple spike trains is proposed by Houghton and Sen (2008). Schreiber et al. (2003) proposed a spike train metric which is derived from the correlation-based measure. Kreuz et al. (2007) proposed a spike train metric which is based on the inter-spike intervals (ISI) of the spike train. Paiva et al. (2009) compare different spike metric measures for the analysis of spike trains. The results reveal that no single measure performs the best or consistently throughout for spike train analysis.

2.2.2 Multiple spike train analysis

There are several methods in the literature for multiple spike train analysis. One of the methods is the unitary event analysis (Grun et al., 2002a). This method is designed to detect coincident spike patterns between two or more simultaneously recorded spike trains and to assess the significance of the patterns. The statistical significance of a pattern is evaluated by comparing the number of occurrences to the number expected on the basis of the firing rates of the spike trains. Proper formulation of the null hypothesis and the derivation of the corresponding count distribution are important steps in this method. This method allows one to analyse correlations not only between pairs of spike trains but also between multiple spike trains, by considering the various spike patterns across the spike trains. In addition, this method allows one to extract the dynamics of correlation between the spike trains by performing the analysis in a time-resolved manner.

Its application has provided important insights into principles of information processing in the cortex in a number of studies: in the visual (Maldonado et al. 2008), the prefrontal (Grun et al. 2002b), and the motor cortex (Riehle et al. 1997, 2000; Grammont and Riehle 1999, 2003; Kilavik et al. 2009). The technique for the detection of unitary events is based on the assumption that the firing rates of the neurons are stationary within the analysis time window. An adjustment for the nonstationary firing rate is also proposed (Grun et al., 2002b). Like nonstationarity of firing rate, experimental data may also exhibit nonstationarity across trials. For this case a nonparametric method was presented (Grun et al., 2003) for the analysis of unitary events in case of experimental data.

Another method based on the estimation of higher order correlations has been suggested by Martignon et al. (1995; 2000) for the analysis of multiple spike trains. This technique is aimed at estimating a huge amount of parameters and, therefore, requires very long recordings and can be applied to a relatively small number of spike trains (about ten spike trains). This approach has been further developed in Staude et al. (2010a) where a Cumulant-Based Inference of higher-order Correlations (CuBIC) method has been presented. This method estimates the low order cumulants and is able to decide whether the high order correlations are needed. Thus, both a direct calculation of higher-order correlations and a requirement of a large sample size might be avoided. Staude et al. (2010b) reported a modified version of the CuBIC method, this version is based on a statistical model which includes the non-stationary compound Poisson process.

Study of functional connectivity using Maximum Likelihood (ML) is a useful method for the analysis of multiple spike trains. This method estimates the probability of a spike occurring as a result of multiple influences from other spike trains. Using the ML function, the algorithm calculates the regression parameters, which characterise the

strength of the influences. Brillinger (1988) and Chornoboy et al. (1988) presented such a method where functional relationships between neurons can be detected and modelled. The method proposed by Chornoboy et al. (1988) is based on a point process model which involves stochastic intensities and an additive rate function. This method is suitable for the investigation of the presence or absence of functional relationships between neurons. Although they restrict their analysis to spike train data, the method can be generalized to include external covariates such as sensory stimuli or motor behaviour other than spike trains. A review of maximum likelihood methods and their validity is presented by Pillow (2007).

There are several models that are used to characterize the functional relationship between external variables and neural spike trains. The simplest model is the linear–nonlinear Poisson (LNP) model (Simoncelli et al., 2004). The LNP neuron model consists of a linear filter, a static nonlinear transfer function and a Poisson spike generating mechanism. To determine the neural response to a given stimulus, the stimulus is first convolved with the linear filter. Subsequently, the filter output is converted into an instantaneous firing rate via a static nonlinear transfer function, and finally spikes are generated from an inhomogeneous Poisson-process according to this firing rate. This model has a number of desirable features, including conceptual simplicity and computational tractability.

More recent work has focused on extending the simple LNP model to include spike-history effects, such as refractoriness, burstiness or adaptation. The extension of the LNP model is known as the generalized linear model (GLM) (Paninski, 2004; Paninski et al., 2007; Truccolo et al., 2005; Okatan et al., 2005; Pillow 2007; Stevenson et al., 2008; Pillow et al., 2008). The GLM generalizes the LNP model to incorporate feedback from the spiking process, allowing the model to account for history-dependent

properties of neural spike trains. In this model each cell's input is described by a set of linear filters: a stimulus filter or spatio-temporal receptive field, a post-spike filter that captures dependencies on spike-train history (for example, refractoriness, burstiness and adaptation), and a set of coupling filters that capture dependencies on the recent spiking of other cells. For each neuron the summed filter responses are exponentiated to obtain an instantaneous spike rate.

Several variations and extensions of the basic GLM have been used to characterize the interaction between stimulus and spike trains (Stevenson et al., 2008). Truccolo et al. (2005) proposed a point process-GLM to relate a neuron's spiking probability to three typical covariates: the neuron's own spiking history, concurrent ensemble activity, and extrinsic covariates such as stimuli or behavior. This parametric model uses the conditional intensity function to define a neuron's spiking probability in terms of the covariates. The discrete time likelihood function for point processes is used to carry out model fitting and model analysis of multivariate single unit activity data simultaneously recorded from the motor cortex of a monkey.

Okatan et al. (2005) introduced a GLM to estimate the functional connectivity of stochastic neural networks based on a discrete time version of the approach developed by Chornoboy et al. (1988). This model is applied to the analysis of simultaneously recorded spike trains of a population of rat hippocampal place cells. Kulkarni and Paninski (2007) developed a multivariate point-process model in which the firing rates of the neurons depend on the experimentally controlled stimulus history, the spiking history of the observed neurons, and the common input from an unobserved population of neurons. Czanner et al. (2008) presented a state-space generalized linear model (SS-GLM) to formulate a point process representation of between-trial and within-trial neural spiking dynamics for analysis of multiple trial neural responses. They illustrate

their model using simulated hippocampal spiking activity and actual neural spiking activity of six hippocampal neurons recorded in a macaque monkey. Eldawlatly et al. (2009) introduced an approach for identifying functionally interdependent neurons that does not consider the interval length over which the functional relationship exists. This approach is specifically tailored to large-scale ensemble recordings obtained across multiple cortical areas where identifying the dynamics of neural circuits is highly desirable. A number of recent studies (Pillow, 2007; Pillow et al., 2008; Paninski et al., 2007) use the GLM to investigate the influence of sensory stimuli to spiking activity of neural population.

There are some limitations of the GLM approach. For example, the result of analysis depends on the window (bin) size. The window is used to find estimates of parameters describing influences to the spike train (Eldawlatly et al., 2009) and these estimates vary with the bin size. Also, GLM might contain many parameters resulting in the optimization problem (finding the maximum of likelihood) having a non-unique solution (Stevenson et al., 2008; Chornoboy et al., 1988). A standard approach to resolve this difficulty is to incorporate a prior knowledge about the nature of the inference. There are different methods to deal with this problem: regularization techniques, Bayesian approach, calculation of the maximum a posteriori (MAP) estimate, etc. (Paninski, 2004; Rigat et al., 2006; Gerwinn et al., 2007; Stevenson et al., 2009).

A number of alternative methods have been developed for modelling multi-neuron spike train data and inferring functional connectivity. Utikal (1997) proposed a stochastic model based on counting process intensities in order to analyse the firing times of an ensemble of neurons. The counting process intensity for a neuron is used to characterize the probability of a spike given the neuron's backward recurrence time as well as the

backward recurrence times of some or all of the other neurons in the ensemble. Another method of measuring associations among simultaneously recorded neural spike trains is the combinatorial method (Lee and Wilson, 2002, 2004). This method explores a particular sequential firing pattern based on the relative firing order. It is useful in applications such as memory of a sequence of experienced events, a sequence of actions in an overall movement, or sequential recruitment of different brain areas during a task. This method is applied to quantify memory traces of sequential experience in the rodent hippocampus during subsequent slow wave sleep (SWS) (Lee and Wilson 2002).

Stuart et al. (2005) presented a visualization technique called correlation grid to analyse synchronous firings of simultaneously recorded multiple spike trains. The fundamental idea of this technique is to arrange spike trains into clusters that are functionally connected and display them in a symmetrical grid. A measure of distance, based on normalized CCF of two spike trains is used to perform the cluster analysis. Nykamp (2007) presented an approach to distinguish between causal connection and common input connection among neurons in a network. This approach is based on modelling the relationship between the activity of neurons, their history dependence and measurable external variables such as stimulus.

2.3 Connectivity study using graph theory

Graph theory is a branch of mathematics that originated with Leonhard Euler's famous 1736 treatment of the Konigsberg bridge problem. Today, its applications are extremely broad, ranging from urban planning and traffic control to epidemiology, financial planning, internet search engines, and the analysis of complex biological systems from ecological to molecular scales (Barabási and Oltvai, 2004). Graph theory has been applied to study the brain connectivity over the past ten years as well as other biological networks, e.g. cellular metabolism, gene regulation, or ecology (Sporns, 2007, Bullmore

and Sporns, 2009). Graph theory is applied to anatomical, functional and effective brain connectivity by a connection matrix or adjacency matrix. This connection matrix is used to identify different graph theoretical measures such as degree, motif, clustering coefficient, path length and centrality measure as well as others (Reijneveld et al., 2007; Bullmore and Sporns, 2009; Rubinov and Sporns, 2010).

The term anatomical connectivity refers to the set of physical or structural connections linking neuronal elements. Functional connectivity captures deviations from statistical independence between neuronal elements (Friston et al., 1993; Friston, 1994). The basis of all functional connectivity is time series data from neural recordings. These data may be extracted from functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG) or multielectrode array (MEA). Functional connectivity can be estimated in a variety of ways, for example through computing CCF, mutual information or spectral coherence (Sporns, 2007). Most of the brain connectivity studies to date have been based on measures of functional connectivity (Bullmore and Sporns, 2009). Effective connectivity describes the network of causal effects between neural elements (Friston, 1994; Buchel and Friston, 2000). Various techniques for extracting effective connectivity have been pursued such as structural causal modelling, dynamic causal modelling and Granger causality (Bullmore and Sporns, 2009).

Watts and Strogatz (1998) studied the anatomical connectivity of the nervous system of C. elegans. Two graph theory measures such as path length and the clustering coefficient were studied in this study. Sporns and Zwi (2004) studied data sets of macaque visual and cat cortex. They computed scaled values of path length, clustering coefficient and cycles. Some of the same data studied in Watts and Strogatz (1998) and Sporns and Zwi (2004) were re-investigated for the presence of motifs by Sporns and

Kotter (2004). Kaiser and Hilgetag (2004) studied the edge vulnerability of macaque and cat cortex, protein-protein interaction networks, and transport networks. The average shortest path length was used as a measure of network integrity and four different measures were used to identify critical connections in the network.

The first attempt to apply graph theoretical concepts to fMRI was a methodological paper by Dodel et al. (2002). In this methodological study, graph theory was used as a new approach to identify functional clusters of activated brain areas during a task. Starting with blood oxygen level dependent (BOLD) time series of brain activity a matrix of correlations between the time series was computed. Eguiluz et al. (2005) and Chialvo (2004) were the first to study clustering coefficients, path length and degree distributions in relation to fMRI data. They studied fMRI in 7 subjects during three different finger tapping tasks and derived matrices of correlation coefficients from the BOLD time series. In the study (Salvador et al., 2005a), fMRI measures were recorded in 12 healthy subjects and a matrix of partial correlations was obtained from the BOLD time series. The clustering coefficient and path length were studied using graph theory. Similar graph theory methods were studied by Salvador et al. (2005b) where the connection matrix was obtained by the partial coherence and normalized mutual information measure. Extensive graph theory methods were performed by Achard et al. (2006) on the data obtained from Salvador et al. (2005b). Here, wavelet analysis was used to study connectivity patterns. The global and local efficiency measures were applied in an fMRI study in 15 healthy young and 11 healthy old subjects (Achard and Bullmore, 2007). The connection matrix was based upon wavelet correlation analysis.

Data derived from fMRI experiments are very suitable for graph analysis because of their high spatial resolution, In contrast, spatial resolution is more problematic with EEG and MEG techniques. However, these techniques do measure directly the electro

magnetic field related to neuronal activity, and have a much higher temporal resolution. The first application of graph theory to MEG was published in 2004 (Stam, 2004). In this experiment correlations between the time series of the 126 artefact-free channels were analysed with synchronization likelihood (SL), a non-linear measure of statistical interdependencies (Stam and van Dijk, 2002; Montez et al., 2006). Clustering coefficient and path length were analysed in this study. Graph theoretical properties of MEG recordings in healthy subjects were studied more extensively by Bassett et al. (2006). The authors applied graph theory techniques such as clustering coefficient, path length and betweenness centrality to MEG recordings in 22 healthy subjects. In this study wavelet analysis was used to obtain correlation matrices. Bartolomei et al. (2006) applied clustering coefficient and path length to MEG resting state recordings in a group of 17 patients with brain tumours and 15 healthy controls.

The first application of graph theory to EEG was published in 2007 (Stam et al., 2007). Here a group of 15 Alzheimer patients was compared to a non-demented control group of 13 subjects. EEG recorded from 21 channels were analysed with the SL. The authors computed the clustering coefficient and path length from the connection matrix. In two related studies Micheloyannis et al. (2006a, 2006b) applied graph theory to 28 channel EEG recorded during a 2-back working memory test. In both studies EEG filtered in different frequency bands was analysed with SL. There are several studies that applied clustering coefficient and path length to the data obtained from EEG recordings (Posthuma et al., 2005; Wu et al., 2006; Ferri et al., 2007; Ponten et al., 2007 and Smit et al., 2008).

2.4 Importance of this research

In this Chapter several techniques for analysing the functional connectivity from simultaneously recorded multiple spike trains have been reviewed. Although these methods are used for studying experimental data and demonstrate some important results, all these methods have a number of limitations which should be taken into account. For example, the majority of the existing methods of analysing functional connectivity require stationarity of spike trains. This assumption may not be met for data simulated by a large neural network and is even less likely to be met for experimental data. A second limitation is that some methods require long spike trains. These two limitations contradict each other: the longer spike trains are, the less likely it is that they will be stationary. Yet another common limitation is that the majority of methods require a small time window (bin) for data analysis. Of course, the result of analysis depends on the bin size. The widely used CCF method is a bin based method and is applied under the assumption of stationary spike trains. Similarly, recently introduced GLM techniques that are used for identifying functional connectivity from experimental data are bin based methods, also GLM as a rule requires long spike train data. Therefore, development of new statistical techniques for analysing multiple spike trains and deriving the functional connectivity is a very actual and important problem. This problem is timely because MEA are now available for experiments in neuroscience and this tool enables researchers to simultaneously record multiple spike trains.

In this thesis a new statistical method is developed to overcome the limitations of the existing spike train analysis methods. This statistical method is based on the mathematical idea of the Modulated Renewal Process (MRP) and is very useful for identifying functional connectivity in generated and experimental data. Functional connectivity identified by this statistical method is a useful tool for the graph theoretical

brain connectivity analysis, while currently graph theoretical brain connectivity analysis is based on the pair-wise correlation methods. In this thesis, another useful method for identification of functional connectivity developed by Stuart et al. (2005) is modified in such a way that this modified method enables researchers to distinguish the direct connections from spurious connections using an automatic algorithm.

Chapter 3

Statistical techniques to improve the analysis of functional connections based on pair-wise cross-correlation function

Although the CCF is widely used in neuroscience to analyze multiple spike trains, it is a difficult problem to identify functional connectivity using a pair-wise method which only considers a pair of spike trains at a time and does not take into account influences from other spike trains. Analyzing pairs of spike trains and finding a statistically significant influence from one spike train to another, it is difficult to distinguish this influence. This influence may occur as a result of direct connection between this pair, due to a common source to both spike trains or due to an intermediate spike train. In this chapter a statistical method is developed for classification of significant influences into three groups: 'direct connection', 'connection due to a common source' and 'indirect connection'. It is shown that this method is efficient and significantly improves a result of study of functional connectivity by a pair-wise method.

3.1 Introduction

Identifying functional connectivity between spike trains is important for understanding how the brain works. In neuroscience a common widely used measure of functional connectivity between spike trains is the cross-correlation function (Perkel et al., 1967). A significant peak in a cross-correlation function can be interpreted as a functional connection between spike trains (Aertsen et al., 1989) and the corresponding time of the peak can be interpreted as a time shift in spike propagation (Nikolic, 2007). A peak in a cross-correlation function indicates that there is a high probability to find a spike in one

spike train due to spike in another spike train with some time shift. This significant peak in the cross-correlation function indicates that the null hypothesis on independence of two spike trains is not supported by data and should be rejected. Consequently there is an influence from one spike train to another. Interpretation of this influence in terms of functional connectivity is not an easy problem because this influence can be considered as a direct connection between two spike trains or this influence can be considered as a result of some common source to both spike trains. Another possibility is to interpret the influence as an indirect connection, i.e. connection via some intermediate spike train.

A 'direct connection' is a connection where one spike train modulates the firing pattern of another spike train directly. An 'indirect connection' is a spurious connection due to connectivity via an intermediate spike train. A 'common source' is a spurious connection due to an influence from the common source to both spike trains in a pair. Analyzing functional connectivity of a large number of spike trains using cross-correlation function, it is difficult to distinguish the direct connection, indirect connection and common source (Dahlhaus et al., 1997; Eichler et al., 2003; Makarov et al., 2005; Nykamp, 2005; Stivenson et al., 2008; Park et al., 2008; Nedungadi et al., 2009). The Correlation Grid (Stuart et al., 2005) is a method which is used for investigating the functional connectivity of a large number of spike trains using cross-correlation function. Correlation grid has been successfully used for study of functional connectivity, however, the correlation grid cannot automatically distinguish direct and spurious (both indirect and common source) connections. The aim of this chapter is to present a method for analyzing functional connectivity of a large number of spike trains using cross-correlation function which can differentiate direct connections from spurious (indirect and common source) connections by an automatic algorithm. This method is called the modified correlation grid (MCG).

3.2 Cross-correlation function (CCF)

The cross-correlation function works on a pair of spike trains A and B, one spike train is considered as the target spike train (A) and another one is considered as the reference spike train (B). A correlation window consists of $(2 * u + 1)$ bins of small time length h is considered. For each spike of the reference spike train B, this correlation window is set in such a way that a centre of the middle bin of the correlation window coincides with this spike of the reference spike train and there are u bins on the left and there are u bins on the right (Fig. 3.1). The counting function $n_{AB}(v)$ counts and accumulates the number of spikes of the target spike train A falling in the correlation window for the spikes of the reference spike train B. The counting function $n_{AB}(v)$ is calculated over the experiment time T.

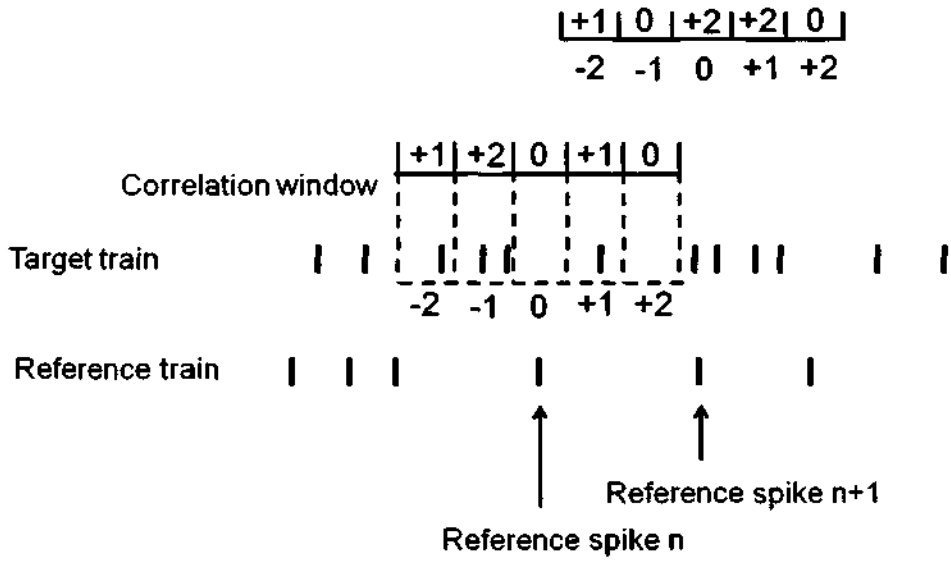


Figure 3.1: Counting procedure of the target spikes that fall within the correlation window for the reference spike.

To test the independence of two spike trains Brillinger (1976c) considers the estimate $\hat{\rho}_{AB}(v) = \sqrt{\hat{p}_{AB}(v)/\hat{p}_A\hat{p}_B}$, where $\hat{p}_{AB}(v) = n_{AB}(v)/2hT$, $\hat{p}_A = n_A/T$ and $\hat{p}_B = n_B/T$ and normalises the counting function $n_{AB}(v)$ accordingly. Here, n_A, n_B denote the number of spikes in the spike trains A and B, respectively. For a large sample size the

random variables $\hat{\rho}_{AB}(v)$ are independent and the distribution of each of them is the normal with the mean $m = \sqrt{\hat{p}_{AB}(v)/\hat{p}_A\hat{p}_B}$ and the standard deviation $s = 1/(2\sqrt{2hT\hat{p}_A\hat{p}_B})$. Therefore, in the case of two independent spike trains the mean of $\hat{\rho}_{AB}(v)$ equals to one (because in independent case $\hat{p}_{AB}(v) = \hat{p}_A\hat{p}_B$). To test the null hypothesis H_0 that two spike trains are independent, the boundaries of the confidence interval at the significance level α are plotted by two horizontal lines at levels $1 \pm Q_{Cr}^{\alpha}/(2\sqrt{2hT\hat{p}_A\hat{p}_B})$, where Q_{Cr}^{α} is a critical value of the normal distribution corresponding to the significance level α. If H_0 is correct then all values of the CCF corresponding to different bins fall inside the confidence interval and the estimated value of the CCF $(\hat{\rho}_{AB}(v))$ is zero. If some value of the CCF exceeds the upper boundary of the confidence interval, then the null hypothesis H_0 is rejected and it is concluded that the two spike trains are not independent. The highest value of the CCF exceeding the upper boundary of the confidence interval can be considered as a measure of influence strength from spike train *A* to spike train *B* and the corresponding time can be considered as a time shift (Δ) in spike propagation (Stuart et al., 2005) (Fig. 3.2).

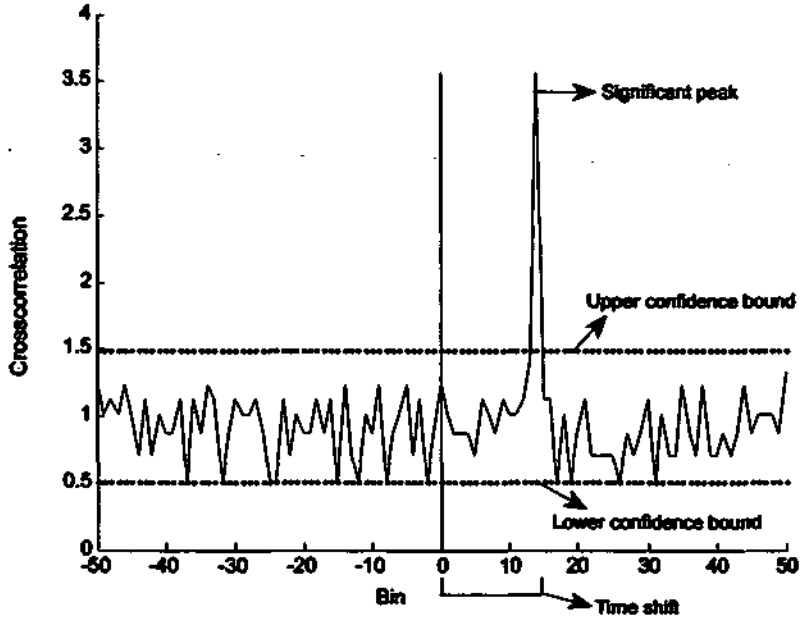


Figure 3.2: An example of a cross-correlation function with confidence interval, significant peak and time shift.

3.3 Classification of functional connection

The interpretation of the CCF results is a difficult task. If the CCF shows that there are no significant peaks, then it can be considered that there is no connection between two spike trains. However, if there is a significant peak of the CCF then there is an influence from one spike train to another. This influence can be interpreted either as a direct coupling or as a spurious connection.

The two measurements from the CCF that can be used for distinguishing direct and spurious connections are the highest significant peak (ρ) (Aertsen et al., 1989) and the corresponding time shift (Δ) (Nikolic, 2007). The following examples show different connection types and corresponding values of ρ and Δ.

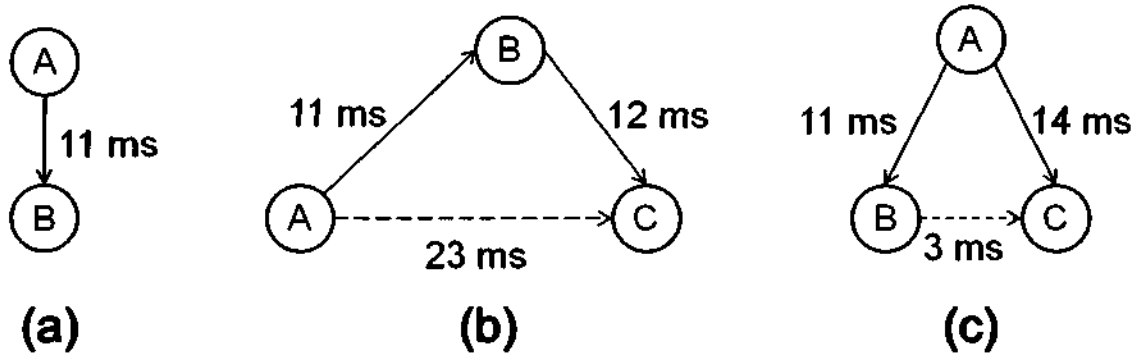


Figure 3.3: Schematic diagram of **(a)** Direct connection where neuron *A* is directly connected to neuron *B*. **(b)** Indirect connection where neuron *A* is connected to neuron *C* through neuron *B*. **(c)** Common source where neuron *A* is connected to both neuron *B* and *C*.

1. **Direct connection**. Fig. 3.3(a) shows an example of a direct connection from neuron *A* to neuron *B* with delay of spike propagation of 11 milliseconds. Two spike trains corresponding to this connectivity have been generated using the Enhanced Leaky Integrate and Fire model (Borisyuk, 2002) with moderate connection strength. Fig. 3.4(a) shows the CCF for generated spike trains *A* and *B*. The value of significant peak of the CCF is 4.51 and the corresponding time

shift is 11 milliseconds. Note, there is a good correspondence between a time delay of spike propagation in ELIF model and a time shift of CCF.

2. **Indirect connection.** An example of connection diagram with indirect connection is shown in Fig. 3.3(b). A time delay of spike propagation from a neuron *A* to neuron *B* is 11 milliseconds and a time delay of spike propagation from neuron *B* to neuron *C* is 12 milliseconds. Three spike trains corresponding to this connection diagram have been generated using the ELIF model with moderate connection strengths. Fig. 3.4(b) shows the result of analyzing spike trains *A* and *C* by the CCF. A height of the significant peak is 1.7 (note, this value is lower than the significant peak of direct connection in previous case) and a corresponding time shift of the CCF is 23 milliseconds (note, this time is longer in comparison with time shift in previous case of direct connection). Thus, a relatively low value of the significant peak and a relatively long time shift compared to direct connections are important characteristics of the 'indirect connection'.
3. **Connection due to a common source**. Fig. 3.3(c) shows a 'common source' connectivity diagram of 3 neurons. Neuron *A* influences both neurons *B* and *C* with delay of spike propagation of 11 and 14 milliseconds, respectively. Three spike trains corresponding to this connection diagram have been generated using the ELIF model with moderate connection strengths. Fig. 3.4(c) shows the CCF for spike trains *B* and *C* where a value of significant peak is 1.93 (note, this value is lower than a significant peak in case of direct connection) and a time shift is 3 milliseconds (note this value of time shift is relatively small).

These examples provide an important guidance how ρ and Δ can be used to justify the result of the CCF analysis corresponding to the direct or spurious connection. The direct connection is characterized by a high value of the significant peak and a relatively short

time shift. 'Indirect connection' is characterized by a relatively low value of the significant peak and a relatively long time shift in comparison with direct connection. 'Common source connection' can be characterized by a relatively low significant peak near zero time shift and both the significant peak and time shift are lower comparison with direct connection. Thus, to derive a diagram of functional connections from the pair-wise CCF analysis, the spurious connections should be ignored and the only direct connections should be included to the diagram.

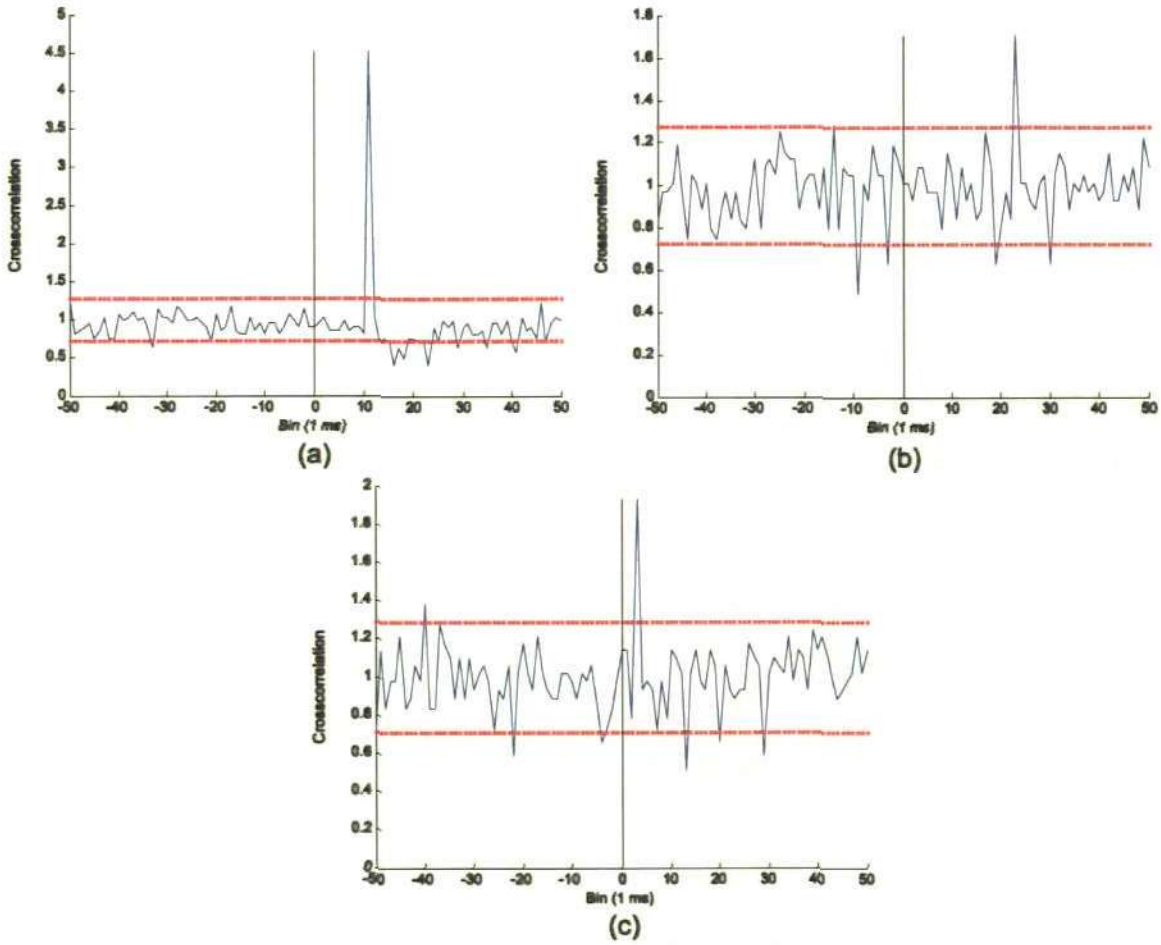


Figure 3.4: Examples of cross-correlation for the three types of connection. **(a)** Direct connection. **(b)** Indirect connection and **(c)** Common source.

All these considerations are relevant to the case of generated data and, of course, they may not completely correct in case of experimental data. Nevertheless, these simple examples provide a useful guidance how to distinguish between direct and spurious connections when deriving the diagram of connectivity from the pair-wise CCF

analysis. Further in this chapter, examples of 10, 15 and 20 generated spike trains with different connection diagrams and distributed parameters are considered and shown that simple ideas which have been described here are useful for deriving a functional connectivity (i.e. connection diagram) from the result of pair-wise CCF analysis.

3.4 Correlation Grid

The fundamental idea of Correlation Grid (Stuart et al., 2005) is to arrange spike trains into clusters of functionally connected spike trains using a measure of similarity. This similarity measure is based on a distance between two spike trains. The distance is calculated using the value of highest significant peak of the cross-correlation function. For multiple spike trains, all pair-wise cross-correlation functions are calculated with a specified bin and correlation window. The main significant peaks are calculated and the results of these cross-correlations are displayed in a matrix format. In the matrix, the magnitudes of the main significant peaks are encoded from white, representing a non-significant peak, to black, representing the largest peak (Fig. 3.5). The rows and columns of the matrix are reordered using a clustering algorithm. A difference between maximum of the main peaks and main peak in the corresponding cross-correlation is considered for clustering the spike trains. After applying the clustering algorithm the most similar spike trains cluster together and show functional connectivity.

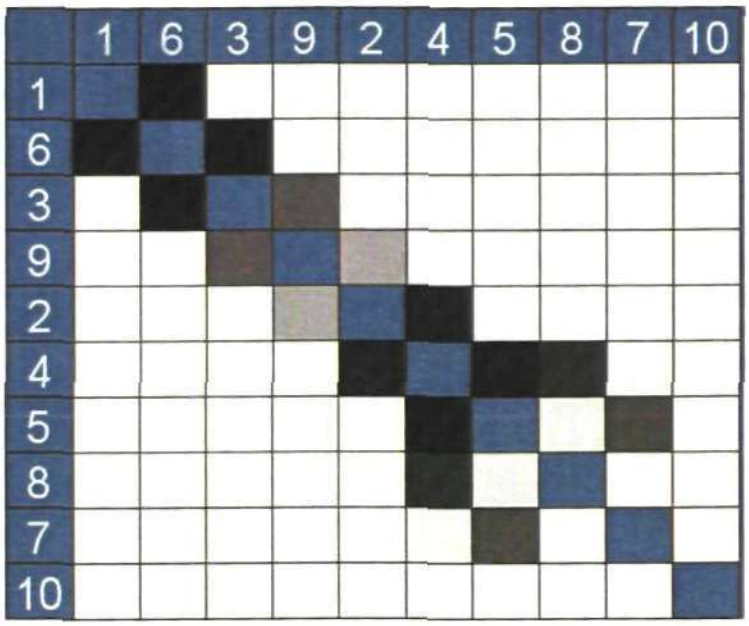


Figure 3.5: An example of a correlation grid to find functional connectivity of multiple spike trains.

Fig. 3.5 shows the functional connectivity of ten spike trains using Correlation Grid method. The functional connectivity obtained by this method is displayed in a symmetrical way. For example, in Fig. 3.5 there is a functional connection between spike trains #1 and #6 and both the cells (1,6) and (6,1) are encoded with black meaning that there is a strong influence between these spike trains and the connection is direct. Similarly, there is a functional connection between spike trains #2 and #9 and the grey color means there is a connection (not strong) between them. It is not easy to identify whether the connection between spike trains #2 and #9 is direct, indirect or due to a common source. To identify this connection the Correlation Grid requires additional in-depth analysis. To overcome this limitation a new statistical method based on the significant peak and time shift of the CCF is developed. This method enables researchers to distinguish the direct connections from spurious (common source and indirect) connections using an automatic algorithm.

3.5 Modified Correlation Grid (MCG)

The steps involved in modified correlation grid method are described below:

3.5.1 Calculation of CCF

Let us consider n simultaneously recorded spike trains and $k = (n^2 - n)/2$ is the number of the CCF which have been calculated. To test the hypothesis of independence of two spike trains, k independent tests are run for the n spike trains. In statistics a multiple comparison problem occurs when a set of simultaneous statistical tests is considered. This problem occurs when all the tests are considered as a family and set the significance level α for the family. For this reason the errors of incorrectly rejecting the null hypothesis are more likely to occur. Several statistical techniques have been developed to prevent this problem. Bonferroni correction is a method that can be used to address the problem. The correction is based on the idea that for testing a set of dependent or independent hypotheses, the significance level α should be adjusted according to the number of tests being performed. So if the significance level for a set of k simultaneous tests is considered to be α, the significance level for each individual tests will be α/k.

Applying Bonferroni correction to the k pair-wise CCFs, the upper and lower boundaries of the confidence interval are calculated for each pair-wise CCF. Any peak that exceeds the upper boundary of the confidence interval is considered as significant. Significant peaks can be found on both positive and negative side of the CCF. Here, significant peak on the positive side of the CCF is considered as a measure of dependence from one spike train to another. For several significant peaks, the highest significant peak is considered as the measure of influence strength. All the highest significant peaks $\rho_{ij}\ (i,j = 1,2,\cdots\cdots,n),\ i \neq j$ and the corresponding time shifts $\Delta_{ij}\ (i,j = 1,2,\cdots\cdots,n),\ i \neq j$ are calculated for n spike trains. Non-significant peaks

ρ_{ij} of the pair-wise CCF supports the null hypothesis (spike trains are independent). Therefore there is no connection from spike train i to spike train j and so ρ_{ij} is not distinguishable from zero. These non significant peaks are not included in the analysis.

3.5.2 Outlier detection

In multiple spike trains there may exist some very strong synaptic connections between neurons. Cross-correlation functions for these neurons show very big significant peaks and they can be considered as outliers. These very big significant peaks are considered outliers because they deviate markedly from the other significant peaks. These outlier connections are considered as direct connections and are not used for the algorithm to distinguish the direct connections from the indirect connections and common source.

In statistics there are two kinds of outlier detection methods, tests of discordancy and outlier labeling methods. Most tests of discordancy need test statistics for hypothesis testing and the tests are usually based on an assumption of some distribution. Some tests of discordancy are for a single outlier and others for multiple outliers. Selection of these tests mainly depends on the numbers and type of target outliers, and type of data distribution (Acuna and Rodriguez, 2004). Though the tests of discordancy are powerful for investigating the outliers, most distributions of real-world data may be unknown or may not follow specific distributions. Another limitation is that the tests of discordancy are susceptible to masking or swamping problems (Acuna and Rodriguez, 2004). Masking problem can occur when few outliers are specified in the test though there are in fact more outliers. These additional outliers may influence the value of the test statistic enough so that no points are declared as outliers. Swamping problem can occur when many outliers are specified in the test though there are in fact few outliers and all the points are declared as outliers.

On the other hand, most outlier labeling methods generate an interval or criterion for outlier detection instead of hypothesis testing, and any observation beyond the interval or criterion is considered as an outlier. There are two reasons for using an outlier labeling method. One is to find possible outliers as a screening device before conducting a test of discordancy. The other is to find the outliers away from the majority of the data regardless of the distribution. When it is difficult to identify the distribution of the data or transform it into a proper distribution, labeling methods can be used to detect outliers (Seo, 2006).

Among several outlier labeling methods, the commonly used method is the Z-score. The Z-score is defined as

$$Z_i = \frac{x_i - \bar{x}}{sd}$$

where $X_i \sim N(\mu, \sigma^2)$, and $\bar{x}$ and sd are the sample mean and sample standard deviation of data. The basic idea of this rule is that if X follows normal distribution with mean μ and variance σ^2, then Z follows standard normal distribution with mean 0 and variance 1. The Z-scores that exceed 3 in absolute value are generally considered as outliers. According to Shiffler (1988), a maximum Z-score is dependent on sample size, and it is computed as $(n-1)/\sqrt{n}$. Since no Z-score exceeds 3 in a sample size less than or equal to 10, the Z-score method is not very good for outlier labeling, particularly in small data sets (Iglewicz and Hoaglin, 1993). Another limitation of this rule is that the standard deviation can be inflated by a few or even a single observation having an outlier value. Thus it can cause a masking problem.

Two estimators are used in the Z-Score, the sample mean and sample standard deviation. These estimators can be affected by a few outlier values or by even a single outlier value. To avoid this problem, another outlier labelling method known as

modified Z-score can be used. In the modified Z-score the median and the median of the absolute deviation of the median (MAD) are employed instead of the mean and standard deviation of the sample, respectively (Iglewicz and Hoaglin, 1993):

$$MAD = median\{|x_i - \tilde{x}|\}$$

where $\tilde{x}$ is the sample median. The modified Z-Score (M_i) is computed as

$$M_i = \frac{0.6745(x_i - \tilde{x})}{MAD}$$

where E(MAD) = 0.6745σ for large data. Iglewicz and Hoaglin (1993) suggested that observations are labeled outliers when $|M_i| > 3.5$. Applying modified Z-score to the significant peaks ρ_{ij} obtained from pair-wise cross-correlation function enables to detect the outliers. In this study, only upper outliers are investigated to indicate that these outliers have very strong connections and can be considered as direct connections. For this reason significant peaks are labelled outliers when $M_i > 3.5$.

3.5.3 Cluster analysis

All the non outlier significant peaks ρ_{ij} and the corresponding time shifts Δ_{ij} are used for the classification of functional connections. For a set of significant peaks ρ_{ij} that do not have outliers with corresponding shifts Δ_{ij}, the typical scatter plot can be observed as in Fig. 3.6. From the Fig. 3.4, it can be assumed that the direct connections can be identified with high significant peak and moderate time shift, common source can be identified with low significant peak and short time shift, and indirect connections can be identified with low significant peak and large time shift. To achieve this classification from the set of significant peaks ρ_{ij} and time shifts Δ_{ij} a clustering algorithm is used.

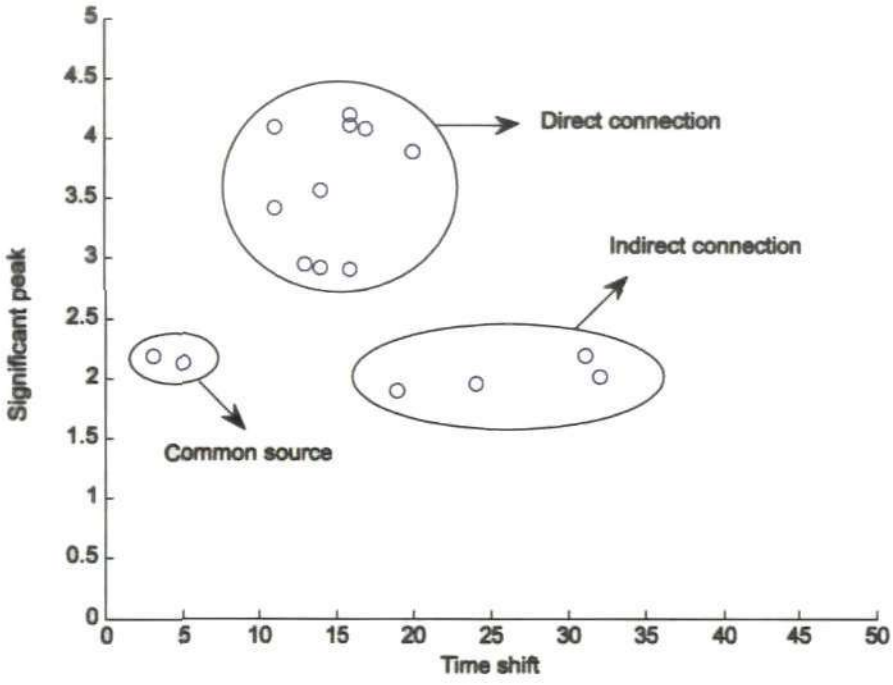


Figure 3.6: An example of a typical scatter plot of a set of significant peaks ρ_{ij} and the time shifts Δ_{ij} where the classification of direct connection, indirect connection and common source are indicated by circles.

A cluster analysis is used for finding clusters of similar objects within a data set. A cluster analysis begins by calculating distances among objects in the data set. For a data set having two or more variables the distances are greatly affected by differences in scale of measurement of the objects. It is a good practice to transform the variables so that they have the similar scales. At the first step, each object represents its own cluster. Clustering begins by finding the two clusters that are most similar, based on the distance, and merging them into a single cluster. The characteristics of this new cluster are based on a combination of all the objects in that cluster. This procedure combining two clusters and merging their characteristics is repeated until all the objects have been joined into a single large cluster.

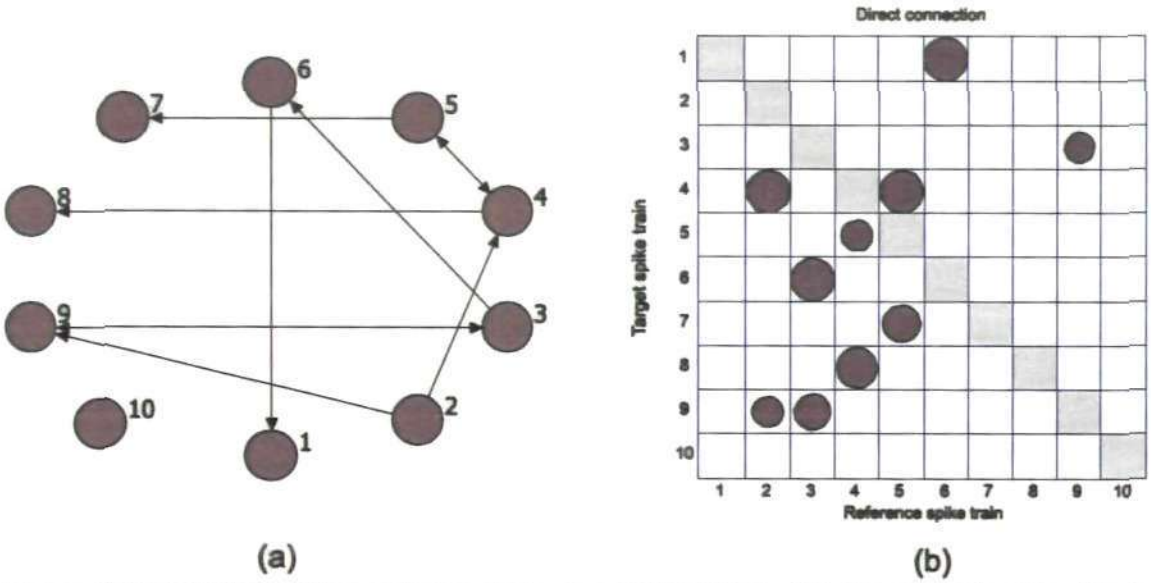
A variety of measures can be used to calculate the distance. For data that show linear relationships, the Euclidean distance is a useful measure. A variety of linkage methods can be used to determine in what order clusters may join. The nearest neighbor or single

linkage method is based on the elements that are most similar. The farthest neighbour or complete linkage method is based on the elements that are most dissimilar. Both of these are based on outliers of distributions, which may not be desirable. The average emphasizes the central tendency of clusters and is less sensitive to outliers.

To apply cluster analysis to the set of significant peaks ρ_{ij} and time shifts Δ_{ij}, both measurements are normalized so that the values of significant peaks and time shifts are in the range between 0 and 1. The normalization is done so that the values of significant peaks and time shifts have the similar scale. A clustering algorithm is applied to these normalized significant peaks and time shifts. In the clustering algorithm the distance between pairs of observations is calculated using the Euclidean distance and the average linkage is used for calculating the distance between two clusters. The average linkage is used due to its less sensitivity to outliers. This clustering algorithm creates three clusters; the cluster of direct connections, the cluster of common source and the cluster of indirect connections.

3.5.4 Classification of significant connections

The outlier connections and the cluster of direct connections are considered as direct connections and displayed in the resulting n-by-n matrix of functional connectivity. In the matrix of functional connectivity rows of the matrix indicate the target spike train and columns of the matrix indicate the reference spike train. In Fig. 3.7, a circle in the cell (i, j) indicates that there is direct connection from spike train j to spike train i. The radius of the circle shows the normalized strength of connection of the pair of spike trains which is the proportional to the height of the significant peak in a pair-wise CCF. Fig. 3.7 shows an example of a matrix of direct functional connectivity for a data set of 10 spike trains generated by the ELIF (Borisyuk, 2002) model. The description, dynamics and parameters of the ELIF model are given in appendix A.



(a) (b)

Figure 3.7: (a) Connection scheme of the ten spike trains. **(b)** An example of a direct functional connectivity matrix of these ten spike trains.

Neuron parameter	**Mean**	**S.D.**
Maximum value of the threshold	44.89	0.45
Threshold decay rate	2.96	0.43
Asymptotic threshold value	14.84	0.65
Amplitude of the noise	4.89	0.41
Noise decay rate	10.00	0.03
Initial value of after spike hyperpolarisation	-28.89	0.43
Soma's membrane potential decay rate	19.76	1.13
External input	-0.05	0.43
Absolute refractory period	5.6	2.7
Connection parameter		
Connection strength	12.18	1.56
Decay rate of postsynaptic potential	4.12	0.47
Time lag of spike propagation	14.8	2.78

Table 3.1: Parameter values of the ELIF model to generate ten spike trains.

The parameter values of the ELIF model for the generation of ten spike trains is given in Table 3.1. From the functional connectivity matrix ten direct functional connections can be observed.

All the significant connections in the common source cluster are further investigated to identify whether these connections are due to common source or not. Two spike trains i and j are taken from the common source cluster and the aim is to identify the connection from spike train i to spike train j. Keeping spike trains i and j fixed all groups of three spike trains $(i, j, k), k = 1,2, \cdots, n$ are detected that have the diagram of connection in Fig. 3.3(c). The triplet time shifts (Δ_{ij}, Δ_{ki}, Δ_{kj}) of these groups are calculated from the CCF. Time shifts (Δ_{ki}, Δ_{kj}) indicate the delay of direct connections from spike train k to spike train i and j, respectively and time shift Δ_{ij} indicates the delay of connection from spike train i to spike train j. The connection from spike train i to spike train j is considered as a common source connection if the time shifts of these spike trains meet the following formula:

$$\Delta_{ij} = |\Delta_{ki} - \Delta_{kj}| \tag{3.1}$$

This procedure is repeated for all significant connections in the common source cluster to identify their connections.

Similarly, all the significant connections in the indirect cluster are further investigated to identify whether these connections are indirect or not. Two spike trains i and j are taken from the indirect cluster and the aim is to identify the connection from spike train i to spike train j. Keeping spike trains i and j fixed all groups of three spike trains $(i, j, k), k = 1,2, \cdots, n$ are detected that have the diagram of connection in Fig. 3.3(b). The triplet time shifts (Δ_{ij}, Δ_{ik}, Δ_{kj}) of these groups are calculated from the CCF. Time shifts (Δ_{ik}, Δ_{kj}) indicate the delay of direct connections from spike train i to

spike train k and from spike train k to spike train j, respectively. Time shift Δ_{ij} indicates the delay of connection from spike train i to spike train j. The connection from spike train i to spike train j is considered as an indirect connection if the time shifts of these spike trains meet the following formula:

$$\Delta_{ij} = \Delta_{ik} + \Delta_{kj} \tag{3.2}$$

This procedure is repeated for all significant connections in the indirect cluster to identify their connections.

To consider the prescribed tolerance (ϵ) in time shifts (ms), common source and indirect connections of formula (3.1) and (3.2) can be written as:

$$|\Delta_{ki} - \Delta_{kj}| - \epsilon \leq \Delta_{ij} \leq |\Delta_{ki} - \Delta_{kj}| + \epsilon \tag{3.3}$$

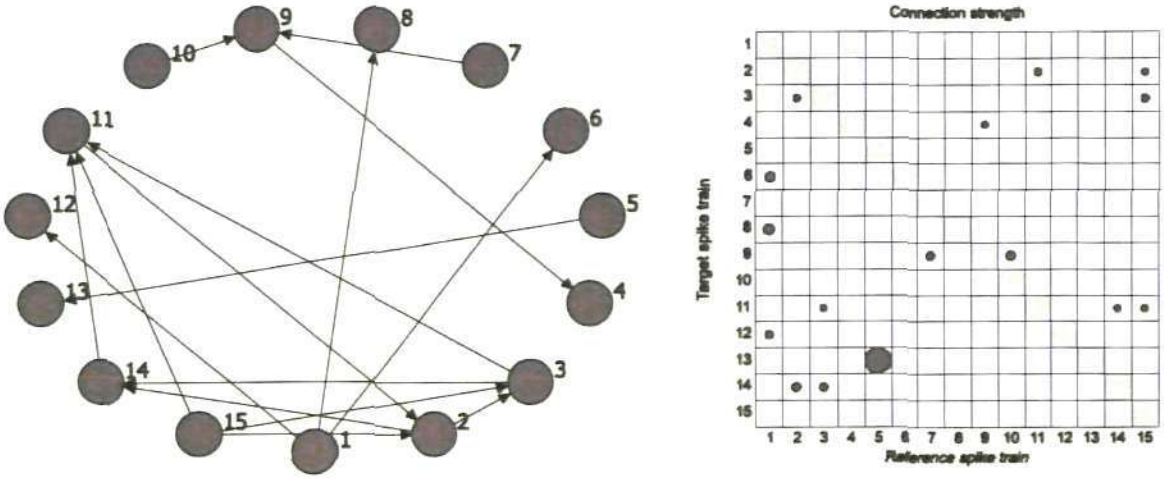
and

$$\Delta_{ik} + \Delta_{kj} - \epsilon \leq \Delta_{ij} \leq \Delta_{ik} + \Delta_{kj} + \epsilon \tag{3.4}$$

The common source and indirect connections obtained from (3.3) and (3.4) are displayed at the n-by-n connectivity matrix in a similar way to direct connection but by different colors.

3.6 Analysis of functional connectivity of fifteen spike trains

The Modified correlation grid method is applied to a simulated fifteen spike trains to identify functional connectivity. The Enhanced Leaky Integrate and Fire (ELIF) model (Borisyuk, 2002) is used for simulations with a given scheme of connections. The parameter values of the ELIF model for the generation of fifteen spike trains is given in Table 3.2. Fifteen spike trains are generated for a period of 30,000 milliseconds using the connection architecture shown in Fig. 3.8(a) with the connection strengths shown in Fig. 3.8(b) in matrix format. The connection from spike train #5 to spike train #13 has the maximum strength.



(a) (b)

Figure 3.8: (a) Connection scheme of fifteen spike trains. There are sixteen non zero connections which are shown by arrows. **(b)** Connection strengths of the sixteen non zero connections in the matrix format.

Neuron parameter	**Mean**	**S.D.**
Maximum value of the threshold	44.82	0.73
Threshold decay rate	2.98	0.38
Asymptotic threshold value	14.24	0.92
Amplitude of the noise	5.13	0.42
Noise decay rate	10.00	0.03
Initial value of after spike hyperpolarisation	-28.98	0.40
Soma's membrane potential decay rate	19.65	0.85
External input	0.01	0.35
Absolute refractory period	5.2	1.61
Connection parameter		
Connection strength	14.38	7.31
Decay rate of postsynaptic potential	2.95	0.75
Time lag of spike propagation	12.5	1.15

Table 3.2: Parameter values of the ELIF model to generate fifteen spike trains.

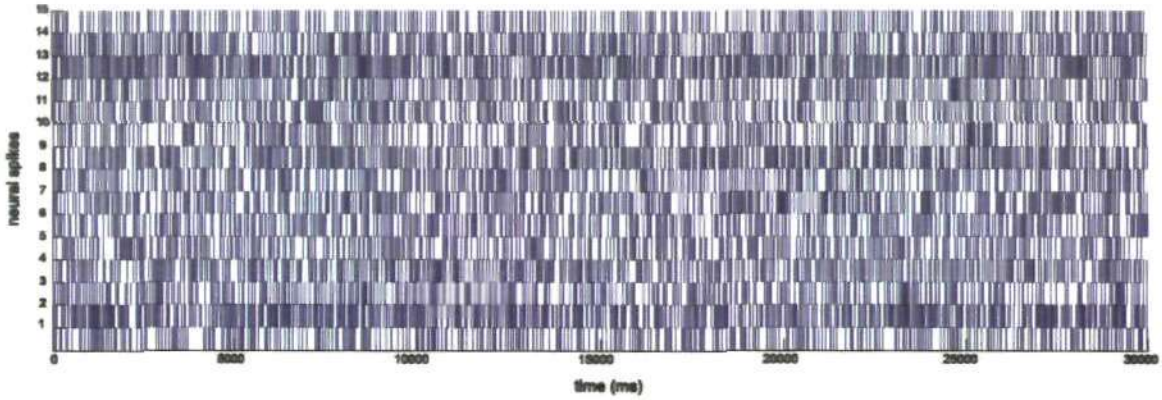


Figure 3.9: Raster plot of fifteen spike trains generated for the connection scheme given in Fig. 3.8 over the duration 30,000 milliseconds.

Fig. 3.9 shows the raster plot of spiking activity of these fifteen spike trains. These spike trains are considered as a data set for analysing the functional connectivity. For analysing the functional connectivity only spike trains are used and assume that the scheme of connections as well as parameter values of neurons and connections are unknown. In this analysis the self coupling is not considered for finding functional connectivity of the fifteen spike trains.

3.6.1 Calculation of CCF

For these fifteen spike trains a total of $(15^2 - 15)/2 = 105$ pair-wise CCF are calculated with a bin size of 1 ms and a correlation window of 100 ms. To test the independence of the pair-wise spike trains the level of significance $\alpha = 0.05$ is considered with the Bonferroni correction. A connection is considered significant if a peak of the cross-correlation function exceeds the upper boundary of the confidence interval. A total of 25 significant connections are found for fifteen spike trains (Table 3.3).

Reference spike train	Target spike train	Peak	Time shift
1	**6**	4.37	13
1	**8**	4.57	15
1	**12**	3.05	11
2	**3**	3.50	12
2	**11**	1.64	27
2	**14**	3.26	14
3	**11**	2.73	12
3	**14**	3.23	12
5	**13**	6.52	12
6	**8**	2.44	2
7	**4**	1.72	23
7	**9**	3.57	11
9	**4**	3.27	12
10	**4**	1.74	25
10	**9**	3.46	13
11	**2**	2.57	13
12	**4**	1.69	23
12	**6**	1.90	2
12	**8**	1.92	4
14	**11**	2.98	13
14	**15**	1.73	40
15	**2**	2.71	12
15	**3**	3.57	14
15	**11**	3.73	11
15	**14**	2.14	26

Table 3.3: Significant connections of the fifteen spike trains with peak and time shift. Connections are indicated from reference spike train to the target spike train.

Table 3.3 shows the significant connections of the fifteen spike trains with corresponding peaks and time shifts. The peaks of the significant connections range from 1.64 to 6.52 and the time shifts range from 2 ms to 40 ms. These significant connections are also shown in a matrix format where the connections are indicated by a cell filled by a circle (Fig. 3.10(a)). The direction of the significant connection in the matrix format is considered from reference spike train to target spike train. For example, spike train #1 has significant connections to spike trains #6, #8, and #12 (first column of the matrix in Fig. 3.10(a)).

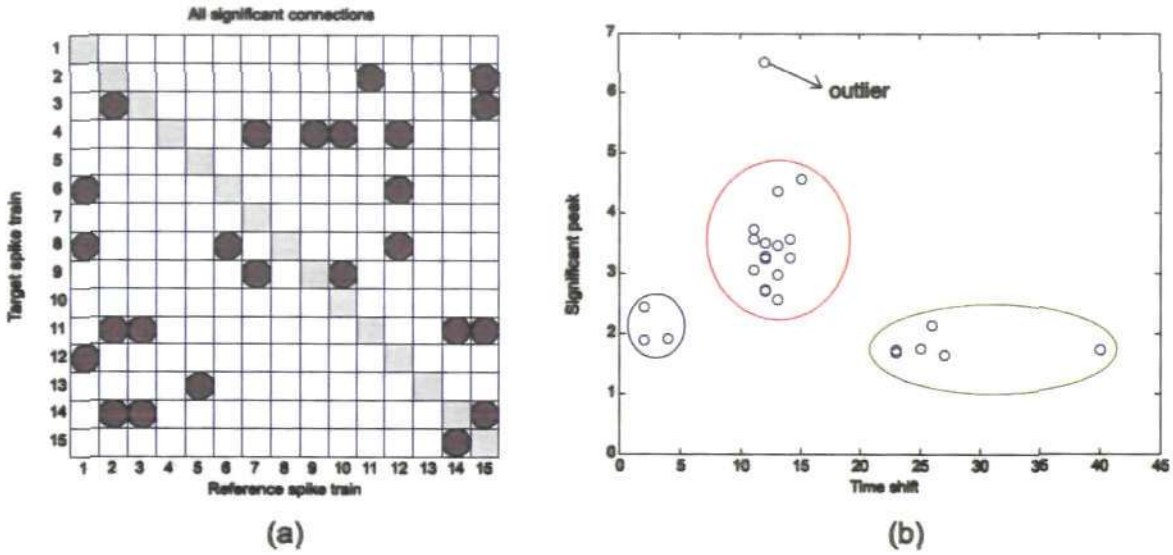


Figure 3.10: (a) Significant connections obtained from pair-wise CCF analysis. A total of 25 significant connections are obtained from the fifteen spike trains. **(b)** Scatter plot of these 25 significant connections showing the significant peak with corresponding time shift. The error connections are shown in blue and green circles.

3.6.2 Outlier detection and cluster analysis

The clustering algorithm is applied to the 25 significant connections for distinguishing direct connections from the indirect connections and common source. Before starting clustering algorithm the modified Z-score is used for searching very strong connections which are considered as outlier connections. Here one outlier is identified from spike train #5 to spike train #13. The peak of this connection is 6.52 and time shift is 12 ms. This outlier connection is considered as a direct connection. A scatter plot of 25 significant connections with the outlier is shown in Fig. 3.10(b).

Reference spike train	Target spike train	Peak	Time shift	Cluster
1	**6**	4.37	13	Direct
1	**8**	4.57	15	Direct
1	**12**	3.05	11	Direct
2	**3**	3.50	12	Direct
2	**11**	1.64	27	Indirect
2	**14**	3.26	14	Direct
3	**11**	2.73	12	Direct
3	**14**	3.23	12	Direct
6	**8**	2.44	2	Common
7	**4**	1.72	23	Indirect
7	**9**	3.57	11	Direct
9	**4**	3.27	12	Direct
10	**4**	1.74	25	Indirect
10	**9**	3.46	13	Direct
11	**2**	2.57	13	Direct
12	**4**	1.69	23	Indirect
12	**6**	1.90	2	Common
12	**8**	1.92	4	Common
14	**11**	2.98	13	Direct
14	**15**	1.73	40	Indirect
15	**2**	2.71	12	Direct
15	**3**	3.57	14	Direct
15	**11**	3.73	11	Direct
15	**14**	2.14	26	Indirect

Table 3.4: Classification of 24 significant connections of fifteen spike train.

All the remaining 24 significant peaks and corresponding time shifts are normalized (between 0 and 1) and the clustering algorithm is applied to these normalized values. In the clustering algorithm the distance between pairs of observations is calculated using the Euclidean distance and the average linkage is used for calculating distance between clusters. Clustering algorithm classifies all the significant connections according to their significant peaks and time shifts. The tree of a dendrogram is cut in such a way that it creates three clusters. A dendrogram of the three clusters is shown in Fig. 3.11 by three different colours. Investigation from dendrogram reveals that the connections of the red colour have high significant peaks and moderate time shifts and are considered as direct connections (Table 3.4). The connections of the blue colour have low significant peaks

and short time shifts and are considered as common source connections. The connections of the green colour have low significant peaks and long time shifts and are considered as indirect connections.

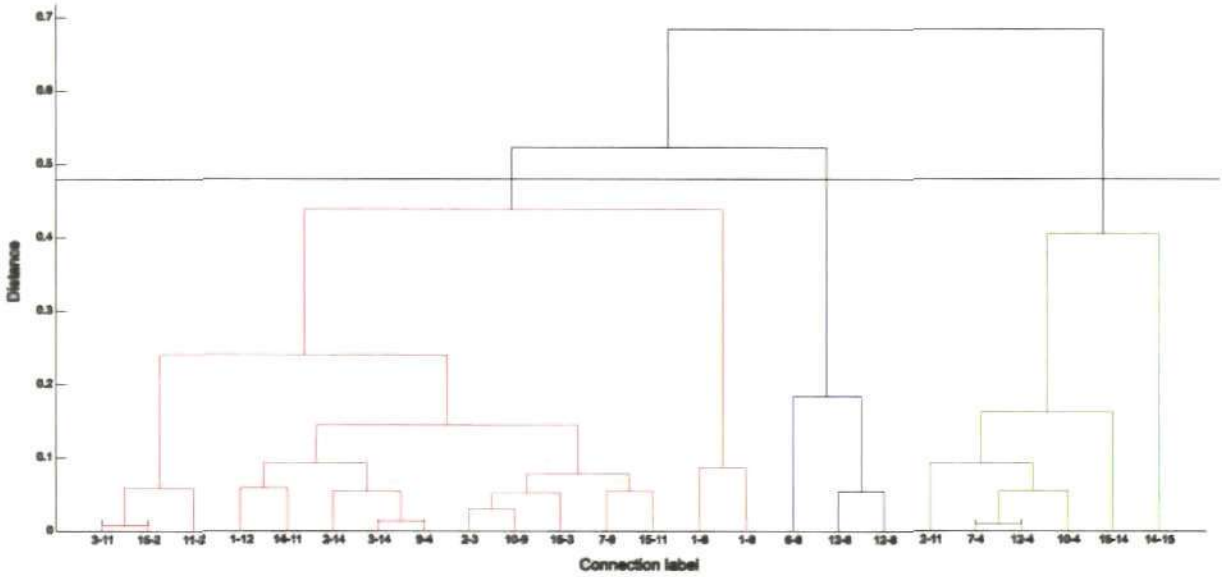


Figure 3.11: Dendrogram of 24 significant connections. Three clusters are indicated by different colours with the connection labels. Red colour indicates direct connection, blue colour indicates common source and green colour indicates indirect connection.

3.6.3 Classification of connection

To investigate the connections obtained from common source cluster, all the groups of three spike trains are detected that have the diagram of connection in Fig. 3.3(c) and also include the connections: from spike train #6 to spike train #8, spike train #12 to spike train #6 and spike train #12 to spike train #8 (Table 3.4). Investigation from Fig. 3.12(a) reveals that the groups of three spike trains are (#1, #8, #6), (#1, #6, #12), and (#1, #8, #12). In the group (#1, #8, #6), spike train #1 influences both spike train #8 and #6 with corresponding time shifts 15 ms and 13 ms. The connection from spike train #8 to spike train #6 is a common source (Eq. 3.1) where spike train #1 connects both spike trains #6 and #8 and the time shift for this connection is 2 ms which is short. The connection from spike train #12 to spike train #6 is a common source where spike train #1 influences both these spike train with corresponding time shifts 11 ms and 13

ms, respectively. Spike train #1 also influences both spike trains #8 and #12 with corresponding time shifts 15 ms and 11 ms. Thus the connection from spike train #12 to spike train #8 is a common source and the time shift for this connection is 4 ms.

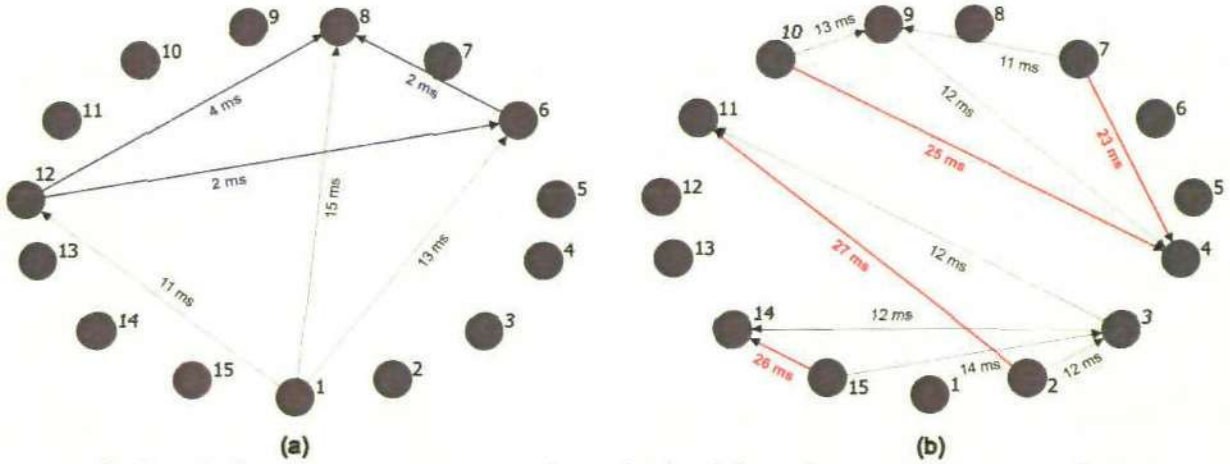


Figure 3.12: (a) Common source connections obtained from the common source cluster in Table 3.4. **(b)** Indirect connections obtained from the cluster of indirect connections in Table 3.4.

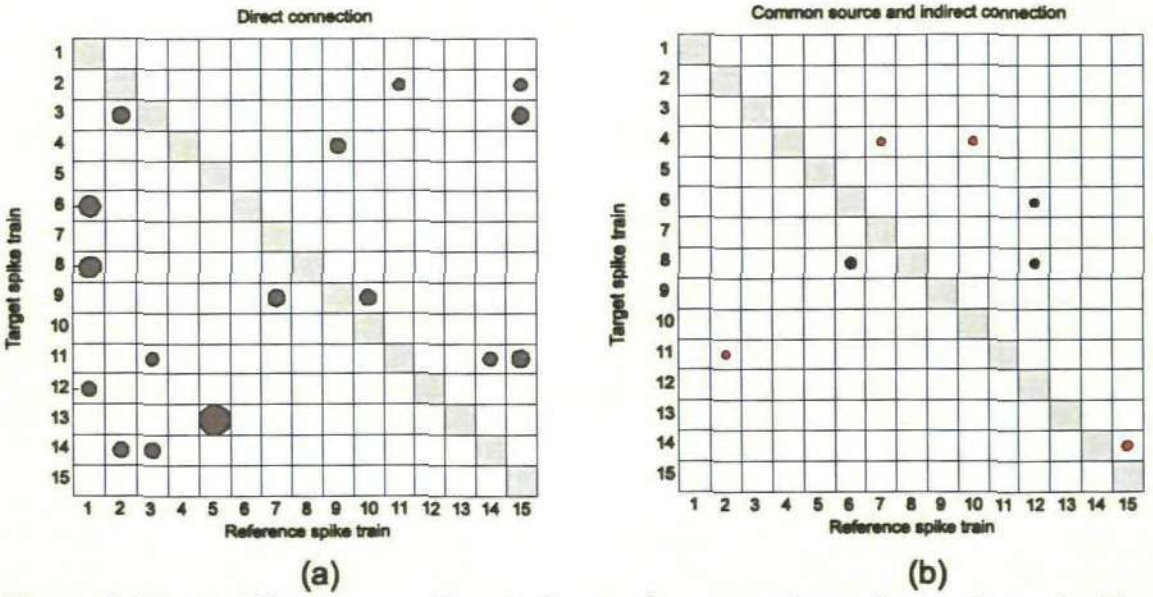


Figure 3.13: (a) Direct connection similar to the connection scheme shown in Fig. 3.8(b). Radius of the circle shows the strength of connection and is proportional to the strength of connection in Fig. 3.8(b). **(b)** Common source and indirect connections. Blue circle shows a common source connection and red circle shows an indirect connection. Radius of the circle shows the strength of connection and the connections are small relative to the direct connection.

Like common source cluster, all connections obtained from indirect cluster are investigated using a group of three spike trains that have the diagram of connection in Fig. 3.3(b) and also include the connections: from spike train #2 to spike trains #11,

spike train #7, #10 and#12 to spike train #4, spike train #14 to spike train #15 and spike train #15 to spike train #14 (Table 3.4). Investigation from Fig. 3.12(b) reveals that the group of three spike trains are (#2, #3, #11), (#4, #7, #9), (#4, #9, #10) and (#3, #14, #15). In the group (#2, #3, #11), spike train #2 influences spike train #3 with time shift 12 ms and spike train #3 influences spike train #11 with time shift 12 ms. The connection from spike train #2 to spike train #11 is an indirect connection (Eq. 3.2) where spike train #3 is an intermediate spike train. The time shift for the connection from spike train #2 to spike train #11 is 27 ms which is long. Connection from spike train #7 to spike train #4 and connection from spike train #10 to spike train #4 are indirect. For both these connections spike train #9 acts as an intermediate spike train. Similarly, connection from spike train #15 to spike train #14 is indirect. For this connection spike train #3 acts as an intermediate spike train. More importantly, all these three indirect connections have long time shifts: 23 ms, 25 ms and 26 ms, respectively.

Two connections (spike train #12 to spike train #4 and spike train #14 to spike train #15) of the indirect cluster reveal that in fact these connections are not indirect. Investigation from Fig. 3.12(b) shows that keeping fixed spike train #4 and #12, there is no spike train in the group which can make the diagram of connection like in Fig. 3.3 (b). Again, to identify the connection from spike train #14 to spike train #15, a group of three spike trains (#3, #14 and #15) is found but the diagram of connection is not the same as Fig. 3.3 (b).

Classifying all the significant connections into direct, indirect and common connections, the functional connectivity of the fifteen spike trains is displayed by two square matrices (Fig. 3.13). One matrix shows the functional connectivity of the direct connections (Fig. 3.13(a)) which reveals the correct connectivity of the diagram used to generate the fifteen spike trains (Fig. 3.8(a)) with corresponding connection strengths

(Fig. 3.8(b)). The other matrix (Fig. 3.13(b)) shows the connections which are not present in the connectivity diagram but are due to common source or indirect coupling. The radius of the circle shows the strength of the connections. More importantly, the common source and indirect connections have small radius compared to the direct connections indicates that they are spurious connections. The red circle shows the indirect connections and the blue circle shows the common source.

3.7 Analysis of functional connectivity of twenty spike trains

In section 3.6 it is shown that in case of very big connections the MCG method is useful for identifying functional connectivity. In this section it is demonstrated that this method is useful for identifying functional connectivity in case of medium strength of connections. Here it is also shown that this method is useful for identifying functional connectivity of a large set of spike trains with large number of connections. A set of twenty spike trains are generated for a period of 30000 milliseconds using the connection architecture shown in Fig. 3.14(a) with the connection strengths shown in Fig. 3.14(b) in matrix format. The parameter values of the ELIF model for the generation of twenty spike trains is given in Table 3.5. The connection from spike train #5 to spike train #7 and spike train #16 to spike train #17 have the same maximum strengths which is 14.92; all other connections have medium strength of connection which range from 10.40 to 14.23. Fig. 3.15 shows the raster plot of spiking activity of these twenty spike trains.

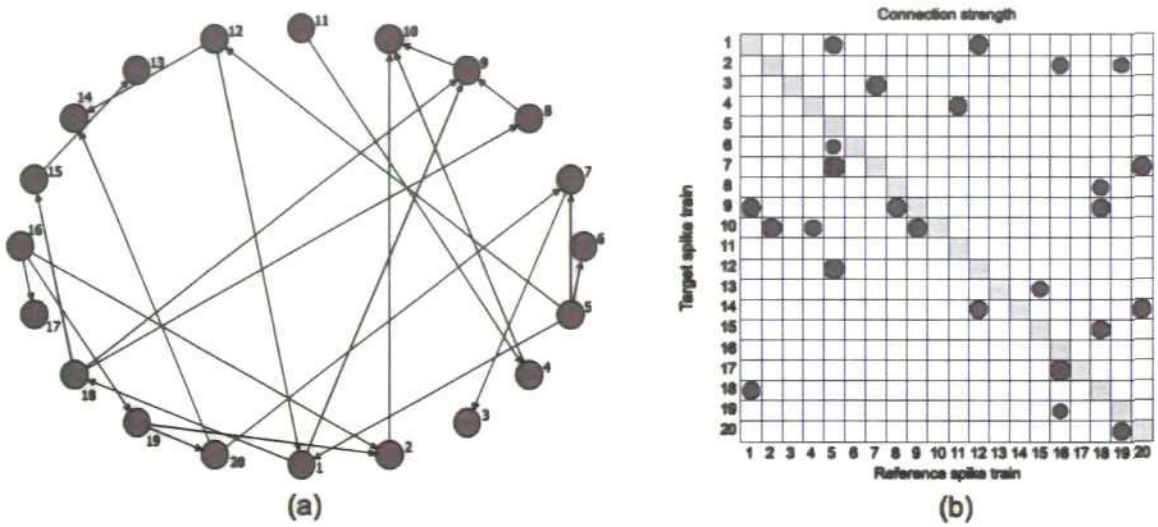


Figure 3.14: (a) Connection scheme of twenty spike trains. There are twenty five non zero connections which are shown by arrows. **(b)** Connection strengths of the twenty five non zero connections in the matrix format.

Neuron parameter	**Mean**	**S.D.**
Maximum value of the threshold	44.80	1.09
Threshold decay rate	2.89	0.36
Asymptotic threshold value	14.15	0.96
Amplitude of the noise	5.07	0.35
Noise decay rate	10.01	0.02
Initial value of after spike hyperpolarisation	-28.86	0.35
Soma's membrane potential decay rate	20.35	0.60
External input	-0.05	0.38
Absolute refractory period	6.35	2.08
Connection parameter		
Connection strength	12.65	1.25
Decay rate of postsynaptic potential	2.83	0.99
Time lag of spike propagation	12.04	1.13

Table 3.5: Parameter values of the ELIF model to generate twenty spike trains.

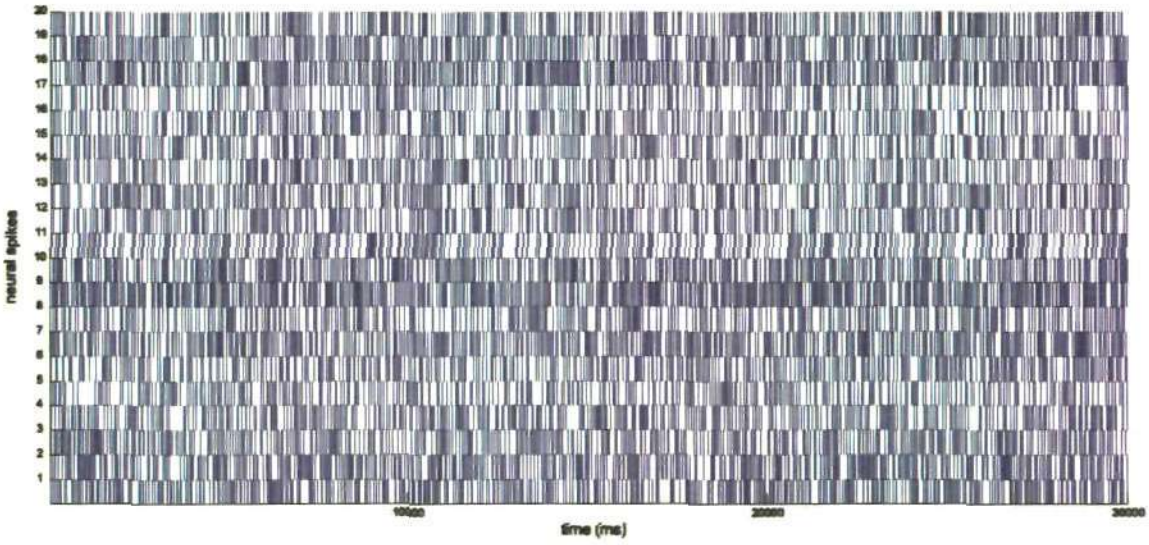


Figure 3.15: Raster plot of twenty spike trains generated for the connection scheme given in Fig. 3.14 over the duration 30,000 milliseconds.

3.7.1 Calculation of CCF

For these twenty spike trains a total of $(20^2 - 20)/2 = 190$ pair-wise CCF are calculated with a bin size of 1 ms and a correlation window of 100 ms. To test the independence of the pair-wise spike trains the level of significance $\alpha = 0.05$ is considered with the Bonferroni correction. A total of 34 significant connections are found for the twenty spike trains (Fig. 3.16(a)). In the Fig. 3.16(a) significant connections are shown by a circle and the direction of connection is considered from the reference spike train to the target spike train. Scatter plot of all significant connections is shown in Fig. 3.16(b). This scatter plot shows the significant peaks with corresponding time shifts of the 34 significant connections. The peaks of the significant connections range from 1.70 to 4.81 and the time shifts range from 2 ms to 26 ms.

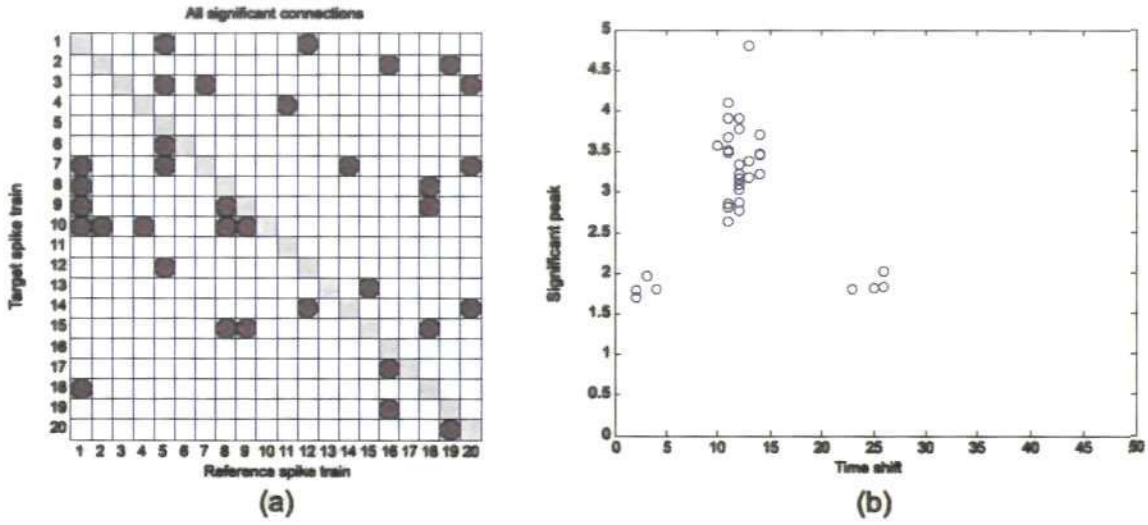


Figure 3.16: (a) Significant connections obtained from pair-wise cross-correlation analysis. **(b)** Scatter plot of these 34 significant connections showing the significant peak with corresponding time shift.

3.7.2 Outlier detection and cluster analysis

Before starting clustering algorithm the modified Z-score is used for searching very strong connections which are considered as outlier connections. Here no outlier is identified from the 34 significant connections. A Clustering algorithm is applied for classifying the 34 significant connections in three clusters according to their significant peaks and time shifts. Fig. 3.17 shows the result of clustering analysis of all significant connections. The cluster of common source is shown in Fig. 3.17(a). In this cluster there are 4 significant connections which are: spike train #1 to spike train #7, spike train #8 to spike train #15, spike train #9 to spike train #15 and spike train #14 to spike train #7. All these significant connections have low significant peaks and short time shifts (Fig. 3.17(d)). Fig. 3.17(b) shows the cluster of direct connections. A total of 25 connections are considered as direct connections. These connections have high significant peaks and moderate time shifts (Fig. 3.17(d)). Fig. 3.17(c) shows the cluster of indirect connections. This cluster includes 5 significant connections which are: spike train #1 to spike trains #8 and #10, spike train #5 to spike train #3, spike train #8 to spike train #10

and spike train #20 to spike train #3. All these connections have low significant peaks and long time shifts (Fig. 3.17(d)).

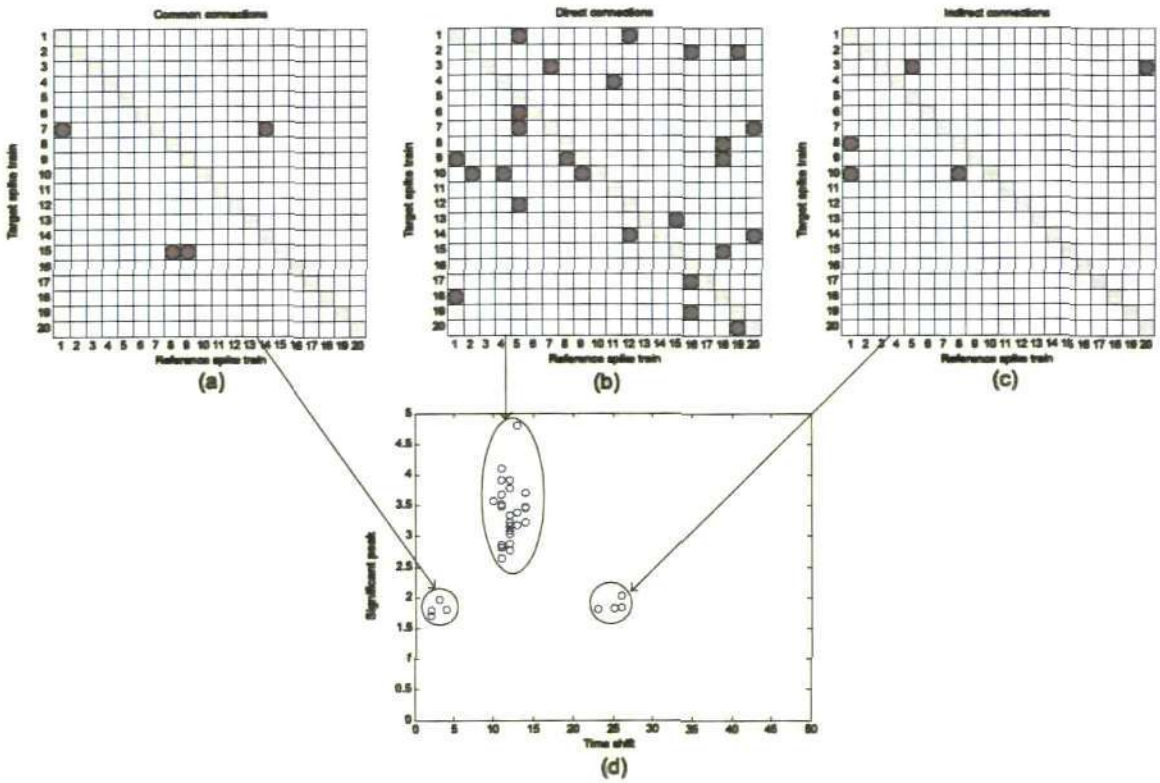


Figure 3.17: Cluster analysis of all significant peaks and time shifts. **(a)** Cluster of common source. **(b)** Cluster of direct connections **(c)** Cluster of indirect connections. **(d)** Scatter plot of the significant peaks and time shifts. The scatter plot shows that low significant peaks and short time shifts constitute the cluster of common source. High significant peaks and moderate time shifts constitute the cluster of direct connections and low significant peaks and long time shifts constitute the indirect connections.

3.7.3 Classification of connection

To investigate the connections obtained from common source cluster (Fig. 3.17(a)) all the groups of three spike trains are detected that have the diagram of connection in Fig. 3.3(c) and include the connections: spike train #1 to spike train #7, spike train #8 to spike train #15, spike train #9 to spike train #15 and spike train #14 to spike train #7. Investigation from Fig. 3.18(a) reveals that the groups of three spike trains are (#1, #5, #7), (#8, #15, #18), (#9, #15, #18) and (#7, #14, #20). In the group (#1, #5, #7), spike train #5 influences both spike train #1 and #7 with corresponding time shifts 10 ms and 14 ms. The connection from spike train #1 to spike train #7 is a common source

connection (Eq. 3.1) where spike train #5 connects both spike trains #1 and #7 and the time shift for this connection is 4 ms. The connection from spike train #8 to spike train #15 is a common source connection where spike train #18 influences both these spike trains with corresponding time shifts 11 ms and 14 ms, respectively. Spike train #18 also influences spike trains #9 and #15 with corresponding time shifts 12 ms and 14 ms. Thus, the connection from spike train #9 to spike train #15 is a common source. Spike train #20 influences both spike trains #14 and spike train #7 with time shifts 12 ms and 14 ms, respectively. Therefore, the connection from spike train #14 to spike train #7 is considered common source connection. More importantly, all these three common source connections have short time shifts 3 ms, 2 ms and 2 ms, respectively.

Like common source cluster, all connections obtained from indirect cluster (Fig. 3.17(c)) are investigated using the group of three spike trains that have the diagram of connection in Fig. 3.3(b) and include the connections: spike train #1 to spike trains #8 and #10, spike train #5 to spike train #3, spike train #8 to spike train #10 and spike train #20 to spike train #3. Investigation from Fig.3.18(b) reveals that the groups of three spike trains are (#1, #8, #18), (#1, #9, #10), (#5, #7, #3), (#8, #9, #10) and (#20, #7, #3). In the group (#1, #8, #18), spike train #1 influences spike train #18 with time shift 12 ms and spike train #18 influences spike train #8 with time shift 11 ms. The connection from spike train #1 to spike train #8 is an indirect connection (Eq. 3.2) where spike train #18 is an intermediate spike train. The time shift for the connection from spike train #1 to spike train #8 is 23 ms, which is long. Connection from spike train #1 to spike train #10 and connection from spike train #18 to spike train #10 are indirect. For both these connections spike train #9 acts as an intermediate spike train. Similarly, connection from spike train #5 to spike train #3 and connection from spike train #20 to spike train #3 are indirect. For both these connections spike train #7 acts as an intermediate spike

train. More importantly, all these four indirect connections have long time shifts 25 ms, 25 ms, and 26 ms and 26 ms respectively.

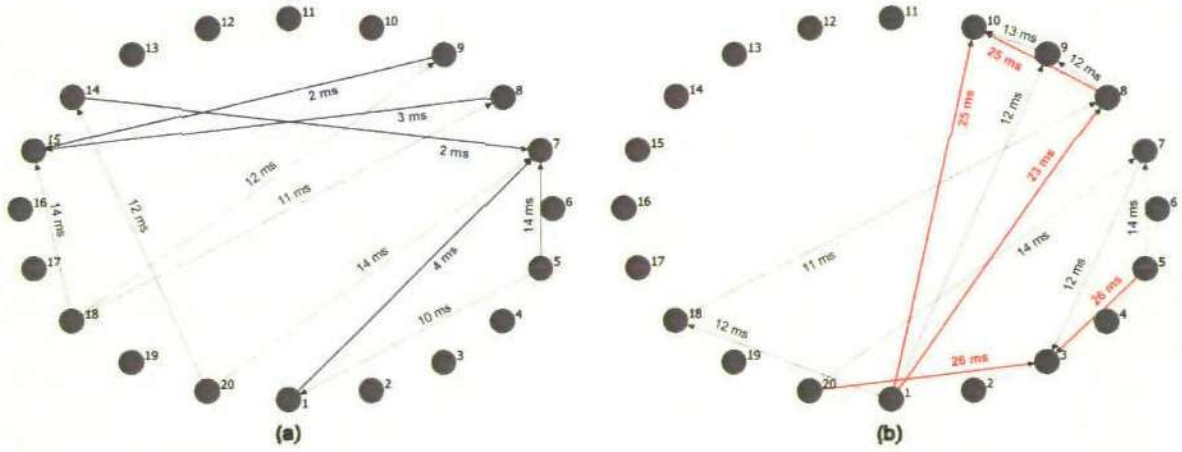


Figure 3.18: (a) Common source connections obtained from the cluster of common source (Fig. 3.17(a)). **(b)** Indirect connections obtained from the cluster of indirect connections (Fig. 3.17(c)).

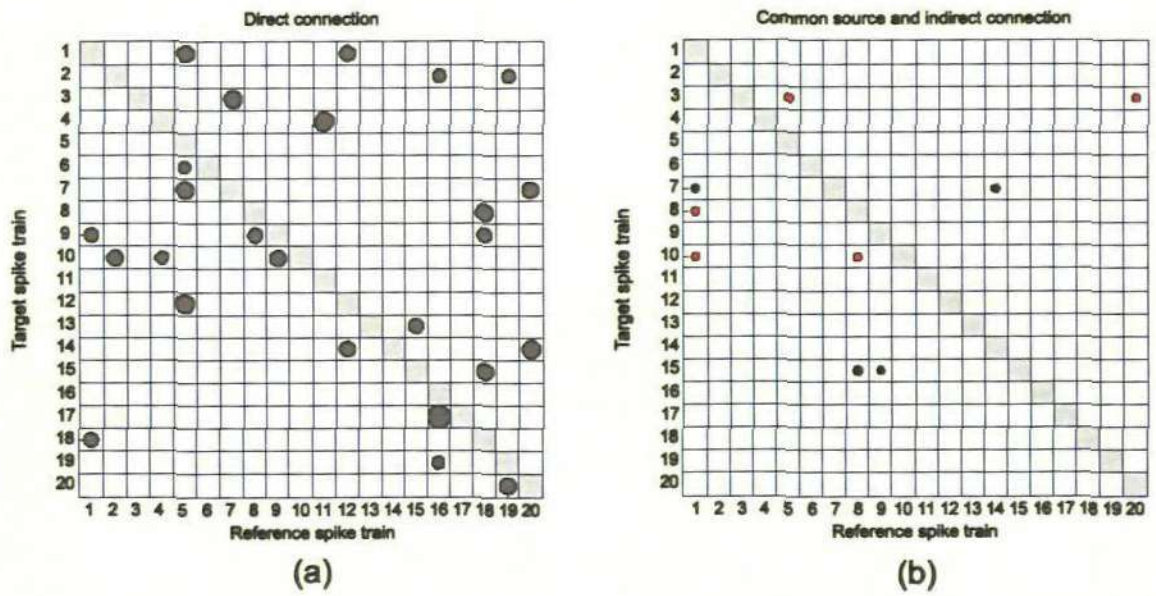


Figure 3.19: (a) Direct connections which are exactly the same as at the connection scheme shown in Fig. 3.14(b). Radius of the circle shows the strength of connection estimated from CCF and is proportional to the strength of connection in Fig. 3.14(b). **(b)** Common source and indirect connections. Blue circle shows a common source connection and red circle shows an indirect connection. Radius of the circle shows the strength of connection and the connections are small relative to the direct connection.

Classification of all the significant connections into direct, indirect and common source connections, the functional connectivity of the twenty spike trains are displayed into two square matrices (Fig. 3.19). One matrix shows the functional connectivity of the

direct connections (Fig. 3.19(a)) which reveals the correct connectivity of the diagram used to generate the twenty spike trains (Fig. 3.14(a)) with corresponding connection strengths (Fig. 3.14(b)). The radii of the circles (Fig. 3.19(a)) show the estimated strength of connections which are very similar to the radii of the circles in the diagram of connections (Fig. 3.14(a)). The other matrix (Fig. 3.19(b)) shows the connections which are not present in the connectivity diagram but are due to common source or indirect coupling. The radius of the circle shows the estimated strength of the connections. More importantly, the common source and indirect connections have small radius compare to the direct connections means that they are spurious connections. The red circle shows the indirect connections and the blue circle shows the common source connection.

Chapter 4

A generator of multiple spike trains based on the Modulated Renewal Process

In this chapter, a new probabilistic model for the generation of dependent spike train is described. The dependency is based upon the Modulated Renewal Process. In order to define realistic parameter values (i.e. in a range corresponding to the experimental data) the probabilistic model is fitted to the Enhanced Leaky Integrate and Fire Model. The probabilistic model based on realistic parameter values is of interest as it can be used to generate spike trains with prescribed stochastic properties. In particular, the emphasis of this research is on the use of this probability model for the generation of spike trains to test the Cox method.

4.1 Introduction

Experimental recordings can be considered to be a series of points in time (a spike train), where each point marks the moment of spike generation. In mathematics the series of points in time is called a point process (Eden, 2008). In neuroscience, the most popular representation of a spike train is based on a Poisson process (Eden, 2008). However, the more general case requires the consideration of the renewal process. In order to introduce dependencies between spike trains, a Modulated Renewal Process (MRP) is introduced (Cox, 1972). In this chapter a probabilistic model based on a MRP for spike train generation is developed. In a MRP, the generation of a spike train is based on the influences it receives from other spike trains in the neural circuit. This probabilistic model is relatively straightforward and it generates spikes similar to biophysically-based integrate-and-fire models.

4.2 Point process

A point process is a set of discrete events that occur in continuous time. There are three useful ways to describe a point process: (a) as a sequence of times at which an event occurs, (b) as a sequence of time elapsed between events occurring, (c) as a counting process. Let us assume that $S_1, S_2, \cdots\cdots$ are random variables describing the event times of a point process. The identification of a process in terms of when it occurred is the event $S_1 = s_1, S_2 = s_2, \cdots\cdots$ for some collection of times $0 < s_1 < s_2 < \cdots\cdots$ (Fig. 4.1(a)). Let us assume that $X_1, X_2, \cdots\cdots$ are random variables describing the interevent times of a point process. Interevent times of a point process can be calculated by evaluating the difference between subsequent event times. Mathematically, $X_1 = S_1$ and $X_i = S_i - S_{i-1}$. The identification of a process as a point process in terms of interevent times is represented in Fig. 4.1(b). Assume that $N(t)$ is the counting process which is defined as the total number of events that have occurred up to and including time t. The identification of a process as a point process in terms of counting process is represented in Fig. 4.1(c). Assume that $N(t_1), N(t_2)$ are the total number of events that have occurred up to and including times t_1, t_2 respectively assuming that the start time $t = 0$ and $N(t_1, t_2)$ is the total number of events observed in an interval (t_1, t_2). The counting process $N(t_1, t_2)$ is calculated as $N(t_2) - N(t_1)$ and this is called the increment of the point process between times t_1 and t_2.

The Poisson process is a simple structure of any point process. It is defined as a point process where the occurrences of events are random, independent and have uniform probability. Furthermore, a Poisson process is a stationary point process and the number of events that occur in any time interval depends only on the length of the time interval but not on the specific time. A Poisson process is also a memoryless process where interevent times of the process have an exponential distribution. A renewal process is a

particular type of point process where the occurrences of events are random and also depend on the last event. A renewal process is a generalization of the Poisson process which supports a variety of interevent distributions.

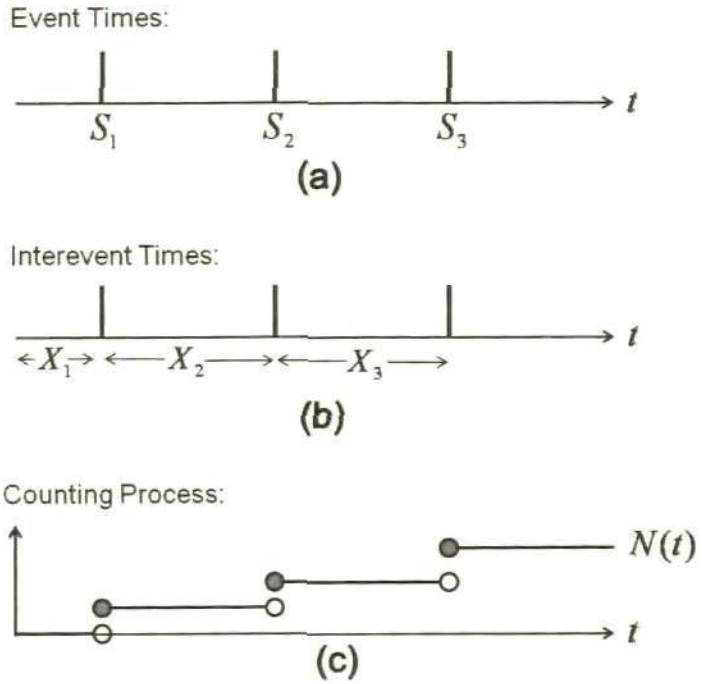


Figure 4.1: Multiple specifications for point process data. **(a)** Point process in terms of event occurrence. **(b)** Point process in terms of interevent occurrence and **(c)** Point process in terms of counting process.

4.3 Influence function

An influence function describes the effect of an incoming spike from a pre-synaptic neuron B on the membrane potential of the post synaptic neuron A. To study neuronal interactions, it is important to choose an appropriate influence function which accurately describes the dynamics of the post-synaptic potential: this influence function increases when a spike arrives at the post-synaptic neuron and the probability of spike generation by the post-synaptic neuron increases; after that the influence function decays to zero. Among different influence functions, it is found that the best influence function can be approximated by the alpha function which is used in neuroscience to describe synaptic connectivity between neurons (Gerstner and Kistler, 2002).

Thus, the influence function can be defined as:

$$Z_B(t) = \frac{g_m}{(\tau_s - \tau_r)}\left(e^{-\frac{t}{\tau_s}} - e^{-\frac{t}{\tau_r}}\right), \quad t > 0 \tag{4.1}$$

where $Z_B(t)$ is the influence function from pre synaptic neuron B to the post synaptic neuron A, where τ_s and τ_r are the characteristic times of decay and rise of post-synaptic potential, respectively. Parameter g_m provides the normalization so that the maximum influence function equals one and is defined by:

$$g_m = \frac{\tau_s - \tau_r}{e^{-\frac{t_m}{\tau_s}} - e^{-\frac{t_m}{\tau_r}}}$$

where

$$t_m = \frac{log(\tau_s/\tau_r)}{\frac{1}{\tau_r} - \frac{1}{\tau_s}}.$$

A simplified version of the influence function, corresponding to the case $\tau_s = \tau_r$, is given by the following formula:

$$Z_B(t) = \frac{t}{\tau_s}\, e^{1-t/\tau_s}, \quad t > 0. \tag{4.2}$$

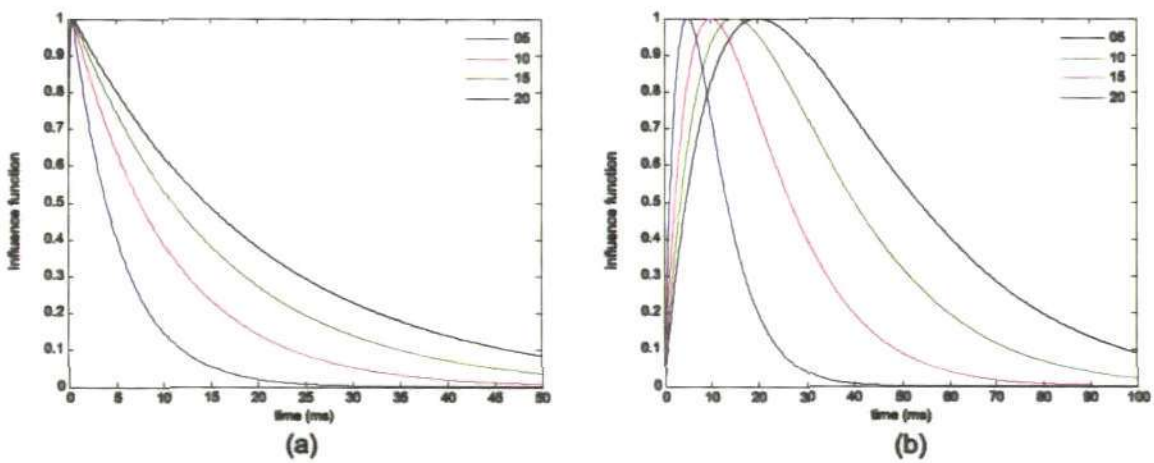


Figure 4.2: Influence function. **(a)** The influence function of (4.1) with rise time of post synaptic potential $\tau_r = 0.1$ ms for different decay times $\tau_s = (5\ ms,\ 10\ ms,\ 15\ ms,\ 20\ ms)$. At rise time $\tau_r = 0.1$ ms the peak value of the influence function is 1 and then it decays to zero. **(b)** The influence function of (4.2) for different decay time $\tau_s = (5\ ms,\ 10\ ms,\ 15\ ms,\ 20\ ms)$. At time $t = 0$, the value of influence function is zero. The influence function has peaks at time $t = \tau_s$ which is 1 and then decays to zero for large values of t.

The influence function of (4.1) for different values of decay time $\tau_s = (5\,ms,\ 10\,ms,\ 15\,ms,\ 20\,ms)$ with a fixed value of rise time $\tau_r = 0.1$ milliseconds is shown in Fig. 4.2(a). This influence function increases to one in a very short time (0.1 milliseconds) and then decays to zero. For a small decay time τ_s, the influence function decays rapidly to zero and for a large decay time τ_s, the influence function takes subsequently large to return back to zero. Fig. 4.2(b) shows the influence function of (4.2) for different values of decay time $\tau_s = (5\,ms,\ 10\,ms,\ 15\,ms,\ 20\,ms)$ where $\tau_r = \tau_s$. This influence function reaches its peak value (which equals one at time τ_s) and then returns back to zero. Like (4.1) this influence function takes a long time to fall back to zero when τ_s is large.

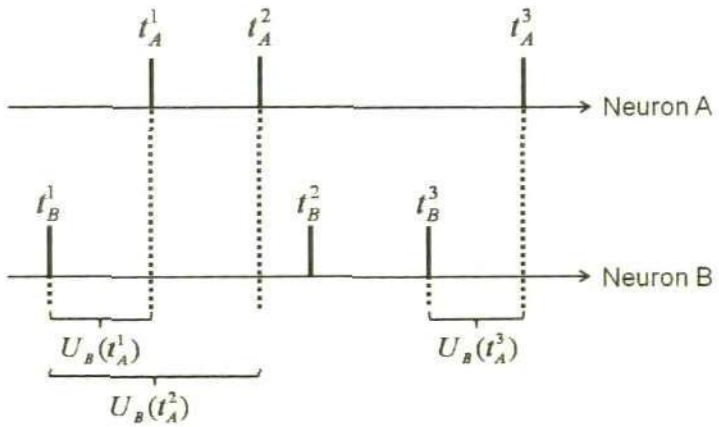


Figure 4.3: Illustration of the backward recurrence time of post-synaptic neuron B. The backward recurrence time is calculated using the difference between the spike time in neuron A and the last spike time in neuron B. Assuming that the spike times in neuron A are t_A^1, t_A^2 and t_A^3 and the spike times in neuron B are t_B^1, t_B^2 and t_B^3, the backward recurrence time for neuron B at time t_A^1 is the time form t_A^1 to the time t_B^1 and is denoted by $U_B(\,t_A^1)$. Similarly the backward recurrence times for the neuron B at times t_A^2 and t_A^3 are calculated and denoted by $U_B(\,t_A^2)$ and $U_B(\,t_A^3)$, respectively.

In influence functions (4.1-4.2) it is assumed that the pre synaptic neuron B has generated a spike at time zero and that this spike arrives to the post synaptic neuron A without any time delay. The functions (4.1) and (4.2) can be rewritten to incorporate the spiking at the pre synaptic neuron B. It is assumed that the influence function depends

on the backward recurrence time of the pre synaptic neuron B, which is denoted as $U_B(t)$ (Fig. 4.3) and can be substituted for this variable as the argument to the function (4.1).

This means that the influence described by function (4.1) starts increasing when the last spike occurs in neuron B before time t. To account for the time delay in spike propagation, the argument in formula (4.1) should be shifted by the time lag Δ. Thus, the influence function is defined by the following formula:

$$Z_B(t) = \frac{g_m}{(\tau_s - \tau_r)}\left(e^{-\frac{U_B(t-\Delta)}{\tau_s}} - e^{-\frac{U_B(t-\Delta)}{\tau_r}}\right), \quad (t-\Delta) > 0 \qquad (4.3)$$

Note that $U_B(t)$ is the backward recurrence time of B; Δ is a time lag corresponding to delay of spike propagation from neuron B to neuron A. Similarly to (4.2), a simplified version of the influence function with time lag Δ is given by the following formula:

$$Z_B(t) = \frac{1}{\tau_s} U_B(t-\Delta) e^{1-\frac{U_B(t-\Delta)}{\tau_s}}, \quad (t-\Delta) > 0 \qquad (4.4)$$

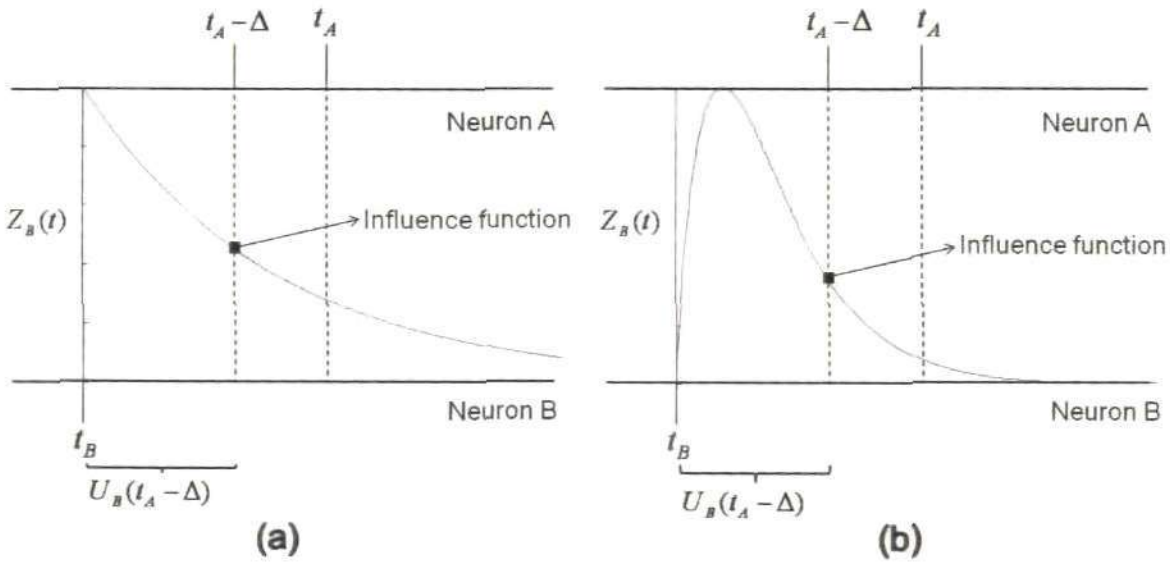


Figure 4.4: Influence function accounting for the time delay Δ of spike propagation. The square indicates the value of influence function when a spike t_A in neuron A is shifted by the time delay Δ. (**a**) Influence function by formula (4.3) **(b)** Influence function by formula (4.4).

Both the influence functions (4.3) and (4.4) are shown in Fig. 4.4 with their corresponding values shown.

In the formulas (4.3-4.4), the influence function is defined by solely taking the last spike in neuron B before the time $(t-\Delta)$. In some cases it is productive to consider accumulation of the post-synaptic potential in time interval T, considering a history of spiking in B over the time interval $(t-T,t)$. This type of the influence function is useful when the decay time of the post-synaptic potential is relatively small in comparison to the mean interspike interval of neuron B. Thus, a generalized influence function is defined as:

$$Z_B(t) = \sum_{j=1}^{k} \frac{g_m}{(\tau_s - \tau_r)} \left(e^{-\frac{U_B^j(t-\Delta)}{\tau_s}} - e^{-\frac{U_B^j(t-\Delta)}{\tau_r}} \right), \qquad (t-\Delta) > 0 \qquad (4.5)$$

here k is an index which denotes the highest order of the backward recurrence time in the history interval (the first order corresponds to a spike which is the most close to the moment $(t-\Delta)$ in backward time, the second order relates to the previous spike in the reward time etc.). Fig. 4.5 shows the generalized influence function over the accumulation time T, thus, $U_B^k(t) < T$ and $U_B^{k+1}(t) > T$. A simplified version of the generalised influence function, corresponding to the case $\tau_s = \tau_r$, is given by the following formula:

$$Z_B(t) = \sum_{j=1}^{k} \frac{1}{\tau_s} U_B^j(t-\Delta) e^{1-\frac{U_B^j(t-\Delta)}{\tau_s}}, \qquad (t-\Delta) > 0 \qquad (4.6)$$

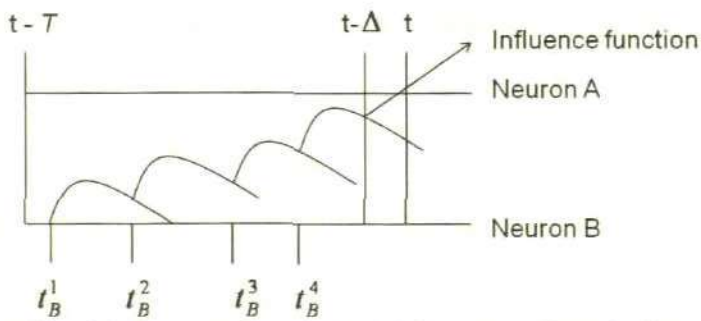


Figure 4.5: Generalized influence function which accumulates influence from previous spikes of neuron B in the time interval $(t-T,t)$ with propagation delay Δ.

4.4 The probabilistic model

One of the simplest, most commonly used classes of neural spiking models is the Poisson process (Tuckwell, 1988; Rieke et al., 1997; Gabbiani and Koch, 1998; Reich et al., 1998). In the Poisson spiking model, the probability distribution of spiking at any point in time is independent of previous spiking activity. Due to this independent spiking structure, the Poisson spiking model cannot be used to accurately model many neural systems (Gabbiani and Koch, 1998; Shadlen and Newsome, 1998). Therefore, it is essential that new models of spiking activity that also incorporate the effect of previous spike events of a neuron are developed. The simplest type of spiking model which depends on previous spike events is a renewal process model. A renewal process model is specified by a distribution function for the inter-spike intervals (ISI). This model includes gamma, lognormal and inverse Gaussian (Rodieck et al., 1962; Tuckwell, 1988; Gabbiani and Koch, 1998) and generalized inverse Gaussian (Iyengar and Liao, 1997) probability distributions. The gamma distribution is a common basis for renewal models (Bishop et al., 1964; Nakahama et al., 1968; Correia and Landolt, 1977; Gabbiani and Koch, 1998; Eden, 2008) that is defined by two values, a scale and a shape parameter. If the shape parameter is equal to one, then the gamma distribution simplifies to an exponential distribution. If the shape parameter is less than one, then the density drops faster than an exponential distribution. This is useful for describing point processes that fire in rapid bursts. If the shape parameter is greater than one, then the gamma density function starts from zero, rises to a maximum at some positive value and then falls back to zero. This is useful when describing the refractory nature of point processes, that is the property that after firing a spike, the process is less likely to fire again immediately afterward. In the next section, a probabilistic model is developed which generates two spike trains, one is the reference spike train based on renewal process and the other is the target spike train based on the modulated renewal process

(Cox, 1972). The target spike train is generated based upon its backward recurrence time, the backward recurrence time of the reference spike train and the strength of influence from the reference spike train.

4.4.1 Modulated renewal process

Let us consider a renewal process where $X_1, X_2, \cdots\cdots$ are independent and identically distributed inter-spike intervals (ISI) with a probability density function $f(x)$. Another approach to the specification of renewal process comes from hazard function. The hazard function or age dependent death rate is defined as the event rate at time t conditional on survival until time t or later:

$$\varphi(t) = \lim_{\Delta t \to 0} \frac{Pr(t \leq X < t + \Delta t | t \leq X)}{\Delta t}.$$

This hazard function can be calculated in terms of the probability density function of the ISI's:

$$\varphi(t) = \frac{f(t)}{1 - F(t)} \tag{4.7}$$

where $F(t)$ is the cumulative distribution function of ISI.

A spike train A is said to be generated by a modulated renewal process if the hazard function of the spike train A can be represented at the moment t as:

$$\varphi(t) = \varphi_A\big(U_A(t)\big)\, exp\{\beta Z_B(t)\}, \tag{4.8}$$

where $\varphi_A(.)$ is the hazard function of the spike train A without influence from the other spike train B, $U_A(t)$ is the backward recurrence time of the spike train A at the moment t, $Z_B(t)$ is the influence function determining how spike train B influences spike train A, and β is the unknown regression coefficient describing the strength of the influence from spike train B to spike train A. The modulated renewal process supports the introduction of interaction between the spike trains considering the strength of influence β and the influence function $Z_B(t)$. The hazard function is a history dependent rate

function which completely characterizes the stochastic structure of a point process (Daley and Vere-Jones, 2003). Therefore, in order to model the neural spike train in terms of a point process, a hazard function (Truccolo et al., 2005) must be defined.

In the case of multiple reference spike trains $B = (B_1, B_2, \cdots\cdots, B_k)$, the influence function should be defined independently for each reference spike train. The hazard of the target spike train A is

$$\varphi(t) = \varphi_A\big(U_A(t)\big)\, exp\big\{\beta_1 Z_{B_1}(t) + \beta_2 Z_{B_2}(t) + \cdots + \beta_k Z_{B_k}(t)\big\}, \qquad (4.9)$$

where $\varphi_A(.)$ is the hazard function of the spike train A without influence from the reference spike trains $B = (B_1, B_2, \cdots\cdots, B_k)$, $U_A(t)$ is the backward recurrence time of the spike train A at the moment t, $Z_{B_i}(t)$ is the influence function determining how the spike train B_i influences the spike train A, and β_i is the parameter describing the strength of the influence from the spike train B_i to A (i=1,2,…,k).

4.4.2 Model description

The probabilistic model presented in this research generates two spike trains, one is the target (A) and the other one is the reference (B). The reference spike train B is a renewal process and spikes of B modulate the probability of spike generation in the target spike train A. This target spike train A is the modulated renewal process with the hazard defined by (4.8). The reference spike train B is generated with gamma-distribution $\gamma(x; a, b)$ of inter-spike intervals (ISI) where a and b are the shape and the scale parameters respectively. The cumulative sum of these inter-spike intervals gives the spike times of the reference spike train B over the spike train generation time [0,T]. The spike train generation time is divided into n small sub intervals of length $\Delta t = T/n$, and $t_k = k\Delta t,\ k = 1,2 \cdots\cdots, n$. For each time t_k, the backward recurrence times of the target spike train A and reference spike train B are calculated. It is assumed that the hazard function of the backward recurrence times of the target spike train A is

distributed as Weibull $W(x; c, d)$ (Cox, 1972), where c and d are the scale and the shape parameters respectively. In fact, the type of the ISI distribution of spike train B and the distribution of backward recurrence times of spike train A can vary. The choice of the gamma distribution and the Weibull distribution was motivated by the fact that both distributions include the exponential distribution which has been observed in experimental neuroscience data. Now the hazard of the backward recurrence time of the target spike train A is calculated using (4.7). The influence function defined by (4.3), or (4.4), is specified with specific values of the characteristic times τ_s, τ_r, time lag Δ and the backward recurrence time of the reference spike train B. Considering all these values, the hazard of the target spike train is calculated using (4.8), with a specified strength of influence β from the reference spike train B to the target spike train A. If this hazard is greater than a randomly chosen number which has the uniform distribution in the range [0,1], then a spike is generated in the target spike train A. Repetition of the procedure for each time t_k over the entire spike train generation time results in the generation of the target spike train A which is a modulated renewal process and the probability distribution of ISI of this spike train is represented by a Weibull distribution.

The same procedure can be used to generate the target spike train A under the influence of multiple independent reference spike trains $B = (B_1, B_2, \cdots\cdots, B_k)$. All the reference spike trains $B = (B_1, B_2, \cdots\cdots, B_k)$ are generated with the gamma distribution of ISI with different scale and shape parameters. The cumulative sum of the inter-spike intervals give the spike times of the reference spike trains over the spike train generation time [0,T]. For each time t_k the hazard of the backward recurrence times of the target spike train is calculated using the Weibull distribution. The influence of each reference spike train B_i with $i = 1,2, \cdots\cdots, k$ to the target spike train A is calculated by (4.3) or (4.4) with a specified characteristic time τ_{si} with $i = 1,2, \cdots\cdots, k$, τ_{ri} with

$i = 1,2,\cdots\cdots,k$ and a time lag Δ_i with $i = 1,2,\cdots\cdots,k$. The hazard of the target spike train is calculated using (4.8) with a specified vector of influence strength $\beta = (\beta_1, \beta_2, \cdots\cdots, \beta_k)$ from the reference spike trains $B = (B_1, B_2, \cdots\cdots, B_k)$ to the target spike train A. If this hazard is greater than a randomly chosen number which has the uniform distribution in the range [0,1], then a spike is generated in the target spike train A. Repetition of the procedure for each time t_k over the entire spike train generation time results in the generation of the target spike train A.

The algorithm for generating spike train A based on the influence of spike train B over the generation time T is given below:

(i) Generate ISI of the spike train B using gamma $\gamma(x; a, b)$ distribution.

(ii) Find spike time of B using cumulative sum of ISI over the time T.

(iii) Divide the spike generation time T into n small sub intervals of length $\Delta t = T/n$.

(iv) Set the time $t_k = k.\Delta t$, $k = 1,2\cdots\cdots,n$.

(v) Compute the hazard of the backward recurrence time of spike train A using (4.7) with Weibull $W(x; c, d)$ distribution.

(vi) Compute the influence function Z_B using (4.3) or (4.4) with specified values of τ_s, τ_r and Δ.

(vii) Set the strength of influence β.

(viii) Compute the hazard of the spike train A using (4.8).

(ix) Pick a random number using uniform distribution over [0,1].

(x) If the hazard in (viii) is greater than the random number in (ix) then a spike is generated in neuron A.

(xi) Go to (iv).

4.5 Goodness of fit

An essential component of the statistical analysis is the assessment of the goodness of fit, i.e. the evaluation of how well the observed data is described by the given probability distribution. Measures of goodness of fit typically summarize the discrepancy between the distribution of observed values and the values expected under the model in question. Two goodness of fit measures are considered, the Q–Q plot and K–S test. These are used to observe how well ISI's of the generated spike trains are described by the probability distribution used in the model.

4.5.1 Q-Q plot (Quantile-Quantile plot)

A Q-Q plot is a graphical display invented by Wilk and Gnanadesikan (1968) to compare an observed dataset with a particular probability distribution. A Q-Q plot is a plot of the quantiles of the observed dataset against the quantiles of the probability distribution. The plotting positions of the quantiles are computed by the following formula

$$p_i = \frac{i - .5}{n}, \quad i = 1,2,\cdots\cdots,n \tag{4.10}$$

where n is the total number of observations in the data. To compare the dataset to the theoretical probability distribution, the empirical quantiles of the data i.e. the order statistics of the data are plotted on the vertical axis (y-axis) and the corresponding quantiles from the assumed probability distribution are plotted on the horizontal axis (x-axis). The quantiles of the assumed probability distribution are computed based on the plotting positions associated with the empirical quantiles of the data. A 45-degree reference line is also plotted. If the dataset and the assumed probability distribution are similar, the points in the Q-Q plot will fall approximately along the reference line. The greater the departure from the reference line, the greater the evidence that the data and the assumed probability distributions are not similar.

4.5.2 K-S test (Kolmogorov-Smirnov test)

The Kolmogorov-Smirnov test is concerned with the degree of agreement between the distribution of a set of observed data and a specified theoretical probability distribution. The Kolmogorov-Smirnov test is based on the empirical distribution function $S(x)$ and a specified theoretical distribution function $F(x)$. Given N ordered data points $X_1, X_2, \cdots\cdots, X_N$, the empirical distribution function is defined as

$$S(x_i) = \frac{n(i)}{N}$$

where $n(i)$ is the number of points less than X_i and the X_i are ordered from the smallest to the largest value. This is a step function that increases by $1/N$ at the value of each ordered data point. Under the null hypothesis, that the set of observed data follow the theoretical distribution, it is expected that every value X_i, $S(x_i)$ should be fairly close to $F(x_i)$. That is when the null hypothesis H_0 is true, it is expected that the differences between $S(x_i)$ and $F(x_i)$ will be small. The Kolmogorov-Smirnov test focuses on the largest deviation and the test statistic is defined as

$$D = max|F(x_i) - S(x_i)| \quad i = 1,2,\cdots\cdots, N$$

The null hypothesis is rejected if the test statistic D is greater than the critical value obtained from the table of Kolmogorov-Smirnov test.

4.6 Fitting with ELIF model

The probabilistic model is easy to understand, requires few parameters and can be used for generating spike trains similar to those produced by the enhanced leaky integrate and fire neuron (ELIF) model (Borisyuk, 2002). In this section, an optimization procedure is discussed which can be used to find the parameters of the probabilistic model in such a way that the spike train generated by the probabilistic model using these optimal parameter values provides the best fit to the spike train generated by the ELIF model. In this optimization procedure, a cost function is used which depends on

the difference between the histogram of ISI of the probabilistic model and the ELIF model. The minimum value of this cost function shows the minimum difference of the histogram of ISI of the probabilistic model and the ELIF model. The minimum value of the cost function suggests that both models generate almost the same shape for ISI distribution.

Two spike trains ($A1$ and $B1$) are generated using the enhanced leaky integrate and fire model (ELIF) with directed coupling from $B1$ to $A1$. It is assumed that spike train $A1$ is given and the aim is to adjust the parameters of the probabilistic model in such a way that the probabilistic model will be able to generate spike train A which is similar to the given spike train $A1$ with directed coupling from B to A. The probabilistic model for the generation of two spike trains requires five parameters ($\beta,\ a,\ b,\ c, d$). Here β is the strength of the influence from spike train B to spike train A, (a, b) are the shape and scale parameters of the gamma distribution which is used to generate the ISI, of the spike train B and (c, d) are the scale and shape parameters of the Weibull distribution which is used to find the hazard of the backward recurrence time of spike train A. To start the optimization procedure some initial parameter values of the probabilistic model ($\beta_0,\ a_0,\ b_0,\ c_0, d_0$) are selected. Using these parameter values two spike trains A and B are generated by the probabilistic model. The cost function Q accounts for the difference between the ISI distributions of the spike train $A1$ generated by the ELIF model and the spike train A generated by the probabilistic model:

$$Q = \sum_{i=1}^{k} \left(h_i^{A1} - h_i^{A}\right)^2 \tag{4.11}$$

where $h_i^{A_1}$ and h_i^{A} are the frequency of appearance of ISI in the bin i of the histogram for $A1$ and A respectively and k is the number of bins. The optimal parameter values are obtained for the minimum value of the cost function. Thus, generation of spike train A

using these optimal parameter values gives the best fit to the spike train $A1$ generated by ELIF model.

4.7 Generation of two spike trains by the probabilistic model

In this section an example of the generation of two spike trains is given. Two spike trains, the target (A) and the reference (B) are generated using the probabilistic model. The spike train generation time T is 2000 milliseconds and the time interval $[0\ T]$ is divided into $n\ (n = 200{,}000)$ small sub-intervals of length $\Delta t = T/n\ (= .01)$. A gamma-distributed set of inter-spike intervals (ISI) of the reference spike train B are generated with the shape parameter $a = 5$ and the scale parameter $b = 3$. The reference spike train B is obtained using the cumulative sum of these ISI's over the interval $[0, T]$. The generation of spikes in the target spike train A depends on the strength of influence from the reference spike train B to the target spike train A, the influence function from spike train B to spike train A and the hazard function of the target spike train A without influence. For the generation of spike train A, the strength of influence is considered as $\beta = 1$, the influence function is determined by the formula (4.4) with $\tau_s = 5$, $\Delta = 0$ and the hazard function of spike train A without influence from B, which depends on the backward recurrence time of the target spike train A, is calculated by formula (4.7). For each time t_k, where $t_k = k\Delta$ where $k = 1,2 \cdots \cdots, n$, the backward recurrence times of the reference spike train B and target spike train A are calculated. The influence function of spike train B to spike train A is calculated using the backward recurrence time of the reference spike train B. A Weibull distribution with the scale parameter $c = 15$ and the shape parameter $d = 5$ is considered for the backward recurrence time of the target spike train A and the hazard function is calculated for the backward recurrence time. Based on all these values, the hazard of the target spike train A is calculated using the formula (4.8). If this hazard is greater than a randomly chosen number which has the

uniform distribution in the range [0,1], then it is concluded that a spike is generated in the target spike train *A*.

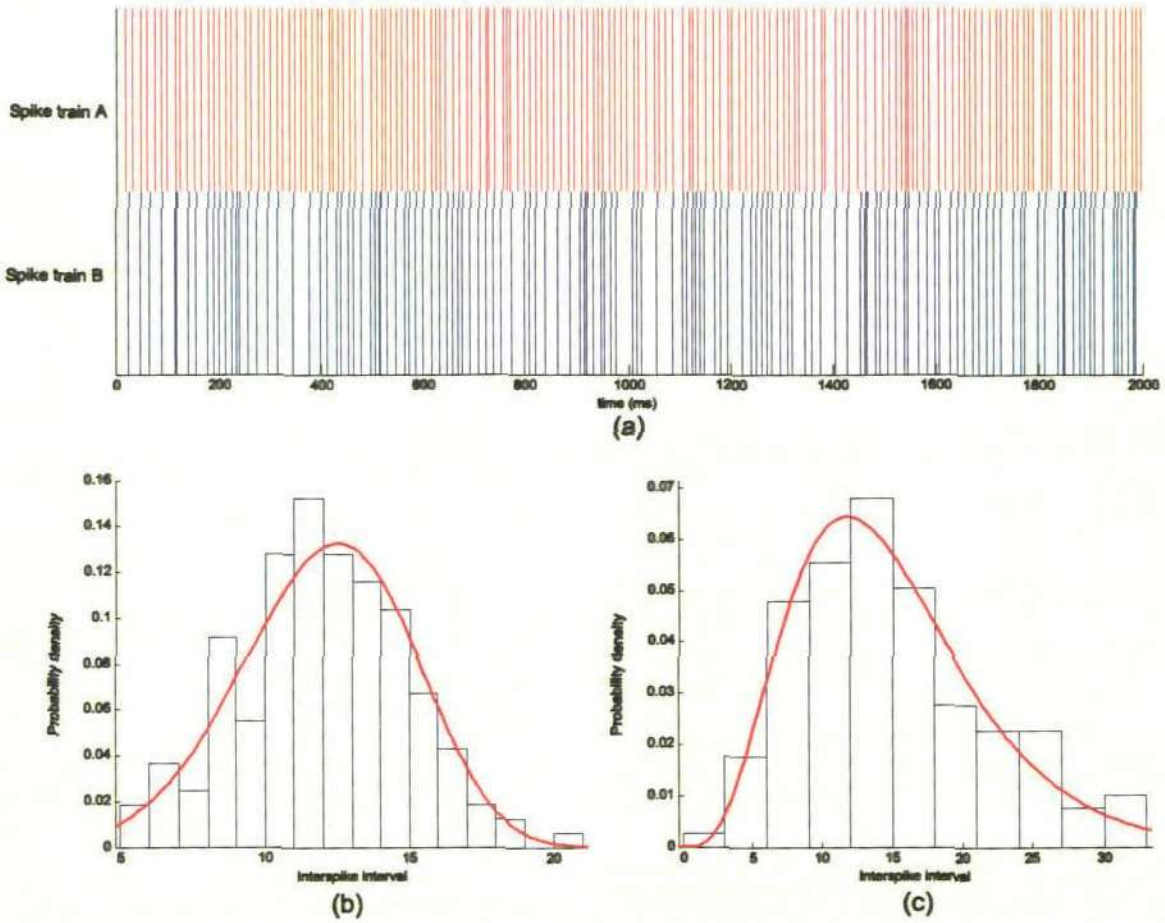


Figure 4.6: (a) Raster plot of the target spike train *A* (red) and the reference spike train *B* (blue). **(b)** ISI histogram of the target spike train *A* superimposed with the fitted Weibull probability density **(c)** ISI histogram of the reference spike train *B* superimposed with the fitted gamma probability density.

The raster plot of the target spike train *A* and reference spike train *B* is shown in Fig. 4.6(a). There are 165 spikes in the target spike train and 133 spikes in the reference spike train. The coefficient of variation of the target spike train *A* is 0.2412 and the reference spike train *B* is 0.4475. Thus, spike train *A* is more regular than the spike train *B*. Fig 4.6(b) shows the interspike interval histogram of the spike train *A*. A fitted Weibull probability density is superimposed on the interspike interval of *A*. The Weibull model under predicts the short ISIs ($\leq$12 ms) and predicts accurately for long ISIs (>12 ms). A fitted gamma probability density is superimposed on the interspike interval

histogram of the reference spike train *B* (Fig. 4.6(b)). The gamma model also underpredicts the short ISIs (<9 ms), the moderate ISIs (between 12 ms and 18 ms) and the long ISIs (>23 ms) but overpredicts the ISIs between 9ms and 11 ms and the ISIs between 19 ms and 22 ms. Fig. 4.6(b) and Fig. 4.6(c) show that there is agreement between the ISI of the generated spike trains and the theoretical probability distribution. To assess how much the probability distribution of ISI of the generated spike trains agrees with the theoretical probability distribution the Q-Q and K-S goodness of fit tests are performed.

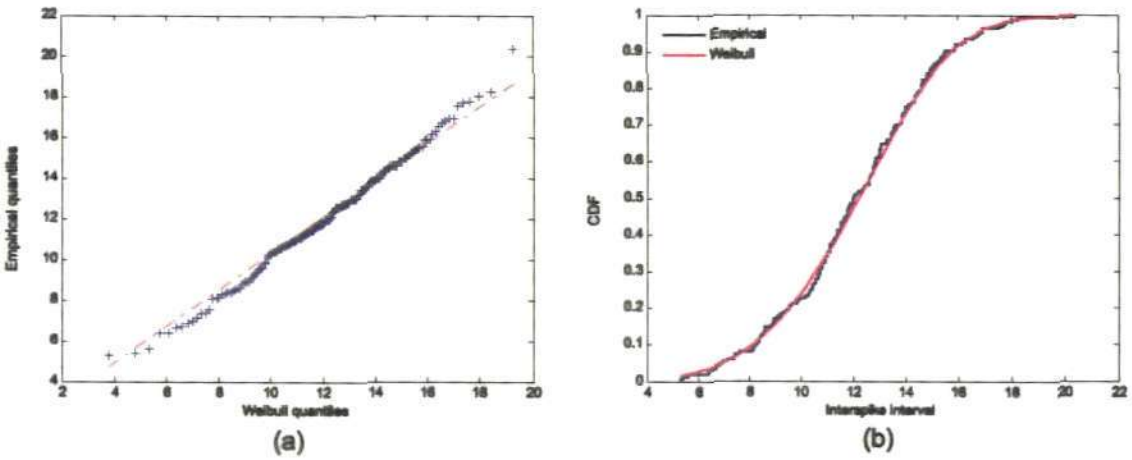


Figure 4.7: (a) Q-Q plot of the ISI of target spike train *A* fitted with the Weibull probability distribution. The red line shows the 45-degree reference line and the blue cross represents points of the empirical and Weibull quantile. **(b)** K-S plot of the ISI of target spike train *A* fitted with the Weibull probability distribution. The red line shows the cumulative distribution function of the fitted Weibull distribution function and the blue line shows the empirical cumulative distribution function of the interspike interval.

The Q-Q plot for the ISI of target spike train *A* is shown in Fig. 4.7(a). The empirical quantiles of the ISI i.e. the order statistics of the ISI are plotted on the vertical axis. The plotting positions of the fitted Weibull quantiles are calculated using the formula (4.10). The Weibull quantiles are computed at the plotting position and are plotted on the horizontal axis. These plotting points are depicted by a blue cross. A 45-degree reference line (red) is shown to aid the comparison of points. All of the plotted points

are very close to the reference line indicating that there is a good agreement between the probability distribution of ISI of the spike train *A* and the Weibull distribution.

Another goodness of fit test, that is K-S test, is used for evaluating the agreement between the probability distribution of ISI of spike train *A* and the Weibull probability distribution. The K-S plot of the ISI is shown in Fig. 4.7(b). The empirical distribution function of the ISI (blue) and the theoretical Weibull distribution function (red) are very similar depicting an accurate estimation of the distribution. The maximum distance of the empirical distribution function and the Weibull distribution function is 0.0355 for which the null hypothesis is accepted at 0.05 significance level. Thus, it can be deducted that the probability distribution of the ISI of spike train *A* follows the Weibull probability distribution.

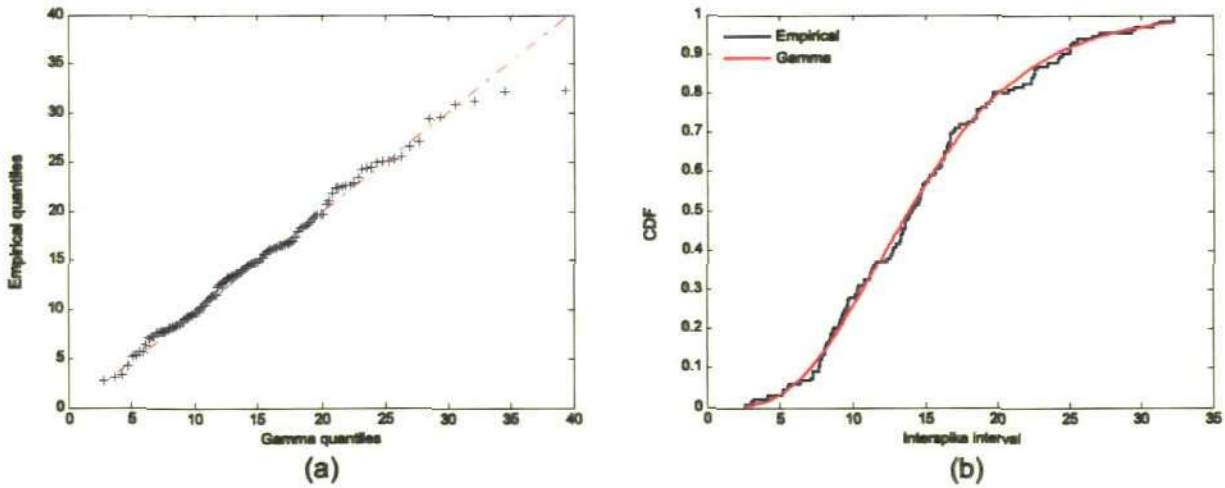


Figure 4.8: **(a)** Q-Q plot of the ISI of reference spike train *B* fitted with a gamma probability distribution. The red line shows the 45-degree reference line and the blue cross represents point of the empirical and gamma quantile. **(b)** K-S plot of the ISI of reference spike train *B* fitted with the gamma probability distribution. The red line shows the cumulative distribution function of the fitted gamma distribution function and the blue line shows the empirical cumulative distribution function of the interspike interval.

A random generator of gamma distribution was used ('gamrnd' is the routine for generating gamma distribution in Matlab). It is anticipated that further testing will reinforce initial conclusion, that the model is accurate that it generates biologically

realistic data. The Q-Q plot for the ISI of target spike train B is shown in Fig. 4.8(a). The empirical quantiles of the ISI i.e. the order statistics of the ISI are plotted on the vertical axis. The gamma quantiles are plotted on the horizontal axis. These plotting points are very close to the 45-degree reference line and indicate that there is an agreement between the probability distribution of ISI of the spike train B and the gamma distribution. The K-S plot of the ISI is shown in Fig. 4.8(b). The empirical distribution function of the ISI (blue) and the theoretical gamma distribution function (red) are very similar. The maximum distance of the empirical distribution function and the gamma distribution function is 0.0408 for which the null hypothesis is accepted at 0.05 significance level. Thus, the probability distribution of ISI of spike train B follows the gamma probability distribution accurately.

4.8 Fitting probabilistic model with ELIF model for two spike trains

In this section an example is demonstrated for the best fitting of the spike trains generated by the ELIF model and the probabilistic model. This fitting derives a set of parameters which enable the generation of spikes which are physiologically realistic. Two spike trains, the target A and the reference B are generated using the ELIF model (Borisyuk, 2002) using a connection from spike train B to spike train A. The connection strength is 12.86 and the time delay of spike propagation is 12 milliseconds. Spike trains A and B are generated for a period of 20,000 milliseconds. It is assumed that the spike train A is given and that the aim is to adjust the parameters of the probabilistic model in such a way that this model will be able to generate spike train $A1$ which is similar to the given spike train A. To generate the spike train $A1$ from a connection of the spike train $B1$ the optimization procedure is initialized using parameter values of the probabilistic model. For the optimization procedure, the initial value of the strength of influence β from $B1$ to $A1$ is considered to be equal to 1. Estimation of the parameters

of ISI of spike train B using the gamma distribution gives the initial value of the parameters (a, b) of gamma distribution which are (3.3191, 23.5735). Estimation of the parameters of the ISI of spike train A using Weibull distribution gives the initial value of the parameters (c, d) of Weibull distribution which are (94.5186, 1.9060). Thus the optimization procedure begins with the initial value $(\beta, a, b, c, d) = (1, 3.31, 23.57, 94.51, 1.90)$. The cost function Q is calculated using (4.11) which takes account for the sum of the squared difference in the frequency corresponding to each bin of ISI of the spike train A and the ISI of spike train $A1$. The number of bins for the optimization procedure is taken as 10 and the minimum value of the cost function Q is 5.8808×10^{-4}. For this cost function, the optimal parameter values are obtained as $(\beta, a, b, c, d) \approx (1.02, 3.23, 21.19, 97.43, 1.96)$.

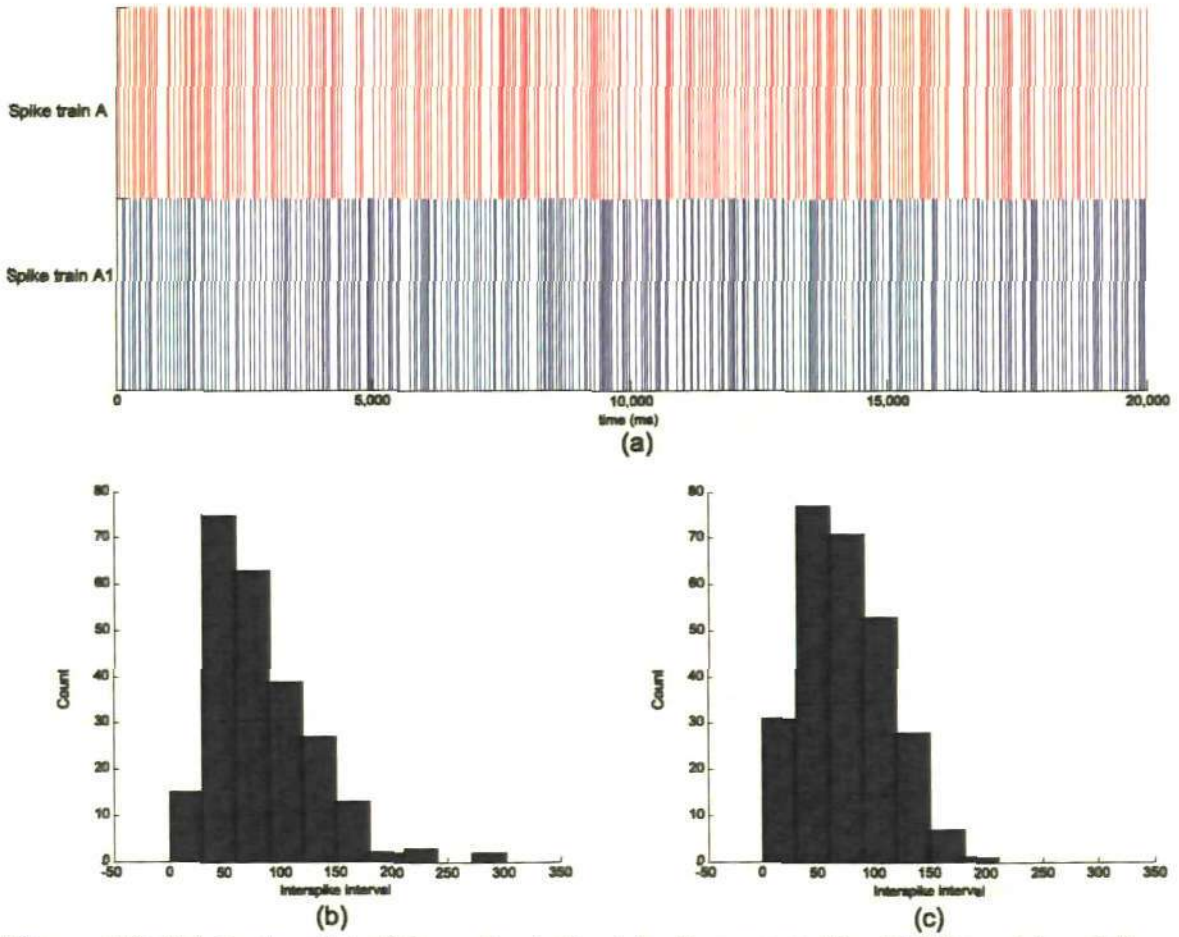


Figure 4.9: (a) Raster plot of the spike train A (red) generated by ELIF model and the spike train $A1$ (blue) generated by probabilistic model for the period of 20,000 ms. **(b)** ISI histogram of the spike train A generated by the ELIF model **(c)** ISI histogram of the spike train $A1$ generated by probabilistic model.

Now using the optimal parameter values spike trains *A1* and *B1* are generated for a period of 20,000 ms with $\tau_s = 5$, time lag $\Delta = 12$. Fig. 4.9(a) shows the raster plot of the spike train *A* generated by ELIF model and spike train *A1* generated by probabilistic model. The number of spikes in spike train *A* is 240 and the number of spikes in spike train *A1* is 269. The coefficients of variation of these spike trains are very similar which are 0.5621 and 0.5004 respectively. Fig. 4.9(b) shows the ISI histogram for the ELIF generated spike train *A* and Fig. 4.9(c) shows the ISI histogram for the spike train *A1* generated by the probabilistic model and these histograms are very similar.

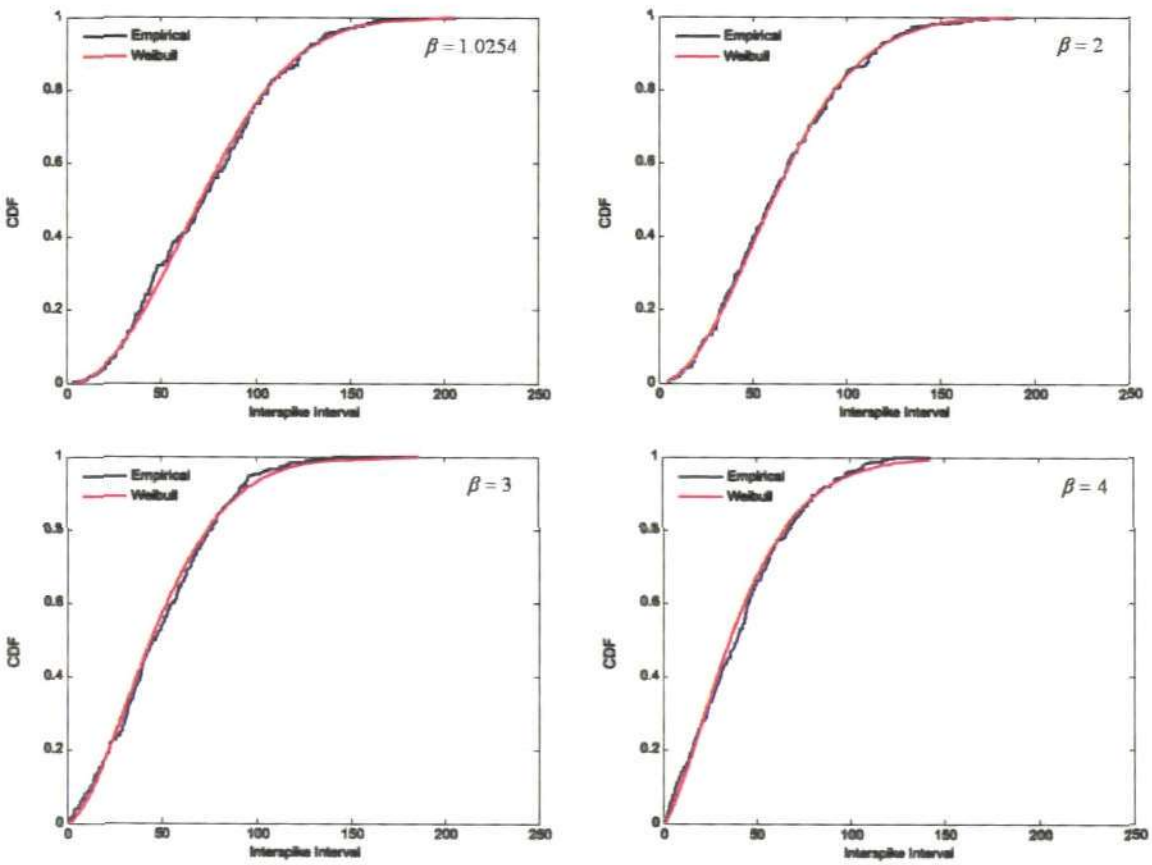


Figure 4.10: K-S plot of the ISI of spike train *A1* generated by the probabilistic model for different strength of influence β. The red line shows the cumulative distribution function of the fitted Weibull probability distribution and the blue line shows the empirical cumulative distribution function of the interspike interval.

It is assumed that, even in the case of influence from spike train, the probability distribution of the ISIs of *A1* can be approximated by the Weibull distribution. To assess how well the ISI of the generated spike train *A1* agrees with the Weibull probability

distribution, the K-S goodness of fit test is used. Spike train *A1* is generated using different strengths of influence $\beta = (1.0254, 2, 3, 4)$ whilst keeping all other parameter values set to their optimal value. The K-S plot of the ISI of spike train *A1* is shown in Fig. 4.10 for each strength of influence β. The empirical distribution function of the ISI (blue) and the theoretical Weibull distribution function (red) show a good agreement for all strengths of influence. The maximum distances of the empirical distribution function and the Weibull distribution function for different strength of influences $\beta = (1.0254, 2, 3, 4)$ are 0.0536, 0.0234, 0.0446 and 0.0558, respectively. For all the strength of influences the null hypothesis is accepted at 0.05 significance level. That is the ISI of spike train *A1* follows the Weibull distribution.

Chapter 5

Statistical technique for the analysis of functional connections of multiple spike trains

A new statistical technique is presented for analyzing functional connectivity of simultaneously recorded multiple spike trains. This method is based on the theory of Modulated Renewal Process (MRP) and it estimates a vector of influence strengths from multiple spike trains (called reference trains) to the selected (target) spike train. Selecting another target spike train and repeating the calculation of the influence strengths from the reference spike trains enables researchers to find all functional connections among multiple spike trains. Application of the Cox method to multiple sets of data generated by the ELIF model (Borisyuk, 2002) with a prescribed architecture of connections suggest that this method is highly successful for analyzing functional connectivity of simultaneously recorded multiple spike trains.

5.1 Introduction

Development of Multi-Electrode Arrays (MEA) enables researchers to record multiple spike trains simultaneously from associated neurons. Simultaneously recorded multiple spike trains are used to study how groups of neurons process information and how they interact with each other. Developing a new statistical method for analyzing multiple spike trains and, in particular, estimating the functional connectivity between spike trains, is a challenging problem that has resulted in substantial research (Brown et al., 2004; Reed and Kaas, 2010).

The standard approach for analyzing functional connectivity is based on the calculation of the CCF (Perkel et al., 1967). Though the CCF is widely used in neuroscience study, this technique has several limitations. This technique assumes that the spike trains are stationary. It also assumes that the number of spikes should be large enough to ensure reliable estimation. This technique is usually applied to characterize the dependencies between pairs of spike trains at a time, ignoring possible effects from other spike trains which can lead to inaccurate results in many cases (Okatan et al., 2005).

In this chapter, a new statistical technique called the Cox method (Cox 1972; Borisyuk et al., 1985) based on MRP is developed for analyzing functional connectivity of simultaneously recorded multiple spike trains. The MRP is considered in terms of spike generation under multiple influences from other spike trains and estimates of the strength of each influence using the Cox method, which itself is based on the conditional likelihood method (Cox, 1972). The MRP model describes the hazard function of spike appearance at the MRP and it includes a modulation which is the exponential of the linear combination of influence functions. In fact, this model is similar to the regression model and the set of influence strengths is similar to the regression coefficients. The definition of the influence function is based on some neurobiological details of spike generation and propagation. This influence function reflects the dynamics of postsynaptic potential under bombardment by spikes from other neurons.

The Cox method (Borisyuk et al., 1985) was developed for analyzing pairs and triplets of spike trains. Here it is further developed to support simultaneous consideration of any possible set of multiple spike trains. The corresponding formulas for the calculation of estimates of the influence strengths and their confidence intervals have been derived. Thus, this new development of the Cox method enables researchers to simultaneously

analyze any number (n) of spike trains. In comparison to existing techniques, the Cox method has the following advantages: it does not require the specification of a bin (binless method); it is applicable in situations where sample sizes are small; it is sufficiently sensitive such that it estimates weak influences; it supports the simultaneous analysis of multiple spike trains and provides statistical estimates of influence strengths and their confidence interval (to test the null hypothesis that the influence is zero); it is able to identify a correct connectivity scheme in difficult cases of 'common source' or 'indirect' connectivity.

5.2 The Cox method

Cox (1972) considered a point process in which the hazard function of a renewal process is expressed by a factor depending on quantities. The quantities are thought to influence the probability of occurrence and each quantity is combined with an unknown regression coefficient. The unknown regression coefficients depict the strength of influence of the probability of occurrence. Cox suggested a statistical method to estimate these unknown regression coefficients using a conditional maximum likelihood principle. Application of the Cox method to analyze influences between two and three spike trains is described by Borisyuk et al. (1985). In this section a generalization of the Cox method for simultaneous analysis of arbitrary number (n) of spike trains is developed.

The Cox method is based on the MRP and the MRP allows introduction of dependencies (influences) between spike trains. It is assumed that for two spike trains A and B, spike generation in spike train A depends on spikes of spike train B and the hazard function of spike train A is a product of two multipliers: one is the own hazard of spike train A without influence from B and another multiplier describes influence from spike train B. Thus, the hazard function at the moment t is:

$$\varphi(t) = \varphi_A\big(U_A(t)\big)\, exp\{\beta Z_B(t)\}, \qquad (5.1)$$

where $\varphi_A(.)$ is the hazard function of the spike train A without influence from the spike train B, $U_A(t)$ is the backward recurrence time of the spike train A at the moment t, $Z_B(t)$ is the influence function determining how the spike train B influences the spike train A, and β is the unknown parameter (Cox coefficient) describing the strength of the influence from spike train B to A. Therefore, given the influence function $Z_B(t)$, the goal is to estimate the parameter β. If $\beta = 0$ then there is no influence from spike train B to A. To test the null hypothesis $H_0: \{\beta = 0\}$ the statistical technique based on conditional maximum likelihood (Cox, 1972) is used.

The Cox method is applied to analyze a set of n simultaneously recorded spike trains. One spike train is selected to be considered as a target spike train and all other $(n-1)$ spike trains are considered to be the reference spike trains. The Cox method allows analyzing of all n spike trains and estimating the $(n-1)$ dimensional vector β of regression coefficients under the assumption (5.1), where $Z_B(t)$ is $(n-1)$ dimensional vector-function of influences from reference spike trains to the target and $\beta Z_B(t)$ is the dot product. Application of the Cox method provides both the estimates of unknown parameter (Cox coefficients) $\left(\hat{\beta}_1, \hat{\beta}_2, \cdots\cdots, \hat{\beta}_{n-1}\right)$ and the corresponding confidence intervals of these estimates $\{[\, lb_i\, , ub_i],\ i = 1,2, \ldots (n-1)\}$, where lb_i and ub_i are lower and upper boundaries respectively of the confidence interval for $\hat{\beta}_i$. The null hypothesis H_0^i: $\beta_i = 0$ is accepted if the corresponding confidence interval contains zero $(0 \in [lb_i, ub_i])$ otherwise the null hypothesis is rejected and the estimate $\hat{\beta}_i$ is considered as a measure of influence strength from the ith reference spike train to the target. To study the complete diagram of functional connectivity this method is applied consequently (n times) selecting the target and estimating the influence strengths from reference spike trains.

5.2.1 Derivation of the formula

A derivation of the formula of the Cox method follows the paper by Cox (1972). The derivation of the formula for two and three spike trains was done by Borisyuk et al. (1985). In this section the derivation of the formula of the Cox method is done for (p+1) spike trains. To describe dependencies and influences among (p+1) spike trains a spike train is selected as a target spike train denoted by A, other p spike trains are selected as reference spike trains and they are denoted by $B = (B_1, B_2, \cdots\cdots, B_p)$. The goal is to estimate the vector of unknown parameters $\beta = (\beta_1, \beta_2, \cdots\cdots, \beta_p)$ which describe the strengths of influences from reference spike trains $B = (B_1, B_2, \cdots\cdots, B_p)$ to the target spike train A. Thus β_m $(m = 1,2, \cdots\cdots, p)$ represents the strength of influence from the reference spike train B_m $(m = 1,2, \cdots\cdots, p)$ to the target A. The main assumption is that the target spike train A is the MRP with the hazard function:

$$C_A(t) = \varphi_A\big(U_A(t)\big) exp\left\{\sum_{m=1}^{p} \beta_m Z_{B_m}(t)\right\}, \tag{5.2}$$

where $C_A(t)$ is the hazard of the MRP A, $\varphi_A(.)$ is the hazard function of the renewal process A without modulation (i.e. without influence from another spike trains), $U_A(t)$ is the backward recurrence time of the renewal process A at the moment t, $Z_{B_m}(t)$ $(m = 1,2, \cdots\cdots, p)$ is the influence function determining how the reference spike train B_m $(m = 1,2, \cdots\cdots, p)$ influences the renewal process A, and β_m is the parameter describing the influence strength from spike train B_m $(m = 1,2, \cdots\cdots, p)$ to the target A. To estimate the parameters $\beta = (\beta_1, \beta_2, \cdots\cdots, \beta_p)$ the method of conditional likelihood is used which eliminates the nuisance function $\varphi_A(.)$ (Cox, 1972). The influence function $Z_{B_m}(t)$ is defined by:

$$Z_{B_m}(t) = \frac{g_m}{(\tau_s - \tau_r)}\left(e^{-\frac{U_{B_m}(t-\Delta)}{\tau_s}} - e^{-\frac{U_{B_m}(t-\Delta)}{\tau_r}}\right), \quad (t - \Delta) > 0 \tag{5.3}$$

where τ_s and τ_r are the characteristic times of decay and rise of postsynaptic potential respectively. Parameter g_m provides the normalization that the maximum of the influence function is one and is defined by:

$$g_m = \frac{\tau_s - \tau_r}{e^{-\frac{t_m}{\tau_s}} - e^{-\frac{t_m}{\tau_r}}}$$

where

$$t_m = \frac{log(\tau_s/\tau_r)}{\frac{1}{\tau_r} - \frac{1}{\tau_s}}.$$

$U_{B_m}(t)$ is the backward recurrence time of spike train B_m and Δ is the time lag corresponding to delay of spike propagation from B_m to A. A simplified version of the influence function corresponding to the case $\tau_s = \tau_r$ is given by the following formula:

$$Z_{B_m}(t) = \frac{1}{\tau_s} U_{B_m}(t-\Delta) e^{1-\frac{U_{B_m}(t-\Delta)}{\tau_s}}, \quad (t-\Delta) > 0 \tag{5.4}$$

It is assumed that spike train A contains n interspike intervals $x_1, x_2, \ldots\ldots, x_n$. For simplicity all intervals $x_1, x_2, \ldots\ldots, x_n$ are considered as of different length. If there are several identical intervals a randomization procedure is used and a small normally distributed random number to the interval length is added. The intervals are arranged in order of increasing size $x_{(1)} < x_{(2)} < \cdots\cdots < x_{(n)}$. For $i > j$, assume that $x_{(i)} = x_k$ and $x_{(j)} = x_l$. Now $Z_{B_{mij}}$ is defined to be the value of $Z_{B_m}(t)$ $(m = 1,2,\cdots\cdots,p)$ where time t is calculated in the following way: the interval x_l is allocated inside of the interval x_k and the left ends of both intervals coincide, time t corresponds to the right end of the interval x_l. Respectively, $Z_{B_{mil}}$ is the value of $Z_{B_m}(t)$ at the right end of the interval x_k.

The likelihood is constructed for the data conditionally on the magnitudes of the intervals, by considering contributions in order starting with the smallest interval. Thus the contribution from the first interval is:

$$\frac{exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{m11}}\right)}{\sum_{l=1}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{ml1}}\right)}.$$

Conditionally on the first interval, the contribution from the second interval is

$$\frac{exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{m22}}\right)}{\sum_{l=2}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{ml2}}\right)}$$

and so on.

The likelihood has the form

$$\prod_{i=1}^{n}\frac{exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mii}}\right)}{\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}.$$

The log likelihood is defined as

$$L(\beta_1, \beta_2, \cdots\cdots, \beta_p) = \sum_{i=1}^{n}\sum_{m=1}^{p}\beta_m Z_{B_{mii}} - \sum_{i=1}^{n} log\left\{\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\right\}.$$

Now the first derivative of the log likelihood is

$$\frac{\partial L}{\partial \beta_m} = \sum_{i=1}^{n} Z_{B_{mii}} - \sum_{i=1}^{n}\left[\frac{\sum_{l=i}^{n} Z_{B_{mli}}\, exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}{\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}\right]; (m = 1,2, \cdots\cdots, p).$$

The estimate for β_m is obtained by setting the first derivative to zero.

The second derivative can be obtained by

$$\frac{\partial^2 L}{\partial \beta_r \partial \beta_s} = -\sum_{i=1}^{n}\left[\frac{\sum_{l=i}^{n} Z_{B_{rli}}\, exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right) Z_{B_{sli}} \sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}{\left[\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\right]^2}\right]$$

$$+\sum_{i=1}^{n}\left[\frac{\sum_{l=i}^{n} Z_{B_{rli}}\, exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\sum_{l=i}^{n} Z_{B_{sli}}\, exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}{\left[\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\right]^2}\right];$$

$$(r, s = 1,2, \cdots\cdots, p).$$

$$= -\sum_{i=1}^{n}\left[\frac{\sum_{l=i}^{n} Z_{B_{rli}} Z_{B_{sli}} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}{\left[\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\right]^2}\right]$$

$$+\sum_{i=1}^{n}\left[\frac{\sum_{l=i}^{n} Z_{B_{rli}}\, exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\sum_{l=i}^{n} Z_{B_{sli}}\, exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}{\left[\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\right]^2}\right];$$

$$(r, s = 1,2, \cdots\cdots, p).$$

$$= -\sum_{i=1}^{n}\left[\frac{\sum_{l=i}^{n} Z_{B_{rli}} Z_{B_{sli}} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}{\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}\right]$$

$$+\sum_{i=1}^{n}\left[\frac{\sum_{l=i}^{n} Z_{B_{rli}}\, exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\sum_{l=i}^{n} Z_{B_{sli}}\, exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)}{\left[\sum_{l=i}^{n} exp\left(\sum_{m=1}^{p}\beta_m Z_{B_{mli}}\right)\right]^2}\right];$$

$$(r, s = 1,2,\cdots\cdots,p).$$

Let $U(\beta) = \left[\frac{\partial L}{\partial \beta_m}\right]_{p\times 1}$ be the score vector and $I(\beta) = \left[-\frac{\partial^2 L}{\partial \beta_r \partial \beta_s}\right]_{p\times p}$ is the Fisher information matrix.

Now to obtain maximum likelihood estimate $\hat{\beta}_q$ a numerical solution of the equation $U(\beta) = 0$ should be obtained. The Newton-Raphson iterative method starting from the initial guess $\beta^{(0)}$ is used. Formula for iterations is:

$$\hat{\beta}_q = \hat{\beta}_{q-1} + \left[I\left(\hat{\beta}_{q-1}\right)\right]^{-1} U\left(\hat{\beta}_{q-1}\right),\ q = 1,2,\ldots,Q. \tag{5.5}$$

The iterations converge to the estimates:

$$\hat{\beta}_q \underset{q\to\infty}{\longrightarrow} \hat{\beta},\ \hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \cdots\cdots, \hat{\beta}_p)\ .$$

To obtain the confidence interval for β_m, it is assumed that $\frac{\partial L}{\partial \beta_m}$ has asymptotically normal distribution $N\left(0, I(\beta_m)\right)$ where $I(\beta_m)$ is the mth diagonal element of $I(\beta)$. The confidence interval with the confidence level γ is obtained by the following equation:

$$\left[\hat{\beta}_m - 1/\sqrt{I\left(\hat{\beta}_m\right)}K_{\frac{1-\gamma}{2}}\ ,\quad \hat{\beta}_m + 1/\sqrt{I\left(\hat{\beta}_m\right)}K_{\frac{1-\gamma}{2}}\right] \tag{5.6}$$

where $K_{(1-\gamma)/2}$ is the upper $(1-\gamma)/2$ quantile of normal distribution. The null hypothesis that spike train $B_m\,; (m = 1,2,\cdots\cdots,p)$ does not influence A is accepted if the confidence interval includes zero.

To obtain the confidence region for β_r and β_s, it is assumed that $U^T(I_{rs})^{-1}U$ has χ^2 distribution with two degrees of freedom. The use of χ^2 distribution to obtain

confidence region is a natural choice in statistics. Here $U = \left(\frac{\partial L}{\partial \beta_r}, \frac{\partial L}{\partial \beta_s}\right)$ and the matrix I_{rs} is

$$I_{rs} = \begin{bmatrix} -\frac{\partial^2 L}{\partial \beta_r^2} & -\frac{\partial^2 L}{\partial \beta_r \partial \beta_s} \\ -\frac{\partial^2 L}{\partial \beta_s \partial \beta_r} & -\frac{\partial^2 L}{\partial \beta_s^2} \end{bmatrix}.$$

The confidence region on the plane (β_r, β_s) with the confidence level γ is defined by the following equation:

$$(\hat{\beta}_r - \beta_r, \hat{\beta}_s - \beta_s)' [I_{rs}(\hat{\beta}_r, \hat{\beta}_s)]^{-1} (\hat{\beta}_r - \beta_r, \hat{\beta}_s - \beta_s) \leq \chi^2_{(1-\gamma,2)} \qquad (5.7)$$

where $\chi^2_{(1-\gamma,2)}$ is the upper $(1-\gamma)$ quantile of chi-square distribution with two degrees of freedom. The null hypothesis that Cox coefficients $(\hat{\beta}_r, \hat{\beta}_s) = 0$ is accepted if the confidence region includes zero.

5.2.2 Parameters of the influence function (τ_r, τ_s and Δ)

To apply the Cox method to multiple spike trains, one spike train should be selected as a target and other spike trains are considered as the reference spike trains. The influence functions should be specified to describe the influence from the reference spike trains to the target. In this section it is assumed that all influences are identical and the influence function is specified by (5.3). This function includes three parameters (τ_s, τ_r, Δ) and their values should be defined for each reference spike train.

In the case of generated data the characteristic times of rise (τ_r) and decay (τ_s) of the postsynaptic potential (PSP) are usually known but they are difficult to determine from experimental data and should be assumed (Lansky and Ditlevsen, 2008). How the result of analyzing the functional connectivity depends on chosen values of PSP characteristic times was studied and it was found that the Cox method has low sensitivity to the selected values of these parameters. In other words, there is no requirement to choose

these times accurately. In fact, these characteristic times can be varied over a broad range and the results of the analysis will be similar.

To specify the time lag Δ corresponding to the delay of spike propagation, the traditional CCF is used which provides both the statistical estimate of dependency and the corresponding time lag Δ. To do this a pair of spike trains, the target and the reference spike train, are considered. The highest value of the CCF exceeding the upper boundary of the confidence interval can be considered as a measure of influence strength from the reference spike train to the target spike train, and the corresponding time shift can be considered the time lag Δ in spike propagation (Stuart et al., 2005). If there are no values of the CCF exceeding the upper boundary of the confidence interval, both the influence strength and the corresponding time lag Δ are considered zero. For example a pair of spike trains, A (target) and B (reference), are generated with a connection from B to A by the ELIF model (Borisyuk, 2002). The connection strength from B to A is considered 18.04 with a time delay $\Delta = 11$ ms of spike propagation. The CCF of this pair of spike trains is shown in Fig. 5.1(c) which indicates the influence strength is 4.673 and the corresponding time shift is 11 ms.

The time lag Δ of spike propagation can also be found by analyzing a pair of spike trains with the Cox method. The same spike trains A and B in the above example are considered and the aim is to find the time delay of spike propagation using the Cox method. It is assumed that two spikes of the reference spike train B appear at times t_B^1 and t_B^2 and there is a time delay δ_0 of spike propagation from the reference spike train B to the target spike train A. This indicates that if there is a spike in train B at time moment t_B^2 then the probability of spike at train A at time moment $t_A = t_B^2 + \delta_0$ is very high. The time delay δ_0 is unknown, therefore the estimation of the Cox coefficient for different values of time lag Δ is repeated. Fig. 5.1(a) shows estimates $\hat{\beta}_{BA}$ versus time

lag Δ. The increment of time lag is 1 ms and the corresponding confidence intervals are shown by vertical bars. The estimate $\hat{\beta}_{BA}$ increases with increase of Δ and reaches its highest value at Δ= 10 ms but for Δ= 11 ms this coefficient drops down to a negative value (Fig. 5.1(a)). The estimation of the Cox coefficient for values of the time lag in the interval [10.95, 11.03] (ms) with an increment of 0.01 ms. is shown in Fig. 5.1(b). The Cox coefficient drops down from a high positive value to a negative value in a small interval [10.99, 11] (ms).

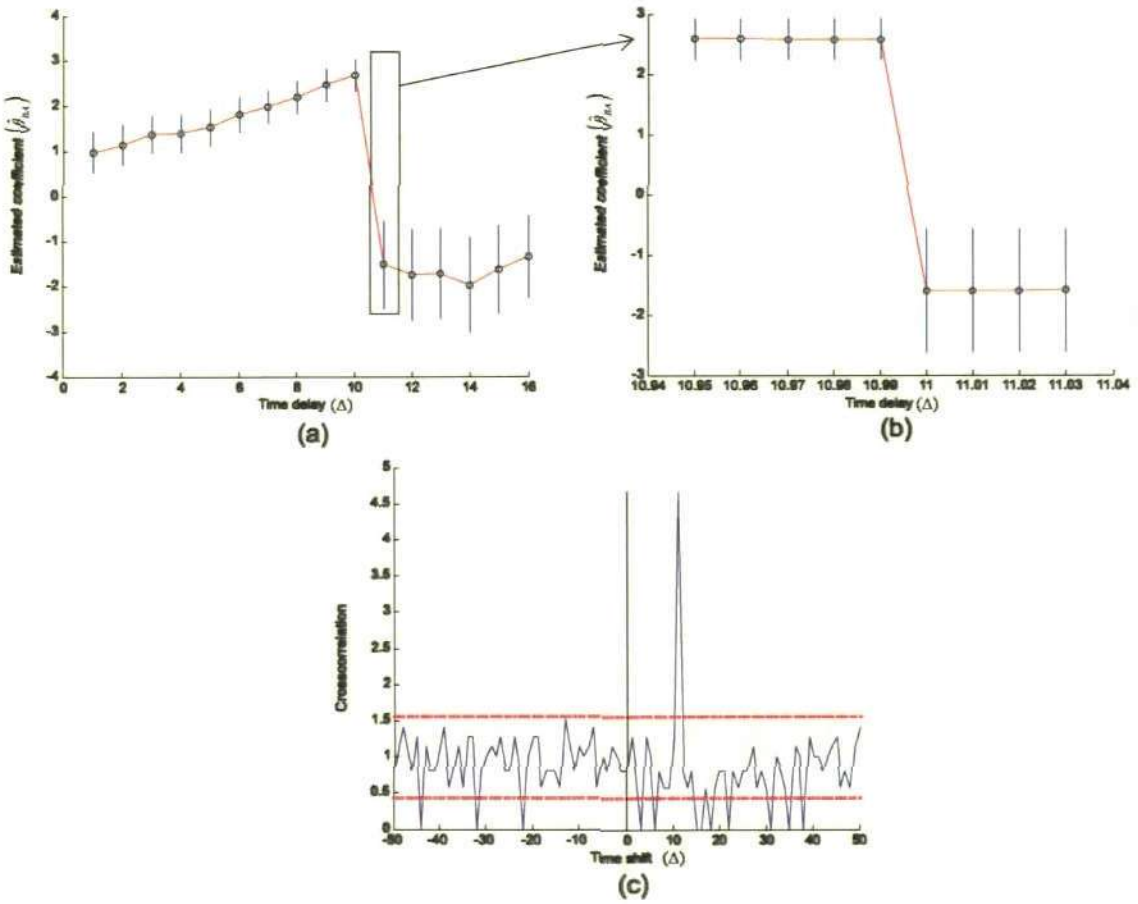


Figure 5.1: Pair-wise analysis of Cox method and the CCF for the analysis of influence strength from spike train *B* to spike train *A*. **(a)** Cox coefficient $\hat{\beta}_{BA}$ with different time delays. **(b)** Cox coefficient $\hat{\beta}_{BA}$ in the interval 10.95 ms to 11.03 ms with a small step of time delays. **(c)** The CCF for the analysis of influence strength from spike train *B* to *A*.

Thus the conclusion is that a time delay of spike propagation is considered to be 11 ms ($\delta_0 = 11$) and an estimate of the Cox coefficient $\hat{\beta}_{BA} = 2.1$. To justify this interpretation of the data analysis, it is assumed that a chosen time lag is smaller than

the time delay of spike propagation: $\Delta_S < \delta_0$. According to formula (5.3), the backward recurrence time is calculated at the moment $(t_A - \Delta_S)$ and this backward recurrence time is smaller than the time delay of spike propagation: $U_B(t_A - \Delta_S) < \delta_0$. Therefore, the value of the influence function depends on the backward recurrence time $Z_B(U_B(t_A - \Delta_S))$, which is shown by the circle in Fig. 5.2. This value is less than the maximum of the influence function and if the time lag Δ_S increases, then the influence function also increases and tends to the maximum of the influence function if the time lag tends to δ_0. A described calculation of the backward recurrence time can be applied in a small vicinity of each spike of the train B under the main assumption that the time delay of spike propagation is δ_0. This consistency in calculation of the backward recurrence time is important for a reliable numerical procedure for calculation of estimate of the Cox coefficient.

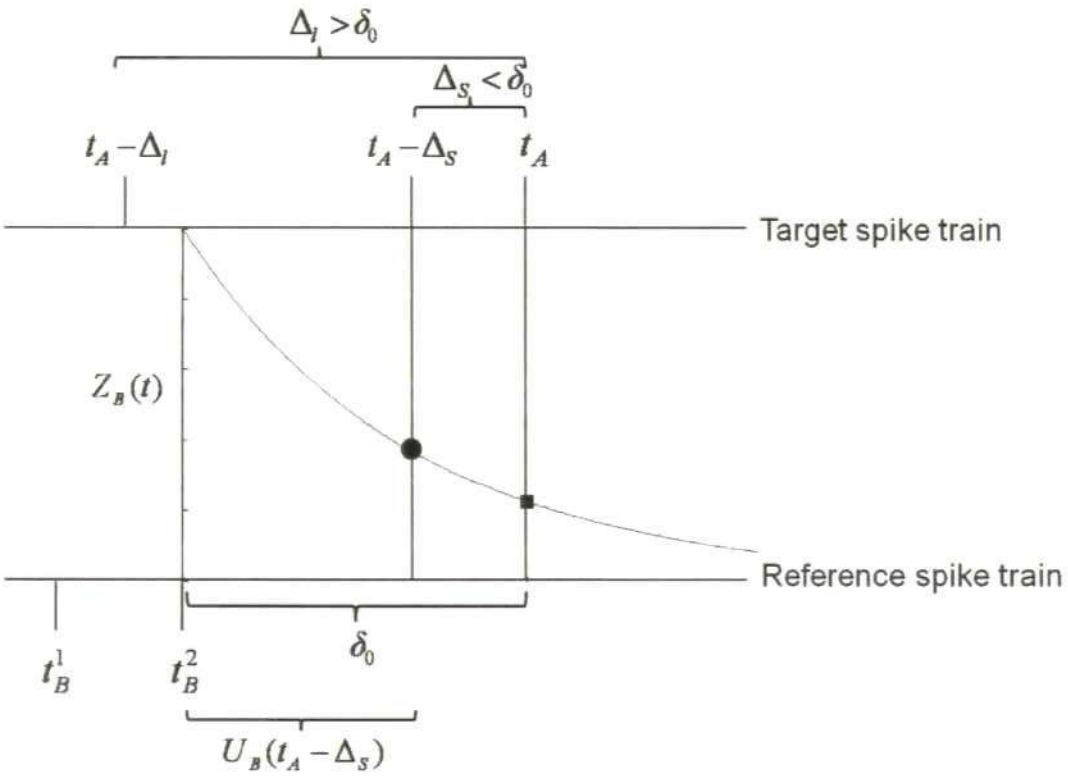


Figure 5.2: Illustration for the calculation of influence function.

Now it is assumed that a time lag Δ_l of the influence function is larger than the time delay of spike propagation $\Delta_l > \delta_0$ (Fig. 5.2) and the calculation of the backward recurrence time is based on the spike at time t_B^1, which is a previous spike according to the spike at moment t_B^2 (Fig. 5.2). Thus, for the time lag $\Delta_l > \delta_0$ the backward recurrence time can get an arbitrary value. The backward recurrence time is calculated for each spike of train B, therefore, the estimate of the Cox coefficient is calculated using arbitrary (random) values of the influence function. Thus, the estimate will be very different from the correct value and it might be zero or a negative number. The time lag Δ obtained by the CCF and pair-wise Cox method is same. In this chapter the lag Δ is calculated by the CCF method for analyzing functional connectivity of multiple spike trains.

5.3 Analysis of functional connectivity by the Cox method

The Cox method is used for the analysis of functional connectivity of two spike trains, three spike trains and the general case of p spike trains. The choice of analysis depends on the nature of the experiment and the desire of the experimenter. In this section the Cox method procedure for the analysis of functional connectivity based on two spike trains, three spike trains and the general case of p spike trains are discussed.

5.3.1 Analysis of two spike trains

To identify the functional connectivity of two spike trains, one spike train is considered as the target spike train and another one is considered as the reference spike train. Assume that spike train A is the target spike train and spike train B is the reference spike train. The Cox method is used to identify the influence strength from reference spike train B to the target spike train A. The hazard function (5.2) of the target spike train A with $p = 1$ can be represented as:

$$C_A(t) = \varphi_A(U_A(t))exp\{\beta Z_B(t)\}.$$

The influence function $Z_B(t)$ from B to A is determined by (5.3) with the specified values of τ_s, τ_r and Δ. The values τ_s, τ_r are assumed and time lag Δ is obtained from pair-wise CCF. The Cox coefficient β is estimated using (5.5) and the corresponding confidence interval is calculated using (5.6). The null hypothesis that the reference spike train B does not influence the target spike train A is tested using the confidence interval. If the confidence interval contains zero, then it is concluded that the corresponding Cox coefficient is not distinguishable from zero; therefore the functional connection from the reference spike train B to the target spike train A is absent. If the confidence interval does not include zero, it is concluded that there is a significant influence from the reference spike train B to target spike train A and the value of the estimate characterizes the strength of this functional connection.

5.3.2 Analysis of three spike trains

To identify the functional connectivity of three spike trains, one spike train is considered as the target spike train and the two other spike trains are considered as the reference spike trains. Assume that three spike trains, A, B and C are used to identify the functional connectivity between them. To understand the procedure of analysing three spike trains, assume that spike train A is the target spike train, and that B and C are the reference spike trains. The Cox method is used to identify the influence strength from the reference spike trains B and C to the target spike train A. The hazard function (5.2) of the target spike train A with $p = 2$ can be represented as:

$$C_A(t) = \varphi_A(U_A(t))exp\{\beta_B Z_B(t) + \beta_C Z_C(t)\}.$$

The influence function $Z_B(t)$ from B to A and $Z_C(t)$ from C to A are determined by (5.3) with the specified values of τ_s, τ_r and Δ. In this chapter it is assumed that for both influence functions $Z_B(t)$ and $Z_C(t)$ the values of τ_s and τ_r are the same, in fact the

values can be different for $Z_B(t)$ and $Z_C(t)$. The parameter time lag Δ is obtained by the CCF method for $Z_B(t)$ and $Z_C(t)$. The Cox coefficients β_B and β_C are estimated using (5.5) and the corresponding confidence intervals are calculated using (5.6). The confidence region for the coefficients β_B and β_C is calculated using (5.7). The confidence region has an elliptical shape and the centre of the confidence region is located at the point $(\hat{\beta}_B, \hat{\beta}_C)$. The null hypothesis that the pair of Cox coefficients $(\hat{\beta}_B, \hat{\beta}_C)$ is equal to zero (i.e. both components of the pair are zero) is tested and the null hypothesis is accepted (i.e. the data does not contradict the null hypothesis) if the origin is inside the confidence region, concluding that both connections are absent (i.e. connection strength is zero). If the null hypothesis is rejected then the null hypothesis that one Cox coefficient is equal to zero is tested. This null hypothesis is tested separately for each coefficient. Two projections of the elliptical confidence region to the coordinate axis (β_B) and (β_C) are considered. If projection to the axis (β_B) contains zero then the null hypothesis is accepted and it is concluded that the connection is absent (i.e. the connection strength is zero), otherwise the null hypothesis is rejected and a centre of the interval $(\hat{\beta}_B)$ is considered as strength of connection. Similarly, if a projection to another axis (β_C) contains zero then the null hypothesis is accepted and it is concluded that the connection is absent, otherwise the null hypothesis is rejected and a centre of the interval $(\hat{\beta}_C)$ is considered as strength of connection. To identify the complete diagram of functional connectivity of spike trains *A*, *B* and *C*, the above procedure is repeated considering *B* is a target spike train and *A* and *C* are reference spike trains, and that *C* is a target spike train and *A* and *B* are reference spike trains. Thus, to identify the functional connectivity of three spike trains using the Cox method, six simultaneous statistical tests are done. For this reason a Bonferroni correction is applied to the significance level α and the corrected significance level is considered as $\alpha = (1 - \gamma)/6$.

5.3.3 Analysis of the general case of p spike trains

Besides the analysis of two spike trains and three spike trains, functional connectivity of p spike trains are analysed considering the effects of all spike trains at once. In this method one spike train is considered as a target spike train and the remaining $p-1$ spike trains are considered as reference spike trains. For the analysis of functional connectivity of p spike trains, assume that A_1 is the target spike train and $A_2, \cdots\cdots, A_p$ are the reference spike trains. The Cox method is used to identify the influence strength from the reference spike trains $A_2, \cdots\cdots, A_p$ to the target spike train A_1. The hazard function (5.2) of the target spike train A_1 with $p = p-1$ can be represented as:

$$C_A(t) = \varphi_{A_1}\left(U_{A_1}(t)\right) exp\left\{\beta_{A_2} Z_{A_2}(t) + \beta_{A_3} Z_{A_3}(t) + \cdots\cdots + \beta_{A_p} Z_{A_p}(t)\right\}.$$

The influence functions $Z_{A_i}(t), i = 2,3, \cdots\cdots, p$ are determined by (5.3) with the specified values of τ_s, τ_r and Δ. Here it is assumed that the values of τ_s and τ_r are same for all influence functions and Δ is calculated from the CCF method. The Cox coefficients $\beta_{A_i}, i = 2,3, \cdots\cdots, p$ are estimated using (5.5) and the corresponding confidence intervals are calculated using (5.6). The null hypothesis that the reference spike train $A_i, i = 2,3, \cdots\cdots, p$ does not influence the target spike train A_1 is tested using the confidence interval. If the confidence interval contains zero, then it is concluded that the corresponding Cox coefficient is not distinguishable from zero, therefore the functional connection from the reference spike train $A_i, i = 2,3, \cdots\cdots, p$ to the target A_1 is absent. If the confidence interval does not include zero, it is concluded that there is a significant influence from the reference spike train $A_i, i = 2,3, \cdots\cdots, p$ to the target A_1 and the values of the estimates characterize the strengths of this functional connections. To identify the complete diagram of functional connectivity of spike trains $A_1, A_2, \cdots\cdots, A_p$, the above procedure is repeated considering A_2 is target and spike trains $A_1, A_3 \cdots\cdots, A_p$ are reference; A_3 is target and spike trains $A_1, A_2, A_4 \cdots\cdots, A_p$ are

reference and so on. Thus, for p spike trains this procedure should be repeated p times and for this repetitive application of Cox method the Bonferroni correction is applied to the significance level α and the corrected significance level is considered as $\alpha = (1-\gamma)/p(p-1)$.

5.4 Simultaneous analysis of p spike trains gives better result than pairs and triplets

In this section using an example it is shown that how the best result of functional connectivity can be obtained by the Cox method for multiple spike trains. A small neural circuit of four spike trains is generated using the ELIF model (Borisyuk, 2002) with a given scheme of coupling. Functional connectivity of four generated spike trains are analyzed using pairs, triplets and all four spike trains at once.

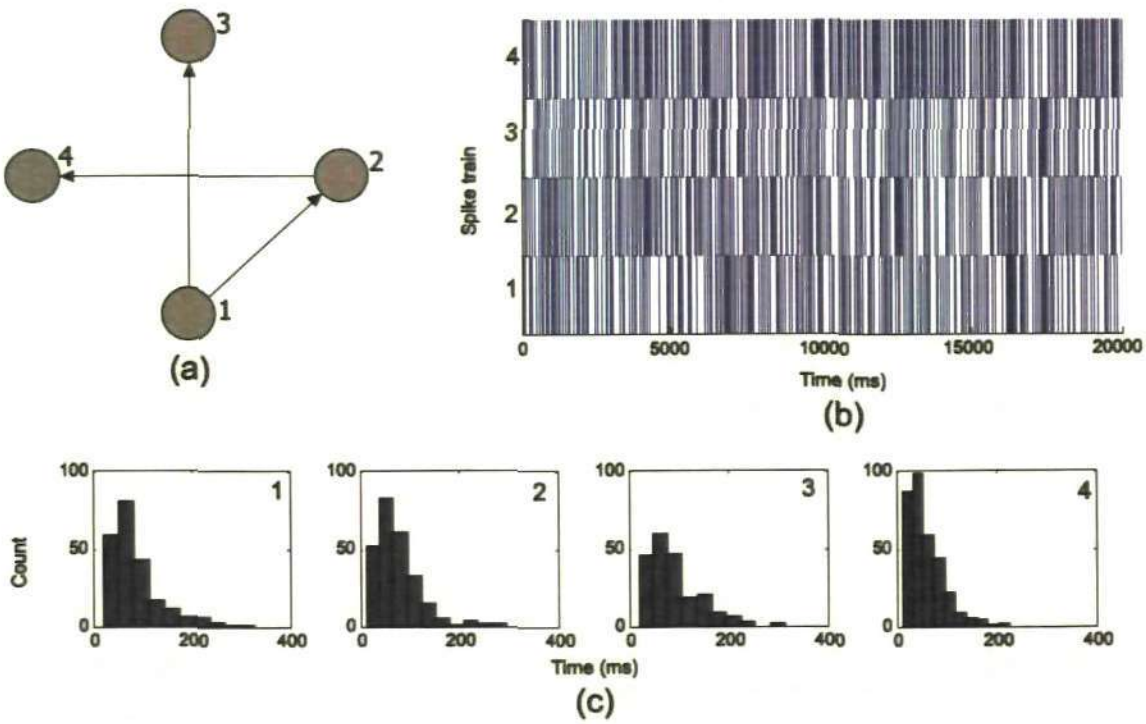


Figure 5.3: **(a)** Connection scheme of the four spike train. There are three non zero connections which are shown by arrows. **(b)** Raster plot of four generated spike trains of the duration 20,000 ms. **(c)** ISI histograms of the generated four spike trains.

The four spike trains are generated using the connection architecture shown in Fig. 5.3(a). The values of connection strength, time delay of spike propagation and time of

decay of postsynaptic potential are given in Table 5.1. The values of the neuron parameters of the ELIF model are given in Table 5.2. Fig. 5.3(b) shows the result of ELIF model generation, i.e. the raster plot of spiking activity of these four spike trains in time interval of 20 seconds.

Connection strength (w)	**Time delay (Δ)**	**Decay time (τ_s)**
$w_{1\to2} = 18.047$	$\Delta_{12} = 11$	2.86
$w_{1\to3} = 17.281$	$\Delta_{13} = 10$	3.08
$w_{2\to4} = 15.764$	$\Delta_{24} = 9$	3.42

Table 5.1: Connection strengths, time delays of spike propagation, and decay times of postsynaptic potential that are used for generating four spike trains.

Neuron parameter	**Mean**	**S.D.**
Maximum value of the threshold	44.87	0.74
Threshold decay rate	3.02	0.22
Asymptotic threshold value	15.20	1.42
Amplitude of the noise	5.03	0.26
Noise decay rate	9.99	0.02
Initial value of after spike hyperpolarisation	-28.77	0.19
Soma's membrane potential decay rate	19.73	0.50
External input	-0.25	0.48
Absolute refractory period	5.25	1.50

Table 5.2: Neuron parameters of the ELIF model of four spike trains.

These spike trains are considered as a data set for analyzing the functional connectivity. It is important to note that for analyzing the functional connectivity only spike trains are used and it is assumed that the scheme of connections is unknown. It is also assumed

that values of neuronal parameters and parameters characterizing connections (connection strength, delay of spike propagation and time of decay of PSP) are unknown. After completing the analysis the results are compared with the parameter values that were used for spike train generation. Fig. 5.3(c) shows the histogram of inter-spike intervals (ISIs) for each spike train.

5.4.1 Pair-wise analysis

To apply pair-wise analysis of the Cox method to the four spike trains a total of twelve possible pairs of spike trains are analyzed, taking one spike train as target and one other spike train as reference. All the influence functions are considered identical and specified by (5.3). In this case the values of PSP decay time are known (Table 5.1); however, it is assumed that these values are unknown. For the analysis of functional connectivity the characteristic times of rise and decay are assumed as $\tau_r = 0.1\ ms$ and $\tau_s = 10\ ms$. The other parameter, time lag Δ is specified by the pair-wise CCF method which is given in Fig. 5.4. For the pair-wise CCF, 6 simultaneous tests are conducted to find the independence of the pairs of spike trains. For these 6 simultaneous tests the Bonferroni correction is applied to the significance level α and the adjusted significance level is considered as $\alpha/6$. In Fig. 5.4 the boundaries of the confidence interval of the pair-wise CCF is calculated using the Bonferroni correction and is considered as $.05/6 = .0083$.

The highest peak outside the confidence interval is interpreted as an indicator of influence and the corresponding time shift of the CCF is considered as a time lag Δ corresponding to delay of spike propagation. Time lags Δ are summarized in the Table 5.3 and these values are used for analyzing the functional connectivity of pairs of spike trains. For example, if the second spike train is selected as a target, then the second row

and first column of the Table 5.3 provides a parameter value of time lag Δ from spike train #1 to spike train #2 and this value is: $\Delta_{12}= 11$.

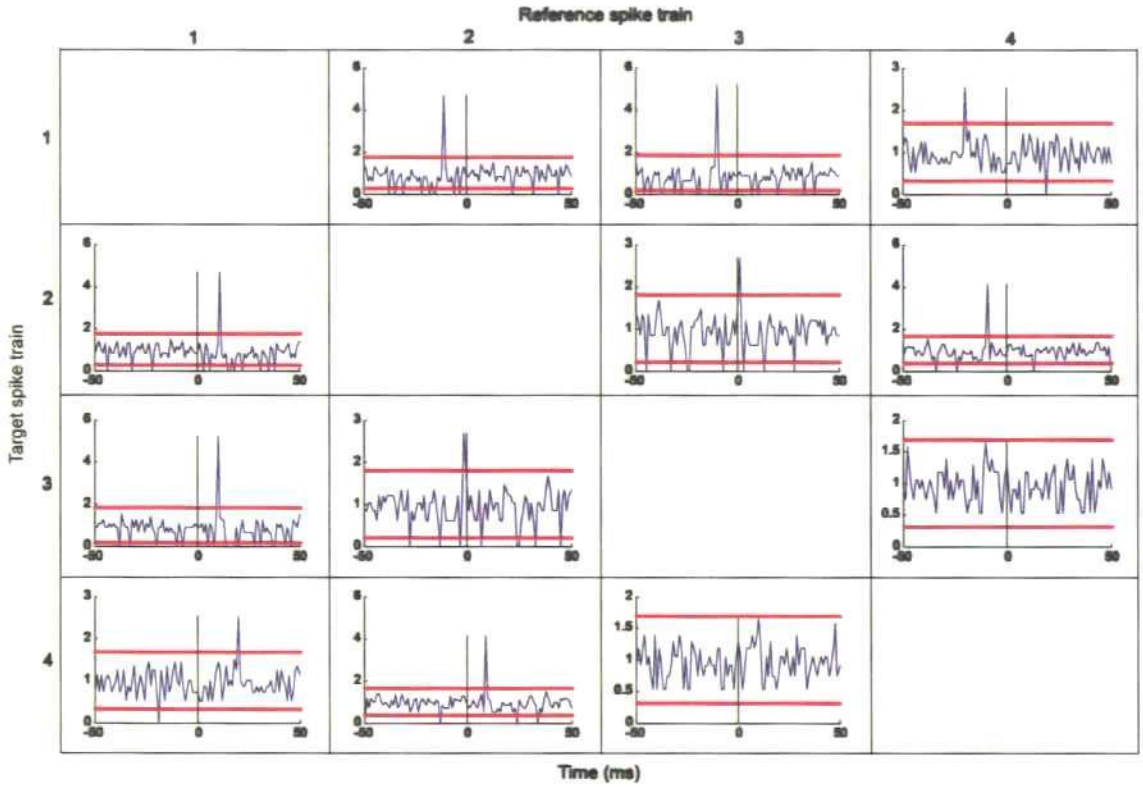


Figure 5.4: Cross correlation function of the four spike trains.

Reference spike train / Target spike train	1	2	3	4
1	0	0	0	0
2	11	0	1	0
3	10	0	0	0
4	20	9	0	0

Table 5.3: Time lags obtained from Fig. 5.4. These time lags are used to get the functional connectivity of four spike trains.

Reference spike train / Target spike train	1	2	3	4
1	0	0.318 (-0.520, 1.157)	-0.029 (-1.117, 1.058)	-0.264 (-1.098, 0.568)
2	**2.678 (2.154, 3.203)**	0	0.715 (-0.020, 1.451)	0.453 (-0.264, 1.172)
3	**3.232 (2.691, 3.773)**	-0.304 (-1.268, 0.659)	0	-0.597 (-1.543, 0.349)
4	**0.967 (0.357, 1.577)**	**2.494 (2.030, 2.959)**	0.225 (-0.482, 0.932)	0

Table 5.4: Results of pair-wise analysis of four spike trains. The estimates of Cox coefficients and corresponding confidence intervals are shown. Cox coefficients which significantly differ from zero (i.e. the confidence interval does not include zero) are in bold.

Now using the parameter values $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$ and time lags obtained from Table 5.3, the parameters are estimated using (5.5) with the corresponding confidence interval from (5.6). Applying Bonferroni correction to the significance level $\alpha = 0.05$ changes the significance level to $\alpha = 0.05/4(4-1) = .0042$. Functional connections can be derived from these estimates and their confidence intervals.

Table 5.4 summarizes the result of analyzing four spike trains with the pair-wise analysis. Each row of the table shows the Cox coefficients that characterize the influence strength from the reference spike trains to the target spike trains. The first row of Table 5.3 corresponds to the case that the first spike train is considered as a target and this row shows the estimates of Cox coefficients characterizing influences to the target spike train (#1) from the reference spike trains (#2 to #4): $\hat{\beta}_{21} = 0.318$, $\hat{\beta}_{31} = -0.029$, $\hat{\beta}_{41} = -0.264$ with the corresponding confidence intervals. In the first row all the confidence intervals include zero, therefore functional connection from spike trains #2,

#3 and #4 to the spike train #1 are concluded to be absent. In the second row there is only one Cox coefficient that significantly differs from zero (shown in bold) which characterizes the influence from spike train #1 to spike train #2. This non-zero influence strength is interpreted as strength of the functional connection from spike train #1 to spike train #2 and the strength of influence is $\hat{\beta}_{12} = 2.7$. All other Cox coefficients at the second row are not distinguishable from zero and the corresponding functional connections to the target spike train #2 are concluded to be absent. Similar to the second row there is only one Cox coefficient in the third row that significantly differs from zero (shown in bold) which characterizes the influence from spike train #1 to spike train #3. This non-zero influence strength is interpreted as strength of the functional connection from spike train #1 to spike train #3 and the strength of influence is $\hat{\beta}_{13} = 3.2$. All other Cox coefficients at the third row are not distinguishable from zero and the corresponding functional connections to the target spike train #3 are concluded to be absent. In the fourth row there are two Cox coefficients that significantly differ from zero (shown in bold) which characterize the influence from spike train #1 to spike train #4 and from spike train #2 to spike train #4. These non-zero influence strengths are interpreted as strength of the functional connection from spike train #1 to spike train #4 and from spike train #2 to spike train #4. The strength of influences of these functional connections are $\hat{\beta}_{14} = 1$ and $\hat{\beta}_{24} = 2.5$, respectively. All other Cox coefficients at the fourth row are not distinguishable from zero and the corresponding functional connections to the target spike train #4 are concluded to be absent. Thus, considering Table 5.4 it is concluded that there are four Cox coefficients that significantly differ from zero; therefore there are four functional connections of the four spike trains. These functional connections are shown by circles in Fig. 5.5(b) and the radius of the circle is proportional to the relative strength of influence: a small radius corresponds to a relatively weak functional connection. The diagonal is shown by filled squares.

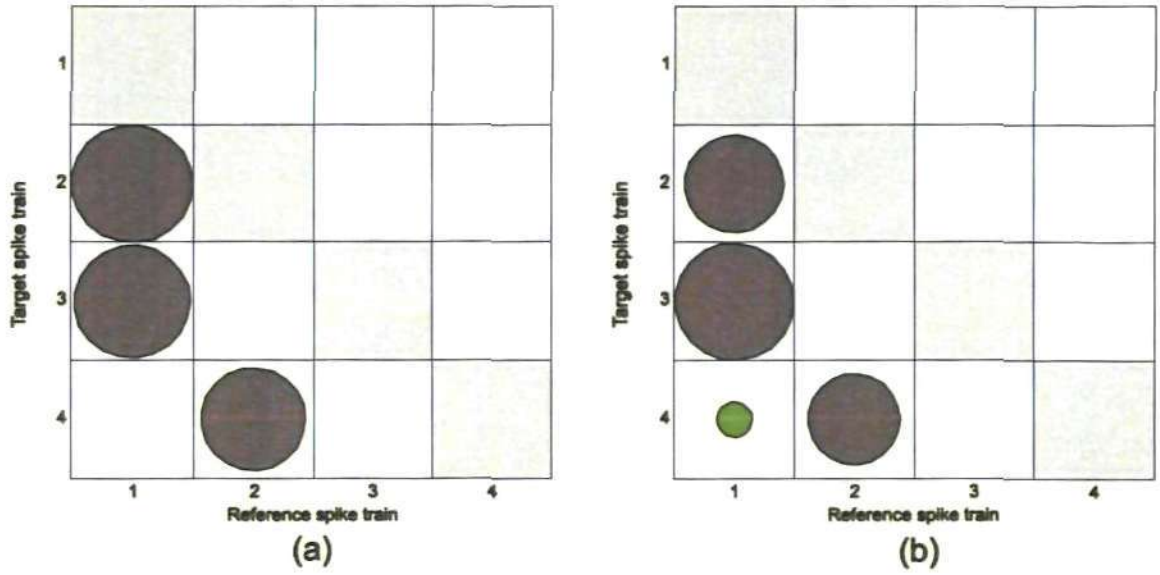


Figure 5.5: (a) Connection scheme of four spike trains in matrix format (the same as the scheme shown in Fig. 5.3(a) in graph format). **(b)** A diagram of functional connections of four spike trains obtained by the pair-wise analysis.

Comparison of the matrix of functional connectivity (Fig. 5.5(b)) with the matrix of connections (Fig. 5.5(a)) used for simulating the spike trains reveals a good correspondence between these two schemes of connections except one connection. The diagram of connectivity in Fig. 5.3(a) contains three direct connections which are shown by arrows; from spike train #1 to spike train #2, from spike train #1 to spike train #3, and from spike train #2 to spike train #4. In Fig 5.3(a) there are some "spurious" connections which are not direct: a connection due to a "common source" and a connection due to "indirect coupling". There is no direct connection between spike trains #2 and #3; however, spike train #1 is a common source, delivering spikes to both spike trains (#2 and #3). Similarly, there is no direct connection between spike trains #1 and #4, however, there is an indirect influence (coupling) from spike train #1 to spike train #4 via spike train #2. Distinguishing the direct connections and these spurious connections is difficult using pair-wise analysis. For instance the connection from spike train #1 to spike train #4 is in fact an indirect connection but incorrectly identified by the pair-wise analysis (green circle in Fig. 5.5(b)).

5.4.2 Triplet analysis

Triplet analysis of the Cox method for the four spike trains considers all possible groups of three spike trains. A total of 4 groups of three spike trains are obtained (Fig. 5.6). All four groups of three spike trains are analyzed to find functional connectivity by the triplet analysis. All the influence functions are considered identical and specified by (5.3) with parameters $\tau_s = 10\ ms$ and $\tau_r = 0.1\ ms$. The time lags Δ are obtained using the CCF method (Fig.5.4 and Table 5.3).

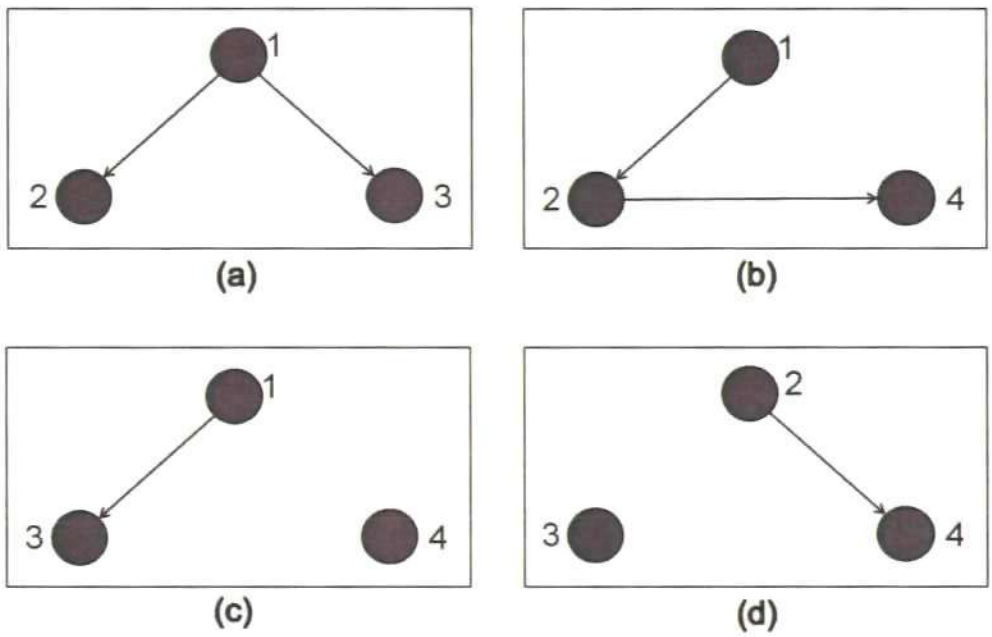


Figure 5.6: All possible groups of three spike trains for the four spike trains. Connection schemes obtained from Fig. 5.3(a). **(a)** For spike trains #1, #2 and #3. **(b)** For spike trains #1, #2 and #4. **(c)** For spike trains #1, #3 and #4. **(d)** For spike trains #2, #3 and #4.

5.4.2.1 Analysis of spike trains {#1, #2, #3}

To apply triplet analysis to the first group consisting of spike trains {#1, #2 and #3} (Fig. 5.6(a)), spike train #1 is considered as a target spike train and influences from reference spike trains #2 and #3 to this target are estimated using (5.5) with confidence intervals using (5.6). A confidence region is also calculated using (5.7) which has an elliptical shape. Similarly, spike train #2 is considered as a target spike train and influences from reference spike trains #1 and #3 to this target are estimated with

confidence intervals and confidence regions. Finally, spike train #3 is considered as a target spike train and influences from reference spike trains #1 and #2 to this target are estimated with confidence intervals and confidence regions. The Bonferroni correction is applied to the significance level α and the corrected significance level is considered as $\alpha = .05/6 = .0083$.

Fig. 5.7 shows the functional connections of this group of spike trains with three confidence regions (Fig. 5.7(b)). The region on the left side corresponds to the target spike train #1; the region in the middle corresponds to the target spike train #2, and the region on the right side corresponds to the target spike train #3. It is shown in Fig. 5.7(b) that the region on the left side contains zero, therefore both connections to spike train #1 are concluded to be absent. This result is shown in Fig. 5.7(c) by two dashed arrows pointing to spike train #1. These dashed arrows indicate an absence of both connections from spike train #2 and #3 to spike train #1.

The region in the middle does not contain the origin and this indicates that the null hypothesis should be rejected. The centre of the confidence region is shown by the blue cross and its coordinates are the estimates $(\hat{\beta}_{12}, \hat{\beta}_{32})$. The projection to the vertical axis β_{32} contains zero, therefore it is concluded that the null hypothesis should be accepted and the connection from spike train #3 to #2 is concluded to be absent. The projection to the horizontal axis β_{12} does not contain zero, therefore it is concluded that the null hypothesis should be rejected and the estimate $\hat{\beta}_{12}$ is the strength of the connection from spike train #1 to #2. This result is shown in Fig. 5.7(c) by two arrows pointing to the spike train #2: the dashed arrow indicates the absence of a connection from spike train #3 to #2, the solid arrow indicates the presence of a connection from spike train #1 to #2, and the value of the connection strength is $\hat{\beta}_{12} = 2.8$.

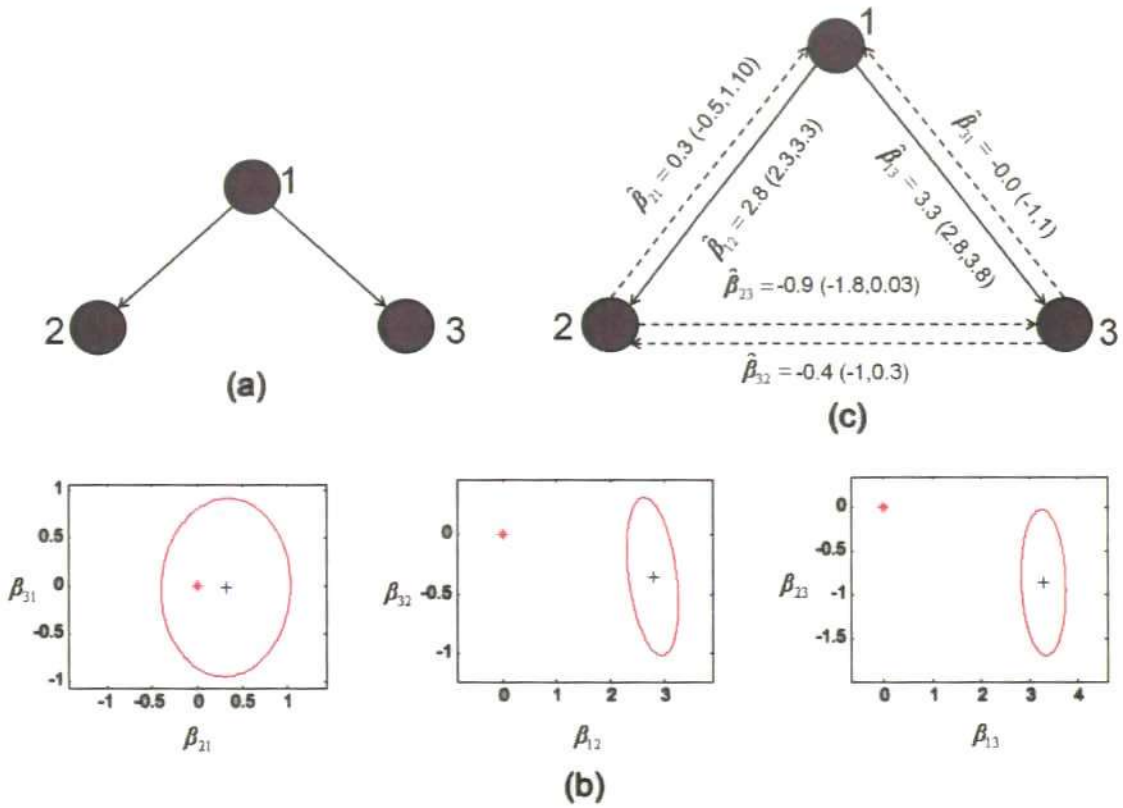


Figure 5.7: (a) Connection scheme of the spike trains #1, #2 and #3. **(b)** Confidence regions of the estimated Cox coefficients in three cases: influences to spike train #1 (left), influences to spike train #2 (middle), influences to spike train #3 (right). **(c)** Estimated coefficients of the Cox method with confidence intervals. Significant connections are indicated by solid arrows.

The region on the right side can be interpreted in a similar way. The result is shown in Fig. 5.7(c) by two arrows pointing to spike train #3: the dashed arrow indicates the absence of a connection from spike train #2 to #3, the solid arrow indicates the presence of a connection from spike train #1 to #3, and the value of the connection strength is $\hat{\beta}_{13} = 3.3$. The results of the analysis are in good agreement with the connection scheme shown in Fig. 5.7(a). The results in Fig. 5.7(c) indicate that there are two significant influences (shown by solid arrows, all others are shown by the dashed arrows): from spike train #1 to spike train #2 and from spike train #1 to spike train #3, the influence strengths are shown with their confidence intervals.

5.4.2.2 Analysis of spike trains {#1, #2, #4}

The second group consisting of spike trains {#1, #2 and #4} (Fig. 5.6(b)) are analyzed by the Cox method and the functional connections of this group of spike trains are shown in Fig. 5.8 with three confidence regions (Fig. 5.8(b)). The region on the left side corresponds to the target spike train #1; the region in the middle corresponds to the target spike train #2, and the region on the right side corresponds to the target spike train #4. It is shown in Fig. 5.8(b) that the region on the left side contains zero, therefore both connections to spike train #1 are concluded to be absent. This result is shown in Fig. 5.8(c) by two dashed arrows pointing to spike train #1. These dashed arrows indicate an absence of both connections.

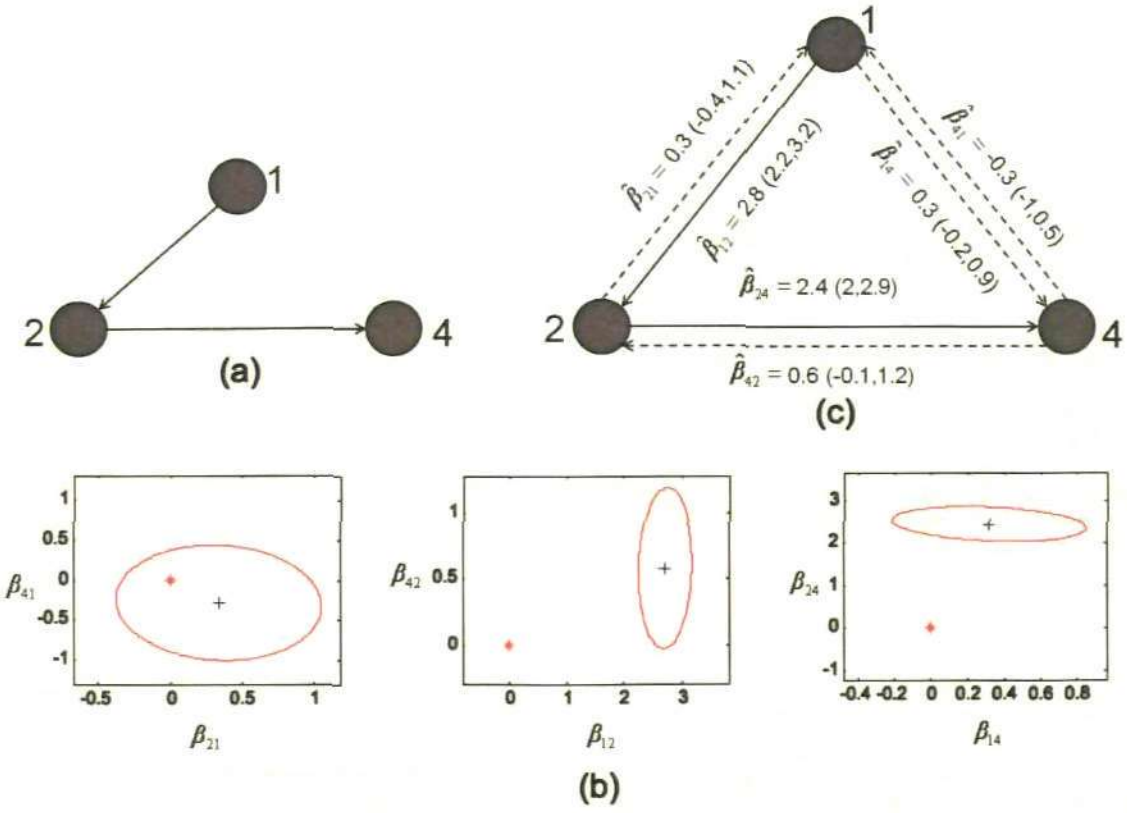


Figure 5.8: (a) Connection scheme of the spike trains #1, #2 and #4. **(b)** Confidence regions of the estimated Cox coefficients in three cases: influences to spike train #1 (left), influences to spike train #2 (middle), influences to spike train #4 (right). **(c)** Estimated coefficients of the Cox method with confidence intervals. Significant connections are indicated by solid arrows.

The region in the middle does not contain the origin and this indicates that the null hypothesis should be rejected. The centre of the confidence region is shown by the blue cross and its coordinates are the estimates $(\hat{\beta}_{12}, \hat{\beta}_{42})$. The projection to the vertical axis β_{42} contains zero, therefore it is concluded that the null hypothesis should be accepted and the connection from spike train #4 to #2 is absent. The projection to the horizontal axis β_{12} does not contain zero, therefore it is concluded that the null hypothesis should be rejected and the estimate $\hat{\beta}_{12}$ is the strength of the connection from spike train #1 to #2. This result is shown in Fig. 5.8(c) by two arrows pointing to the spike train #2: the dashed arrow indicates the absence of a connection from spike train #4 to #2, the solid arrow indicates the presence of a connection from spike train #1 to #2, and the value of the connection strength is $\hat{\beta}_{12} = 2.8$.

The region on the right side can be interpreted in a similar way. The result is shown in Fig. 5.8(c) by two arrows pointing to spike train #4: the dashed arrow indicates the absence of a connection from spike train #1 to #4, and the solid arrow indicates the presence of a connection from spike train #2 to #4, and the value of connection strength is $\hat{\beta}_{24} = 2.4$. The results of the analysis are in good agreement with the connection scheme shown in Fig. 5.8(a). The results in Fig. 5.8(c) indicate that there are two significant influences (shown by solid arrows, all others are shown by the dashed arrows): from spike train #1 to spike train #2, from spike train #2 to spike train #4, and the influence strengths are shown with their confidence intervals.

5.4.2.3 Analysis of spike trains {#1, #3, #4}

The third group consisting of spike trains {#1, #3 and #4} (Fig. 5.6(c)) are analyzed by the Cox method and the functional connections of this group of spike trains are shown in Fig. 5.9 with three confidence regions (Fig. 5.9(b)). The region on the left side corresponds to the target spike train #1; the region in the middle corresponds to the

target spike train #3, and the region on the right side corresponds to the target spike train #4. It is shown in Fig. 5.9(b) that the region on the left side contains zero, therefore both connections to spike train #1 are concluded to be absent. This result is shown in Fig. 5.9(c) by two dashed arrows pointing to spike train #1. These dashed arrows indicate an absence of both connections.

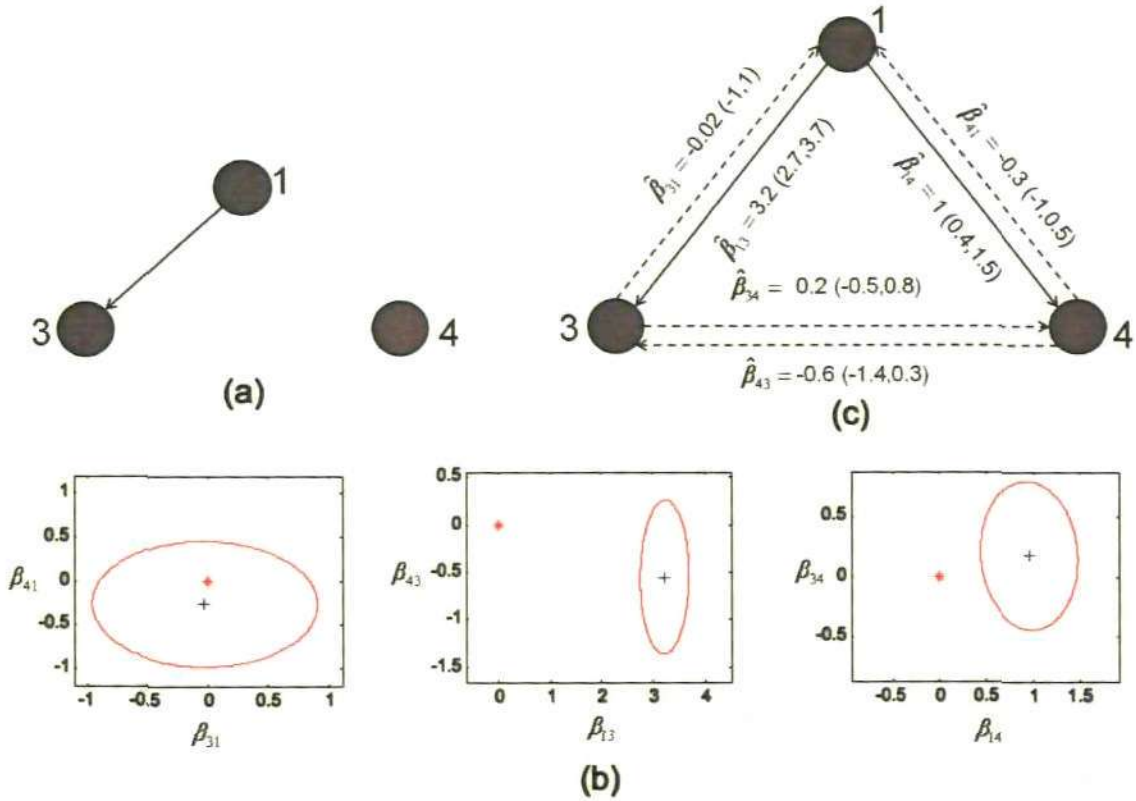


Figure 5.9: (a) Connection scheme of the spike trains #1, #3 and #4. **(b)** Confidence regions of the estimated Cox coefficients in three cases: influences to spike train #1 (left), influences to spike train #3 (middle), influences to spike train #4 (right). **(c)** Estimated coefficients of the Cox method with confidence intervals. Significant connections are indicated by solid arrows.

The region in the middle does not contain the origin and this indicates that the null hypothesis should be rejected. The centre of the confidence region is shown by the blue cross and its coordinates are the estimates $(\hat{\beta}_{13}, \hat{\beta}_{43})$. The projection to the vertical axis β_{43} contains zero, therefore it is concluded that the null hypothesis should be accepted and the connection from spike train #4 to #3 is absent. The projection to the horizontal axis β_{13} does not contain zero, therefore it is concluded that the null hypothesis should

be rejected and the estimate $\hat{\beta}_{13}$ is the strength of the connection from spike train #1 to #3. This result is shown in Fig. 5.9(c) by two arrows pointing to the spike train #3: the dashed arrow indicates the absence of a connection from spike train #4 to #3, the solid arrow indicates the presence of a connection from spike train #1 to #3, and the value of connection strength is $\hat{\beta}_{13} = 3.2$. The region on the right side can be interpreted in a similar way. The result is shown in Fig. 5.9(c) by two arrows pointing to spike train #4: the dashed arrow indicates the absence of a connection from spike train #3 to #4, the solid arrow indicates the presence of a connection from spike train #1 to #4, and the value of the connection strength is $\hat{\beta}_{14} = 1$.

The result of analyzing functional connectivity is not in good agreement with the connection scheme shown in Fig. 5.9(a). The reason is that out of four spike trains, only three spike trains are analyzed at a time without considering the effect of the other spike train. In the connection scheme there is no connection from spike train #1 to spike train #4, however, triplet analysis shows this spurious connection. The result in Fig. 5.9(c) indicates that there are two significant influences (shown by solid arrows, all others are shown by the dashed arrows): from spike train #1 to spike train #3, from spike train #1 to spike train #4, and the influence strengths are shown with their confidence intervals.

5.4.2.4 Analysis of spike trains {#2, #3, #4}

The fourth group consisting of spike trains {#2, #3 and #4} (Fig. 5.6(d)) are analyzed by the Cox method. Functional connections of this group of spike trains are shown in Fig. 5.10 with three confidence regions (Fig. 5.10(b)). The region on the left side corresponds to the target spike train #2; the region in the middle corresponds to the target spike train #3, and the region on the right side corresponds to target spike train #4.

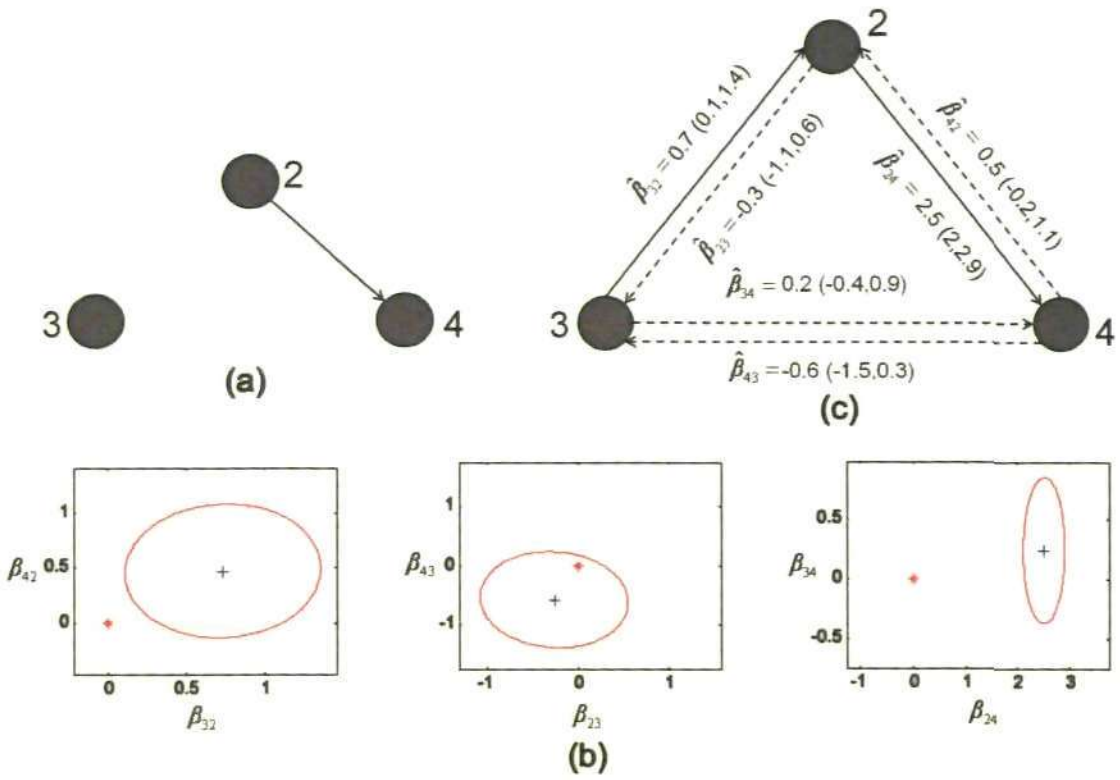


Figure 5.10: (a) Connection scheme of the spike trains #2, #3 and #4. **(b)** Confidence regions of the estimated Cox coefficients in three cases: influences to spike train #2 (left), influences to spike train #3 (middle), influences to spike train #4 (right). **(c)** Estimated coefficients of the Cox method with confidence intervals. Significant connections are indicated by solid arrows.

The region on the left side does not contain the origin and this indicates that the null hypothesis should be rejected. The centre of the confidence region is shown by the blue cross and its coordinates are the estimates $(\hat{\beta}_{32}, \hat{\beta}_{42})$. The projection to the vertical axis β_{42} contains zero, therefore it is concluded that the null hypothesis should be accepted and the connection from spike train #4 to #2 is absent. The projection to the horizontal axis β_{32} does not contain zero, therefore it is concluded that the null hypothesis should be rejected and the estimate $\hat{\beta}_{32}$ is the strength of the connection from spike train #3 to #2. This result is shown in Fig. 5.10(c) by two arrows pointing to the spike train #2: the dashed arrow indicates absence of a connection from spike train #4 to #2, the solid arrows indicate the presence of a connection from spike train #3 to #2, and the value of the connection strength is $\hat{\beta}_{32} = 0.7$. The region on the right side can be interpreted in a

similar way. The result is shown in Fig. 5.10(c) by two arrows pointing to spike train #4: the dashed arrow indicates the absence of a connection from spike train #3 to #4, the solid arrow indicates the presence of a connection from spike train #2 to #4, and the value of connection strength is $\hat{\beta}_{24} = 2.5$. It is shown in Fig. 5.10(b) that the region in the middle contains zero, therefore both connections to spike train #3 are concluded to be absent. This result is shown in Fig. 5.10(c) by two dashed arrows pointing to spike train #3. These dashed arrows indicate an absence of both connections.

The result of analyzing functional connectivity is not in good agreement with the connection scheme shown in Fig. 5.10(a). The reason is that out of four spike trains, only three spike trains are analyzed at a time without considering the effect of the other spike train. In the connection scheme there is no connection from spike train 3 to spike train 2, however triplet analysis shows this spurious connection. The result in Fig. 5.10(c) indicates that there are two significant influences (shown by solid arrows, all others are shown by the dashed arrows): from spike train #3 to spike train #2, from spike train #2 to spike train #4, and the influence strengths are shown with their confidence intervals.

As with pair-wise analysis, triplet analysis of Cox method identifies spurious connections from the four spike trains. Analyzing the first two groups of three spike trains {#1, #2, #3}, {#1, #2, #4} (Fig. 5.6(a)-5.6(b)) reveals the correct connectivity pattern but analysis of the other two groups of three spike trains {#1, #3, #4}, {#2, #3, #4} (Fig. 5.6(c)-5.6(d)) reveals two spurious connections which are from spike train #1 to spike train #4 and spike train #3 to spike train #2. The reason for finding these spurious connections is that, in triplet analysis, effects of all four spike trains are not considered for identifying functional connectivity as with pair-wise analysis.

5.4.3 Analysis considering all four spike trains

To identify the functional connectivity of the four spike trains by the Cox method, effects of all four spike trains are considered in finding influence strengths from the reference spike trains to the target spike train.

Reference spike train / Target spike train	1	2	3	4
1	0	0.338 (-0.496, 1.174)	-0.014 (-1.098, 1.070)	-0.287 (-1.121, 0.546)
2	**2.790 (2.264, 3.316)**	0	-0.326 (-1.053, 0.400)	0.564 (-0.145, 1.274)
3	**3.275 (2.738, 3.812)**	-0.818 (-1.775, 0.138)	0	-0.488 (-1.435, 0.457)
4	0.307 (-0.292, 0.907)	**2.440 (1.975, 2.905)**	0.228 (-0.495, 0.952)	0

Table 5.5: Result of analysis of four spike trains by the Cox method considering the effects of all four spike trains. The estimates of Cox coefficients and corresponding confidence intervals are shown. Cox coefficients which significantly differ from zero (i.e. the confidence interval does not include zero) are in bold.

To apply the Cox method one spike train is considered as a target spike train and the other three spike trains are considered as reference spike trains. The influence functions are considered identical for all reference spike trains and are specified by the formula (5.3). The characteristic times are considered as $\tau_s = 10\ ms$ and $\tau_r = 0.1\ ms$ and the time lags Δ are obtained from the CCF method (Fig.5.4 and Table 5.3). First, spike train #1 is considered as the target and influences from reference spike trains #2, #3 and #4 to this target are estimated using (5.5) with confidence intervals using (5.6). Second, spike train #2 is considered as the target and influences from reference spike trains #1, #3 and #4 to this target are estimated with confidence intervals. Third, spike train #3 is

considered as the target and influences from reference spike trains #1, #2 and #4 to this target are estimated with confidence intervals and fourth, spike train #4 is considered as the target spike train and influences from reference spike trains #1, #2 and #3 to this target are estimated with confidence intervals. The Bonferroni correction is applied to the significance level α and the corrected significance level is considered as $\alpha \approx 0.05/4(4-1) = .0042$.

Table 5.5 summarizes the result of analyzing four spike trains by the Cox method considering all spike trains. Each row of the table shows the Cox coefficients characterizing the influence strength to the target spike trains. The first row of Table 5.5 corresponds to the case that the first spike train is considered as a target and this row shows the estimates of Cox coefficients characterizing influences to the target spike train (#1) from the reference spike trains (#2 to #4): $\hat{\beta}_{21} = 0.338$, $\hat{\beta}_{31} = -0.014$, $\hat{\beta}_{41} = -0.287$ with the corresponding confidence intervals. On the first row all the confidence intervals include zero, therefore functional connection from spike trains #2, #3 and #4 to the spike train #1 are concluded to be absent.

On the second row there is only one Cox coefficient that significantly differs from zero (shown in bold) which characterizes the influence from spike train #1 to spike train #2. This non-zero influence strength is interpreted as strength of the functional connection from spike train #1 to spike train #2 which is $\hat{\beta}_{12} = 2.8$. All other Cox coefficients on the second row are not distinguishable from zero and the corresponding functional connections to the target spike train #2 are concluded to be absent. On the third row there is only one Cox coefficient that significantly differs from zero. This non-zero influence strength is interpreted as strength of the functional connection from spike train #1 to spike train #3 which is $\hat{\beta}_{13} = 3.3$. All other Cox coefficients on the third row are not distinguishable from zero and the corresponding functional connections to the target

spike train #3 are concluded to be absent. On the fourth row there is only one Cox coefficient that significantly differs from zero. This non-zero influence strength is interpreted as strength of the functional connection from spike train #2 to spike train #4 which is $\hat{\beta}_{24} = 2.5$. All other Cox coefficients on the fourth row are not distinguishable from zero and the corresponding functional connections to the target spike train #4 are concluded to be absent. Thus, considering Table 5.5, it is concluded that there are three Cox coefficients that significantly differ from zero; therefore there are three functional connections of four spike trains. These functional connections are shown by circles in Fig. 5.11(b) and a radius of the circle is proportional to the relative strength of influence: a small radius corresponds to a relatively weak functional connection. The diagonal is shown by filled squares.

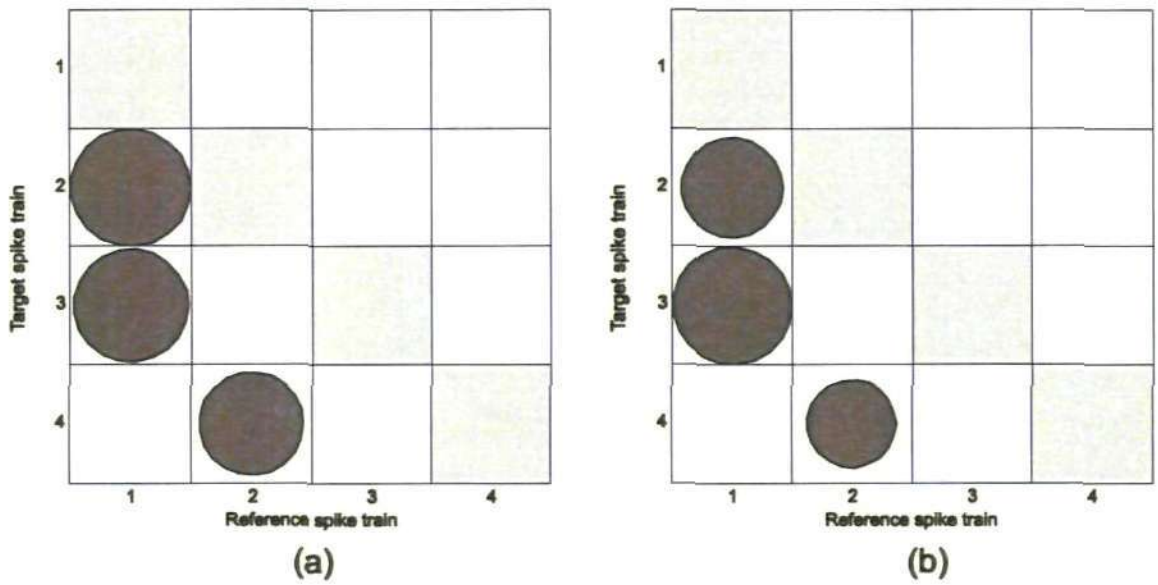


Figure 5.11: (a) Connection scheme of four spike trains in matrix format (the same as the scheme shown in Fig. 5.3(a) in graph format). **(b)** A diagram of functional connections of four spike trains obtained by the Cox method considering all spike trains at once.

A comparison of the matrix of functional connectivity (Fig. 5.11(b)) with the matrix of connections (Fig. 5.11(a)), which has been used for simulation of spike trains, reveals a good correspondence between these two schemes of connections. Moreover, relative connection strengths have been correctly identified: circles with smaller radius

correspond to weaker connections (Table 5.1). The Cox method, considering the effects of all spike train at a time, ignores all 'spurious' connections and correctly finds the direct connections which have been used for data generation. For example, there is no direct connection between spike trains #1 and #4; however there is an indirect connection from spike train #1 to spike train #4 via spike train #2. This connection is identified by pair-wise analysis and triplet analysis of the Cox method but is not identified by the Cox method considering all four spike trains at once. Thus, it is shown that the Cox method can distinguish between 'direct connection' and the connectivity due to a 'common source' (or similarly, to distinguish 'direct' and 'indirect' connections) if all spike trains are included in the analysis.

From this section using the analysis of four spike trains, it can be assumed that to identify the correct functional connectivity of multiple (p) spike trains, all spike trains (p) should be analyzed at once. Though there are limitations of the pair-wise and triplet-wise analysis, these methods are useful in certain conditions. For example, the similarity and dissimilarity measures are useful in order to identify the relationship among pairs of spike trains (Paiva et al., 2009). A traditional similarity or dissimilarity measure depends on the CCF method which is bin based. The use of pair-wise analysis to identify the relationship among the pairs of spike trains is useful and it shows a good result which is discussed in section 5.8. Motif analysis (Milo et al., 2002; Sporns and Kotter, 2004) is a useful technique to identify the patterns of interconnections among the spike trains. A motif is a connected graph consisting of a certain number of vertices (neurons) with different patterns of edges. A motif with three vertices with different patterns of edges is useful in finding the patterns of interconnections among the spike trains. To find these patterns, the triplet analysis of Cox method is a useful technique which is discussed in section 5.9.

5.5 Analysis of functional connectivity of five spike trains

In this section an application of Cox method considering the effects of all spike trains at once is demonstrated for the identification of functional connectivity. A small neural circuit of five spike trains is considered, generated using the connection architecture shown in Fig. 5.12(a). The values of connection strength, time delay of spike propagation and time of decay of postsynaptic potential are given in Table 5.6. The values of the neuron parameters of the ELIF model are given in Table 5.7. Fig. 5.12(b) shows the raster plot of spiking activity of these five spike trains over a time interval of 20 seconds and the histogram of inter-spike intervals (ISIs) for each spike train are shown in Fig. 5.12(c).

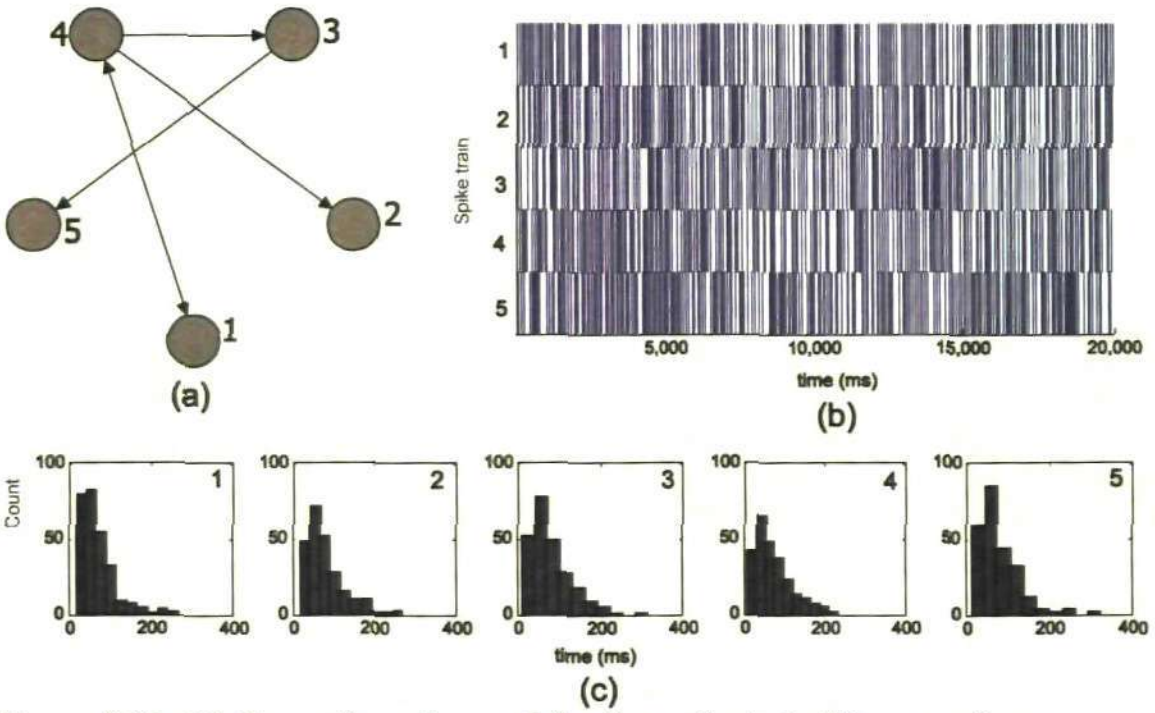


Figure 5.12: **(a)** Connection scheme of the five spike train. There are five non zero connections which are shown by arrows. **(b)** Raster plot of five spike trains generated for the neural circuit (a) of the duration 20,000 ms. **(c)** ISI histograms of the generated five spike trains.

Connection strength (w)	Time delay (Δ)	Decay time (τ_s)
$w_{4\to1} = 10.786$	$\Delta_{41} = 12$	2.09
$w_{4\to2} = 11.081$	$\Delta_{42} = 10$	1.63
$w_{4\to3} = 8.973$	$\Delta_{43} = 10$	4.66
$w_{1\to4} = 7.354$	$\Delta_{14} = 10$	4.35
$w_{3\to5} = 6.901$	$\Delta_{35} = 6$	4.35

Table 5.6: Connection strengths, time delays of spike propagation and decay times of postsynaptic potential that are used for generating five spike trains.

Neuron parameter	Mean	S.D.
Maximum value of the threshold	45.00	0.68
Threshold decay rate	2.87	0.21
Asymptotic threshold value	14.83	1.07
Amplitude of the noise	4.99	0.25
Noise decay rate	9.97	0.04
Initial value of after spike hyperpolarisation	-28.87	0.17
Soma's membrane potential decay rate	20.00	0.00
External input	0.10	0.46
Absolute refractory period	5.20	2.38

Table 5.7: Neuron parameters of the ELIF model of five spike trains.

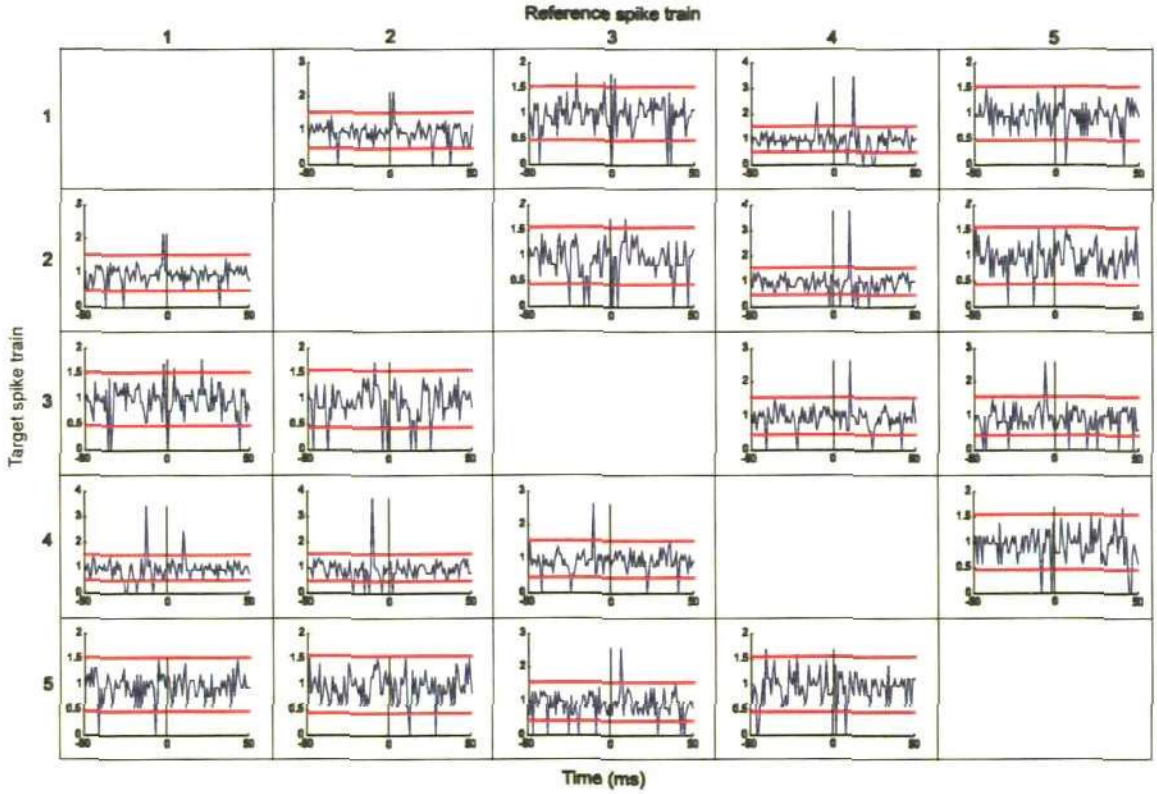


Figure 5.13: Cross correlation function of the neural circuit of five spike trains.

Reference spike train / Target spike train	1	2	3	4	5
1	0	2	2	12	0
2	0	0	9	10	0
3	21	0	0	10	0
4	10	0	0	0	41
5	0	0	6	0	0

Table 5.8: Time lags obtained from Fig. 5.13. These time lags are used to get the full functional connectivity of neural circuit of five spike train.

To apply the Cox method to this neural circuit of five spike trains, all the influence functions are considered identical and specified by formula (5.3) with the characteristic times $\tau_s = 10\ ms$ and $\tau_r = 0.1\ ms$. The time lags Δ are obtained from the CCF method (Fig. 5.13). The highest peak outside the confidence interval is interpreted as an indicator of influence and the corresponding time shift of the CCF is considered as a time lag corresponding to delay of spike propagation. Time lags are summarized in the Table 5.8 and these values are used for analyzing the functional connectivity.

Functional connections of these five spike trains can be derived by estimating the parameters of the Cox method with their confidence intervals using (5.5) and (5.6), taking one spike train as the target and finding the influence from the other four spike trains to this target. To define a complete diagram of functional connectivity the procedure is repeated for each target spike train, each time estimating the parameters of the Cox method and confidence intervals to the target. Thus, for these five spike trains this procedure should be repeated five times.

Table 5.9 summarizes the result of analyzing spike trains by the Cox method. Each row of the table shows the Cox coefficients characterizing the influence strength to the target spike trains. The first row of Table 5.9 corresponds to the case that the first spike train is considered as a target and this row shows the estimates of Cox coefficients characterizing influences to the target spike train (#1) from the reference spike trains (#2 to #5): $\hat{\beta}_{21}$ =0.5, $\hat{\beta}_{31}$ =0.2, $\hat{\beta}_{41}$ =1.6, $\hat{\beta}_{51}$ =0.09. Also the corresponding confidence intervals (the confidence level here is 0.95) are shown. These intervals are used to test the null hypothesis that the Cox coefficient is zero: if the confidence interval includes zero then the null hypothesis should be accepted and it is concluded that there is no influence from the reference spike train to the target (i.e. influence strength is zero). On the first row there is only one Cox coefficient that significantly differs from

zero (shown in bold) which characterizes the influence from spike train #4 to spike train #1. This non-zero influence strength is interpreted as strength of the functional connection from spike train #4 to spike train #1 and the strength is $\hat{\beta}_{41} = 1.6$. All other Cox coefficients on the first row are not distinguishable from zero and the corresponding functional connections to the target spike train #1 are absent. This procedure of estimation of Cox coefficients is repeated for the target spike train #2 and the result is shown in row 2, etc.

Reference spike train / Target spike train	**1**	**2**	**3**	**4**	**5**
1	0	0.5 (-0.09, 1.1)	0.2 (-0.4, 0.9)	**1.6 (1.1, 2.1)**	0.09 (-0.6, 0.8)
2	-0.2 (-1.0, 0.4)	0	0.4 (-0.1, 1.1)	**1.9 (1.4, 2.4)**	0.3 (-0.3, 1.0)
3	0.5 (-0.06, 1.2)	-0.3 (-1.0, 0.4)	0	**1.1 (0.5, 1.7)**	-0.2 (-1.1, 0.5)
4	**1.2 (0.7, 1.8)**	-0.1 (-0.9, 0.6)	0.1 (-0.5, 0.8)	0	-0.2 (-1.0, 0.5)
5	-0.1 (-0.8, 0.6)	-0.2 (-1.1, 0.5)	**1.2 (0.7, 1.8)**	0.003 (-0.7, 0.7)	0

Table 5.9: Result of analysis of five spike trains by the Cox method. The estimates of Cox coefficients and corresponding confidence intervals are shown. Cox coefficients which significantly differ from zero (i.e. the confidence interval does not include zero) are in bold.

Thus, considering Table 5.9 it is concluded that there are five Cox coefficients that significantly differ from zero; therefore there are five functional connections between spike trains, these are: $\hat{\beta}_{41} = 1.6$, $\hat{\beta}_{42} = 1.9$, $\hat{\beta}_{43} = 1.1$, $\hat{\beta}_{14} = 1.2$, $\hat{\beta}_{35} = 1.2$. These functional connections are shown by circles in Fig. 5.14(b) and the radius of the circle is proportional to the relative strength of influence: a small radius corresponds to a relatively weak functional connection. The diagonal is shown by filled squares.

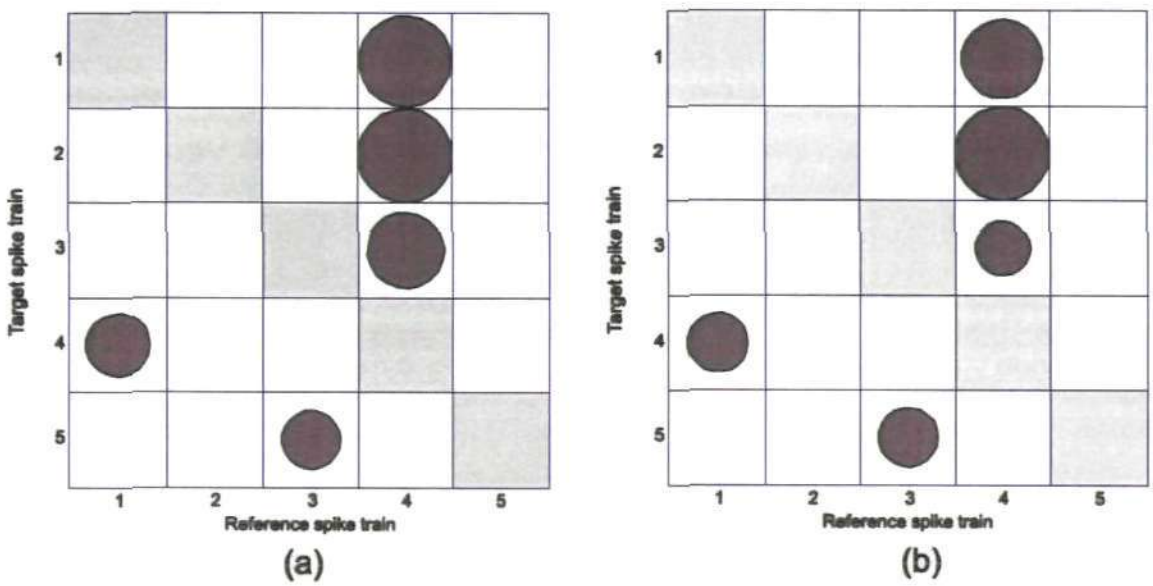


Figure 5.14: (a) Connection scheme of five spike trains in matrix format (the same as the scheme shown in Fig. 5.12(a) in graph format). **(b)** A diagram of functional connections of five spike trains obtained by the Cox method.

Comparison of the matrix of functional connectivity (Fig. 5.14(b)) with the matrix of connections (Fig. 5.14(a)), which has been used for simulation of spike trains, reveals a good correspondence between these two schemes of connections. Moreover, relative connection strengths have been correctly identified: circles with smaller radius correspond to weaker connections (Table 5.6).

To emphasize the importance of this result, it is noted that the diagram of connectivity in Fig. 5.12(a) contains direct connections shown by arrows (e.g. from spike train #4 to spike train #3) and some 'spurious' connections: i.e., connections due to a 'common source' and connections due to 'indirect coupling'. For example, there is no direct

connection between spike trains #1 and #2; however, spike train #4 is a common source which delivers spikes to both spike trains (#1 and #2). Again, there is no direct connection between spike trains #1 and #3, however, there is an indirect influence (coupling) from spike train #1 to spike train #3 via spike train #4. Thus, from this example it is shown that the Cox method can distinguish between 'direct connection' and the connectivity due to a 'common source' (or similarly, to distinguish 'direct' and 'indirect' connections).

5.6 Analysis of functional connectivity of twenty spike trains

In this section a relatively large set of twenty spike trains is analyzed which are generated by the ELIF model with twenty elements and with forty two connections. A diagram of connections is shown in Fig. 5.15(a).

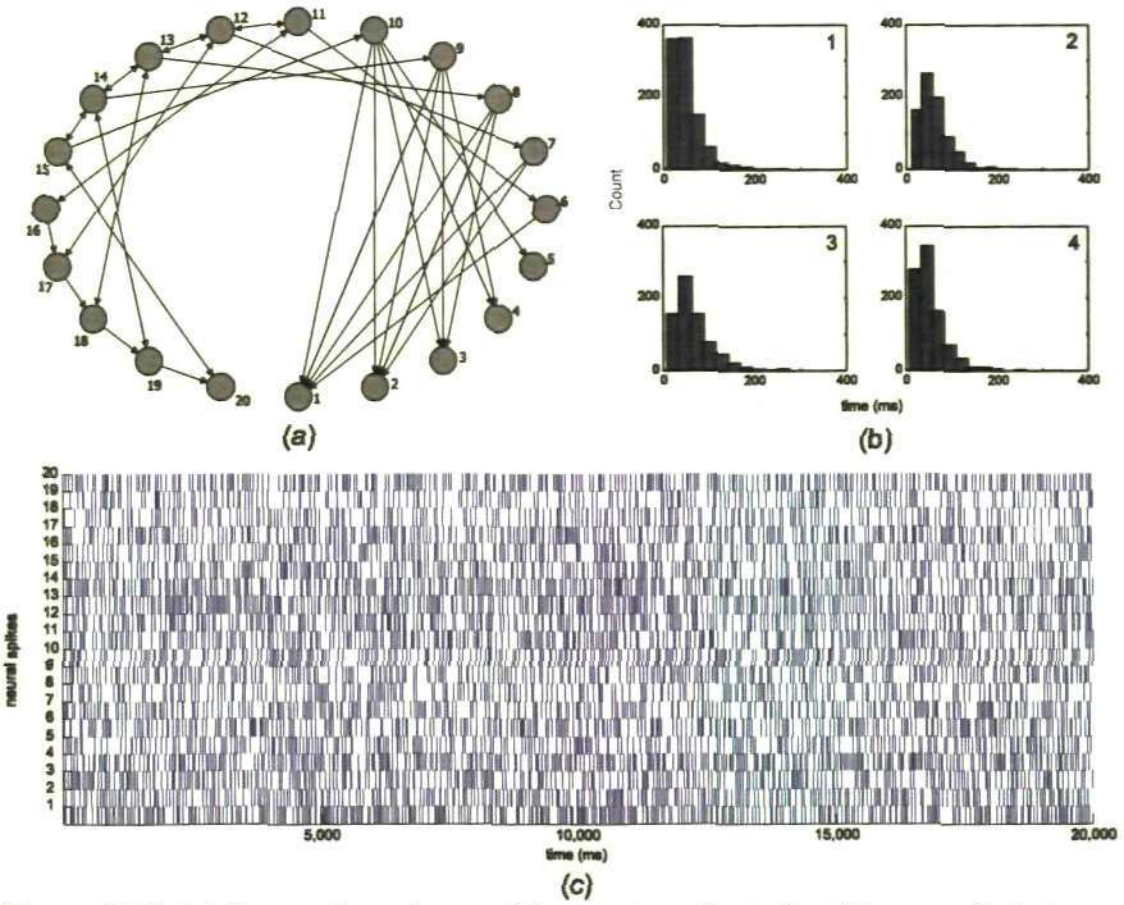


Figure 5.15: (a) Connection scheme of the twenty spike trains. There are forty two non zero connections which are shown by arrows. **(b)** ISI histograms of the first four generated spike trains. **(c)** Raster plot of twenty spike trains generated for the neural circuit. This raster plot shows a portion of time (20,000 ms) of the duration 50,000 ms.

Parameter values of the ELIF model are similar to the parameter values which have been used in the previous example of five spike trains. Parameter values of the ELIF model are given in Table 5.10. A simulation has been run over a time interval of 50 seconds and Fig. 5.15(c) shows an initial part of the raster plot of twenty spike trains generated by the model (from zero to twenty seconds). Fig. 5.15(b) shows an example of the four ISI histograms of spike trains #1 to #4.

Neuron parameter	**Mean**	**S.D.**
Maximum value of the threshold	45.12	0.97
Threshold decay rate	3.02	0.30
Asymptotic threshold value	14.47	1.02
Amplitude of the noise	5.06	0.35
Noise decay rate	10.01	0.03
Initial value of after spike hyperpolarisation	-29.10	0.41
Soma's membrane potential decay rate	20.03	0.78
External input	0.009	0.40
Absolute refractory period	4.75	1.51
Connection parameter		
Connection strength	10.44	1.85
Decay rate of postsynaptic potential	2.96	0.78
Time lag of spike propagation	10.14	2.26

Table 5.10: Parameter values of the ELIF model to generate twenty spike trains.

The procedure for analyzing the functional connectivity is the same as with the neural circuit of five spike trains. The target spike train is selected and the other nineteen spike trains are considered as reference spike trains. All the influence functions are considered

identical and specified by the formula (5.3). The characteristic times are the same as in the case of five spike trains $\tau_s = 10$ ms and $\tau_r = 0.1$ ms. To select a proper time lag for the influence function the pair-wise CCF between the reference spike train and the target spike train is calculated and the highest significant peak is identified. The corresponding time shift of the CCF is used as the value of time lag Δ. Thus, the influence function is defined and the estimates $(\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_{19})$ of Cox coefficients and their confidence intervals are calculated.

To simplify, the result of data analysis is compared with the connection scheme used for data generation, both the connection schemes in matrix format are shown in Fig. 5.16. Fig. 5.16(a) shows connections of a neural circuit of twenty ELIF elements used for data generation (the same scheme is shown in Fig. 5.15(a) in a graph format). Fig 5.16(b) shows a diagram of functional connections in matrix format which have been identified by the Cox method. The results from twenty repetitive applications of the Cox method give a matrix of functional connectivity in Fig. 5.16(b). The first row of the matrix corresponds to the case when the first spike train is selected to be a target; the second row corresponds to the case when the second spike train is the target, etc. A circle indicates that there is a significant influence (functional connection) to the target spike train and the radius of the circle shows the relative strength of the influence. Comparison of the connectivity matrix in Fig. 5.16(a) with the matrix in Fig. 5.16(b) shows that the Cox method correctly identifies all forty two direct connections between spike trains. The connectivity matrix is derived from the repetitive testing of the null hypothesis that there are no dependencies between the target and reference spike trains, using the Cox method. In the hypothesis testing, two null hypotheses of independence are incorrectly rejected. These two false positive connections are shown by green circles

(Fig. 5.16(b)) and these erroneous connections are not present in the circuit of the spike train generation.

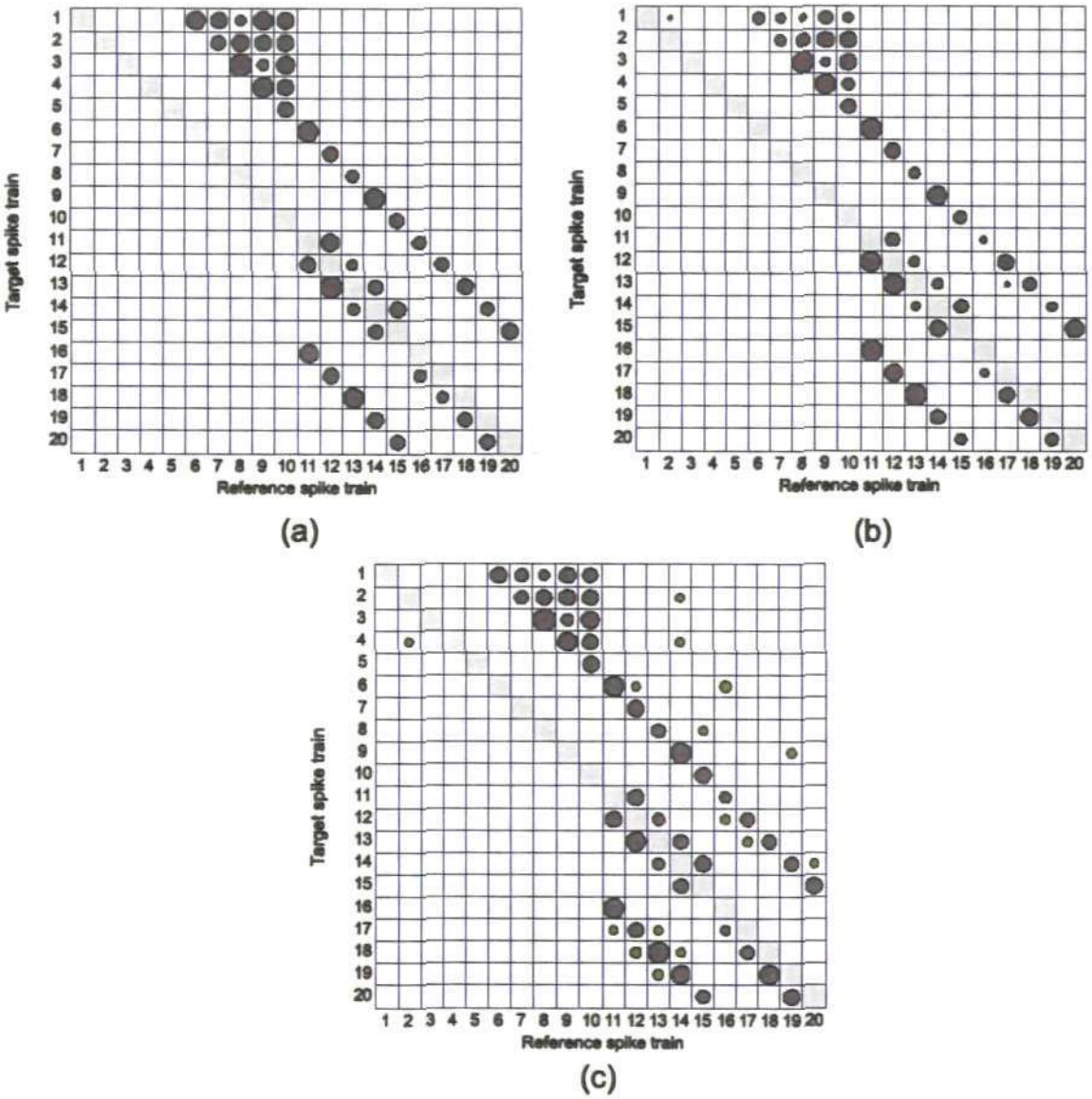


Figure 5.16: **(a)** Connection scheme of the neural circuit of twenty spike trains in matrix format (the same as a scheme of connections in a graph format in Fig. 5.15(a)). **(b)** Functional connections identified by the Cox method. **(c)** Functional connections obtained by the CCF method.

Fig. 5.16(c) shows the matrix of functional connectivity which has been constructed by using the pair-wise CCF technique. Comparison of this matrix with the matrix of connections which have been used for data generation shows that all forty two null hypotheses of independency are correctly rejected. It means that all forty two non-zero connections have been correctly identified. However, in addition, fifteen null

hypotheses of independency are incorrectly rejected. Thus, there are fifteen type I errors (false positives) and the corresponding non-zero erroneous connections are shown by green circles (Fig. 5.16(c)). The radius of circles corresponding to these erroneous non-zero connections is relatively large; therefore, strength of erroneous influence is also relatively large. From this example it can be shown that the Cox method has some advantages over the CCF technique.

5.7 Cox method versus CCF

In this section the Cox method is compared with a traditional technique based on the CCF and the advantages of the Cox method are shown, especially in cases which are difficult for analysis. The CCF is a pair-wise method and therefore in this section it is mainly the connection of two spike trains that are analyzed. The main assumption of the Cox method is that the target process is the MRP with the hazard function described by formula (5.1). A probabilistic model has been developed to generate the MRP which is discussed in chapter 4. It is expected that for this data the estimate $\hat{\beta}$ of the Cox coefficient equals the influence strength β in formula (5.1). Of course, in the general case of data generation using the ELIF model, it is not expected that the target spike train is a MRP. However, in this section it is demonstrated that the Cox method can be successfully applied to analyse functional connectivity and that the estimate $\hat{\beta}$ monotonically increases with increase of connection strength in generated data. Also, in this section connectivity of three spike trains generated by the ELIF model with 'common source' connections is studied. This connection scheme is very difficult for analysing by pair-wise methods and in particular by the CCF. It is shown that for the neural circuit of three spike trains, the Cox method can analyse three spike trains at once to identify functional connectivity. In a similar way another set of three spike trains with 'indirect' connections, is studied.

5.7.1 Analysis of two spike trains

In this section two spike trains, A (target) and B (reference), are generated using the probabilistic model discussed in chapter 4 with a connection from B to A denoted by β_{BA}. The reference spike train B is generated with a gamma-distribution $\gamma(5,3)$ of inter spike intervals. The influence function from reference spike train B to target spike train A is specified by (5.4) with characteristic times $\tau_s = \tau_r = 5$ and time lag $\Delta = 0$. The backward recurrence time of spike train A is calculated using a Weibull distribution $W(15,5)$. Using these values the target spike train A can be generated with the formula (5.1). After generating the target spike train A and the reference spike train B, the Cox method and the CCF method are used for estimating the connection strength β_{BA}.

To apply the Cox method, the influence function is specified by (5.4) with characteristic times $\tau_s = \tau_r = 5$ ms and zero time lag ($\Delta = 0$). The Cox coefficient $\hat{\beta}_{BA}$ is estimated using (5.5) and the confidence interval using (5.6). Also, the CCF has been calculated and the value of the highest peak outside of the confidence interval $\hat{\rho}_{BA}$ is considered as an estimate of the connection strength. Of course, if there are no peaks outside of the upper bound of the confidence interval, the connection strength is considered to be zero and $\hat{\rho}_{BA} = 0$. In this section some advantages of the Cox method, both in the case of short spike trains and in the case of a weak coupling, are demonstrated.

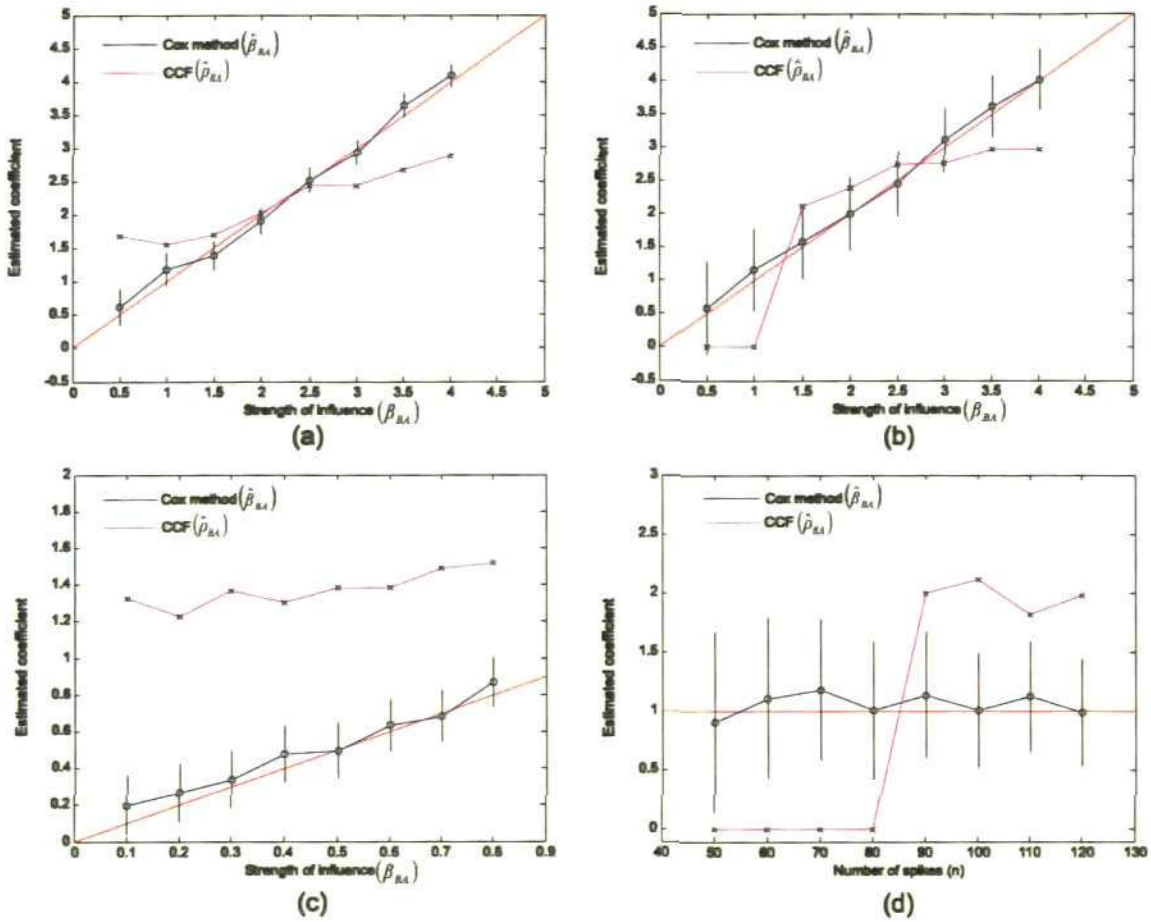


Figure 5.17: Estimate of the Cox coefficient and CCF measure for two spike trains. Estimated Cox coefficients are shown by black circles and the confidence interval of the estimates are shown by black vertical lines. Estimated measures of independency using CCF are shown by black cross sign. **(a) Moderate and strong influence.** Eight pairs of spike trains are generated using the probabilistic model taking the strength of influence from the range from 0.5 to 4: $\beta_{BA} = (0.5,1,1.5,2,2.5,3,3.5,4)$. The average number of spikes in the reference spike train B is about 400. Estimated Cox coefficients $\hat{\beta}_{BA}$ identify accurately all the strengths of influences (blue line with circle markers and vertical black bars for confidence intervals) and are monotically increasing. The highest peaks $\hat{\rho}_{BA}$ of the CCF (independency measure) are shown by the magenta line (with cross markers), they also can identify functional connectivity but do not demonstrate a monotonic increase. **(b) Short spike train.** A short version of eight pairs of spike trains described in (A) are considered. The average number of spikes in the reference spike train B is about 70. The estimated Cox coefficients $\hat{\beta}_{BA}$ identify accurately all the strengths of influences β_{BA} except for one ($\beta_{BA} = 0.5$) and demonstrate a monotonic increase. The independency measure of CCF ($\hat{\rho}_{BA}$) show connection for large strength but they fail to identify connection for $\beta_{BA} = (0.5,1)$. Also these values do not demonstrate a monotonic increase. **(c) Weak influence.** Eight pairs of spike trains are generated with weak influences $\beta_{BA} = (0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8)$. The number of spikes in the reference spike train B is about 1400. Estimated Cox coefficients ($\hat{\beta}_{BA}$) identify accurately all these strengths of influences (β_{BA}) and are monotonically increasing. Independency measures of CCF ($\hat{\rho}_{BA}$) identify functional connectivity though they do not indicate an increase of influence. **(d) Length of spike train.** Eight pairs of spike trains of a different length are generated keeping the same connection strength $\beta_{BA} = 1$. The length n of the reference spike train B increases: $n =$

50,60, ⋯⋯ ,120. Estimated Cox coefficients ($\hat{\beta}_{BA}$) are almost constant for all lengths but independency measures of CCF ($\hat{\rho}_{BA}$) fail to identify strengths of influences for shorter lengths of reference spike trains (n= 50, 60, 70, 80).

Moderate and high strengths of influence. Varying the strength of influence β_{BA} in a range from moderate to high, eight pairs of spike trains are generated and for each pair connectivity is analyzed. The average number of spikes in the reference spike train B is about 400 and the target spike train A has a larger number of spikes. The blue line in Fig. 5.17(a) shows the estimated Cox coefficient $\hat{\beta}_{BA}$ (with its corresponding confidence interval which is shown by a black vertical bar) versus the strength of influence β_{BA}. This plot shows that the estimated values are close to the values which have been used for data generation. The magenta line shows the independency measure of the CCF($\hat{\rho}_{BA}$) versus the strength of influence β_{BA}. It is clear from Fig. 5.17(a) that the CCF also identifies this connection, however, the plot of the CCF($\hat{\rho}_{BA}$) is not monotonically changing and fails to indicate the increase of influence.

Short spike trains. To test sensitivity of the Cox method in the case of short spike trains the same pairs of spike trains are used as in the previous example with the epoch time shortened and considering only a part where the independent spike train has about 70 spikes. Thus a time epoch is about six times shorter than in the previous example. Fig. 5.17(b) shows the estimate of Cox coefficient $\hat{\beta}_{BA}$ versus the strength of influence β_{BA} for the moderate and high strengths of influence (shown by the blue line with the black vertical bars of confidence intervals). This line shows that the estimated values are similar to the strengths which have been used for data generation with only one exception: for $\beta_{BA} = 0.5$ the confidence interval contains zero and therefore the Cox coefficient is not distinguishable from zero. The magenta line in Fig. 5.17(b) shows that the CCF fails to identify functional connections for the moderate influences $\beta_{BA} = (0.5, 1)$. For the higher influences $\beta_{BA} = (1.5, 2, 2.5, 3, 3.5, 4)$ the CCF measure is

nearly constant and fails to indicate the increase of influence. Therefore, the Cox method has some advantages in the case of short spike trains.

Weak influence. To test the efficiency of methods and to identify weak connection strength the same probabilistic model is used for generating another eight pairs of spike trains with weak influence: $\beta_{BA} = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8)$. In this case the time epoch should be long enough (about 1400 spikes in the reference spike train *B*) to allow distinguishing of weak influence connections. Fig. 5.17(c) shows that both methods demonstrate good results and identify the connection. The Cox coefficient increases with the connection strength increase but the CCF measure is not monotonically increasing.

Sensitivity to the length of spike trains. How sensitivity of the methods depends on the length of spike trains, under a constant value of the influence strength, is studied. The conclusion is that for shorter spike trains the Cox method identifies the connection but the CCF method fails. The strength of influence $\beta_{BA} = 1.0$ is relatively small. The value of influence strength is fixed and eight pairs of spike trains are generated with different numbers of spikes in the reference spike train *B*: $n = 50, 60, \cdots\cdots, 120$. Fig. 5.17(d) shows that the estimated Cox coefficient is almost constant ($\hat{\beta}_{BA} = 1$) and does not depend on the length of spike train. The CCF measure ($\hat{\rho}_{BA}$) identifies the connection for the larger spike trains (n= 90, 100, 110, 120) but fails to identify a strength of influence for the shorter lengths of reference spike trains (n= 50, 60, 70, 80).

5.7.2 Analysis of three spike trains

Here it is shown that the Cox method is very effective in analyzing connections which are not direct such as 'common source' circuits (Fig. 5.18(a)) and 'indirect connection' circuits (Fig. 5.20(a)). Usually it is very difficult to analyze these types of connections

using the pair-wise CCF technique. The Cox method is multivariate and can analyze three spike trains at once, making this method more sensitive than pair-wise CCF. For example, this advantage enables the Cox method to distinguish between 'direct' connections and connections due to a 'common source' in case of a moderate influence from the common source.

5.7.2.1 Common source connection

Three spike trains ({#1, #2, #3}) are generated using ELIF with the parameters given in Table 5.11 and with a 'common source' connection (Fig. 5.18(a)). The 'common source' circuit includes two connections from spike train #1 to spike trains #2 and #3. Connection strengths are 12.6 and 10.6; delays of spike propagation are 11 ms and 14 ms, respectively.

Neuron parameter	**Mean**	**S.D.**
Maximum value of the threshold	44.70	0.36
Threshold decay rate	2.83	0.46
Asymptotic threshold value	14.40	0.47
Amplitude of the noise	4.59	0.49
Noise decay rate	10.01	0.02
Initial value of after spike hyperpolarisation	-28.68	0.68
Soma's membrane potential decay rate	19.85	0.48
External input	0.16	0.23
Absolute refractory period	7.00	1.73

Table 5.11: Neuron parameters of the ELIF model of two spike trains with common source connection.

These three spike trains are analyzed by the Cox method with the influence function given by formula (5.3), characteristic times are $\tau_s = 10\,ms$ and $\tau_r = 0.1\,ms$. To prescribe the time lags the CCF function is calculated for all pairs of spike trains (Fig. 5.19): Δ_{12}= 11ms, Δ_{13}= 14ms, Δ_{23}= 3ms, and all other time lags are zero. It is assumed that the target spike train is $\#k, (k = 1,2,3)$. The estimates $(\hat{\beta}_{ik}, \hat{\beta}_{jk})$ $(i = 1,2,3; j = 1,2,3; k = 1,2,3,\ i \neq j, i \neq k, j \neq k)$ of two Cox coefficients have been calculated using formulas (5.5) and (5.6) as well as a confidence region on the plane (β_{ik}, β_{jk}) using formula (5.7). The confidence region ($\alpha = 0.05$) has an elliptic shape and the centre of the confidence region is located at point $(\hat{\beta}_{ik}, \hat{\beta}_{jk})$.

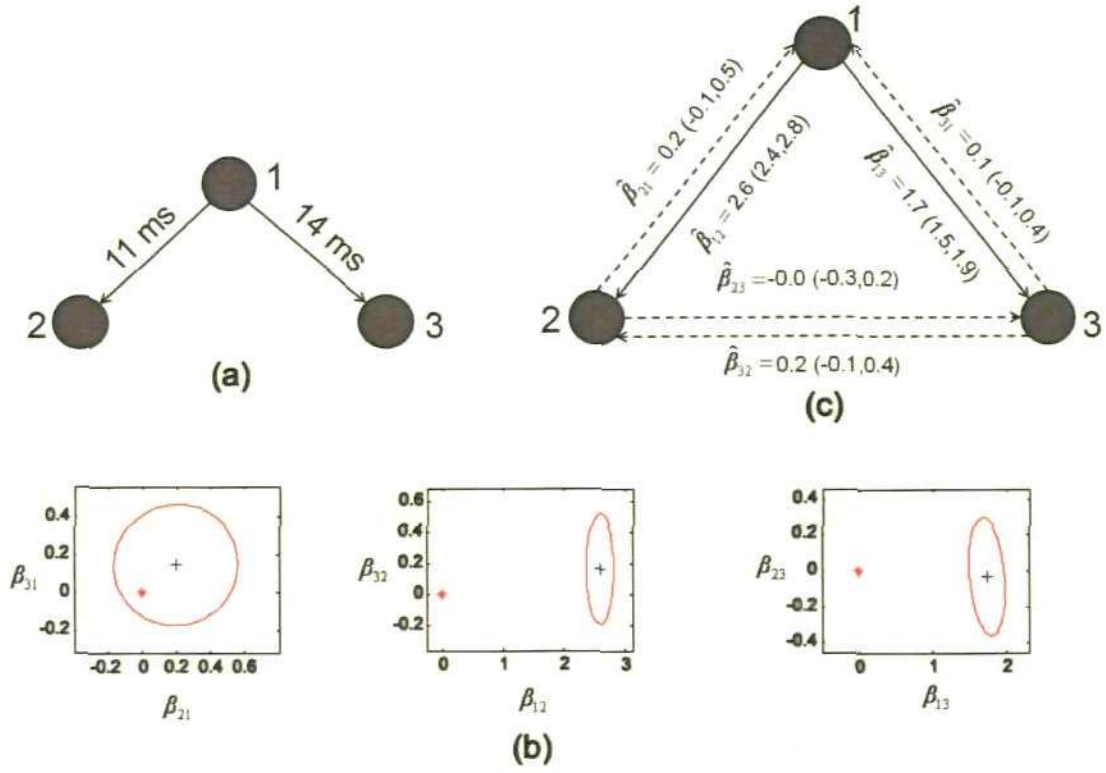


Figure 5.18: **(a)** Connection scheme of three spike trains which have common source. Spike train #1 influences both spike train #2 and spike train #3 with time delays 11 ms and 14 ms, respectively. **(b)** Confidence regions of the estimated Cox coefficients in three cases: influences to spike train #1 (left), influences to spike train #2 (middle), influences to spike train #3 (right). **(c)** Estimated coefficients of the Cox method with confidence intervals. Significant connections are indicated by solid arrows.

Fig. 5.18(b) shows three confidence regions. The region on the left side corresponds to the target spike train #1; the region in the middle corresponds to target spike train #2, the region on the right side corresponds to target spike train #3. It is shown in Fig. 5.18(b) that the region on the left side contains zero, therefore it is concluded that both connections to spike train #1 are absent. This result is shown in Fig. 5.18(c) by two dashed arrows pointing to #1. These dashed arrows indicate an absence of both connections.

The region in the middle does not contain the origin and this indicates that the null hypothesis should be rejected. The centre of the confidence region is shown by the blue cross and its coordinates are the estimates $(\hat{\beta}_{12}, \hat{\beta}_{32})$. The projection to the vertical axis β_{32} contains zero, therefore, it is concluded that the null hypothesis should be accepted and that the connection from #3 to #2 is absent. The projection to the horizontal axis β_{12} does not contain zero, therefore, it is concluded that the null hypothesis should be rejected and the estimate $\hat{\beta}_{12}$ is the strength of the connection from #1 to #2. This result is shown in Fig. 5.18(c) by two arrows pointing to #2: the dashed arrow indicates the absence of a connection from #3 to #2 and the solid arrow indicates presence of a connection from #1 to #2 and the value of connection strength is $\hat{\beta}_{12} = 2.6$. The region on the right side can be interpreted in a similar way. The result is shown in Fig. 5.18(c) by two arrows pointing to #3: the dashed arrow indicates the absence of a connection from #2 to #3 and the solid arrow indicates the presence of a connection from #1 to #3 and the value of the connection strength is $\hat{\beta}_{13} = 1.7$.

The results of analysis are in good agreement with the architecture of connections which were used to generate the data (compare Fig 5.18(a) with Fig. 5.18(c)). For example, for data generation, a higher connection strength was selected for connection from #1 to #2

and the estimated connection strength from #1 to #2 is also higher than the estimated connection strength from #1 to #3.

Thus, the result in Fig. 5.18(c) indicates that there are two significant influences only (shown by solid arrows, all others are shown by the dashed arrows): from spike train #1 to spike train #2, from spike train #1 to spike train #3, and the influence strengths are shown with their confidence intervals ($\alpha = 0.05$).

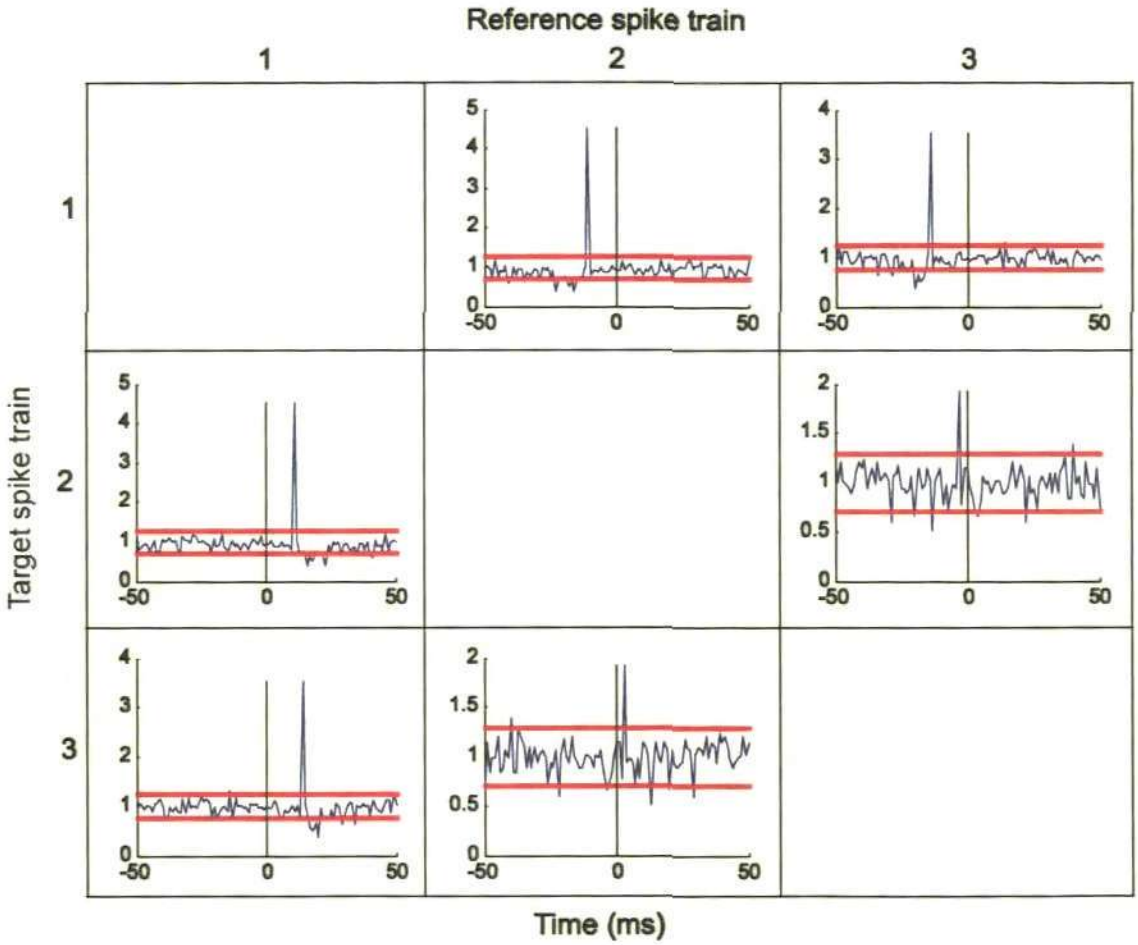


Figure 5.19: Pair-wise cross correlation functions of three spike trains. Each CCF is shown for selected pairs of spike trains (called target and reference). Diagram of connections (common source) is shown in Fig. 5.18(a).

Fig. 5.19 illustrates the result analyzing of the same three spike trains by the pair-wise CCF method. Each row of the figure shows two CCF corresponding to the selected target spike train – spike train #1 is the target for the first row, spike train #2 is the target for the second row, etc. The CCF analysis reveals three connections: from spike

train #1 to spike train #2 (second row, first column), from spike train #1 to spike train #3 (third row, first column), and from spike train #2 to spike train #3 (third row, second column). The first two of these connections correspond well to the diagram of connectivity (Fig. 5.18(a)) but the third one is erroneous and this connection appears to be due to the common source to spike trains #2 and #3. Thus, the Cox method is able to distinguish the common source from the direct connections but the CCF method fails.

5.7.2.2 Indirect connection

Similar to the previous example, a set of three spike trains ({#1, #2, #3}) are generated using the ELIF model with the parameters given in Table 5.12 and with indirect connections (Fig. 5.20(a)). The 'indirect connection' circuit includes two direct influences: from spike train #1 to spike train #2 with the time lag 11 ms, and from spike train #2 to spike train #3 with the time lag 12 ms. The connection strengths are 11.2 and 9.1, respectively.

Neuron parameter	**Mean**	**S.D.**
Maximum value of the threshold	45.20	0.74
Threshold decay rate	2.93	0.51
Asymptotic threshold value	15.20	1.06
Amplitude of the noise	4.67	0.50
Noise decay rate	10.00	0.009
Initial value of after spike hyperpolarisation	-28.92	0.64
Soma's membrane potential decay rate	19.56	0.78
External input	0.08	0.33
Absolute refractory period	6.00	3.46

Table 5.12: Neuron parameters of the ELIF model of two spike trains with indirect connection.

To analyse functional connectivity by the Cox method with the influence function given by formula (5.3), the characteristic times are specified as $\tau_s = 10\,ms$ and $\tau_r = 0.1\,ms$. To specify a time lag the CCF is calculated for all pairs of spike trains (Fig. 5.21): Δ_{12}= 11ms, Δ_{23}= 12ms, Δ_{13}= 23 ms, and all other lags are zero. It is assumed that the target spike train is $\#k, (k = 1,2,3)$. The estimates $(\hat{\beta}_{ik}, \hat{\beta}_{jk})$ $(i = 1,2,3; j = 1,2,3; k = 1,2,3, i \neq j, i \neq k, j \neq k)$ of two Cox coefficients have been calculated using formulas (5.5) and (5.6) as well as a confidence region on the plane (β_{ik}, β_{jk}) using formula (5.7). The confidence region ($\alpha = 0.05$) has an elliptic shape and the centre of the confidence region is located at point $(\hat{\beta}_{ik}, \hat{\beta}_{jk})$.

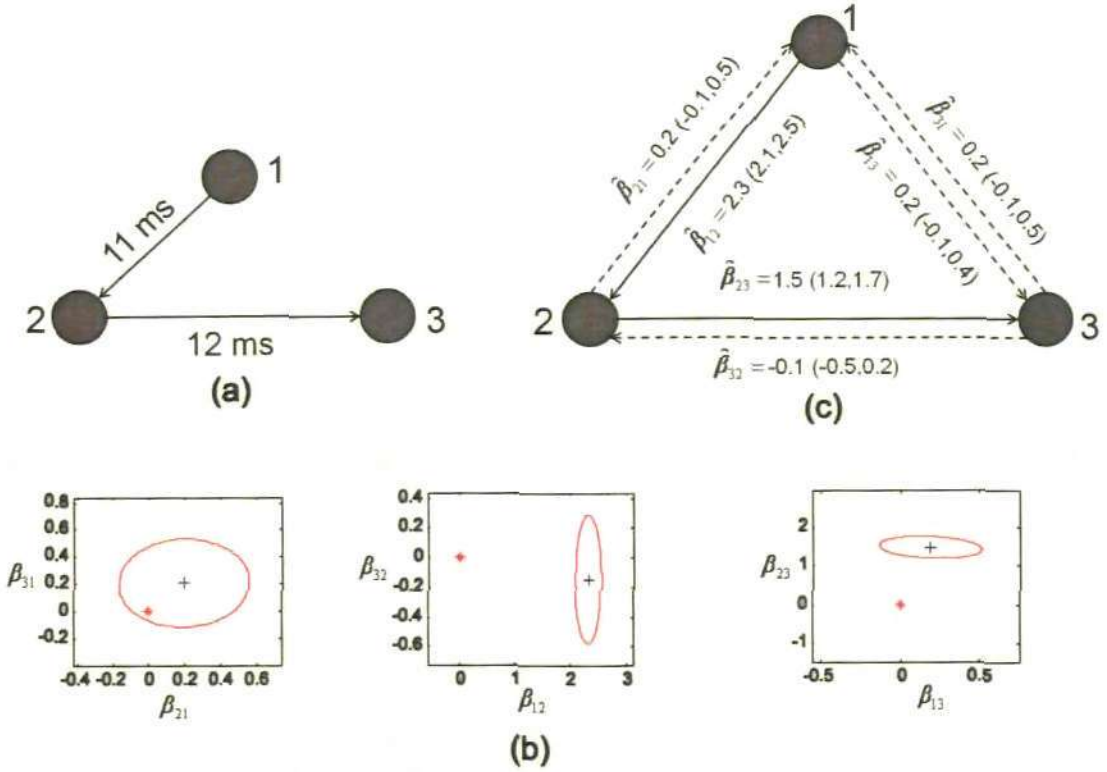


Figure 5.20: **(a)** 'Indirect connection' scheme of three spike trains. Spike train #1 influences spike train #2, which influences spike train #3, with time delays 11 ms and 12 ms, respectively. **(b)** Confidence regions of the estimated Cox coefficients in three cases: influences to spike train #1 (left), influences to spike train #2 (middle), influences to spike train #3 (right). **(c)** Estimated coefficients of the Cox method with confidence intervals. Significant connections are indicated by solid arrows.

Fig. 5.20(b) shows three confidence regions. The region on the left side corresponds to the target spike train #1; the region in the middle – target spike train #2, the region on right side – target spike train #3. The region on the left side contains zero, therefore both connections to spike train #1 are concluded to be absent. This result is shown in Fig. 5.20(c): two dashed arrows pointing to #1 indicates the absence of both connections.

The region in the middle does not contain the origin and this indicates that the null hypothesis should be rejected. The centre of the confidence region is shown by the blue cross and its coordinates are the estimates $(\hat{\beta}_{12}, \hat{\beta}_{32})$. The projection to the vertical axis $\hat{\beta}_{32}$ contains zero, therefore, it is concluded that the null hypothesis should be accepted and the connection from #3 to #2 is absent. The projection to the horizontal axis $\hat{\beta}_{12}$ does not contain zero, therefore, it is concluded that the null hypothesis should be rejected and the estimate $\hat{\beta}_{12}$ is the strength of connection from #1 to #2. This result is shown in Fig. 5.20(c) by two arrows pointing to the #2: the dashed arrow indicates the absence of a connection from #3 to #2, the solid arrow indicates the presence of a connection from #1 to #2, and the value of the connection strength is $\hat{\beta}_{12} = 2.3$.

The region on the right side can be interpreted in a similar way. The result is shown in Fig. 5.20(c) by two arrows pointing to #3: the dashed arrow indicates the absence of a connection from #1 to #3, the solid arrow indicates the presence of a connection from #2 to #3, and the value of the connection strength is $\hat{\beta}_{23} = 1.5$.

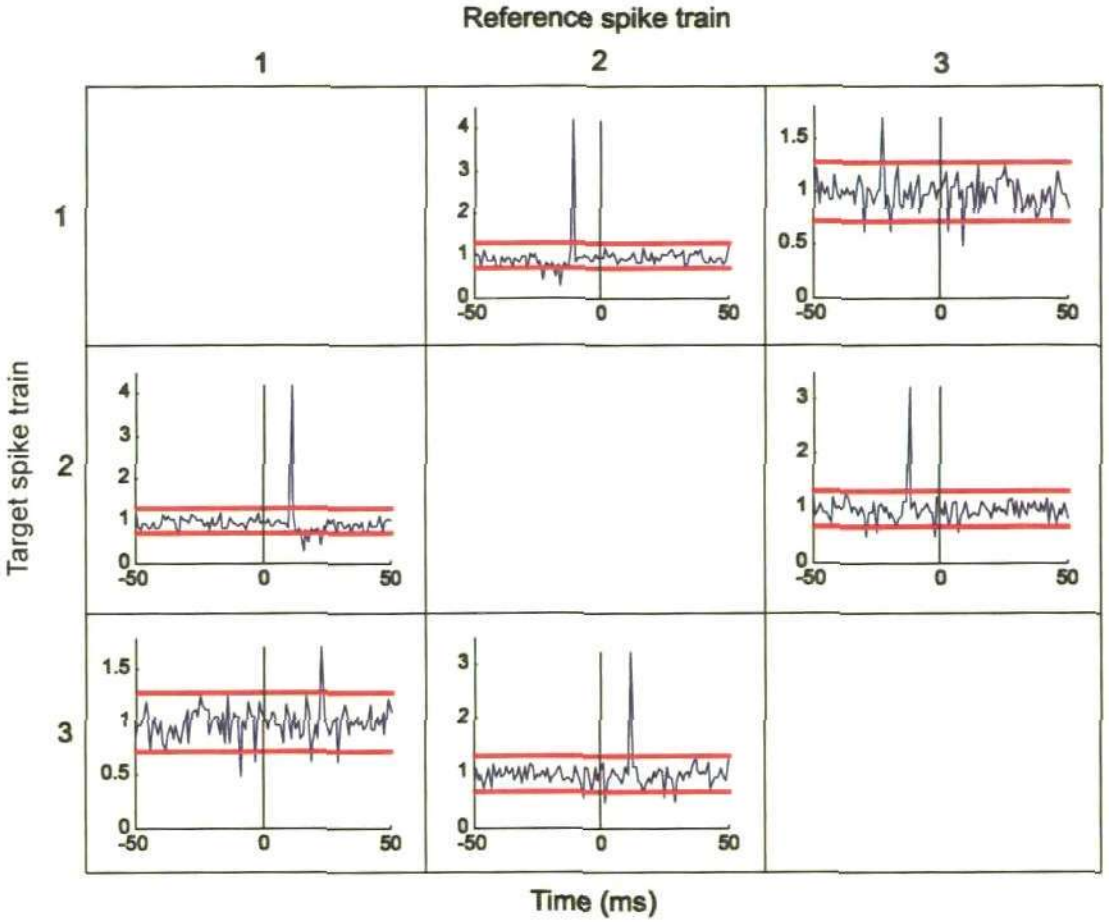


Figure 5.21: Pair-wise cross correlation functions of three spike trains. Each CCF is shown for selected pair of spike trains (called target and reference). Diagram of connections (indirect connection) is shown in Fig. 5.20(a).

The results of analysis are in good agreement with the architecture of connections which were used to generate the data (compare Fig 5.20(a) with Fig. 5.20(c)). For example, for data generation, a higher connection strength was selected for connection from #1 to #2 and the estimated connection strength from #1 to #2 is also higher than the estimated connection strength from #2 to #3.

Thus, the result in Fig. 5.20(c) indicates that there are two significant influences only (shown by solid arrows, all others are shown by the dashed arrows): from spike train #1 to spike train #2, from spike train #2 to spike train #3, and the influence strengths are shown with their confidence intervals ($\alpha = 0.05$).

Fig. 5.21 illustrates the result of analyzing the same three spike trains by the pair-wise CCF method. Each row of the figure shows two CCF corresponding to the selected target spike train – spike train #1 is the target for the first row, spike train #2 is the target for the second row, etc. The CCF analysis reveals three connections: from spike train #1 to spike train #2 (second row, first column), from spike train #2 to spike train #3 (third row, second column), and from spike train #1 to spike train #3 (third row, first column). The first two of these connections correspond well to the diagram of connectivity (Fig. 5.20(a)) but the third one is spurious (from #1 to #3) and this connection appears to be due to the 'indirect' connection from trains #1 to train #3. Thus, the Cox method is able to distinguish the 'indirect' connection from the direct connections but the CCF method fails.

5.8 Cox metric

Spike train similarity or dissimilarity measures are important tools to quantify the relationship among pairs of spike trains (Paiva et al., 2009). Such measures are essential for classification, clustering, or other forms of spike train analysis. By using a distance measure it is possible to identify the similarity or dissimilarity of the pair of spike trains. If the distance between two spike trains is small enough then it can be assumed that these spike trains are 'identical'. A traditional measure of similarity between two spike trains is based on the calculation of the pair-wise cross-correlation function (CCF) (Perkel et al., 1967). It is a bin based method which requires stationarity and a sufficiently large length of spike trains. There are several binless spike train similarity measures (Victor and Purpura, 1997; van Rossum, 2001; Schreiber et al, 2003; Hunter and Milton, 2003; Kurtz et al, 2007; and Houghton, 2009). In this section a new binless spike train metric is proposed which is based on the pair-wise analysis of the Cox method and hence the name Cox metric. The Cox metric is used to identify the groups

of similar spike trains, between multiple spike trains, with a clustering algorithm. This is a mathematical metric which has the following properties:

1. $d(SPT_A, SPT_B) \geq 0$ (non-negativity)
2. $d(SPT_A, SPT_A) = 0$
3. $d(SPT_A, SPT_B) = d(SPT_B, SPT_A)$ (symmetry)
4. $d(SPT_A, SPT_C) \leq d(SPT_A, SPT_B) + d(SPT_B, SPT_C)$ (triangle inequality)

where SPT_A, SPT_B are pairs of spike trains.

The Cox metric is based on the pair-wise estimate of the Cox coefficient (β_{ij}) from i th spike train to j th spike train $(i = 1,2,\cdots\cdots,p,\ j = 1,2,\cdots\cdots,p,\ i \neq j)$, where p is the total number of spike trains. Estimation of the pair-wise Cox coefficient (β_{ij}) and the confidence interval are described in section 5.3.1 and an example is given in section 5.4 for four generated spike trains. If the confidence interval contains zero, then it is concluded that the corresponding Cox coefficient is not distinguishable from zero and the influence from spike train i to spike train j is absent. If the confidence interval does not include zero, it is concluded that there is a significant influence from spike train i to spike train j and the value of the estimate characterizes the strength of this influence.

In fact, the Cox coefficient (β_{ij}) from i th spike train to j th spike train can also be estimated considering all spike trains at once instead of pair-wise analysis. The analysis based on all spike trains at once identifies better functional connections than the pair-wise analysis which is demonstrated in section 5.4. Analysis based on all spike trains at once takes very long computation time. For example, analysing 30 spike trains at a time each have 1000 spikes (average) needs 16 minutes in FORTRAN on a PC computer with 2.8 dual core processor. On the other hand, the pair-wise analysis is faster than the analysis based on all spike trains at once. For example, the same 30 spike trains needs

only 26 seconds in FORTRAN. More importantly, the Cox metric based on pair-wise analysis and all spike trains at once produces similar result.

To formulate the Cox metric, it is assumed that Cox coefficient from spike train i to spike train j (β_{ij}) is the same as the Cox coefficient from spike train j to spike train (β_{ji}); i.e., they are symmetric. To get the symmetric Cox coefficient, the maximum value of (β_{ij}) and (β_{ji}) is taken as a measure of functional connectivity between spike trains i and j. Thus, a symmetric matrix of functional connectivity is obtained from p spike trains. This matrix is considered as a matrix of similarities between p spike trains. Thus, the pair of spike trains whose Cox coefficient has the largest value in the similarity matrix is considered to be the most similar pair of spike trains. The distance between two points i and j in the clustering algorithm is considered as a difference between the maximum Cox coefficient in the similarity matrix and the Cox coefficient between spike trains i and j. Thus, a metric (distance) between spike train i and spike train j is calculated:

$$D_{ij} = \max_{i,j} \beta_{ij} - \max(\beta_{ij}, \beta_{ji}), \qquad i,j = 1,2,\cdots\cdots,p, \;\; i \neq j \tag{5.8}$$

In combination with a clustering algorithm this metric is used for finding functional connectivity and identifying the groups of interacting neurons. In the clustering algorithm the average linkage is used for calculating a distance between clusters.

5.8.1 Application to twenty spike trains

To apply the Cox metric to generated data, a set of twenty spike trains is generated with the ELIF model. The connection diagram includes forty nine connections and the connections are coupled into five groups. A diagram of connections is shown in Fig. 5.22. Parameter values of the ELIF model are given in Table 5.13. A simulation has been run over the time interval of 50 seconds.

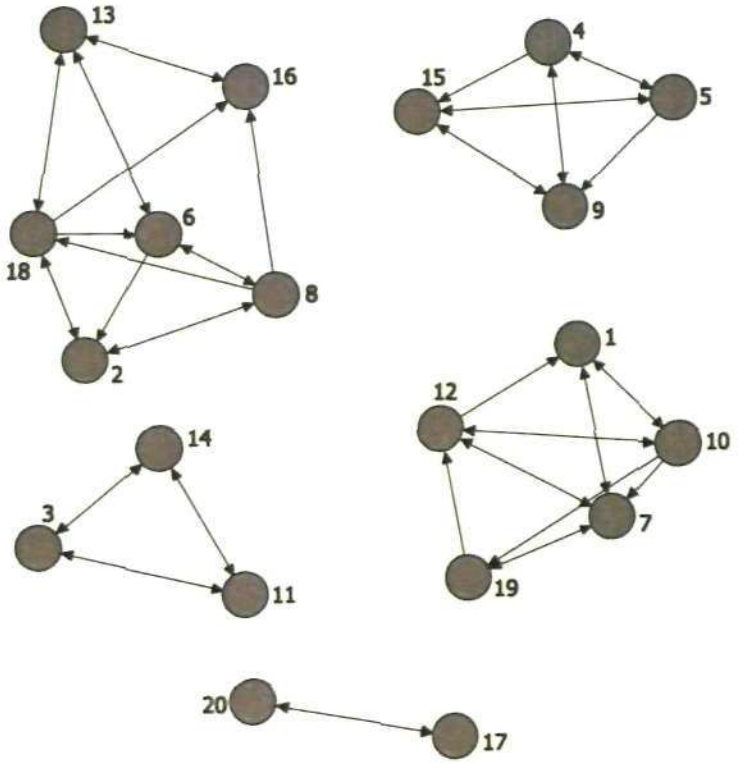


Figure 5.22: Connection scheme of the twenty spike train. There are forty nine non zero connections which are shown by arrows. These forty nine connections are coupled in five groups.

In order to apply the Cox metric to the twenty spike trains the pair-wise analysis of the twenty spike trains is conducted. A total of 380 possible pairs of spike trains are analyzed by the pair-wise Cox method taking one spike train as a target (j) and the other spike train as a reference (i). All the influence functions are considered to be identical and specified by (5.3) with the parameters $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$. Another parameter, time lag Δ of the influence function, is specified by the pair-wise CCF method. Now using the parameter values $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$, and time lags Δ obtained from CCF, the parameters of pair-wise analysis of the Cox method are estimated using (5.5) with the corresponding confidence interval by (5.6). Fig. 5.23(a) shows the connection of the twenty spike trains used for data generation (the same scheme is shown in Fig. 5.22 in a graph format). Fig 5.23(b) shows a diagram of functional connections in matrix format which has been identified by the pair-wise Cox

method. A circle indicates that there is a significant influence (functional connection) from the reference spike train (i) to the target spike train (j) and the radius of the circle shows the relative strength of the influence. Comparison of the connectivity matrix in Fig. 5.23(a) with the matrix in Fig. 5.23(b) shows that the pair-wise analysis of the Cox method correctly identifies all forty nine direct connections. This result is accurate enough, however, there are eight erroneous connections shown by small red circles.

Neuron parameter	**Mean**	**S.D.**
Maximum value of the threshold	44.93	1.17
Threshold decay rate	2.92	0.46
Asymptotic threshold value	14.47	1.08
Amplitude of the noise	4.94	0.38
Noise decay rate	10.02	0.03
Initial value of after spike hyperpolarisation	-28.84	0.42
Soma's membrane potential decay rate	20.15	0.76
External input	-0.03	0.38
Absolute refractory period	5.20	1.60
Connection parameter		
Connection strength	12.35	1.11
Decay rate of postsynaptic potential	2.80	0.73
Time lag of spike propagation	11.73	1.33

Table 5.13: Parameter values of the ELIF model to generate twenty spike trains.

The Cox metric is also applied to the connectivity matrix obtained by the Cox method considering all spike train at once. Using the same parameter values of the pair-wise analysis, the parameters of the Cox coefficient and their confidence intervals are

estimated considering all spike trains at once. Fig 5.23(c) shows a diagram of functional connections in matrix format which has been identified by the Cox method considering all spike trains at once. Comparison of the connectivity matrix in Fig. 5.23(a) with the matrix in Fig. 5.23(c) shows that the Cox method considering all spike trains at once correctly identifies all forty nine direct connections with only three erroneous connections shown by the small red circles.

In the Cox metric it is assumed that the Cox coefficients (β_{ij}) and (β_{ji}) are symmetric, for example, in Fig. 5.23(b) there is a connection from spike train #12 to spike train #1 but there is no connection from spike train #1 to spike train #12. In the Cox metric it is also assumed that there is a connection from spike train #1 to spike train #12 with the same strength as from spike train #12 to spike train #1. Again, in the connection matrix there is a connection from spike train #1 to spike train #10 and from spike train #10 to spike train #1, but the connection strength from spike train #1 to spike train #10 is stronger than the connection strength from spike train #10 to spike train #1. In the Cox metric it is assumed that both connections are the same and the connection strength is considered to be the maximum of these two connection strengths, which in this case is from spike train #1 to spike train #10. After having the symmetric matrix of functional connections the Cox metric is obtained by (5.8). A clustering algorithm is applied to the Cox metric which enables the identification of the groups of similar spike trains in the set of twenty spike trains. The result of the clustering algorithm applied on the Cox metric is presented in Fig. 5.24.

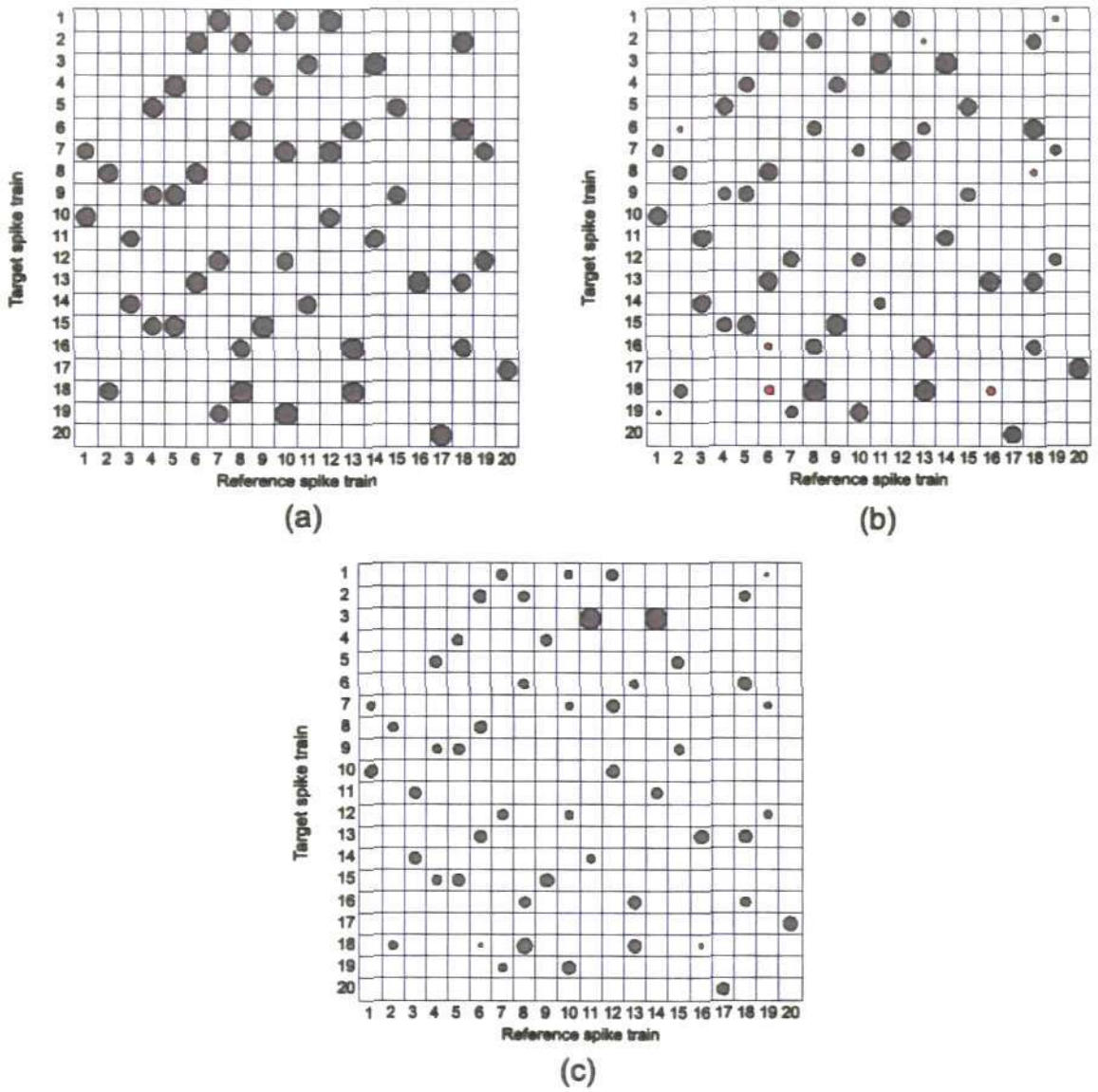


Figure 5.23: (a) Connection scheme of twenty spike trains in matrix format (the same as the scheme shown in Fig. 5.22 in graph format). **(b)** A diagram of functional connections of twenty spike trains obtained by the pair-wise analysis of Cox method. **(c)** A diagram of functional connections of twenty spike trains obtained by the Cox method considering all spike trains at once.

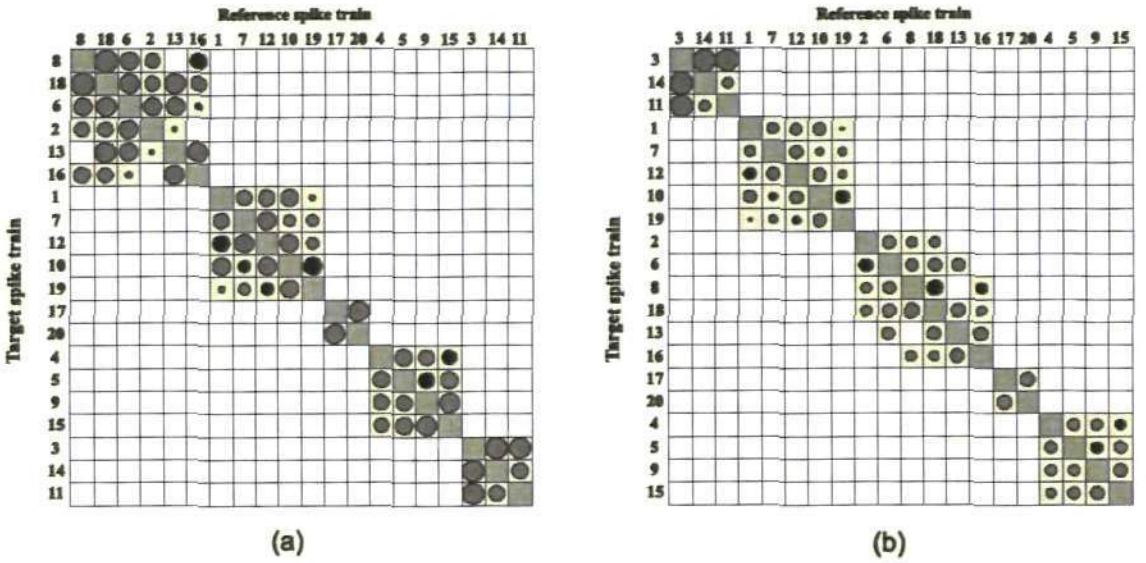


Figure 5.24: Groups of similar spike trains revealed by the Cox metric of twenty spike trains shown in Fig. 5.22. **(a)** Cox metric using pair-wise analysis. **(b)** Cox metric considering all spike trains at once.

In Fig. 5.24(a) the grey circles show the connections revealed by the pair-wise analysis of Cox method. The black circles are not identified by pair-wise analysis of the Cox method but are shown in the matrix to keep symmetry. Similarly, in Fig. 5.24(b) the grey circles show the connections revealed by the Cox method considering all spike trains at once. The black circles are not identified by the Cox method considering all spike trains at once but are shown in the matrix to keep symmetry. From Fig. 5.24, it is observed that Cox metric based on pair-wise analysis and considering all spike trains at once correctly identifies all the five coupled groups of spike trains. The ordering of the groups of similar spike trains in Fig. 5.24 is different because of clustering. In Fig. 5.24(a) there are some connections which do not exist in Fig. 5.24(b). For example, there is a connection from spike train #6 to spike train #16 in Fig. 5.24(a), but this connection does not exist in Fig. 5.24(b). Investigation from the diagram (Fig. 5.22) that is used for the generation of the spike trains shows that there is no connection from spike train #6 to #16. This indicates that Cox method based on all spike trains at once seems to give better functional connectivity than the pair-wise analysis of Cox method.

Though there are differences to identify functional connectivity, these two methods enable to identify groups of coupled spike trains which are similar. Thus the Cox metric is useful for ordering and visualising of spike train couplings, as well as for finding the groups of mutually coupled spike trains.

5.9 Motif analysis using Cox method

Patterns of interconnections among multiple spike trains are important to understand their relationships. A pattern of interconnections is usually meant as a connected m-vertex graph which is a subgraph of a larger graph. To find these patterns of interconnections a motif analysis is used (Milo et al., 2002; Sporns and Kotter, 2004). A motif is a connected subgraph of m vertices occurring in a directed graph at a number significantly higher than in randomized versions of the graph. That is, in graphs with the same number of vertices, edges and degree distribution as the original one, but where the edges are distributed at random.

A directed graph is a configuration whose figures are ordered pairs of points. In this context, the content of a figure is one or zero in respective accordance with the existence or non-existence of a directed line from the first member of the figure to its second member. Hence the figure counting series is $1 + x$. Let $d_m(x)$ is the counting polynomial for directed graphs with m vertices. The counting polynomials $d_m(x)$ for $m = 1$ to 5 is provided by Harary and Palmer (1973):

$$d_1(x) = 1$$

$$d_2(x) = 1 + x + x^2$$

$$d_3(x) = 1 + x + 4x^2 + 4x^3 + 4x^4 + x^5 + x^6$$

$$d_4(x) = 1 + x + 5x^2 + 13x^3 + 27x^4 + 38x^5 + 48x^6 + 38x^7 + 27x^8 + 13x^9 + 5x^{10} + x^{11} + x^{12}$$

$$d_5(x) = 1 + x + 5x^2 + 16x^3 + 61x^4 + 154x^5 + 379x^6 + 707x^7 + 1155x^8$$
$$+1490x^9 + 1670x^{10} + 1490x^{11} + 1155x^{12} + 707x^{13} + 379x^{14}$$
$$+154x^{15} + 61x^{16} + 16x^{17} + 5x^{18} + x^{12} + x^{20}$$

Using the counting polynomial $d_m(x)$ the numbers of directed graphs for $m = 2$ and 3 vertices are shown in Fig. 5.25(a) and Fig. 5.26(a). To obtain the number of motifs from the directed graph, all vertices must have either outdegree or indegree of at least one. Thus the number of motifs for $m = 2$ and 3 vertices are 2 and 13 which are identified from Fig. 5.25(a) and Fig. 5.26(a). The motif ID for $m = 2$ and 3 vertices are shown in Fig. 5.25(b) and Fig. 5.26(b). For $m = 4$ and 5 the corresponding numbers of directed graphs are 218 and 9608; and the motif ID's are 199 and 9,364 (Harary and Palmer 1973). In this section motifs of size $m = 3$ are considered. There are some connected motifs that form a strongly connected graph. For $m = 3$, motifs with ID = 7, 9, 10, 12, and 13 are connected motifs. In a connected motif all vertices can be reached from all other vertices.

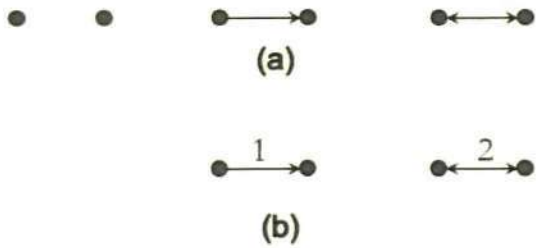
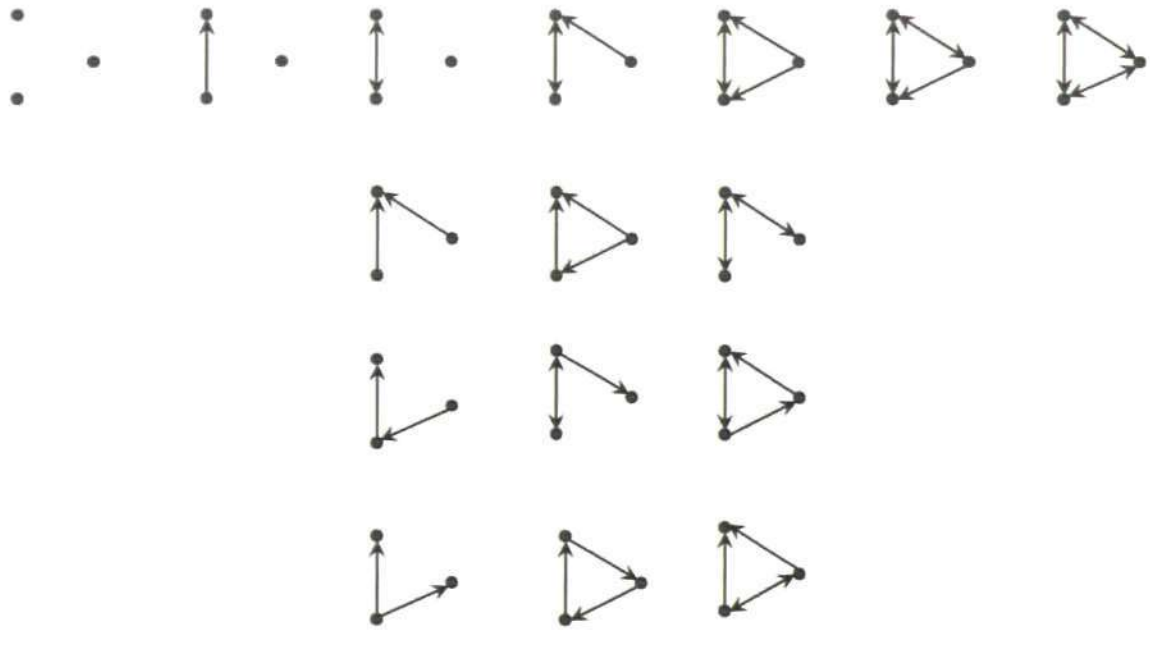
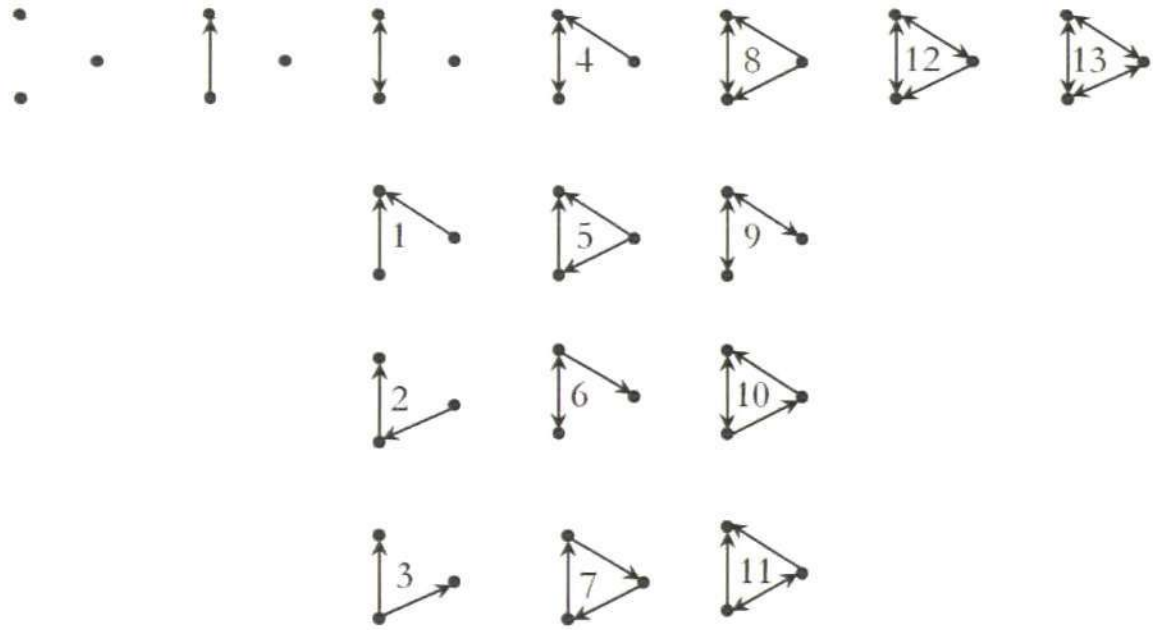


Figure 5.25: (a) The 3 directed graph of 2 vertices **(b)** The 2 motif ID of 2 vertices.



(a)



(b)

Figure 5.26: (a) The 16 directed graph of 3 vertices **(b)** The 13 motif ID of 3 vertices.

For multiple spike trains the diagram of functional connectivity is identified by the Cox method considering all spike trains at once. From this diagram of connectivity a structural motif count can be obtained by counting the number of distinct motif ID. Given the structural motif count for distinct motif ID, the bar diagram of structural

motif count is obtained. The diagram of functional connectivity can also be obtained using the triplet analysis of Cox method. For multiple spike trains all possible triplets are analysed and a diagram of functional connectivity can be identified. An example of triple analysis of Cox method is demonstrated in section 5.4.2. Structural motif can also be identified by the diagram of connectivity obtained from triplet analysis. The advantage of triplet analysis is that it requires less computational time than the analysis of all spike trains at once. The motif obtained from the triplet analysis is also very similar to the motif obtained from the analysis of all spike trains at once.

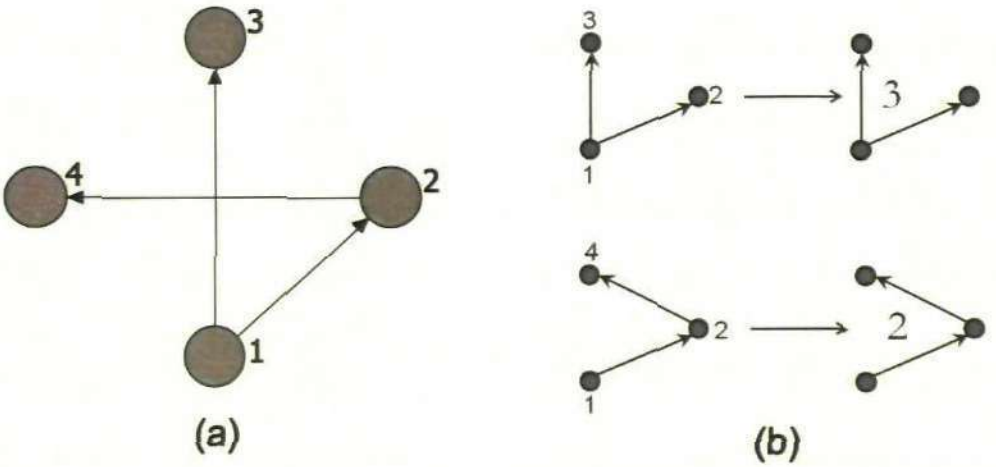


Figure 5.27: (a) Diagram of functional connectivity of four spike trains identified by the Cox method considering all spike trains at once. **(b)** Structural motif identified from the diagram of connectivity in (a). Numbers represent motif ID.

To understand the motif analysis a small set of four spike trains is considered, which was used in section 5.4. The diagram of functional connectivity of these four spike trains, obtained by the Cox method considering all spike trains at once, is given in Fig. 5.27(a). The diagram of functional connectivity of all possible triplets, obtained by the triplet-wise analysis of Cox method, is given in Fig. 5.28(a). From the diagram of connectivity (Fig. 5.27(a)), two possible structural motifs are identified (Fig. 5.27(b)). Similarly, from the diagram of connectivity (Fig. 5.28(a)), four possible structural motifs are identified (Fig. 5.28(b)).

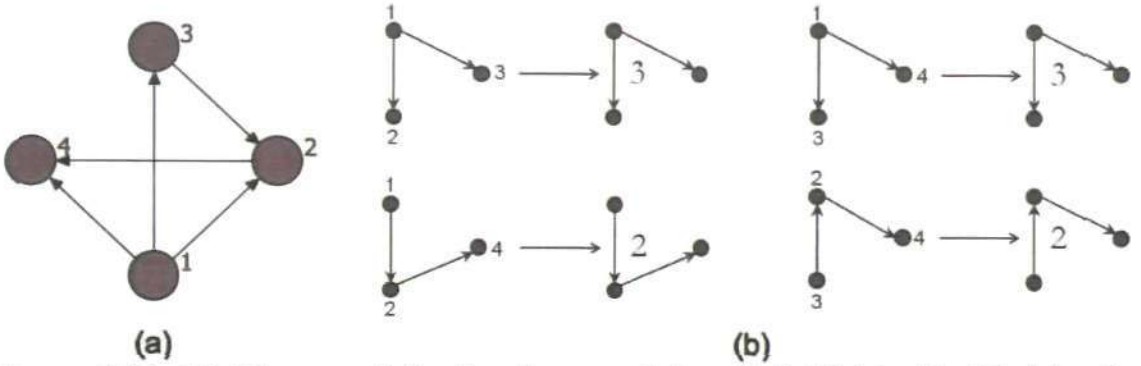


Figure 5.28: (a) Diagram of functional connectivity of all triplets identified by the triplet-wise analysis of Cox method. **(b)** Structural motif identified from the diagram of connectivity in (a). Number represents motif ID.

To search the significant structural motifs from the diagram of functional connectivity, a large number of randomized diagrams (n=100 or 1000) are generated keeping the same number of vertices and edges as the original diagram but distributing the edges at random. In order to quantify the significance of a given motif ID i, its Z-score can be computed (Boccaletti et al., 2006). If $n_i^{(real)}$ is the number of times that a motif ID i appears in the real diagram of functional connectivity, $< n_i^{(rand)} >$ and $\sigma_i^{(rand)}$ are the average and standard deviation of the motif ID i obtained from the randomized diagrams, then its Z-score can be computed as

$$Z_i = \frac{n_i^{(real)} - < n_i^{(rand)} >}{\sigma_i^{(rand)}} \tag{5.9}$$

A structural motif is considered to be significant if the Z-score of this motif is higher than 2 (Sporns et al., 2007).

5.9.1 Application to twenty spike trains

To identify the patterns of interconnections using structural motifs, a set of twenty spike trains is considered. The connection scheme of these twenty spike trains is given in Fig.5.15(a). The diagram of functional connectivity of these twenty spike trains is given in Fig. 5.29(a) which is obtained by the Cox method considering all spike trains at once.

This diagram of functional connectivity is similar to the Fig. 5.16(b) which is in matrix form. Fig. 5.30(a) shows the structural motif count of size $m = 3$ found within the diagram of connectivity in Fig. 5.29(a). To identify the significant structural motif, 1000 randomized diagrams are generated, with the same number of twenty vertices, forty four edges and degree distribution as the original one, but the edges are distributed at random. The structural motif count of size $m = 3$ found within the randomized diagrams is shown in Fig. 5.30(b). The motif ID 6, 9 and 12 appear more than the randomized diagrams, where as all other motif ID's appear less than the randomized diagrams. The Z-score is calculated using (5.9) and the result shows that motif ID 6 ($Z_6 > 2.15,\ p < 0.04$), motif ID 9 ($Z_9 > 14,\ p < 0.0001$) and motif ID 12 ($Z_{12} > 3.70,\ p < 0.0001$) are significant. An example of the significant patterns of interconnections (motif ID 9) that appears more than the randomized diagram is shown in the Fig. 5.29(b) with blue arrows.

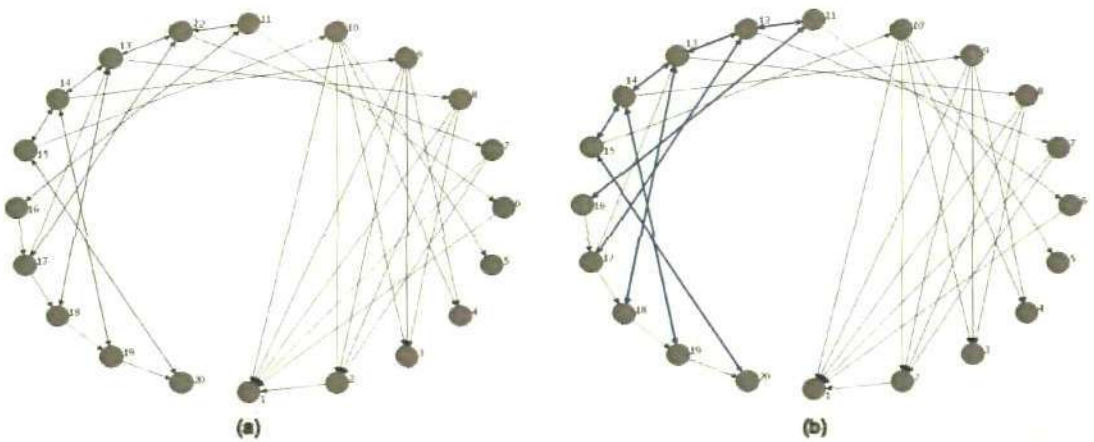


Figure 5.29: (a) Diagram of functional connectivity of the twenty spike trains by the Cox method obtained from the neural circuit of twenty spike train in section 5.6 (Fig. 515(a)). **(b)** Significant structural motifs from this diagram of functional connectivity. Here only motif ID 9 is shown by blue arrows for illustration.

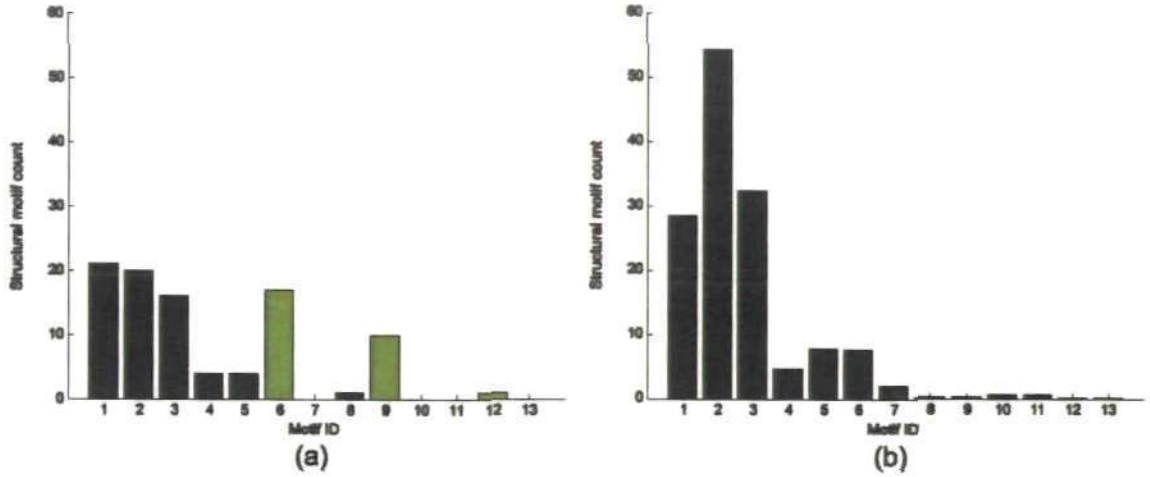


Figure 5.30: (a) Structural motif count of size $m = 3$ for the diagram of connectivity of twenty spike trains. Significant motif ID's are displayed as green. **(b)** Structural motif count of size $m = 3$ for the randomized diagrams.

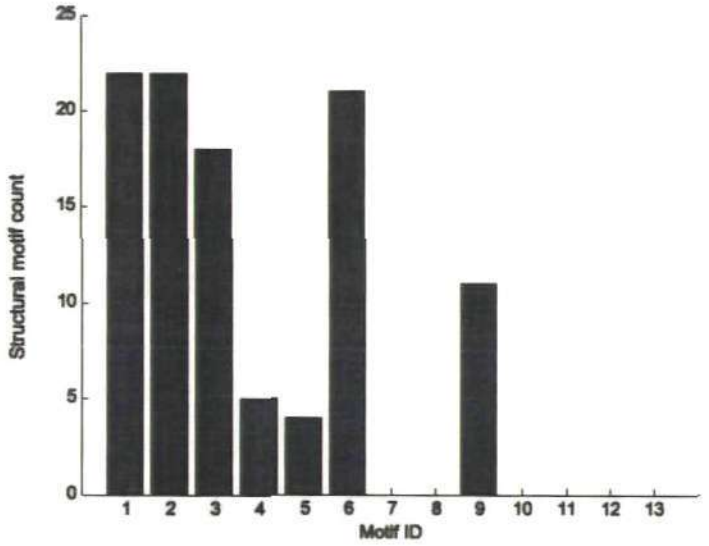


Figure 5.31: Structural motif count of all possible triplets of the twenty spike trains.

Structural motifs of the twenty spike trains can also be identified by applying the triplet analysis of the Cox method to the all possible triplets. For the set of twenty spike trains a total of 1140 groups of triplets are analysed. Diagram of connectivity of each triplet is used to identify the structural motif. A structural motif count for different motif ID's is shown in Fig. 5.31. Out of 1140 triplets, 103 triplets have different structural motif ID's. Among these 103 triplets, 22 triplets have structural motif ID's 1 and 2, which are the highest among the motif ID's. There are no triplets which have structural motif ID's 7, 8, 10, 11, 12 and 13. There are a very low proportion of connected motifs (10.68%) in

the triplets which indicates that the connection from every spike train to every other spike trains in the triplet is very low. More importantly, the structural motif count in Fig. 5.30(a) is mostly similar to that in the Fig. 5.31.

Chapter 6

Application of the methods to the experimental data

In previous chapters, the MCG method and the Cox method for studying the functional connectivity have been described and tested thoroughly using simulated spike trains. In this chapter these two methods are applied to analyse experimentally recorded multiple spike trains and derive a functional connectivity. These recordings from the cat's visual cortex (Nikolic, 2007; Schneider et al., 2006) have been kindly provided for analysing by Dr. Danko Nikolic (Max Planck Institute for Brain Research, Frankfurt, Germany). Also, the Cox method has been used to identify groups of similar spike trains (Cox metric) and reveal the patterns of interconnections among spike trains (motif analysis). The connections obtained by the MCG method and the Cox method are analysed.

6.1 Methods for data acquisition

6.1.1 Preparation

The cat was initially anesthetized with ketamine, and the anaesthesia was maintained with a mixture of 70% N_2O and 30% O_2 supplemented with halothane (0.4–0.6%). The animal was paralysed with pancuronium bromide (Pancuronium, Organon, 0.15 mg kg^{-1} h^{-1}). All the experiments were conducted according to the guidelines of the Society for Neuroscience and German law for the protection of animals, approved by the local government's ethical committee and overseen by a veterinarian.

6.1.2 Recording

Multi-unit activity (MUA) was recorded from a region of area 17 corresponding to the central part of the visual field by using a SI-based multielectrode probe (16 channels per

electrode) supplied by the Centre for Neural Communication Technology at the University of Michigan (Michigan probes) with inter-contact distance 200 μm (0.3–0.5 MΩ impedance at 1000 Hz). Signals were filtered between 500 and 3.5 kHz for extracting multi-unit activity (MUA), digitized with 32 kHz sampling frequency and stored in computer memory. All analyses were made on the basis of discrete spike events detected by a threshold that was set to a value of about two times the noise level. The probe was inserted in the cortex approximately perpendicular to the surface and allowed simultaneous recording from neurons at different cortical depths and along an axis tangential to the cortical surface. Fourteen MUA signals showed good responses to visual stimuli, orientation selectivity and overlapping receptive fields (RF). This resulted in a cluster of overlapping RFs that were stimulated simultaneously by a single visual stimulus.

6.1.2 Visual stimulation

Stimuli were presented binocularly on a 21 inch computer screen (HITACHI CM813ET) with 100 Hz refresh rate. To obtain binocular fusion the optical axes of the two eyes were first determined by mapping the borders of the respective RFs and then aligned on the computer screen with adjustable prisms placed in front of one eye. The software for visual stimulation was a commercially available stimulation tool, ActiveSTIM (http://www.ActiveSTIM.com). The stimuli consisted either of one white bar moving over a black background or consisted of two bars moving in different directions (60° difference). The bars always appeared at about 3° eccentricity of the centre of the cluster of RFs and moved with a speed of 1°/s such that they completely covered the cluster of RFs. In the stimuli with two bars, the bars crossed their paths at the centre of the RF cluster. At each trial the stimulus was presented in total for 5 s, but only

2 s with strongest rate response were used for the analysis. In the six stimulation conditions the bars moved in the following directions (1) 30° and 330°; (2) 0°; (3) 150° and 210°; (4) 180°; (5) 30° and 150°; (6) 210° and 330°. Each stimulation condition was presented 20 times, different conditions being presented in a randomized order.

6.2 Data description

To identify the functional connectivity of multiple spike trains, the MCG method and the Cox method are applied to a set of experimental data recorded from cat's visual cortex (Nikolic, 2007; Schneider et al., 2006). The experimental condition includes application of six different stimuli (different orientations of the moving grid). Each stimulus is repeated 20 times resulting in 120 applications of all stimuli. The order of stimuli presentations is random. Thus for 120 stimuli the spiking activity of 32 channels is recorded. From each channel one spike train is selected – the one with a medium firing rate. This spike train is used to prepare six spike trains corresponding to six stimuli. The twenty time intervals (each of six seconds duration), where the stimulus 1 is presented, have been selected to represent a total interval (120 seconds) of the application of stimulus 1. All spikes from this interval are considered continuously despite the gaps between the intervals of stimulus 1 representation. There are 32 channels in the experiment which results in 32 spike trains for this stimulus. The same operation has been repeated for stimulus 2: all subintervals of six seconds corresponding to application of stimulus 2 have been selected, considered continuously and all spikes have been taken for analysis of the functional connectivity. The same selection is done for each channel and 32 spike trains have been constructed for stimulus 2. Repeating this operation for other stimuli, six sets of 32 simultaneous spike trains have been constructed. Each set corresponds to application of one stimulus. For

each stimulus, 32 spike trains are analysed to identify functional connectivity by the MCG method and the Cox method.

6.3 Analysis of functional connectivity

For each stimulus the analysis of functional connectivity includes the following 4 procedures:

1. MCG method
2. Cox method
3. Cox metric
4. Motif analysis

Analysis of functional connectivity by the MCG method is based on the calculation of CCF which is discussed in chapter 3. This technique is able to differentiate direct connections from spurious connections (common source and indirect connections). In this chapter only direct connections are considered for analysing functional connectivity. Analysis of functional connectivity by the Cox method is based on the modulated renewal process and the procedure is discussed in chapter 5. This method considers the simultaneous effect of all spike trains and identifies only direct functional connectivity. Result of functional connectivity obtained by this method is analysed and compared to the result of functional connectivity obtained by the MCG method.

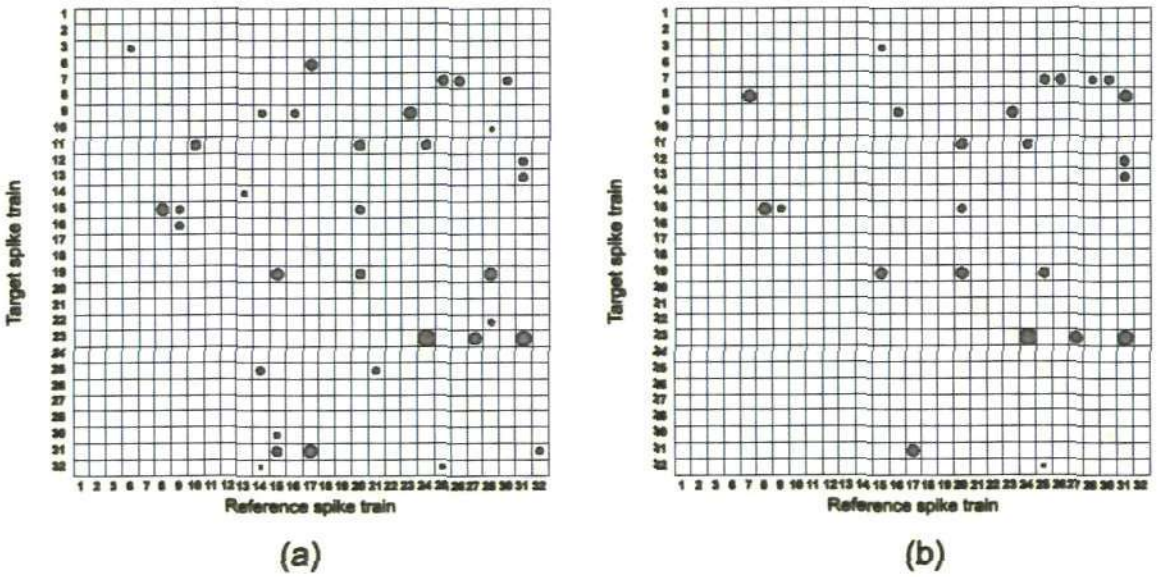


Figure 6.1: Functional connectivity of the 29 spike trains of stimulus 5 identified by the Cox method. **(a)** In the small interval [60000 ms, 66000 ms]. **(b)** In the small interval [66000 ms, 72000 ms].

A simple study using the Cox method shows that this method is useful for the application to experimental data. Two small intervals of stimulus 5, each of duration of 6 seconds ([60000 ms, 66000 ms] and [66000 ms, 72000 ms]) are chosen to identify the functional connectivity of the spike trains. For this study, three outlier spike trains (#4, #5 and #29) are not considered. These three spike trains have higher spiking rates than those of other spike trains. Functional connectivity of the 29 spike trains is identified by the Cox method. The functional connectivity of these two small intervals identified by the Cox method is shown in Fig. 6.1. In the Fig. 6.1, the functional connections are shown by circles and the direction of the connections is considered from reference spike train to target spike train. The radius of the circle shows the relative strength of influence. A big radius corresponds to a strong functional connection, whereas a small radius corresponds to a weak functional connection. The functional connections of the 29 spike trains in the two small intervals show a good agreement. In the interval [60000 ms, 66000 ms], 34 functional connections are identified by the Cox method. Similarly, 24 functional connections are identified by the Cox method in the interval [66000 ms,

72000 ms]. Investigation from these diagrams reveals that there are 18 connections which are common to both intervals. As these two diagrams of functional connectivity show a good agreement in the small time interval, it can be said that Cox method is useful for analysis of this experimental data. On the other hand MCG method is not applied to these small intervals, because the number of spikes is too small for the application.

The Cox method is further applied to identify the groups of similar spike trains by clustering algorithm. To identify the groups of similar spike trains, Cox metric is applied which is discussed in section 5.8. Another application of the Cox method is the motif analysis which is based on the analysis of triplet spike trains at once and is described in section 5.9. Motif analysis is useful in obtaining the patterns of interconnections among the spike trains.

6.4 Analysis of functional connectivity of stimulus 1

The raster plot of 32 spike trains is shown in Fig. 6.2. Raster plot reveals that three spike trains (#4, #5 and #29) have high spiking rates compared to all other spike trains. Therefore, these three spike trains are considered to be outliers and they are excluded from analysis. All the 29 spike trains have similar spiking pattern. Spiking rates of these 29 spike trains are high over time interval [78000 ms, 95000 ms]. These 29 spike trains are analysed to identify functional connectivity keeping the original numeration of the spike trains.

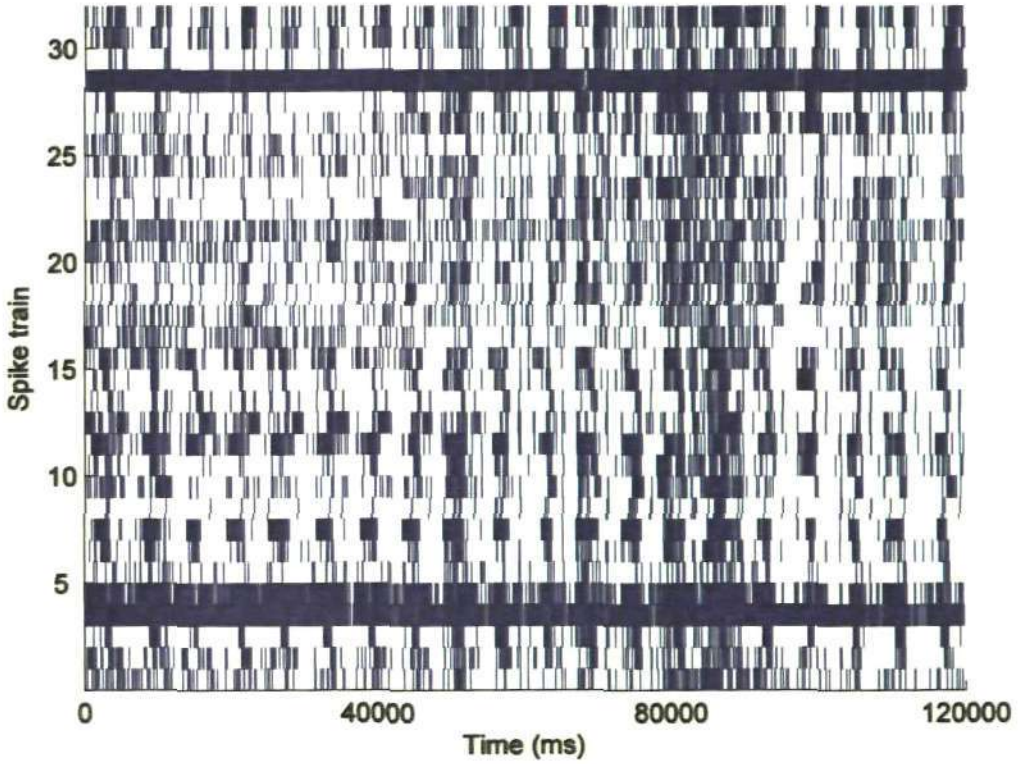


Figure 6.2: Raster plot of 32 spike trains of stimulus 1. Spike trains #4, #5 and #29 have high spiking rates and are not considered for analysing functional connectivity.

6.4.1 MCG method

For the 29 spike trains a total of $(29^2 - 29)/2 \approx 406$ pair-wise CCF are calculated with a bin size of 1 ms and a correlation window of 100 ms. To test the independence of two spike trains the significance level $\alpha = 0.05$ is used with the Bonferroni correction. A connection is considered to be significant if a peak of the CCF exceeds the upper boundary of the 'confidence interval'. A total of 100 significant connections are found for 29 spike trains. These significant connections are shown in a matrix format in Fig. 6.3(a) where the connections are indicated by circles. The direction of connection is considered from the reference spike train to the target spike train. For example, spike train #10 has a connection to spike train #28. Among the 29 spike trains, spike train #28 has 14 outgoing connections to other spike trains which is the highest among 29 spike trains and similar 15 incoming connections from other spike trains.

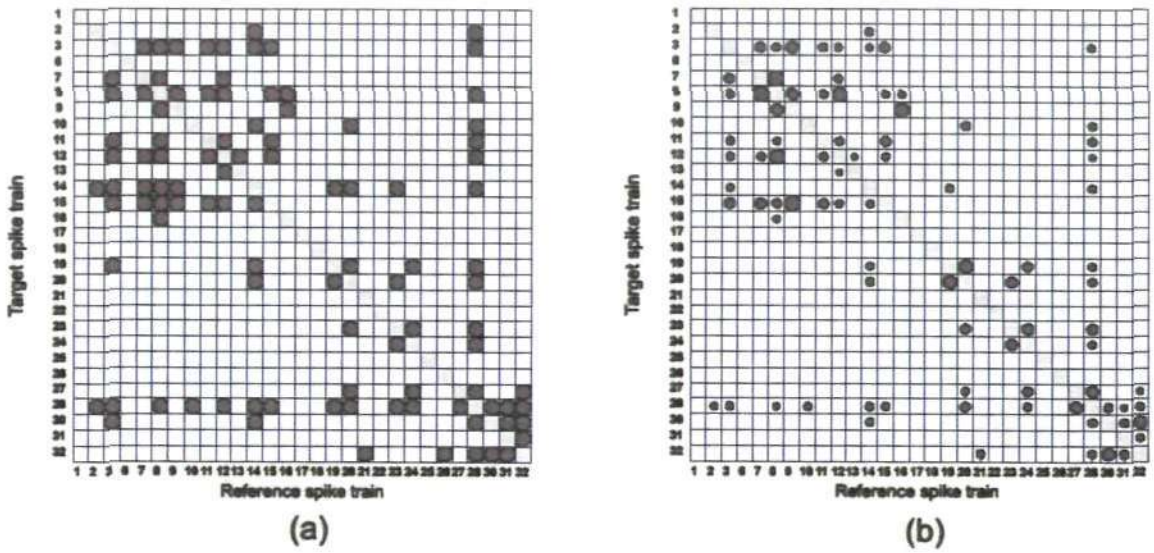


Figure 6.3: (a) Significant connections obtained from pair-wise CCF analysis of the 29 spike trains of stimulus 1. **(b)** Direct connections obtained from the significant connections in (a) after the clustering algorithm. The radius of the circle indicates the relative strength of the connections.

A clustering algorithm is applied to the 100 significant connections for distinguishing direct connections from spurious ones (i.e., indirect connections and common source). Here only direct connections are considered to identify functional connectivity. After clustering, the 85 connections are identified as direct connections (Fig. 6.3(b)). The strengths of direct connections are normalized between 0 and 1 according to their significant peaks. These strengths are shown by circles in Fig. 6.3(b). Big radius indicates strong functional connection and small radius indicates weak functional connection. Among 85 connections 10 connections have strong strength compared to others. These connections are: (#7, #8), (#8, #7), (#8, #12), (#9, #15), (#12, #8), (#16, #9), (#19, #20), (#20, #19), (#30, #32) and (#32, #30). All other connections have medium strength. Spike train #28 has 12 outgoing connections and 12 incoming connections, which are the highest among 29 spike trains. There are 31 pairs of connections where both spike trains have functional connectivity to each other. For example the pair (#7, #8); where there is a connection from spike train #7 to spike train #8 and vice versa.

6.4.2 Cox method

To analyse functional connectivity of the 29 spike trains, one spike train is considered as the target spike train and other 28 spike trains are considered as the reference spike trains. The influence function (and its parameters) which determines how the reference spike train influences the target spike train should be specified. Here it is assumed that all the influence functions are identical. The inter spike interval (ISI) histogram of three spike trains, spike train #1, #6 and #18 are given in Fig. 6.4. These histograms have high count for the short ISI and the ISI count decreases with increase of the ISI length. That suggests that the influence function should be specified by the formula (5.3). The parameters of the influence function (5.3) are $\tau_r = 0.1\,ms$, $\tau_s = 10\,ms$. Another parameter, the time lag Δ is specified from pair-wise CCF analysis. Thus, the influence functions are defined and the Cox coefficients and the corresponding confidence intervals are calculated using formulas (5.5) and (5.6). This procedure is repeated 29 times to obtain the full functional connectivity of the 29 spike trains. The confidence intervals are calculated using the significance level $\alpha = 0.05$ with Bonferroni correction.

The 71 connections, identified by the Cox method, are shown by circles in Fig. 6.5(a). The radius of the circle indicates the strength of functional connection. The direction of functional connection is from the reference spike train to the target spike train. Among the 71 connections, the 9 connections have stronger strength compared to others. These connections are: (#8, #9), (#16, #9), (#20, #19), (#24, #19), (#24, #23), (#25, #1), (#26, #25), (#28, #27) and (#32, #30). 8 connections have a small strength compared to others. These connections are: (#2, #14), (#7, #28), (#10, #28), (#14, #28), (#15, #3), (#15, #28), (#27, #28) and (#28, #2). All other connections have a medium strength. Spike train #8 has 6 outgoing connections and spike train #28 has 6 incoming

connections, which are the highest among 29 spike trains. There are 16 pairs of connections where both spike trains have functional connectivity to each other.

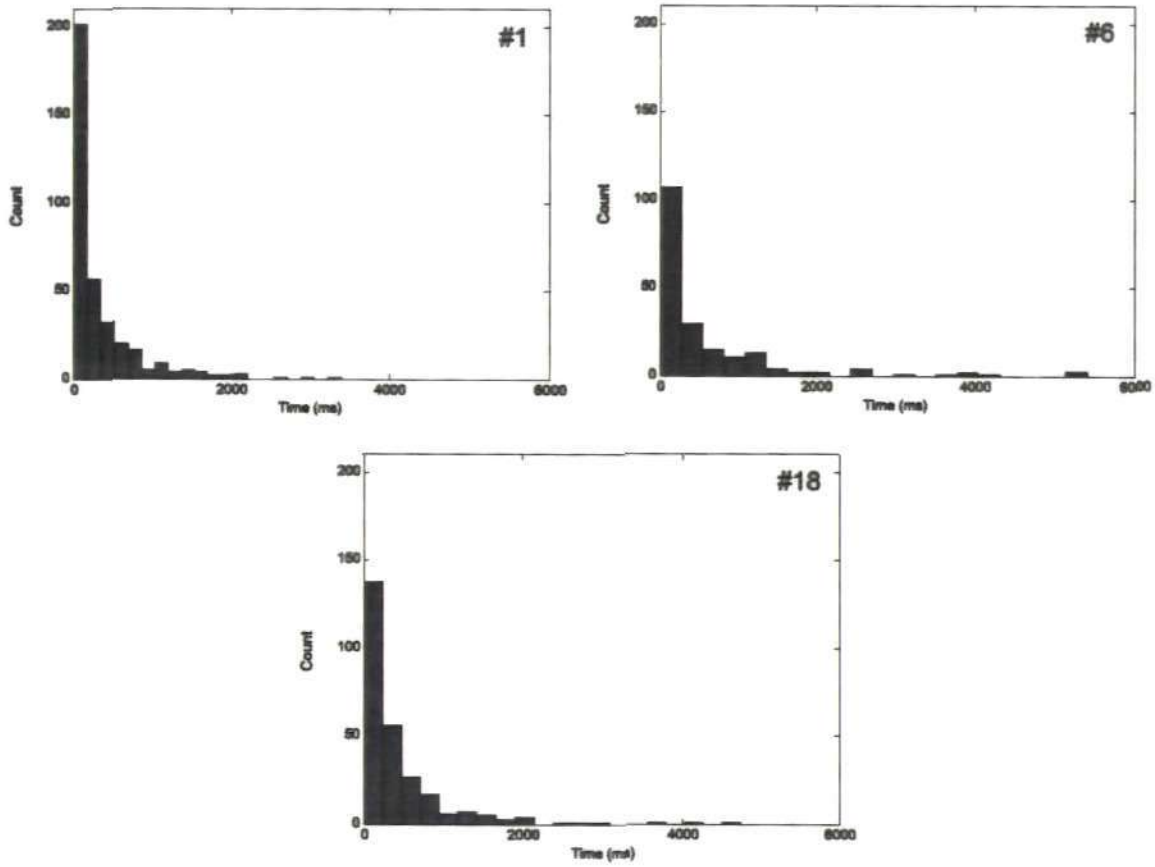


Figure 6.4: Inter spike interval histogram of the spike trains #1, #6 and #18 of stimulus 1.

Functional connectivity obtained by the MCG method and the Cox method show a good agreement between them (Fig. 6.5(b)). There are 43 connections which are common in both techniques. Among the common connections spike train #8 has 6 outgoing connections to other spike trains and spike trains #3, #15 and #28 have 4 incoming connections which are the highest among the 29 spike trains. There are 9 pairs of connections where both spike trains have functional connectivity to each other. These pairs of connections are: (#3, #7), (#3, #15), (#8, #12), (#8, #15), (#19, #20), (#23, #24), (#27, #28) and (#30, #32).

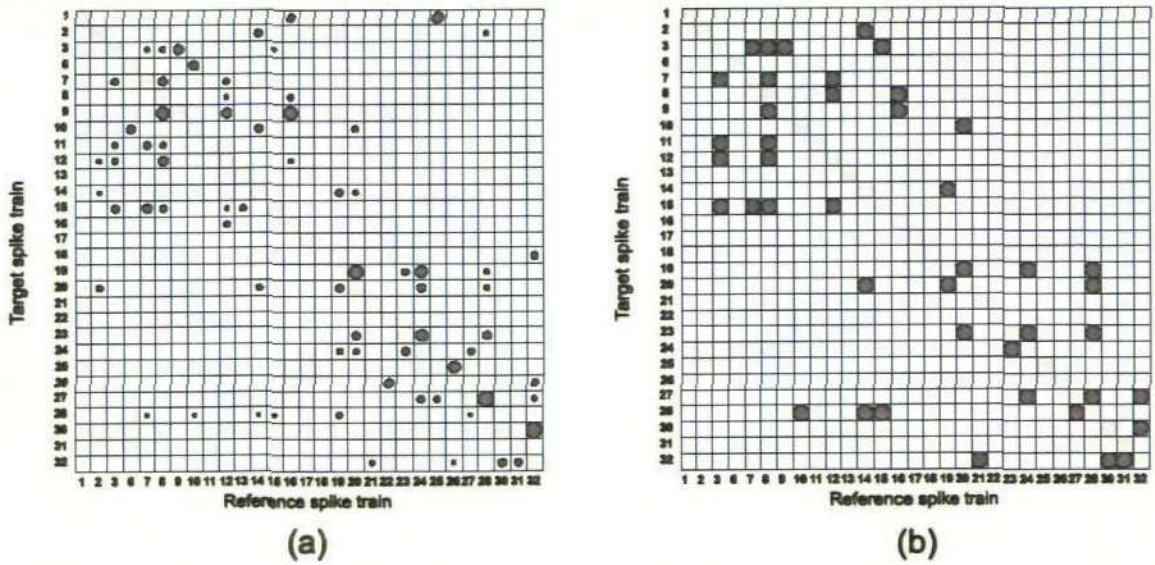


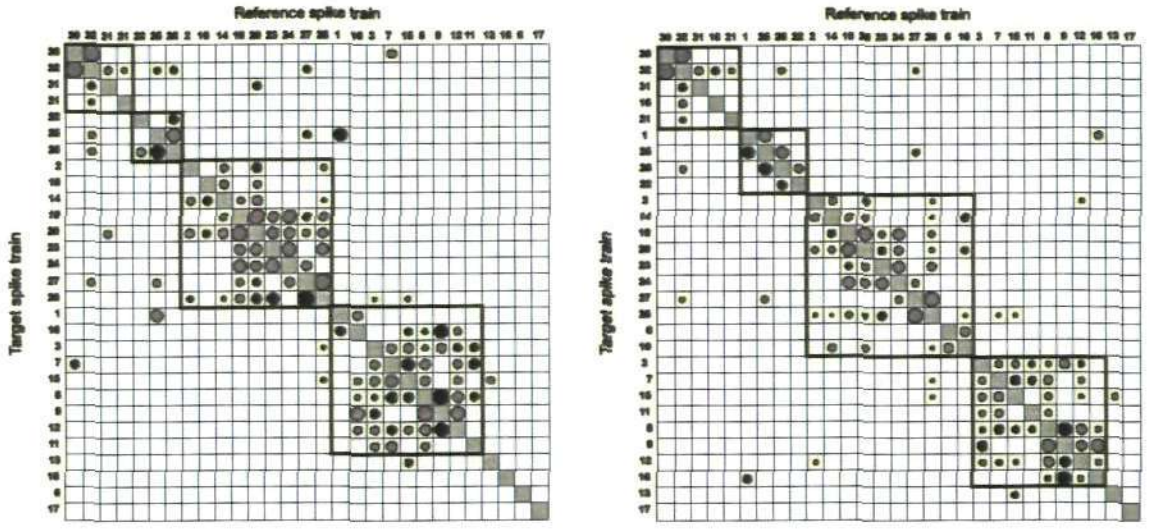
Figure 6.5: (a) Functional connectivity of the 29 spike trains identified by the Cox method of stimulus 1. Radius of the circle indicates the relative strength of connection. **(b)** Connections that are identified both by the MCG method and the Cox method.

6.4.3 Cox metric

To apply the Cox metric to the 29 spike trains a total of 812 possible pairs of spike trains are analysed by pair-wise Cox method where one spike train is taken as the target and the other spike train as the reference. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with the corresponding confidence interval by (5.6). The Cox metric is applied to the significant connections to reveal the groups of similar spike trains. The result of the Cox metric is shown in Fig. 6.6(a) where the grey circles indicate the significant connections obtained by the pair-wise analysis of the Cox method. The black circles indicate symmetric of the grey circles but not identified by the pair-wise analysis of the Cox method. Similarly, the Cox metric is applied to the functional connections identified by the Cox method considering all spike trains at once. This functional connection is shown in Fig. 6.5(a) and the result of the

Cox metric is shown in Fig. 6.6(b). In the figure the grey circles indicate the significant connections obtained by the Cox method considering all spike trains at once. The black circles indicate symmetric of the grey circles but not identified by the Cox method considering all spike trains at once.

From Fig. 6.6(a) four groups of similar spike trains are identified. The first group consists of 4 spike trains, these are: spike trains #30, #32, #31 and #21. In this group, the connection from spike train #30 to #32 and spike train #32 to #30 have the highest strength among other connections. The second group consists of 3 spike trains, these are: spike trains #22, #25 and #26. The third group consists of 9 spike trains; these are: spike trains #2, #10, #14, #19, #20, #23, #24, #27 and #28. In this group, five connections (#19, #20), (#20, #19), (#23, #24), (#24, #23) and (#28, #27) have big strength compared to others. Spike train #19 has 6 outgoing connections to spike trains #14, #20, #23, #24, #27 and #28, and 5 incoming connections from spike trains #14, #20, #23, #24 and #28. As this spike train has the highest outgoing connections, this spike train can be considered as the most influential spike train of this group. The forth group consists of 8 spike trains, these are: spike trains #16, #3, #7, #15, #8, #9, #12 and #11. In this group there are 3 big strengths of connections: (#16, #9), (#8, #9) and (#12, #9). Spike train #8 has 6 outgoing connections to spike trains #3, #7, #15, #9, #11 and #12 and can be considered as the most influential spike train of this group. Spike trains #13, #18, #6 and #17 do not follow any group.



(a) (b)

Figure 6.6: Groups of similar spike trains revealed by the Cox metric of the 29 spike trains of stimulus 1. **(a)** Cox metric using pair-wise analysis. **(b)** Cox metric considering all spike trains at once.

Similarly, from Fig. 6.6(b) four groups of similar spike trains are identified. The first group consists of 5 spike trains, these are: spike trains #30, #32, #31, #18 and #21. In this group, the connection from spike train #30 to #32 and spike train #32 to #30 have the highest strength among other connections. The second group consists of 4 spike trains, these are: spike trains #1, #25, #26 and #22. The third group consists of 10 spike trains; these are: spike trains #2, #14, #19, #20, #23, #24, #27, #28, #6 and #10. In this group, four connections (#19, #20), (#20, #19), (#27, #28) and (#28, #27) have big strength compare to others. Spike train #20 and #28 have 5 outgoing connections to other spike trains of this group. As these spike trains have the highest outgoing connections, these spike trains can be considered as the most influential spike trains for this group. The forth group consists of 8 spike trains, these are: spike trains #3, #7, #15, #11, #8, #9, #12 and #16. In this group there are 2 big strength of connections: (#8, #9) and (#16, #9). Spike train #8 has 6 outgoing connections to other spike trains and can be considered as the most influential spike train of this group. Spike trains #13 and #17 do not follow any group.

Investigation from Fig. 6.6(a)-(b) reveal that the Cox metric identified by the pair-wise analysis and considering all spike trains at once show a good agreement. For example, the first group of both figures consists the same four spike trains (#30, #32, #31 and #21) except one (#18) in Fig. 6.6(b). Similarly, the other three groups of Fig. 6.6(a) and Fig. 6.6(b) consist same spike trains except a few. From this analysis it can be said that application of the Cox metric to this experimental data using pair-wise analysis and analysis of all spike trains at once enables to create similar result.

6.4.4 Motif analysis

To find the patterns of interconnections among the 29 spike trains, a structural motif analysis is conducted using triplet-wise analysis of Cox method. For 29 spike trains, a total of 3654 triplets are analysed. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval using (5.6). Functional connectivity of each triplet spike trains is used to identify the structural motif. The structural motif count is obtained by analysing all 3654 triplets of spike trains which is shown in Fig. 6.7.

Out of 3654 triplets, 753 triplets have different structural motif ID's. Among the 753 triplets, 160 triplets have motif ID 2 which is the highest among other motif ID's. Only 6 triplets have motif ID 7 which is the lowest. Motif ID's 1, 3, 4 and 6 have similar number of triplets and motif ID's 8, 9, 10, 11 and 12 have similar number of triplets. A total of 96 triplets have connected motifs (connected motifs are motif ID 7, 9, 10, 12 and 13). On the other hand, 657 triplets have unconnected motifs. Thus, there are low

proportions of connected motifs (14.61%) in the groups of triplet spike trains which indicate that connection from every spike train to every other spike trains is very low.

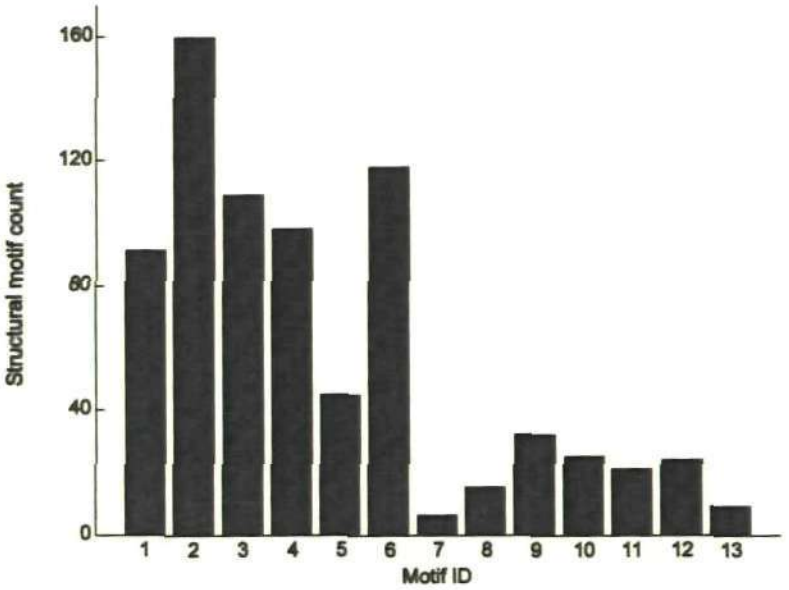


Figure 6.7: *Structural motif count of all 3654 triplets of the 29 spike trains of stimulus* 1.

The same procedure is applied to the other five stimuli to analyse functional connectivity of the spike trains. The results of analysing the five stimuli are presented in Appendix B.

6.5 Summary of functional connectivity of all stimuli

In this chapter functional connectivity of 29 spike trains are analysed for six different stimuli. To analyse the functional connectivity the MCG method and the Cox method are used. In stimulus 1, 85 direct connections are identified by the MCG method (Fig. 6.3(b)). Investigation from this diagram reveals that spike train #8 and #28 have the highest outgoing connections to other spike trains. That means these spike trains are the most influential spike trains. Application of the Cox method to this stimulus reveals 71 connections (Fig. 6.5(a)). Investigation from this diagram of functional connectivity reveals that similar to MCG method, spike trains #8 and #28 have the highest outgoing

connections. Importantly, spike train #8 has 7 out going connections to spike trains #3, #7, #9, #11, #12 #15 and #28 which are the same in both the MCG method and the Cox method. Also, spike train #28 has 4 outgoing connections to spike trains #19, #20, #23 and #26 are same in both the MCG method and the Cox method. From the analysis, it can be concluded that spike trains #8 and #28 are the most influential spike trains in stimulus 1.

Application of MCG method to stimulus 2 reveals 129 direct connections (Fig. A.2(b)). From this diagram it is observed that spike trains #28 and #32 have the highest outgoing connections and can be considered as the most influential spike trains. These two spike trains are also the most influential spike trains identified by the Cox method. The diagram of functional connectivity identified by the Cox method revels 62 connections (Fig. A.4(a)). Spike trains #28 and #32 have the highest 8 outgoing connections to other spike trains. Like stimulus 1, spike train #28 has 7 outgoing connections to spike trains #2, #23, #24, #25, #26, #27 and #32 which are same in both the MCG method and the Cox method. Also, all the outgoing connections from spike train #32 to spike trains #19, #20, #24, #25, #26, #27, #30 and #31 are same in both the MCG method and the Cox method.

In stimulus 3, a small number of 16 direct connections are identified by the MCG method (Fig. A.8(b)). Here spike train #8 has only 3 outgoing connections and can be considered as the most influential spike trains. A different result is obtained from the diagram of functional connectivity identified by the Cox method (Fig. A.10(a)) where 95 connections are identified. Investigation reveals that similar to stimuli 1 and 2, spike train #32 has the highest outgoing connections. This spike train can be considered as the most influential spike train.

Diagram of functional connectivity identified by the MCG method reveals that there are 51 direct connections in stimulus 4 (Fig. A.14(b)). Investigation shows that spike train #32 has the highest outgoing connections and can be considered as the most influential spike train. The similar result is identified by the Cox method. The diagram of functional connectivity (Fig. A.16(a)) shows that there are 71 connections and spike train #32 has the highest outgoing connections. Investigation shows that spike train #32 has 4 outgoing connections to spike trains #20, #25, #27 and #30 which are the same in both the MCG method and the Cox method.

Like stimulus 3, in stimulus 5 a different result is obtained from the diagram of functional connectivity identified by the MCG method (Fig. A.20(b)) and by the Cox method (Fig. A.22(a)). Only 49 connections are identified by the MCG method where spike train #23 has the highest outgoing connections. This spike train can be considered as the most influential spike train. On the other hand, 116 connections are identified by the Cox method. Here, spike train #28 has the highest outgoing connections and can be considered as the most influential spike train.

In stimulus 6, the diagram of functional connectivity identified by the MCG method reveals that there are 154 direct connections (Fig. A.26(b)). Investigation shows that spike train #32 has the highest outgoing connections and can be considered as the most influential spike train. The similar result is identified by the Cox method. The diagram of functional connectivity (Fig. A.28(a)) shows that there are 76 connections and spike train #32 has the highest outgoing connections. Investigation shows that spike train #32 has 8 outgoing connections to spike trains #9, #19, #23, #25, #27, #28, #30 and #31 which are same in the MCG method and the Cox method.

From this analysis it can be concluded that spike trains #28 and #32 are the most influential spike trains among the 29 spike trains, also the most of the information are

transmitted through these spike trains. Though the results of identification of most influential spike train are different in stimuli 3 and 5 by the MCG method and the Cox method, the results are same for other stimuli. Thus, there is a good agreement of these two methods.

The Cox metric is applied to identify the groups of similar spike trains. Importantly, for all stimuli the Cox metric based on pair-wise analysis and the analysis considering all spike trains at once show similar results. Investigation from the Cox metric based on the pair-wise analysis reveal that there are some groups of spike trains which are same in different stimuli. In stimulus 1, the forth group (Fig. 6.6(a)) consists of 9 spike trains (#1, #3, #7, #8, #9, #11, #12, #15 and #16). This group of spike trains is very similar to the first group of stimulus 2 (Fig. A.5(a)) and the second group of stimulus 3 (Fig. A.11(a)). In stimulus 2, the first group consists of 10 spike trains (#3, #6, #7, #8, #9, #11, #12, 14, #15 and #16) where 8 spike trains are same to fourth group of stimulus 1. In stimulus 3, the second group consists of 7 spike trains (#1, #3, #7, #9, #11, #15 and #16) where all the spike trains are same to fourth group of stimulus 1. Similarly, in stimulus 2, the third group consists of 11 spike trains (#19, #20, #23, #24, #25, #26, #27, #28, #30, #31 and #32) which is very similar to the first group of stimulus 3 and the third group of stimulus 6 (Fig. A.29(a)). In stimulus 3, the first group consists of 12 spike trains (#14, #19, #20, #21, #23, #24, #25, #26, #27, #28, #30 and #32) where 10 spike trains are same to third group of stimulus 2. In stimulus 6, the second group consists of 13 spike trains (#18, #19, #20, #21, #23, #24, #25, #26, #27, #28, #30, #31 and #32) where 11 spike trains are same to third group of stimulus 2 and 10 spike trains are same to first group of stimulus 3. Fourth group of stimulus 5 (Fig. A.23(a)) consists spike train #13 and #16 which is exactly same to the fourth group of stimulus 6.

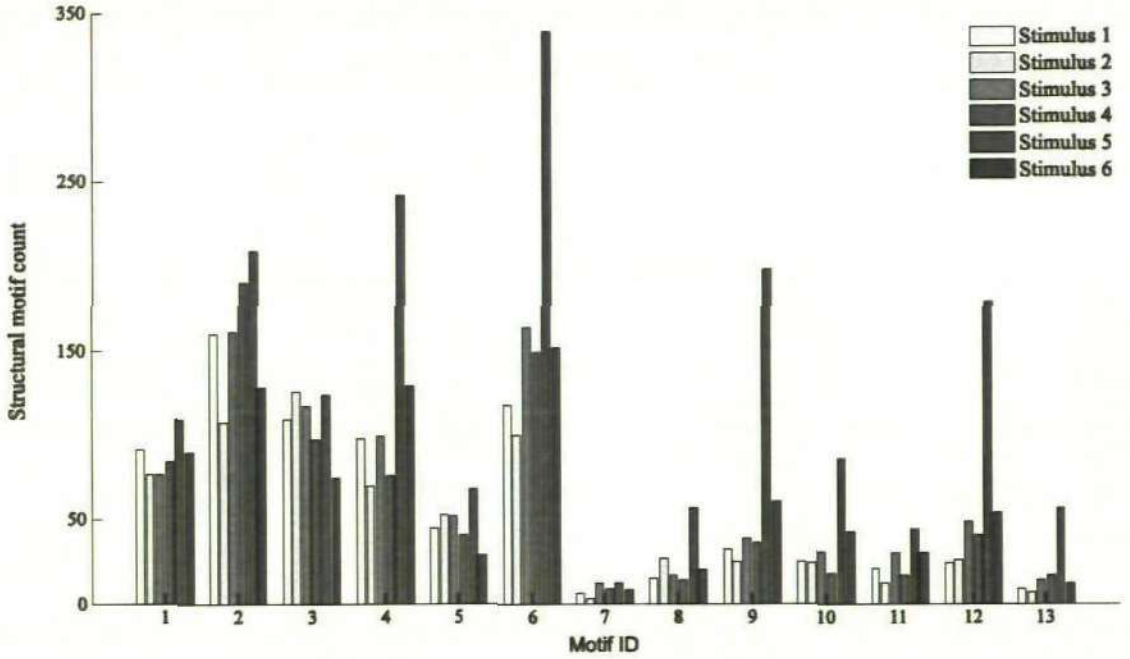


Figure 6.8: Structural motif count of all 6 stimuli.

Patterns of interconnections among the 29 spike trains are identified by the structural motif analysis. Structural motif analysis for all stimuli shows similar results. Among all stimuli, stimulus 5 has the highest structural motif count for all motif ID's except motif ID 3. Structural motif count (Fig. 6.8) shows that motif ID's 2 and 6 have the highest structural motif count whereas motif ID 7 has the lowest structural motif count for all stimuli. Motif ID 1, 3 and 4 have the similar structural motif count for all stimuli. Structural motif count for connected motifs (motif ID 7, 9, 10, 12 and 13) are less than that for the unconnected motif which indicates that the spike trains are weakly connected to each other for all stimuli.

Functional connectivity of all stimuli obtained by the MCG method and the Cox method shows a good agreement. Among the stimuli, stimulus 6 has the highest number of connections which are common to the MCG method and the Cox method (Fig. 6.9). That is these connections exist both in the MCG method and the Cox method. In this stimulus, 76 connections are identified by the Cox method and 154 connections are identified by the MCG method which is the highest among all stimuli. There are 57

connections which are common to the both methods. On the other hand, stimulus 3 has the lowest connections which are common to both the methods. In this stimulus, 95 connections are identified by the Cox method and only 16 connections are identified by the MCG method. A total of 14 connections are identified which are common to both methods. Stimuli 4 and 5 have the same number of 31 connections identified by the both methods.

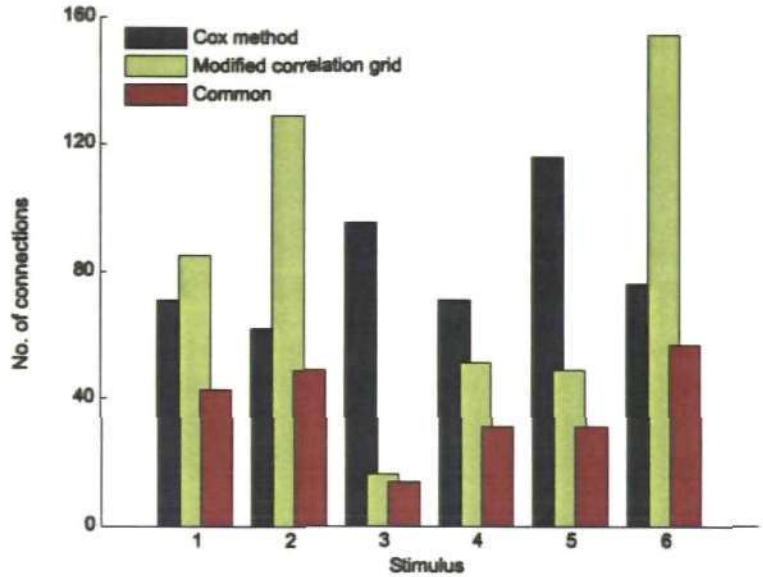


Figure 6.9: Total number of connections identified by the Cox method and the MCG method with the number of connections common to these methods.

In stimulus 4, 71 connections are identified by the Cox method where as 51 connections are identified by the MCG method. In stimulus 5, the highest number of 116 connections is identified by the Cox method; where as 49 connections are identified by the MCG method which is similar to the stimulus 4. In stimulus 1, 71 connections are identified by the Cox method whereas 85 connections are identified by the MCG method. There are 43 connections which are common to the MCG method and the Cox method. Similar to stimulus 1, 49 connections are identified in stimulus 2 which are common to both methods. Here 62 connections are identified by the Cox method and a large number of 129 connections are identified by the MCG method.

Although there is a good agreement between the MCG method and the Cox method, there are some differences between them. The reason is that, the MCG method is based on the pair-wise CCF and the Cox method is based on the analysis of all spike trains at once. The MCG method takes short computation time and can be used as a screening method to derive functional connectivity. Among the six stimuli, stimulus 3 shows the highest difference between these two methods. On the other hand, stimulus 1 and 4 show low differences between these two methods.

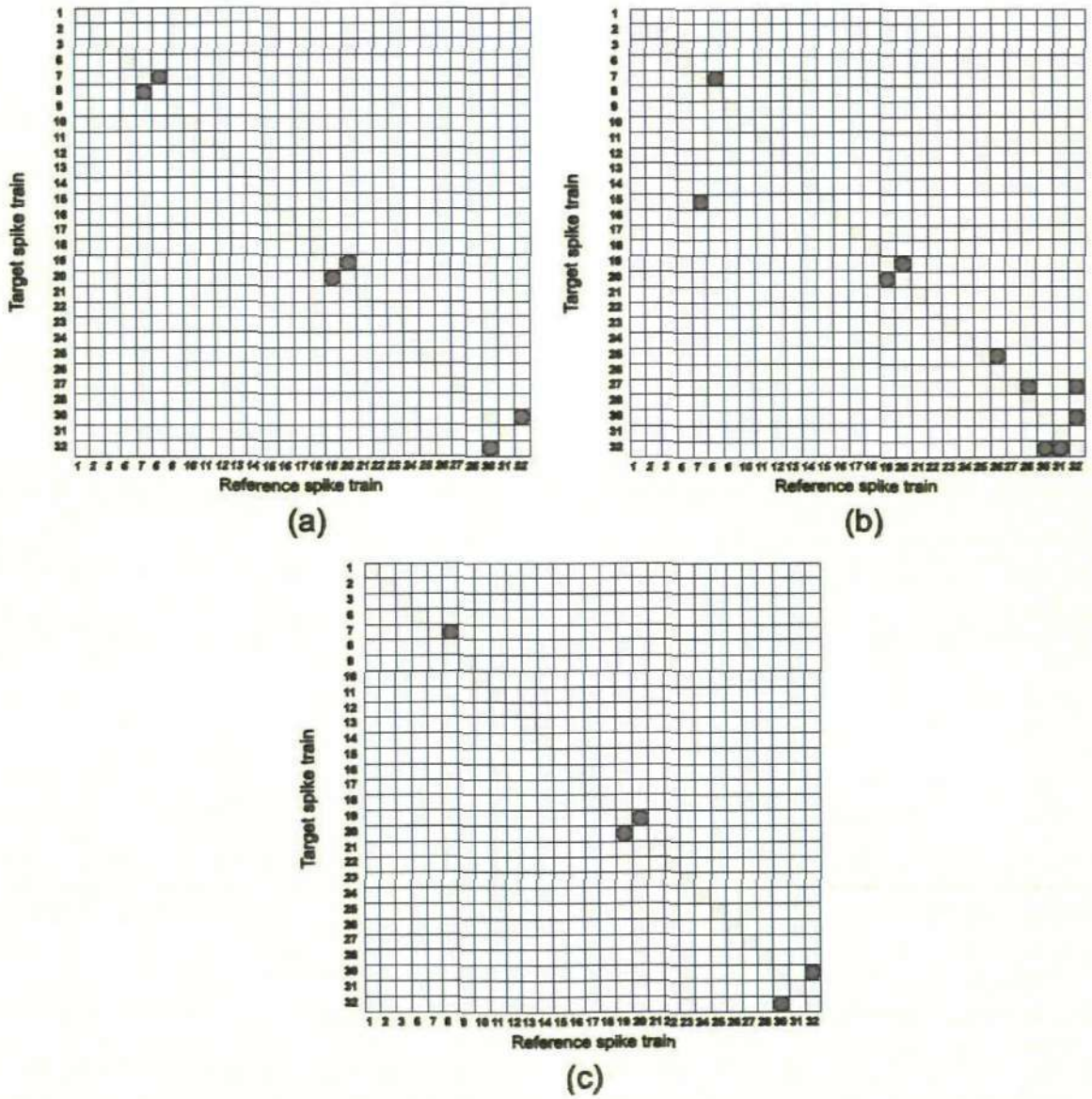


Figure 6.10: (a) Common connections identified by the MCG method for all 6 stimuli. **(b)** Common connections identified by the Cox method for all 6 stimuli. **(c)** Common connections identified both by the Cox method and the MCG method for all 6 stimuli.

Fig. 6.10 shows the common connections identified by the Cox method and the MCG method for all stimuli. Total 6 common connections are identified by the MCG method for all stimuli (Fig. 6.10(a)). These connections are: (#7, #8), (#8, #7), (#19, #20), (#20, #19), (#30, #32) and (#32, #30). Investigation of these connections from each stimulus shows that these connections have strong strength. There are 10 common connections identified by the Cox method for all stimuli (Fig. 6.10(b)). These connections are: (#7, #15), (#8, #7), (19, #20), (#20, #19), (#26, #25), (#28, #27), (30, #32), (#31, #32), (32, #30) and (#32, #31). Like the MCG method, investigation of these connections shows that the connections have strong strength. There are 5 connections identified both by the Cox method and the MCG method which are common to all stimuli (Fig. 6.10(c)). These connections are: (#8, #7), (#19, #20), (#20, #19), (#30, #32) and (#32, #30).

Chapter 7

Methods of graph theory for analysing the connectivity

This chapter presents different graph theory methods that are useful for the comprehensive analysis of connectivity of multiple spike trains. The connectivity matrix is calculated from the experimental data of cat's visual cortex (Nikolic, 2007; Schneider et al., 2006) by the Cox method. The graph theory methods are then applied to analyse the matrix of connections. This experimental data are organized in six sets of spike trains. Therefore, the statistical technique provides six graphs. To analyse the structure of a graph and to compare the graphs, some measures from the graph theory are calculated.

7.1 Introduction

Due to the recent advances in neuroscience and neuroinformatics, an increasing number of neuronal connectivity datasets in the brain areas are available for analysis. The availability of such data sets requires the development of appropriate computational tools for their comprehensive analysis (Kotter, 2001). That is the connection density of the spike trains, the average number of steps required to pass information from one spike train to another, the attractive and influential spike trains, the degree of the spike trains to make cluster with other spike trains, spike trains which pass the most of the information to other spike trains and the significant patterns of interconnections of the spike trains. The methods discussed in chapter 3 and 5 are used to derive functional connectivity of multiple spike trains. In addition to this functional connectivity, it is necessary to study the comprehensive analysis of the connectivity. One way for such a

comprehensive analysis is provided by graph theory, a branch of mathematics which has many applications in diverse fields such as physics, communication science, genetics, linguistics and sociology (Sporns, 2002).

The graph theory methods are based on the connection matrix. In neuroscience this connection matrix can be derived from the analysis of simultaneously recorded multiple spike trains. Most of the connection matrices are derived by the pair-wise analysis of spike trains using cross-correlation. This technique produces symmetrical connection matrix and undirected graph (Bullmore and Sporns, 2009) while functional connectivity of spike trains are characterized as directed graphs (Hilgetag et al., 2002). The Cox method described in chapter 5 can be used for deriving the asymmetrical connection matrix and the directed graph. The distinction between undirected and directed graphs is especially important as different graph measures are computed slightly differently for these two major classes of graphs.

There are several graph theory methods (Rubinov and Sporns, 2010) that are of special relevance to the comprehensive analysis of connectivity. In this chapter, some of them are discussed for the directed graph such as: graph density, nodes degree, characteristic path length, efficiency of a graph, clustering coefficient and betweenness centrality. Motif analysis is another useful graph method which is discussed in section 5.9. Another important method of graph theory is the P1 model (Holland and Leinhardt, 1981), which is used in social science network to find the influential and attractive people in the network. Application of P1 model to the neuroscience is useful in finding the influential and attractive neurons in the brain regions. All these graph methods are applied to the connection matrices obtained by the Cox method from the experimental datasets discussed in chapter 6.

7.2 Graph theory methods

A graph is a mathematical representation of a system that is composed of interconnected elements, comprising a set of nodes and edges. Nodes represent the fundamental elements of the system, such as neurons in the brain region and edges represent connections between pairs of nodes. Edges can be undirected or directed from origin to destination. Here directed edges are considered and all the graph theory methods are discussed for the directed graph. In graph theoretical terms, a directed graph G_{nl} is composed of n nodes and l edges, with l ranging from 0 (null graph) to $n^2 - n$ (complete or fully connected graph excluding self connections). The graph's adjacency matrix (connection matrix), $A(G)$, is composed of binary entries a_{ij} (Fig. 7.1(b)), with $a_{ij} = 1$ indicating the presence of connection from node i to node j, and $a_{ij} = 0$ indicating the absence of connection from node i to node j. The diagonal elements of adjacency matrix a_{ii} are considered as zero and a_{ij} does not necessarily equal to a_{ji}.

7.2.1 Density

The density k_{den} of an adjacency matrix $A(G)$ is the number of all its non-zero entries, divided by the maximal possible number of connections. The density ranges from 0 to 1, 0 indicates null graph and 1 indicates fully connected graph. In the neural network the highest levels of connection density are found at the level of cortical areas and the pathways interconnecting them. Matrices of connection pathways linking cortical areas tend to have $k_{den} \sim 0.2 - 0.4$ (Sporns, 2002).

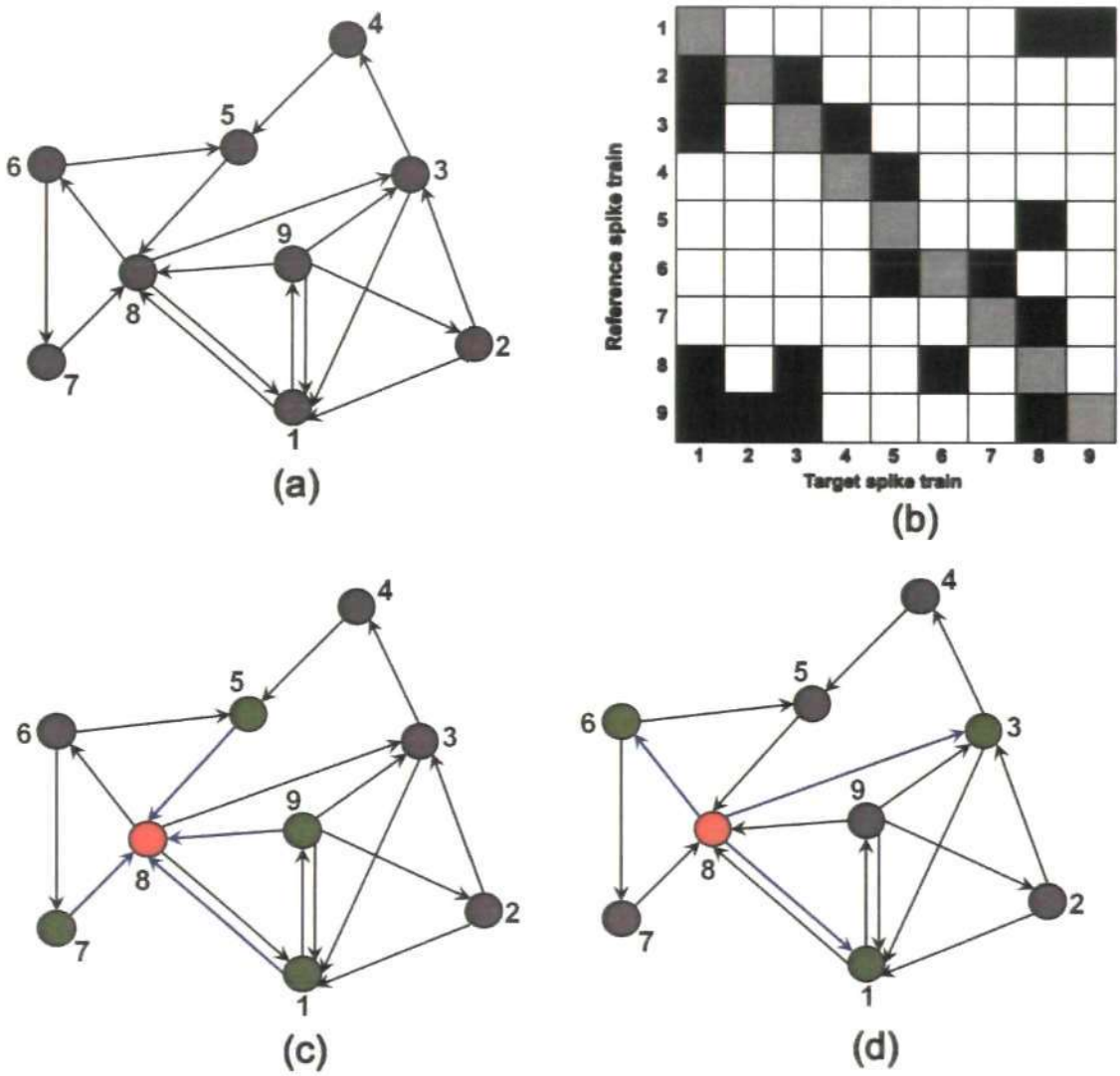


Figure 7.1: (a) Directed graph composed of 9 nodes and 18 directed edges. The graph has 72 ($9^2 - 9$) possible connections among the nodes. The density of this directed graph is 18/72 =0.25. **(b)** Adjacency matrix represents the presence (black square) and absence (white square) of the connections between the nodes. Main diagonals are indicated in grey and self-connections are excluded. **(c)** Indegree of node #8 (orange circle). This node has 4 indegree, from nodes #1, #9, #5 and #7 (green circles). **(d)** Outdegree of node #8 (orange circle). This node has 3 outdegree, to nodes #1, #3 and #6 (green circles).

7.2.2 Degree

The adjacency matrix allows the derivation of one of the most fundamental graph measures, the degree. In directed graph the indegree and outdegree corresponds to the number of incoming and outgoing edges, respectively (Fig. 7.1(c-d)). A node with high

indegree is influenced by many other nodes, while a node with high outdegree has many potential functional targets. The indegree and outdegree of a node i can be calculated as

$$k_i^{in} = \sum_{j \in N} a_{ji}$$
$$k_i^{out} = \sum_{j \in N} a_{ij}$$

where N is the set of all nodes in the network.

7.2.3 Characteristic path length

Nodes can be linked directly by single edges or indirectly by sequences of intermediate nodes and edges. Ordered sequences of unique edges and intermediate nodes are called paths (Fig 7.2). If a finite path between two nodes exists, then one node can be reached by traversing a sequence of edges starting at the other node. In directed graph, the length of a path is equal to the number of edges it contains. Paths of various lengths record possible ways by which signals can travel indirectly between two nodes. Longer paths are likely to have less of an effect than shorter paths. Most analyses focus on shortest possible paths (distances) between nodes since these paths are likely to be most effective for inter node communication. The directed distance from node i to node j, (d_{ij}) is the length (number of edges) of a shortest path (possible one of several) from node i to node j.

The distance between two nodes is often of particular interest. The structure of the adjacency and distance matrices (Fig. 7.3(b)) together describe the pattern of communication within the nodes of the graph.

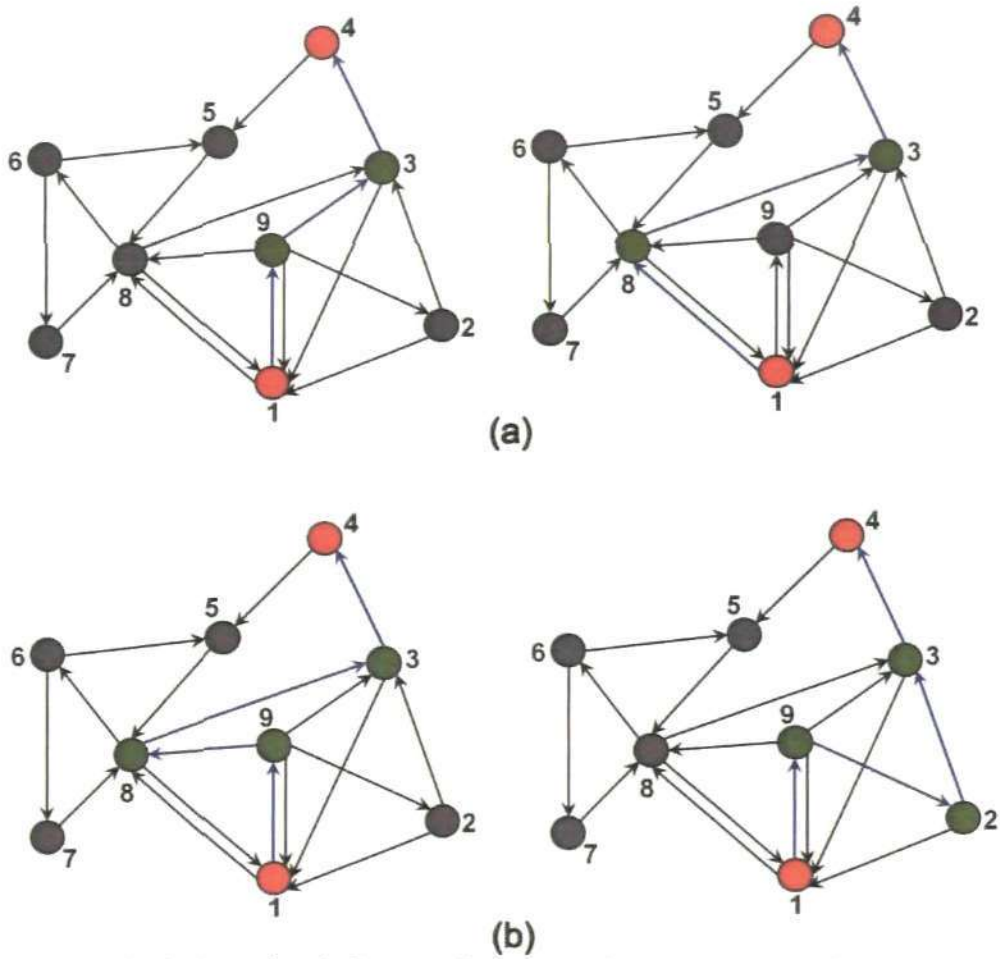


Figure 7.2: Calculation of path from node #1 to node #4 (orange circles). **(a)** Path from #1 to #4 of length 3, denoted by {#1, #9, #3, #4} nodes of green circle, containing the directed edges (blue) (#1, #9), (#9, #3) and (#3, #4). An alternative path of the same length 3 is denoted by {#1, #8, #3, #4} nodes of green circle. **(b)** Path from #1 to #4 of length 4, denoted by {#1, #9, #8, #3, #4} nodes of green circle, containing the directed edges (blue) (#1, #9), (#9, #8), (#8, #3) and (#3, #4). An alternative path of the same length 4 is denoted by {#1, #9, #2, #3, #4} nodes of green circle. The shortest possible path length from node #1 to #4 is 3, hence the distance from node #1 to node #4 is 3.

One of the most commonly used measures in the brain network is the characteristic path length. This is computed as the global average of the graph's distance matrix (Watts and Strogatz, 1998). The characteristic path length is calculated as

$$L = \frac{1}{n}\sum_{i\epsilon N}\frac{\sum_{j\epsilon N, j\neq i} d_{ij}}{n-1}$$

The characteristic path length is a global characteristic; it indicates how well integrated a graph is, and how easy it is to transport information to other entities in the network.

Generally, shorter paths are thought to be more effective in passing information. Thus the average path length for a network can provide an indication of its capacity for global information exchange.

7.2.4 Efficiency

A related to characteristic path length and often more robust method, the global efficiency (Latora and Marchiori, 2001), is computed as the average of the inverse of the distance matrix. A fully connected graph has maximal global efficiency since all distances are equal to one (all pairs of nodes are linked by an edge); while a fully disconnected graph has minimal global efficiency since all distance between nodes are infinite. A high efficiency indicates that pairs of nodes, on average, have short communication distances and can be reached in a few steps. The efficiency of a graph should be compared to the efficiency of a random network keeping the same indegree and outdegree of the nodes. The global efficiency is calculated as

$$E = \frac{1}{n}\sum_{i \in N} \frac{\sum_{j \in N, j \neq i} (d_{ij})^{-1}}{n-1}$$

7.2.5 *Clustering coefficient*

A clustering coefficient is a measure of degree to which nodes in a graph tend to cluster together. It is one of the most elementary measures of local segregation of the network (Watts and strogatz, 1998). There are two versions of this measure, the local and the global clustering. The local clustering coefficient of an individual node measures the density of connections between the node's neighbours. Neighbours are those nodes that are connected, either through an incoming or outgoing connection, to the node (Fig. 7.3(a)). Densely interconnected neighbours form a cluster around the node, while sparsely interconnected neighbours do not. Clustering of a node is high if the node's neighbour's are also neighbour's of each other. The average of the clustering

coefficients for each individual node is the clustering coefficient of the graph known as global clustering coefficient.

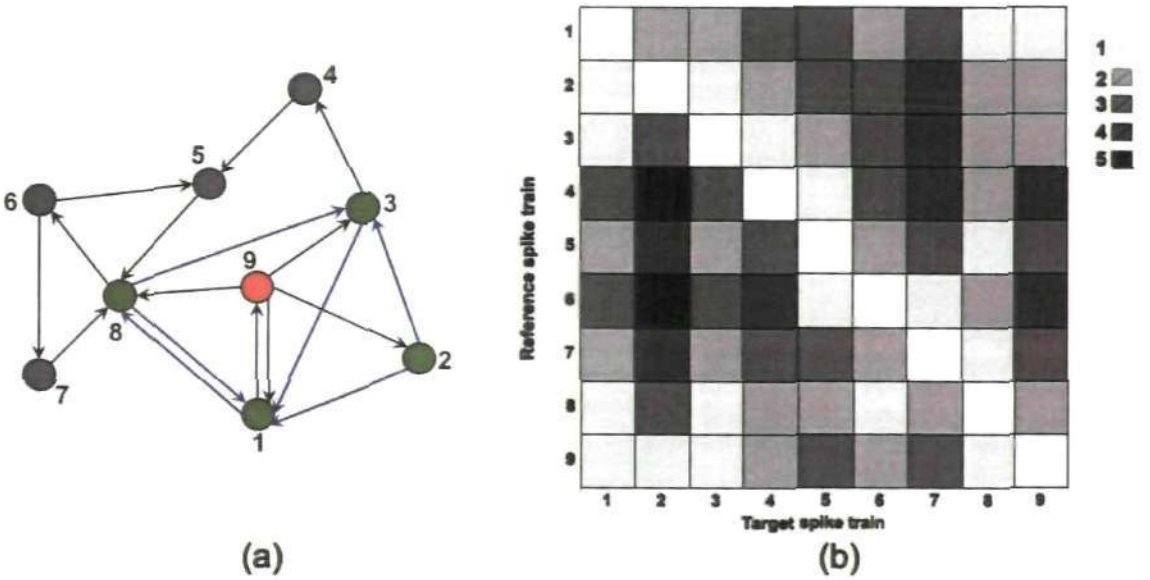


Figure 7.3: (a) Clustering coefficient of node #9 (orange circle). This node's neighbours are #1, #2, #3 and #8 (green circle), which maintain 6 connections (blue edges) among them out of 12 possible $(4^2 - 4)$. Thus the clustering coefficient of this node is 6/12=0.5. **(b)** Distance matrix of the 9 node, indicates the shortest path from node i $(1,2\cdots,9)$ to node $(1,2\cdots,9)$ $i \neq j$. Pairwise distances are integers ranging from 1 to a maximum of 5.

The clustering coefficient C_i of a node i with indegree k_i^{in} and outdegree k_i^{out} is usually calculated as (Fagiolo, 2007)

$$C_i = \frac{\frac{1}{2}\sum_{\substack{j,h\epsilon N \\ h\neq j, h\neq i, j\neq i}} (a_{ij} + a_{ji})(a_{ih} + a_{hi})(a_{jh} + a_{hj})}{(k_i^{out} + k_i^{in})(k_i^{out} + k_i^{in} - 1) - 2\sum_{j\epsilon N} a_{ij}a_{ji}}$$

The clustering coefficient C_i ranges between 0 and 1. Usually C_i is averaged over all vertices to obtain a mean C of the graph

$$C = \frac{1}{n}\sum_{i\epsilon N} C_i$$

7.2.6 Betweenness centrality

Centrality measures of a graph determine the relative importance of a node within the graph. Measures of centrality are based on the notion of the shortest paths. Betweenness centrality of a node is a useful measure to know how much information passes through certain nodes of the graph. A node with high betweenness centrality can control information flow because it is at the intersection of many shortest paths. The betweenness centrality of an individual node is defined as the fraction of all shortest paths in the network that pass through the node. The betweenness centrality of a node i is calculated as (Freeman, 1978)

$$C_i^B = \frac{1}{(n-1)(n-2)} \sum_{\substack{h,j \in N \\ h \neq j, h \neq i, j \neq i}} \frac{\rho_{hj}(i)}{\rho_{hj}}$$

where ρ_{hj} is the number of shortest paths between h and j, and $\rho_{hj}(i)$ is the number of shortest paths between h and j that pass through i.

7.2.7 The P1 model

The P1 model (Holland and Leinhardt, 1981) of a graph determines the relationships between nodes of the graph. For any pair of nodes in a graph, there are three possible relationships between them: no ties (no edges in either direction between the nodes), an asymmetric tie (an edge between the two nodes going in one direction for the other but not both), or a mutual tie (two edges between the nodes, one going in one direction an the other going in the opposite direction). These relationships are known as dyadic relationships and is denoted by

$$D_{ij} = (A_{ij}, A_{ji}), i \neq j$$

where A is the adjacency matrix.

A mutual relationship between node i and node j exists when $i \rightarrow j$ and $j \rightarrow i$ in the dyad which is denoted by $i \leftrightarrow j$. A mutual relationship is obvious when both the (i,j) and (j,i) cells are unity; that is $A_{ij} = 1$ and $A_{ji} = 1$, so that the dyad $D_{ij} = (1,1)$. The asymmetric dyad can occur in two ways, either $i \rightarrow j$ or $j \rightarrow i$, but not both. Specifically, $D_{ij} = (1,0)$ or $D_{ij} = (0,1)$. In null dyad the (i,j) and (j,i) symmetrically placed off-diagonal cells of A are both 0; that is, $A_{ij} = A_{ji} = 0$, implying that $D_{ij} = (0,0)$. Thus the three possible dyadic relationships can be represented as

$$D_{ij} = \begin{cases} (0,0) \\ (1,0) \text{ or } (0,1) \\ (1,1) \end{cases}$$

A dyad with measurements on a directional relation consists of two nodes, i and j, and the possible ties between these two nodes. The ties between the nodes are viewed from the perspective of either node i or node j. From the perspective of i the relational variable A_{ij} records the possible 'choice' of j by i, while the relational variable A_{ji} records the possible 'choice' received by i from j. From the perspective of node j the relational variable A_{ij} records the possible choice of node i by node j, while the relational variable A_{ij} records the possible choice received by node j from node i.

For a pair of nodes the ties in the dyad for both nodes is represented in a 2×2 array. There are two variables of this array. The first variable is indexed with an k, which can be either 0 or 1, codes the value of the tie sent by the row node i to the column node j. The second variable, also with just two levels is indexed with an l, codes the value of the tie sent by the column node j to the row node i. So, the ties for each and every dyad can be presented in one of these 2×2 arrays. The new indices k and l equal to either 0 or 1, depending on the state of the dyad.

Considering all dyads and the single dichotomous relation, there will be $g \times g$ binary matrix for g nodes. If each entry is replaced with the appropriate 2×2 table, a new contingency table is obtained of size $g \times g \times 2 \times 2$. The first two dimensions of this table are indexed by the nodes. The size of the third and fourth dimensions is 2, are coded $k, l = 0$ or 1.

The $g \times g \times 2 \times 2$ matrix is denoted by Y, and its entries are defined as follows:

$$Y_{ijkl} = 1 \text{ if the dyad } D_{ij} \text{ takes on the values } (A_{ij} = k, A_{ji} = l)$$

$$= 0 \text{ otherwise}$$

The Y-array is a cross-classification of four variables and thus, its entries have four subscripts: the nodes as senders (i), the nodes as receivers (j), and the relational variables A_{ij} (indexed by the third subscript, k) and A_{ji} (indexed by the fourth subscript, l). The (i, j) th cell of Y is not a single quantity, but rather a 2×2 submatrix. In this 2×2 submatrix, there is a single 1 found in the (k, l) th cell. The remaining 2^2 - 1 elements are 0. Thus, one can view these submatrices as simply indicator matrices, giving the 'state' of each dyad.

To understand the Y matrix an example of two nodes is given. The matrix in Table 7.1 represents the friendship of the nodes. The data show that node 2 does not name node 1 as a friend he likes, but node 1 nominates node 2.

Node	1	2
1	-	1
2	0	-

Table 7.1: Friendship of two nodes.

From 1's perspective, the relational variable sent is $A_{12} = 1$, implying that node 1 likes node 2 as a friend, and the relation received is $A_{21} = 0$, implying that node 1 is not liked as a friend by node 2. From 2's perspective, the relation sent is $A_{21} = 0$, node 2 does not choose node 1, and the relation received is $A_{12} = 1$, node 2 is chosen by node 1. The recorded data for nodes 1 and 2 in this pair would be $D_{12} = (A_{12}, A_{21}) = (1,0)$, so that $y_{1210} = 1$, while $y_{1200} = y_{1201} = y_{1211} = 0$. Similarly, $D_{21} = (A_{21}, A_{12}) = (0,10)$, so that $y_{2101} = 1$, while $y_{2100} = y_{2101} = y_{2111} = 0$. Now the Y array is presented

		j				
i		$l = A_{ji}$				
	$k = A_{ij}$		0	1	0	1
		0	-	-	0	0
		1	-	-	1	0
		0	0	1	-	-
		1	0	0	-	-

Table 7.2: Y matrix for the friendship of two nodes.

The P1 model is presented by a 4-dimentional Y-array. For a single, directional relation, the effects that represent the 'expansiveness' of nodes, the 'popularity' of their partners, and the 'reciprocation' of the ties within the dyads are focused. The P1 model consists primarily of three sets of parameters: one set of parameters describes the nodes' sending behaviour, one set describes the nodes' receiving behaviour, and one set describes the interactions between pairs of nodes within a dyad. The first set of parameters are called expansiveness effects which reflect the tendency of each node to nominate others as friends. The second set of parameters are called popularity effects which reflect the tendency for a node to be nominated by others as friends. Positive values of these parameters increase the probability of having ties. The final set of parameters are those

that reflect the reciprocation, or mutuality, between two nodes, independent of the expansiveness or popularity of either node. Reciprocity is the extent to which a dyad exhibits mutual, as opposed to asymmetric ties. Positive reciprocity parameters increase the likelihood that the dyad is mutual.

The P1 model is expressed in four statements. Each of the four statements represents one of the four possible states of any given dyad: the null dyad ($A_{ij} = A_{ji} = 0, or\ Y_{ij00} = 1$), the mutual dyad ($A_{ij} = A_{ji} = 1, or\ Y_{ij11} = 1$), and two cases of asymmetric dyads ($A_{ij} = 1, A_{ji} = 0, or\ Y_{ij10} = 1, and\ A_{ij} = 0, A_{ji} = 1, or\ Y_{ij01} = 1$). In order to specify P1, the natural logarithm of the probabilities of each of these four dyadic states is represented as a function of several parameters:

$$\begin{aligned} logP(Y_{ij00} = 1) &= \lambda_{ij} \\ logP(Y_{ij10} = 1) &= \lambda_{ij} + \theta + \alpha_i + \beta_j \\ logP(Y_{ij01} = 1) &= \lambda_{ij} + \theta + \alpha_j + \beta_i \\ logP(Y_{ij11} = 1) &= \lambda_{ij} + 2\theta + \alpha_i + \alpha_j + \beta_i + \beta_j + (\alpha\beta) \end{aligned} \tag{7.1}$$

The $\{\lambda_{ij}\}$ parameters are mathematical necessities included in the model to ensure these four probabilities sum to one for each dyad. Thus these parameters appear in all four statements, regardless of the state of the dyad. The θ parameter is interpreted as an overall choice effect (analogous to a grand mean), reflecting the overall volume of choices sent and received. If one tie is reciprocated, two θ's appear. Note that, θ does not appear in the model statement when ties are not present, and $(\alpha\beta)$ is present only when the dyad is mutual. No substantive parameters appear in the first statement of the model which represents a null dyad. For asymmetric dyads, the log probabilities depend on parameters reflecting only one of the two possible ties in the dyad: dyads in which node i chooses node j without reciprocation (so an α_i but not an α_j is relevant, and a β_j

but not a β_i is included) and dyads in which node j chooses node i with no reciprocated choice (so the relevant parameters are α_j and β_i, but not α_i or β_j or $(\alpha\beta)$). All the parameters appear together only for mutual dyads (the last statement of the model). The $(\alpha\beta)$ (some times denoted by ρ), is also called a mutuality parameter. The parameter will be positive and large when the relation tends to be mutual.

In the equations (7.1), the α parameters are interpreted as expansiveness measures for each node. If α is positive and large of the corresponding node, it may be said that there is a high probability that the node will influence the other nodes. The β parameters are interpreted as attractiveness measures. If β is positive and large of the corresponding node, it can be said that there is a high probability that the node is influenced by other nodes. The parameters are estimated using the principle of maximum likelihood method. To estimate the parameters the package *UCINET IV* is used. In this chapter, only expansiveness and attractiveness parameters are considered for finding the relationships of the nodes.

7.3 Analysis of connectivity of stimulus 1

The connection matrix (Fig. 7.4(a)) of 29 spike trains shows only 71 connections out of 812 possible connections. The connection matrix contains a low density (0.0874) which means that under the stimulus 1 the spike trains are not densely connected. The sum of indegree and outdegree known as degree is shown in Fig. 7.4(b). The degree of the spike trains varies widely from 0 to 11 showing the same number of degrees for certain spike trains. For example, spike trains #3, #8, #19, #24 and #32 have 8 degrees each. On the other hand spike train 17 has no degree. Some spike trains have very few connections known as low-degree spike trains (#13, #18, #21, #22 and #31); whereas some spike trains have large connections known as high-degree spike trains (#28, #20

and #12). A high degree spike train is a spike train whose degree is greater than the mean plus one standard deviation of all the spike trains (Sporns et al., 2007).

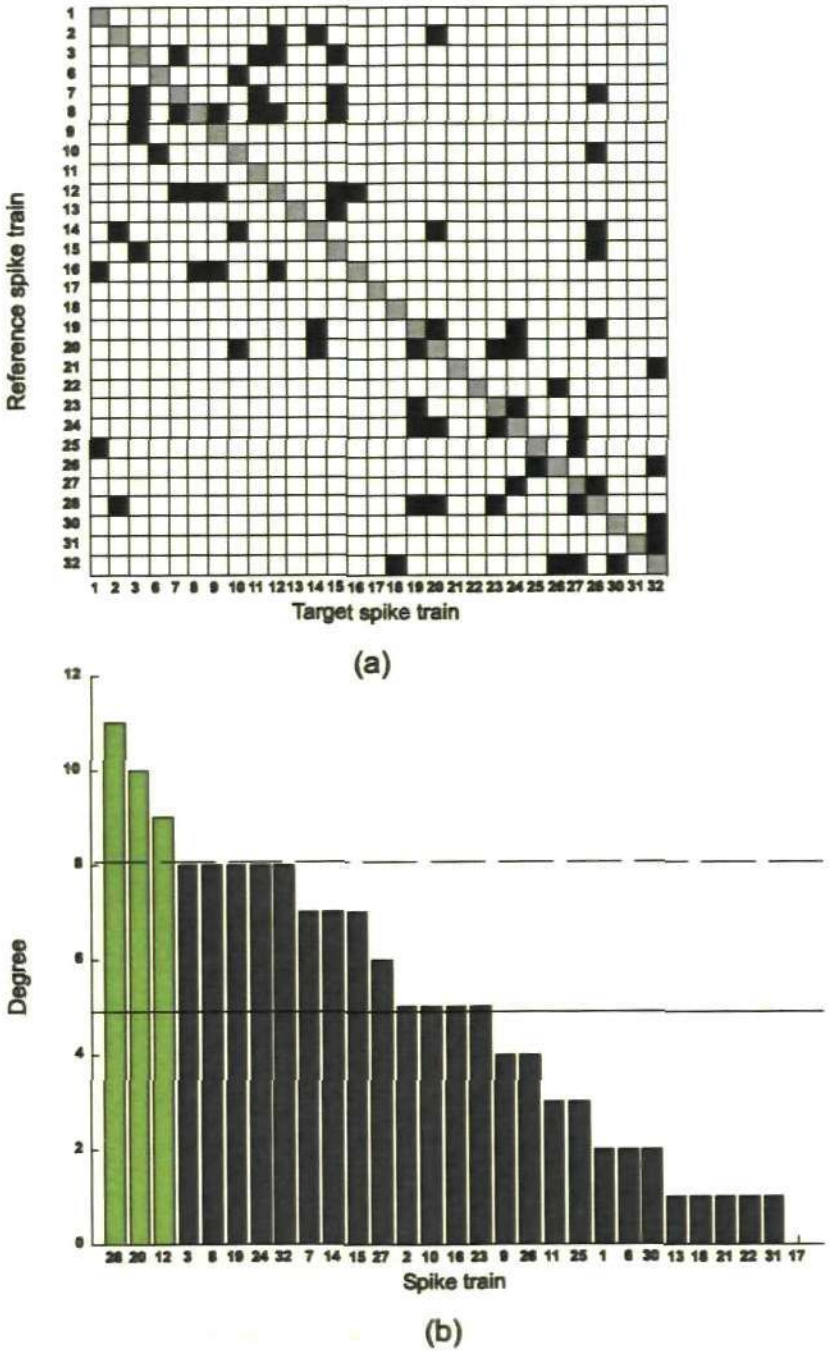


Figure 7.4: **(a)** Connection matrix of the 29 spike trains of stimulus 1. Connection patterns are represented by the presence of connection (black square) and absence of connection (white square). Main diagonals are indicated in grey and self-connections are excluded. **(b)** Degree of the spike trains is displayed in descending order. The solid horizontal line indicates the mean degree of the spike trains and the dashed horizontal line indicates the mean plus one standard deviation of the spike trains. High-degree spike trains are displayed as green.

The characteristic path length of the connection matrix (3.1634) is greater than the characteristic path length obtained from a random network (2.6411). This characteristic path length indicates that, on average, to pass information from one spike train to another spike train, it takes approximately 3 edges. Similarly, the global efficiency of the connection matrix (0.2456) (random: 0.3040) indicates that pairs of spike trains, on average, have long communication distances.

There is a variation of the clustering coefficient of the spike trains (Fig. 7.5(a)) ranging from 0 to 0.72. Some spike trains have high clustering coefficient (#23, #11, #9 and #19) indicating that the neighbours of these spike trains are also neighbours of each other. There are some spike trains which have a low clustering coefficient (#10 and #28), in fact, below the mean of all spike trains. Among the low clustering coefficients, the spike train #28 has the highest degree of all the spike trains which indicates that this spike train communicate to other neighbour spike trains but the neighbours are not connected to each other. There are 13 spike trains that do not form any cluster to their neighbour spike trains. The clustering coefficient of these spike trains is zero. The global clustering coefficient (0.2276) (random: 0.1068) also indicates that many spike trains do not have neighbours which are connected to each other. Fig. 7.5(b) shows the betweenness centrality of the spike trains. There are some central spike trains (#28, #12, #2 and #27) which transfer most of the information to the other spike trains. Among the central spike trains, spike trains #28 and #12 have the highest degree. This means that these two spike trains communicate to other spike trains through incoming and outgoing connections. There are 10 spike trains which do not pass any information to other spike trains. That means the betweenness centrality of these spike trains is zero.

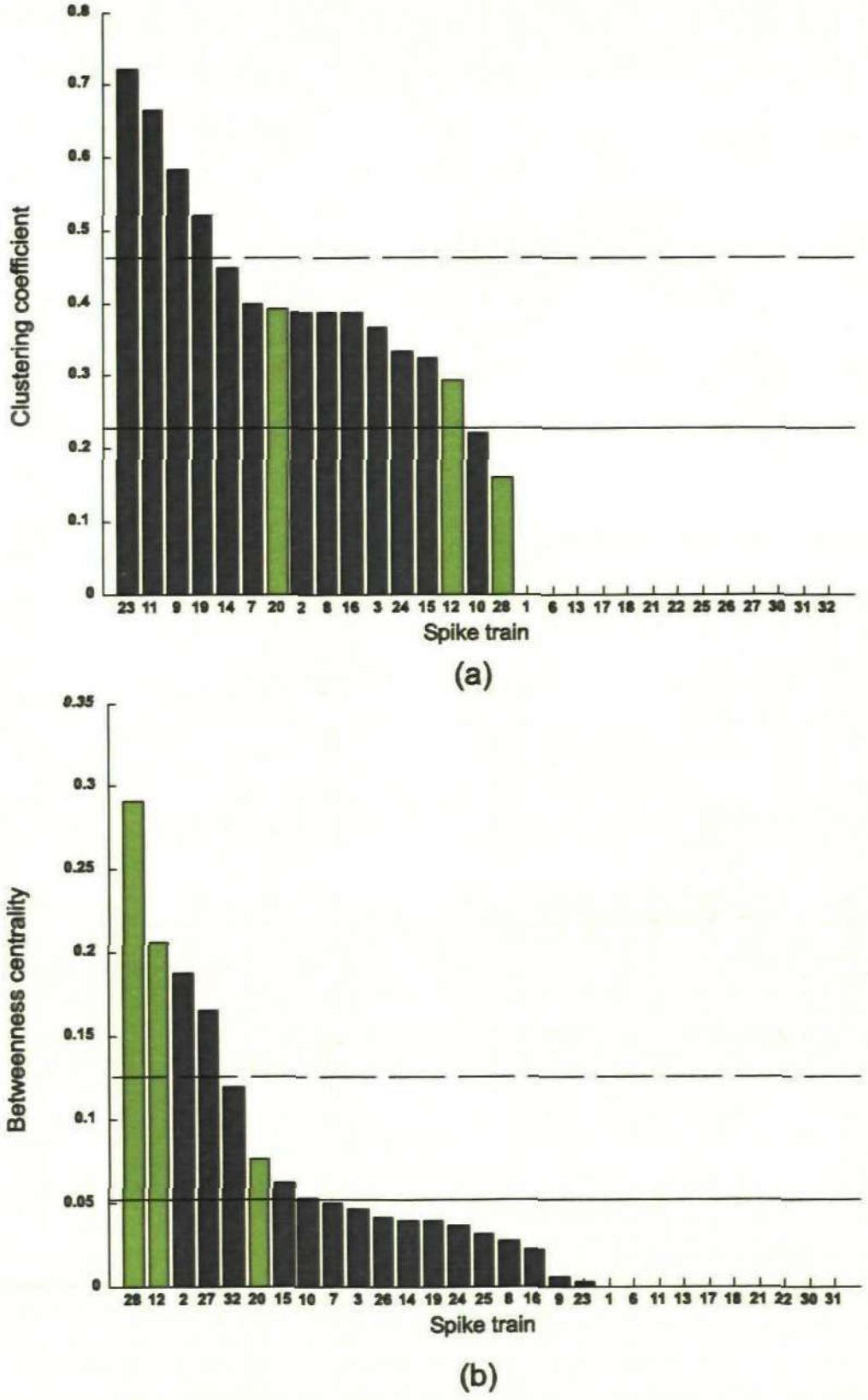


Figure 7.5: Clustering coefficient and betweenness centrality of the 29 spike trains of stimulus 1. The solid horizontal line indicates the mean and the dashed horizontal line indicates the mean plus one standard deviation. High-degree spike trains are displayed as green. **(a)** Clustering coefficient of 29 spike trains is displayed in descending order. **(b)** Betweenness centrality of the 29 spike trains is displayed in descending order.

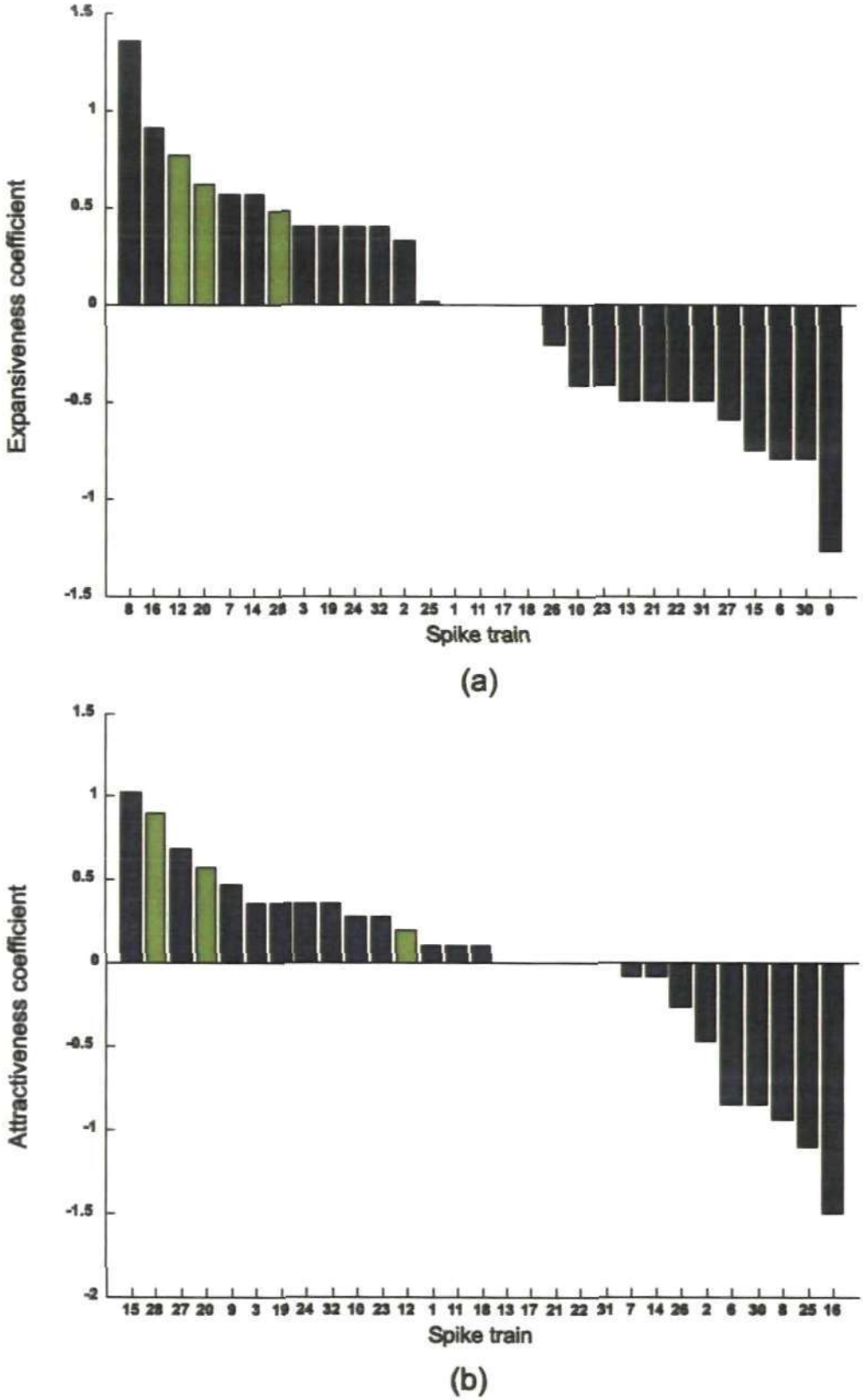


Figure 7.6: Expansiveness and attractiveness coefficient of the P1 model of the 29 spike trains of stimulus 1. High-degree spike trains are displayed as green. **(a)** Expansiveness coefficient displayed in descending order. **(b)** Attractiveness coefficient displayed in descending order.

Among the 29 spike trains, spike trains #8, #16, #12 and #20 are the most influential (Fig. 7.6(a)) spike trains. Investigation from the connection matrix reveals that these spike trains have the high outdegree (6, 4, 5 and 5 respectively). Among the influential spike trains there are two high degree spike trains (#12 and #20). Some of the spike trains (#1, #11, #17 and #18) do not show any expansiveness because their outdegrees are zero. Negative expansiveness coefficient indicates that the indegree is more than the outdegree. Spike train #9 has the most negative expansiveness coefficient. Investigation from the connection matrix reveals that this spike train has 3 indegree and 1 outdegree. Spike trains #15, #28, #27 and #20 are the most attractive spike trains (Fig. 7.6(b)) as they have high indegree (5, 6, 4, and 5 respectively) and two of them have the high degree (#28 and #20). Similar to expansiveness coefficient some of the spike trains (#13, #17, #21, #22 and #31) do not show any attractiveness as their indegrees are zero. Negative attractiveness coefficient indicates that the outdegree is more than the indegree. Spike train #16 has the most negative attractiveness coefficient (4 outdegree and 1 indegree). Spike train #20 has the high indegree and outdegree and simultaneously considered as influential and attractive spike train.

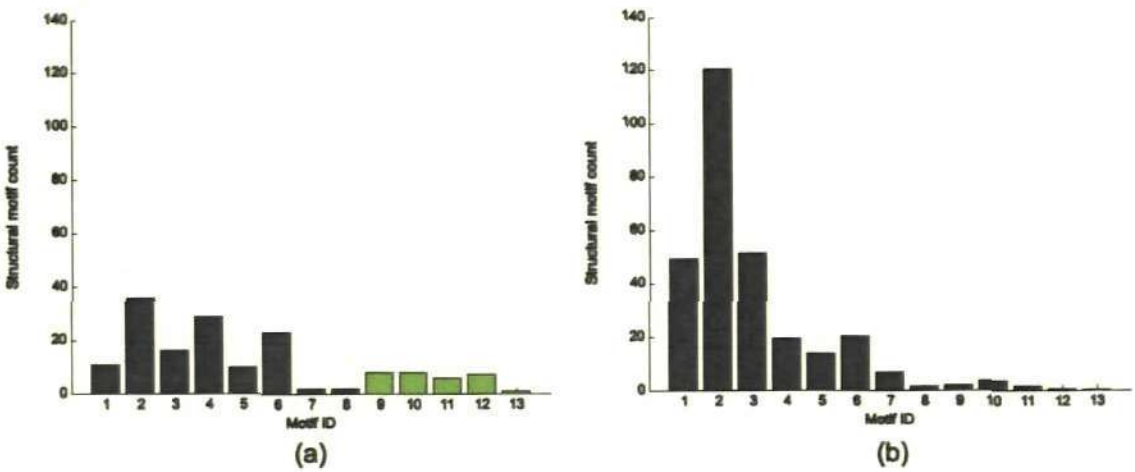


Figure 7.7: (a) Structural motif count of size m = 3 of the 29 spike trains of stimulus 1. Significant motif ID's are displayed as green. **(b)** Structural motif count of size m = 3 for the randomized diagram.

To find the significant interconnections of the spike trains a motif analysis is applied to the connection matrix. Fig. 7.7(a) shows the structural motif count of size m = 3 found in the connection matrix of 29 spike trains. Motif ID 2 appears 36 times which is the highest among the motif IDs. Motif ID 13 appears only 1 time in the connection matrix and it is the lowest. To find the significant motif, 1000 random networks are generated keeping the same indegree and outdegree of the spike trains. The structural motif count of size m = 3 for the random network is shown in Fig. 7.7(b). The motif ID's 9, 10, 11, 12 and 13 appear more than the random network. The Z-score of these motif ID's ($Z_9 = 3.66,\ p = .0005;\ Z_{10} = 2.60,\ p = .013;\ Z_{11} = 3.97,\ p = .0001;\ Z_{12} = 8.10,\ p < .0001; Z_{13} = 6.51,\ p < .0001$) indicate that they are significant. There are a low proportion of connected motifs (25.20%) in the connection matrix indicating that the spike trains are not strongly connected.

The same procedure is applied to the other five stimuli to analyse connectivity of the spike trains. The results of analysing the five stimuli are presented in Appendix C.

7.4 Summary of connectivity of all stimuli

Comprehensive analysis of connectivity of all stimuli shows that all the connection matrices have low density ranging from 0.0764 to 0.1429 (Table 7.3). It is observed that stimulus 5 has the highest density. There are two stimuli (1 and 4) which have the same density. The characteristic path length of all stimuli shows that on an average, pairs of spike trains have long communication distances. The lowest characteristic path length is observed in stimulus 5 which indicates that in this stimulus, the communication distances between spike trains are less than all other stimuli. A related measure to characteristic path length, the global efficiency also shows that on an average, there is a long communication distances between spike trains. Like characteristic path length, stimulus 5 has the highest global efficiency indicating lowest communication distances

among the six stimuli. The global clustering coefficient indicates that most of the spike trains in all stimuli do not form any cluster to their neighbouring spike trains and hence the spike trains are not strongly connected to each other. In stimulus 3, the global clustering coefficient is the highest which indicates that in this stimulus spike trains form some clusters to their neighbouring spike trains.

Stimulus	Density	Characteristic path length	Global efficiency	Global clustering coefficient
1	0.0874	3.1634	0.2456	0.2276
2	0.0764	2.7747	0.1871	0.2213
3	0.1170	2.8137	0.3428	0.3749
4	0.0874	3.1204	0.2278	0.2408
5	0.1429	2.3377	0.4041	0.2715
6	0.0936	3.0560	0.2939	0.2139

Table 7.3: Four graph theory measures for six stimuli.

The degree distribution of individual spike trains shows that some spike trains have high degrees which are common to different stimuli (Fig. 7.8). Spike train #32 has the highest degree in the case of five stimuli except stimulus 1. Similarly, spike train #28 has high degree in the cases of three stimuli (stimulus 1, stimulus 2 and stimulus 5) and spike train #12, #24 and #27 have high degree in the case of two stimuli. Spike trains #3, #9, #14, #20, #21 and #30 have the high degree in the case of only one stimulus.

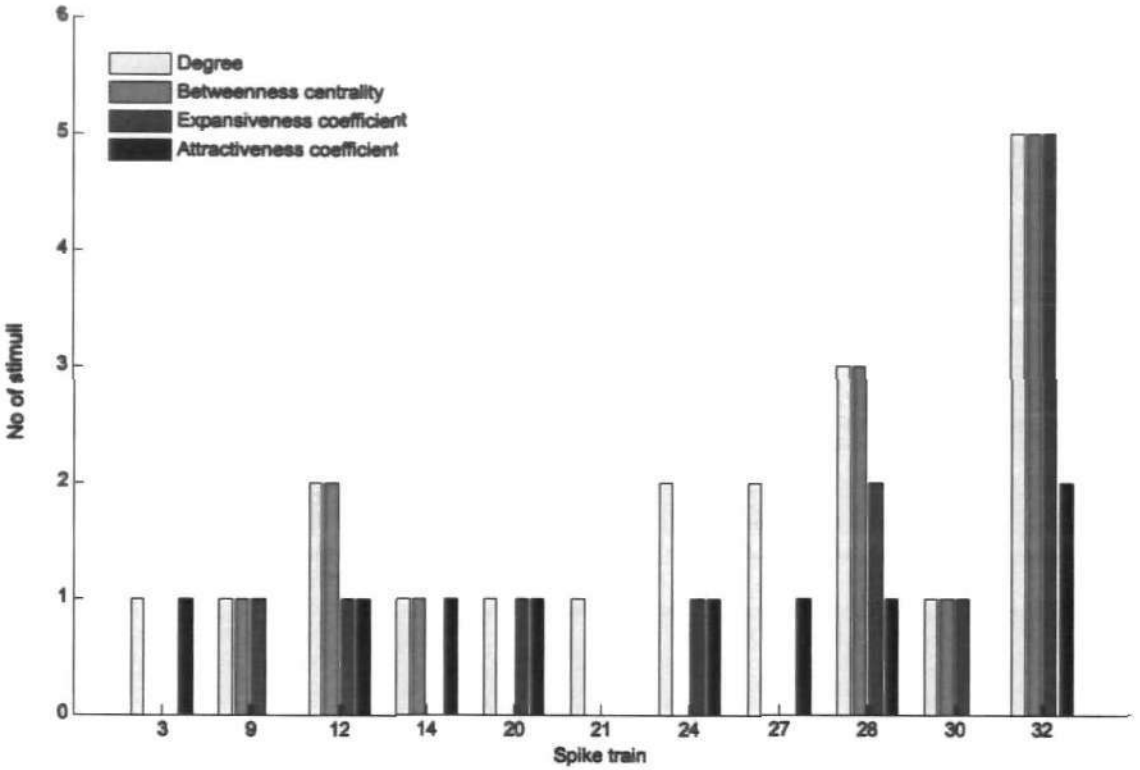


Figure 7.8: Spike trains which have high degree, high betweenness centrality, high expansiveness coefficient and high attractive coefficient among all the spike trains in different stimuli.

Investigation from the clustering coefficients reveals that all the high degree spike trains have low clustering coefficient. Some spike trains have clustering coefficient below the mean of all the spike trains. On the other hand, all these high degree spike trains have high betweenness centrality. Similar to high degree, spike train #32 has the high betweenness centrality in five stimuli except stimulus 1 and spike train #28 has high betweenness centrality in three stimuli (stimulus 1, stimulus 2 and stimulus 5). Spike train #12 has high betweenness centrality in two stimuli which are stimulus 1 and stimulus 5. These high degree spike trains have also high expansiveness coefficient and high attractiveness coefficient. Spike train #32 has high expansiveness coefficient in five stimuli and high attractiveness coefficient in two stimuli. Spike train #28 has high expansiveness coefficient in two stimuli (stimulus 2 and stimulus 5) and high attractiveness coefficient in stimulus 1.

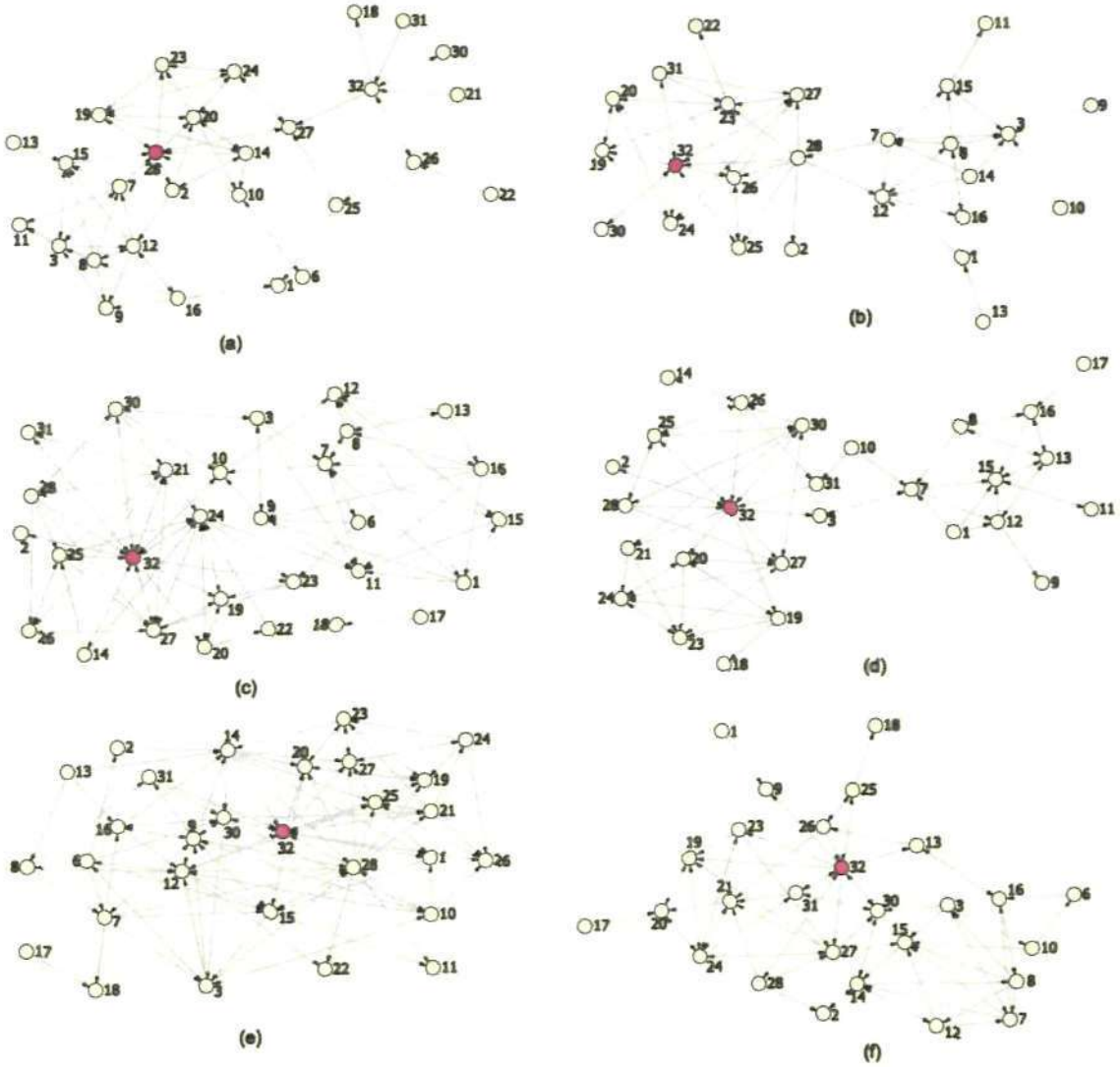


Figure 7.9: Hub spike train in six stimuli. The hub spike train is shaded in magenta colour. **(a)** Spike train #28 is the hub spike train in stimulus 1. **(b)** Spike train #32 is the hub spike train in stimulus 2. **(c)** Spike train #32 is the hub spike train in stimulus 3 **(d)** Spike train #32 is the hub spike train in stimulus 4 **(e)** Spike train #32 is the hub spike train in stimulus 5 and **(f)** Spike train #32 is the hub spike train in stimulus 6.

An important property of the graph theory measures is the identification of hubs. In graph theory measures, highly influential nodes are often referred to as hubs. These hubs have the capacity to transfer or process information to the other nodes. Hubs can be identified either on the basis of the number of degrees or betweenness centrality (Sporns, 2010). From the analysis of high degrees and betweenness centrality, it can be concluded that there are some hub spike trains in all stimuli. Fig. 7.9 shows that there is a hub spike train in stimulus 1 which is spike train #28, and spike train #32 is the hub

spike train for all other stimuli. From the analysis it can be concluded that spike train #28 and #32 are the main spike trains which transfer information to the other spike trains and as well receive information from the other spike trains.

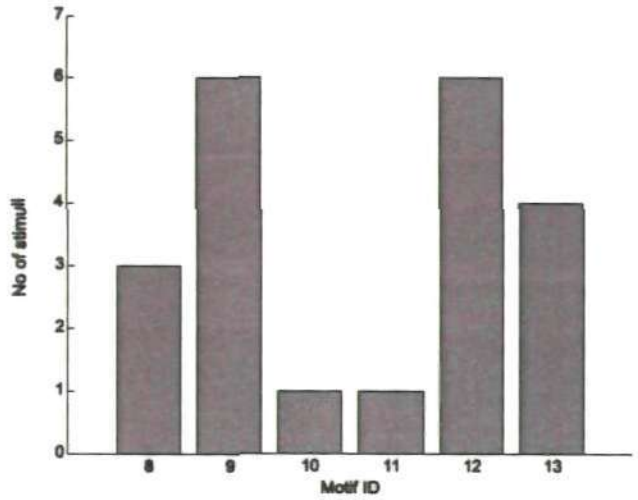


Figure 7.10: Significant motif ID in different stimuli.

Lastly, Fig. 7.10 shows the significant motif ID obtained from different stimuli. Out of 13 motifs ID, only 6 motifs ID are found significant. Motif ID's 9 and 12 are significant in all 6 stimuli. Motif ID 13 is significant in stimuli 1, 3, 4 and 6. Motif ID 8 is significant in 3 stimuli (stimulus 2, stimulus 3 and stimulus 5). Both motif ID's 10 and 11 are significant in stimulus 1.

Chapter 8

Contribution and conclusion

In this chapter the contributions of this thesis are highlighted. The main findings of the statistical methods are reiterated to highlight the significance of the methods for the analysis of functional connectivity.

8.1 Contribution

Here the major contributions of this thesis are summarized:

1. A new statistical method known as the modified correlation grid (MCG) is developed to analyse functional connectivity of multiple spike trains. The MCG method enables researchers to distinguish the direct connections from the spurious (common source and indirect connection) connections using an automatic algorithm.
2. Another statistical method known as the Cox method is developed to analyse functional connectivity of multiple spike trains. The original paper on the application of the Cox method (Borisyuk et al., 1985) to neuroscience data was used to analyse only pairs and triplets of spike trains.
3. A probabilistic model for the generation of dependent spike train is developed. In this probabilistic model an influence function is described which is used in neuroscience to study synaptic connectivity between neurons.
4. A technique based on the pair-wise analysis of the Cox method is developed. This technique is called the Cox metric. The Cox metric is used to identify the groups of similar spike trains among multiple spike trains.

5. Another technique based on the triplet-wise analysis of the Cox method is developed. This technique is used to identify the patterns of interconnections of multiple spike trains.
6. In social science network the P1 model is used to identify influential and attractive people. In this thesis this P1 model is used to the connection matrix of multiple spike trains to identify the influential and attractive spike trains.

8.2 Conclusion

The MCG, developed in this thesis, is a useful method to identify functional connectivity of multiple spike trains. In chapter 3, this method is applied to analyse functional connectivity generated by the ELIF model. To generate spike trains the architecture of connections between ELIF elements is prescribed and in the analysis of these spike trains it is assumed that connections are unknown. A diagram of functional connection is derived as a result of spike train analysis. This diagram of functional connection is compared with the connection of ELIF model which is used for spike train generation. Two examples are considered for comparison, a neural circuit of 15 spike trains and a neural circuit of 20 spike trains. The neural circuit of 15 spike trains has sixteen connections with one very strong connection known as outlier, and fifteen other medium strengths of connections. The neural circuit of 20 spike trains has twenty five connections where all the connections have medium strength.

Pair-wise analysis of the CCF of the 15 spike trains reveals 25 significant connections. Application of the MCG to these 25 significant connections shows 16 direct connections and 9 spurious (common source and indirect) connections. Comparing this diagram of functional connectivity with the diagram of connectivity of ELIF elements used for spike train generation reveals that all the direct connections identified by the MCG are the same as the prescribed connections of the ELIF model.

Similarly, the pair-wise analysis of the CCF of the 20 spike trains reveals 34 significant connections. Application of the MCG to these 34 significant connections shows 25 direct connections and 9 spurious (common source and indirect) connections. Comparing this diagram of functional connectivity with the diagram of connectivity of ELIF elements used for spike train generation reveals that all the direct connections identified by the MCG are the same as the prescribed connections of the ELIF model. Thus using the MCG it is possible to distinguish the direct connection from the spurious connections and can reveal the functional connectivity of the multiple spike trains.

In this thesis a statistical technique known as the Cox method is presented for analysing dependencies of point processes for application to neuroscience data. It is found that the Cox method is an efficient tool to study functional connectivity. Comparison with the CCF which is traditionally used in neuroscience shows significant advantages of the Cox method.

In chapter 5, the Cox method is applied to the multiple spike trains generated by the ELIF model to analyse functional connectivity. To generate spike trains the architecture of connections between ELIF elements is prescribed and in the analysis of spike trains it is assumed that connections are unknown. The diagram of functional connections which is obtained by the Cox method is compared with connections of ELIF model which is used for spike train generation. Two examples are considered for comparison of the result of functional connectivity of the Cox method and the connection of ELIF model. The result of analysing five spike trains shows that the Cox method accurately identifies all functional connections. The result of analysing a large neural circuit of twenty spike trains also accurately identifies all forty two connections but also finds two erroneous connections which are relatively weak.

Comparison with the CCF shows that the Cox method has advantages over the CCF technique. In particular, the Cox method is more accurate in difficult situations such as a weak strength or short spike trains. One important advantage of the Cox method is that this method allows analysing all simultaneously recorded spike trains. To demonstrate this advantage the Cox method is applied to analyse three spike trains coupled according to 'common source' scheme and shown that couplings can be correctly identified, but the pair-wise CCF fails to distinguish between the direct connection and the connection due to a common source. A similar example of three spike trains with 'indirect connection' also demonstrates an advantage of the Cox method over the CCF.

For comparison with CCF, a probabilistic model is used for generating two spike trains which satisfy the assumption of the MRP. In this case the estimated Cox coefficient equals the prescribed strength of influence for spike train generation whereas the CCF fails to show the proper strength of influence. In chapter 4, it is shown that the probabilistic MRP model can be fitted to a wide range of spike trains generated by the integrate-and-fire model. To analyse functional connectivity of multiple spike trains it is not known whether the analysed spike trains satisfy the assumption of MRP. However application of the Cox method to both probabilistic model and ELIF model shows that this method is robust and can be successfully used for finding functional connectivity for a wide range of point processes.

A new technique known as the Cox metric, based on the pair-wise analysis of the Cox method is developed in this thesis. This method is applied to a neural circuit of twenty spike trains generated by the ELIF model to identify the groups of coupled spike trains. The architecture of connections between ELIF elements is prescribed and it is assumed that connections are unknown. The groups of mutually coupled spike trains are derived

by the Cox metric and results are compared with connections of ELIF model which is used for generating groups of coupled spike trains. The result shows that the Cox metric accurately identifies the groups of coupled spike trains.

Another technique known as motif analysis, based on the triplet-wise analysis of the Cox method is developed to identify patterns of interconnections among the spike trains. Application of this method is presented for a neural circuit of twenty spike trains. In addition to this technique, motif analysis is conducted based on the connection matrix derived by the Cox method considering all spike trains at once. It is found that the patterns of interconnections obtained by the triplet-wise analysis of the Cox method are very similar to the patterns of interconnections obtained by the Cox method considering all spike trains at once.

8.3 Application of the methods to the experimental data

8.3.1 Analysis of functional connectivity

The MCG method and the Cox method are applied to a set of experimental data recorded from cat's visual cortex to identify functional connectivity. Experimental data includes six stimuli. For each stimulus a set of 29 spike trains are analysed to identify functional connectivity. A simple study is conducted to understand how the Cox method works for experimental data. Two small time intervals are considered for the analysis of functional connectivity of 29 spike trains. Analysis of functional connectivity of 29 spike trains for these intervals show very similar results, which indicate that Cox method is useful for analysis of experimental data. The MCG is not applied because the number of spikes is too small for the application.

For each stimulus the diagram of functional connectivity obtained by the MCG method is compared with the diagram of functional connectivity obtained by the Cox method.

Although there are differences in the functional connectivity, importantly many functional connections are common in both techniques. Application of Cox metric to the functional connectivity of each stimulus has identified some groups of spike trains which are common in different stimuli. The patterns of interconnections of the spike trains in different stimuli obtained by motif analysis are also similar.

8.3.2 Graph theoretical methods for analysing connectivity

Matrix of connectivity obtained in chapter 6 by the Cox method for each stimulus is further analysed by the graph theoretical methods in chapter 7. Importantly, the results of graph theoretical methods found in chapter 7 have similar characteristics as in previous studies. The average connection density of all the connection matrices is obtained 0.100; this result is similar to the result of macaque cortex done by Young (1993) where the density was 0.152. The average of the characteristic path length (2.877) is found similar to the characteristic path length of macaque cortex (2.312). The degree distribution of individual spike trains obtained in chapter 7 reveals that some spike trains have high degrees which are common to different stimuli. Importantly, these high degree spike trains have found low clustering coefficient even below the mean of all spike trains. The low clustering coefficient of the high degree spike trains is common in each stimulus. The similar result is obtained from a previous study (Sporns et al., 2007) where the high degree areas of macaque cortex and cat cortex show low clustering coefficient. Betweenness centrality of the spike trains obtained from different stimulus reveals that the high degree spike trains have high betweenness centrality. Importantly, all the high degree spike trains which are common to different stimuli also have high betweenness centrality. Previous study (Sporns et al., 2007) also shows that the high degree areas of macaque cortex and cat cortex have high betweenness centrality. Connection matrices of all five stimuli except stimulus 1 reveal that spike

train #32 has high degree and high betweenness centrality and this spike train can be considered as the hub spike train which transmits and receives information from other spike trains. Analysis of the matrices reveals that spike train #32 shows the high expansiveness coefficient which is common in five stimuli except stimulus 1. This spike train has also high attractiveness coefficient for stimuli 3 and 5. For all stimuli structural motif ID's 9 and 12 are found significant. Studies on the macaque cortex (Sporns and Koetter, 2004; Sporns et al., 2007) have also revealed that structural motif ID 9 significantly appear more than the random network.

Appendix A

A.1 Enhanced Leaky Integrate and Fire model

The description of the Enhanced Leaky Integrate and Fire (ELIF) model follows the paper by Borisyuk (2002). A discrete-time version of the model neuron is used with the time increment equal to 1 ms. The state of each neuron at the moment t is characterised by a threshold and the total potential which is the sum of postsynaptic potentials and the noise. If the value of the total potential has reached the threshold, the neuron generates a spike. The spike propagates to other neurons with a time delay. The diagram of connection should be defined as well as connection strengths, time delays, and time decays of postsynaptic potentials. When the spike reaches another neuron, the postsynaptic potential jumps up or down depending on whether the spike is from an excitatory or inhibitory neuron, respectively. The value of the connection strength controls the jump height. The postsynaptic potential exponentially decays to the resting potential if there are no incoming spikes. After spike generation, the neuron is unable to generate a spike during an absolute refractory period. When this period expires, the threshold gets the highest value and then exponentially decays to the asymptotic threshold value. This decay is used to model a relative refractory period. To model a spontaneous background activity, the random noise is added to the membrane potential. The amplitude of the noise exponentially decays with time and a normally distributed random variable with zero mean and a fixed variance is added to the noise at each time step. The noise is independent random process for each element. If the amplitude of noise is large enough, then the element can be spontaneously active even without influences from other neurons.

A.2 Dynamics of the ELIF model

The dynamics of the ELIF is governed by the following equations:

1. The threshold:

$$r(t+1) = (r_{max} - r_{\infty})exp\big(-(t - t_{sp})/\alpha_{th}\big) + r_{\infty}$$

where r_{max} is the maximum value of the threshold

r_{∞} is the asymptotic threshold value when $t \to \infty$

α_{th} is the threshold decay rate

t_{sp} is the last spike moment before t.

2. The post-synaptic potential for the input of the neuron:

$$PSP^j(t+1) = PSP^j(t)exp\big(-1/\alpha^j_{PSP}\big) + a,$$

$$a = \begin{cases} w^j, if\ t^j_{sp} + \tau^j = t+1 \\ 0, \qquad otherwise \end{cases}$$

where w^j is the connection strength, positive for the excitatory connection and negative for the inhibitory one

τ^j is the time delay

α^j_{PSP} is the *j*th neuron PSP decay rate

t^j_{sp} is the last spike moment of the *j*th neuron before t.

3. The noise:

$$N(t+1) = N(t)\exp(-1/\alpha_N) + \xi,\ \ \xi \in N(0,\sigma)$$

where α_N is the noise decay rate

ξ is a random variable with the normal distribution

4. The soma's membrane potential:

$$V(t+1) = V_{AHP}exp\big(-(t - t_{sp})/\alpha_V\big)$$

where V_{AHP} is the value of after spike hyperpolarization

α_V is the soma's membrane potential decay rate

t_{sp} is the last spike moment before t.

5. The total potential:

$$P(t+1) = \sum_j PSP^j(t+1) + N(t+1) + V(t+1) + I_{ext}(t+1)$$

where I_{ext} is the value of the external input.

6. Spike generation:

If $P(t+1) > r(t+1)$, then $t_{sp} = t+1$.

7. The absolute refractory period:

There is no spike generation for the time interval $t \in (t_{sp}, t_{sp} + ref)$.

A.3 Parameters for spike train generation by ELIF model

An ELIF model can be simulated using software from the following website: http://www.tech.plymouth.ac.uk/infovis. To run the simulation, the parameters of ELIF neurons and their coupling should be specified.

Neural parameters:

Neural parameters describe the parameters of each neuron. The parameters are:

a) Maximum value of the threshold, b) Threshold decay rate, c) Asymptotic threshold value, d) Amplitude of the noise (i.e. the standard deviation of the normally distributed random variable), e) Noise decay rate, f) Initial value of after spike hyperpolarisation, g) Soma's membrane potential decay rate, h) External input, i) Absolute refractory period, and j) Type of the neuron (0- non-pacemaker, 1 – pacemaker).

Connection parameters:

Connection parameters contain the parameters describing non-zero connections between neurons. The parameters are:

a) List of the numbers of those neurons which send their connections to the current neuron, b) Connection strengths for these connections (positive for excitatory connection and negative for inhibitory), c) Decay rates of postsynaptic potential for each connection respectively and d) Time lag of spike propagation for each incoming connection (ms).

Appendix B

In chapter 6 the MCG method and the Cox method are applied to analyse functional connectivity of 29 spike trains. The results of analysing the first stimulus are presented in chapter 6. In this appendix the results of analysing the five stimuli are presented.

B.1 Analysis of functional connectivity of stimulus 2

The raster plot of 32 spike trains is shown in Fig. B.1. Like stimulus 1, spike trains #4, #5 and #29 have high spiking rates compared to other spike trains and considered as outliers. These three spike trains are not considered in the analysis of functional connectivity. All other 29 spike trains have similar spiking pattern. Spiking rates of these spike trains are high over the time interval [34000 ms, 78000 ms]. Importantly, there are no spikes over the time interval [78000 ms, 84000 ms] of the 29 the spike trains except spike train #1. These 29 spike trains are analysed to identify functional connectivity keeping the original numeration of the spike trains.

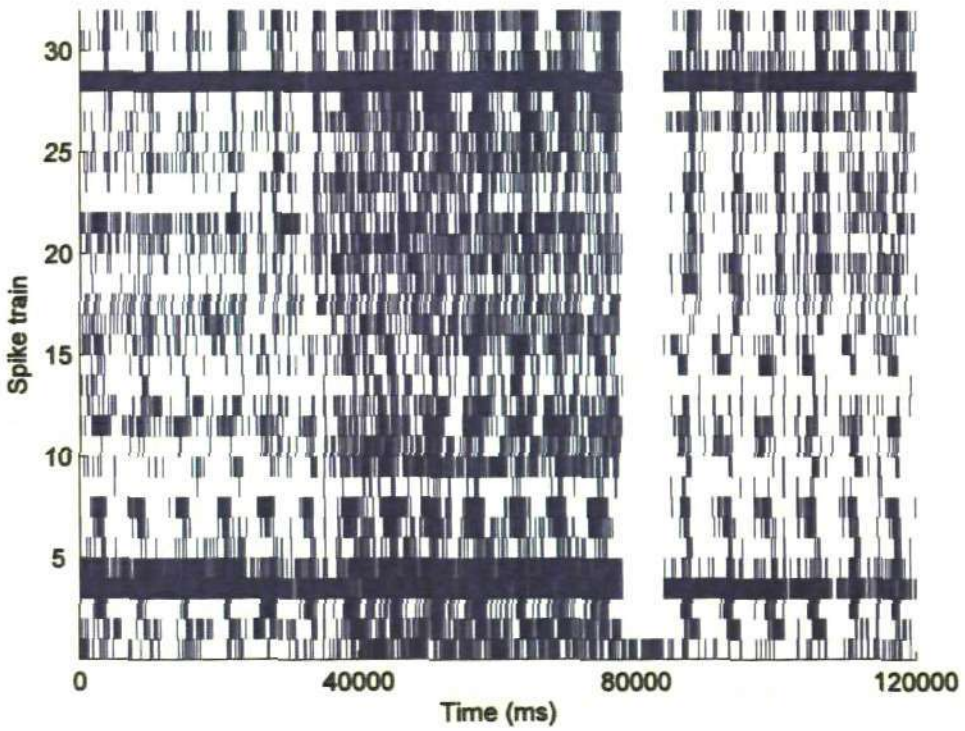


Figure B.1: Raster plot of 32 spike trains of stimulus 2. Spike trains #4, #5 and #29 have high spiking rates and are not considered for analysing functional connectivity.

B.1.1 MCG method

To identify functional connectivity, 406 pair-wise CCF are calculated with a bin size of 1 ms and a correlation window of 100 ms for the 29 spike trains. To test the independence of two trains the significance level $\alpha = 0.05$ is used with the Bonferroni correction. A connection is considered to be significant if a peak of the CCF exceeds the upper boundary of the 'confidence interval'. A total of 199 significant connections are found for 29 spike trains. These significant connections are shown in a matrix format in Fig. B.2(a) where the connections are indicated by circles. The direction of connection is considered from the reference spike train to the target spike train. Among the 29 spike trains, spike train #28 has 18 outgoing connections to other spike trains which is the highest among 29 spike trains and similar 16 incoming connections from other spike trains.

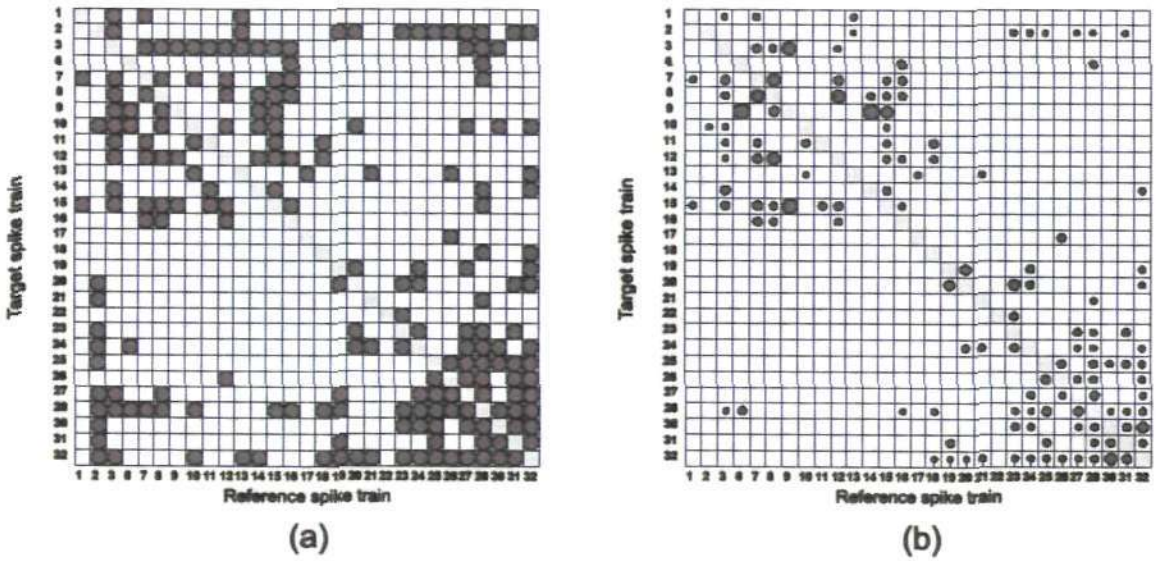


Figure B.2: (a) Significant connections obtained from pair-wise CCF analysis of the 29 spike trains of stimulus 2. **(b)** Direct connections obtained from the connections in (a) after the clustering algorithm. The radius of the circle indicates the strength of the connections.

To identify the direct connections from the 199 significant connections, a clustering algorithm is applied. After clustering, a total of 129 connections are identified as direct connections (Fig. B.2(b)). The radius of the circle indicates the strength of connection.

Among 129 connections, 9 connections have strong strength compared to others. These connections are: (#6, #9), (#7, #8), (#8, #7), (#8, #12), (#9, #3), (#9, #15), (#12, #8), (#14, #9) and (#15, #9). All other connections have medium strength of connection. Spike train #28 has 11 outgoing connections and spike train #32 has 12 incoming connections, which are the highest among 29 spike trains. There are 38 pairs of connections where both spike trains have functional connectivity to each other.

B.1.2 Cox method

To identify functional connectivity of the 29 spike trains using Cox method, one spike train is considered as the target spike train and other 28 spike trains are considered as the reference spike trains. The influence function (and its parameters) which determines how the reference spike train influences the target spike train should be specified. The inter spike interval (ISI) histogram of three spike trains, spike train #7, #10 and #13 are given in (Fig. B.3). These histograms have high count for the short ISI and the ISI count decreases with increase of the ISI length. That suggests that the influence function should be specified by the formula (5.3). The parameters of the influence functions are $\tau_r = 0.1\ ms, \tau_s = 10\ ms$. Another parameter, the time lag Δ is specified from pair-wise CCF analysis.

Thus, the influence functions are defined and the Cox coefficients and the corresponding confidence intervals are calculated using formulas (5.5) and (5.6). This procedure is repeated 29 times to obtain the full functional connectivity of the 29 spike trains. The confidence intervals are calculated using the significance level $\alpha = 0.05$ with Bonferroni correction.

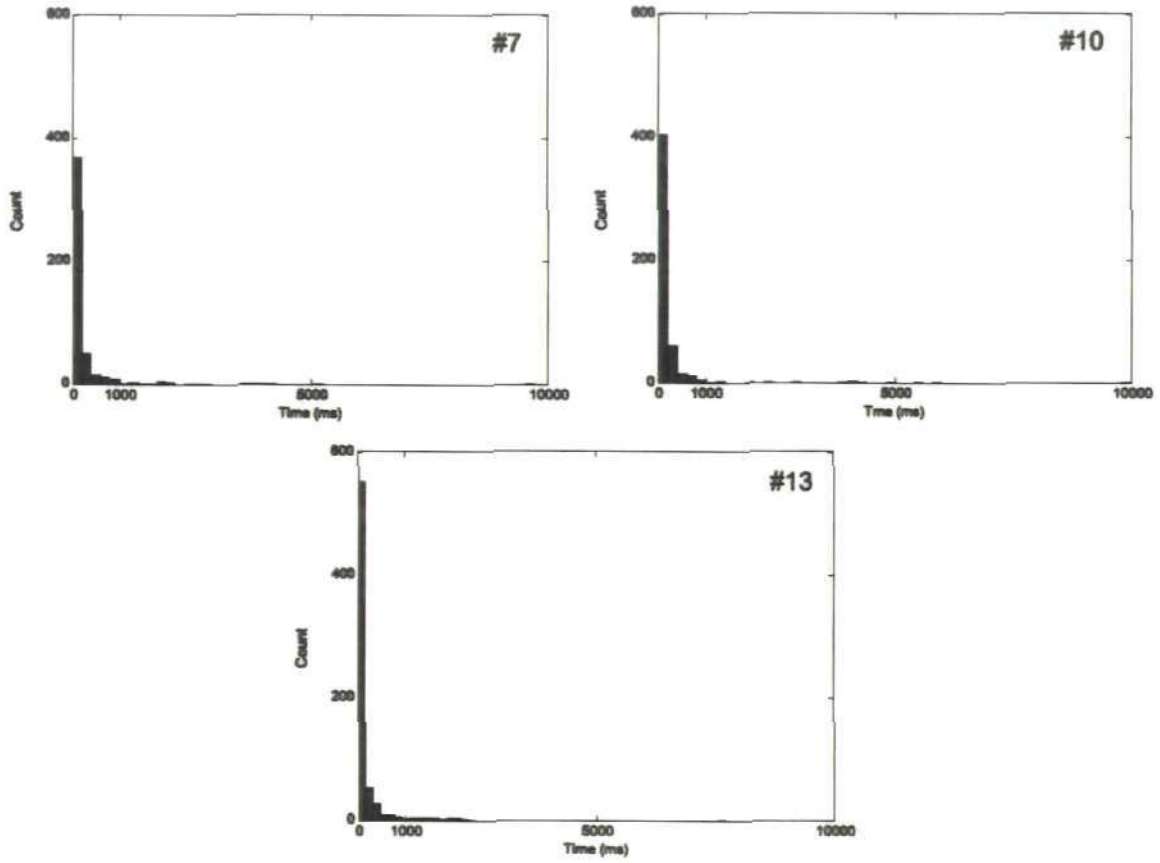


Figure B.3: Inter spike interval histogram of the spike trains #7, #10 and #13 of stimulus 2.

A total of 62 connections are identified by the Cox method which are shown by circles in Fig. B.4(a). The radius of the circle indicates the strength of functional connection. The direction of functional connection is from the reference spike train to the target spike train. Among the 62 connections, the 3 connections have stronger strength compared to others. These connections are: (#8, #7), (#20, #19) and (#32, #30). 9 connections have a small strength compared to others. These connections are: (#3, #15), (#7, #3), (#7, #28), (#8, #3), (#25, #32), (#27, #32), (#28, #12), (#28, #32) and (#32, #31). All other connections have a medium strength of connection. Spike trains #8 and #32 have 8 outgoing connections and spike train #3 has 6 incoming connections, which are the highest among 29 spike trains. There are 13 pairs of connections where both spike trains have functional connectivity to each other.

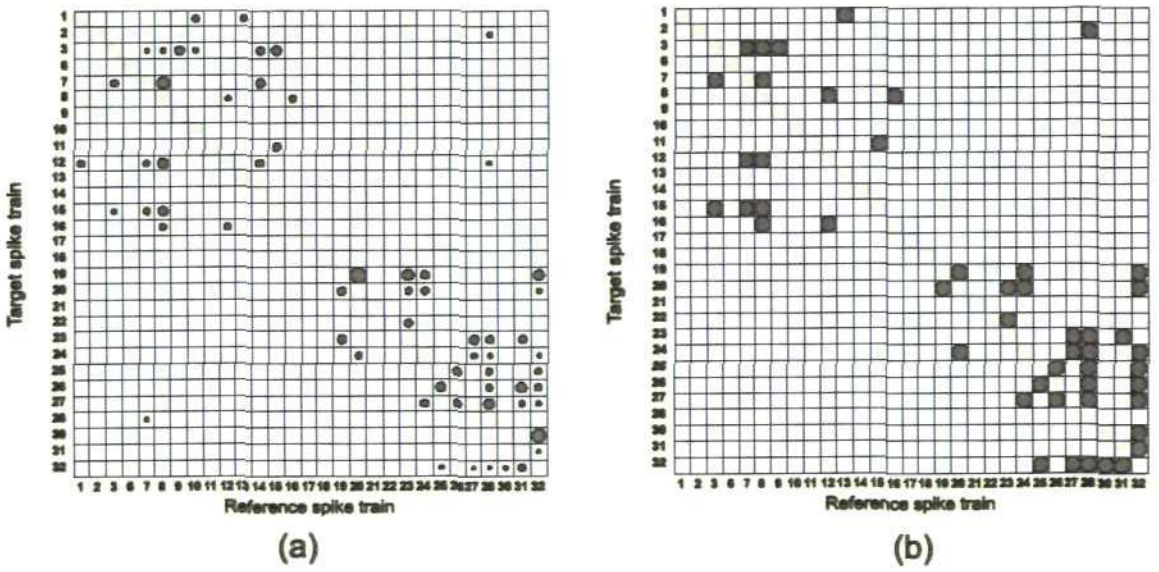


Figure B.4: (a) Functional connectivity of the 29 spike trains identified by the Cox method of stimulus 2. Radius of the circle indicates strength of connection. **(b)** The connections that are identified both by the MCG method and the Cox method.

Functional connectivity obtained by the MCG method and the Cox method show a good agreement between them (Fig. B.4(b)). There are 49 connections which are common in both techniques. Among the common connections spike train #32 has 8 outgoing connections to other spike trains and spike train #32 has 5 incoming connections which are highest among the 29 spike trains. There are 11 pairs of connections where both spike trains have functional connectivity to each other.

B.1.3 Cox metric

All possible pairs of spike trains are analysed by pair-wise Cox method where one spike train is taken as a target and the other spike train as a reference. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\ ms$ and $\tau_s = 10\ ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval by (5.6). The Cox metric is applied to the significant connections to reveal the groups of similar spike trains. The result of Cox metric is

shown in Fig. B.5(a) where the grey circles indicate the significant connections obtained by the pair-wise Cox method. The black circles indicate symmetric of the grey circles but not identified by the pair-wise analysis of Cox method. There are three groups of coupled spike trains. Similarly, the Cox metric is applied to the functional connections identified by the Cox method considering all spike trains at once. This functional connection is shown in Fig. B.4(a) and the result of Cox metric is shown in Fig. B.5(b). In the figure the grey circles indicate the significant connections obtained by the Cox method considering all spike trains at once. The black circles indicate symmetric of the grey circles but not identified by the Cox method considering all spike trains at once.

From Fig. B.5(a) three groups of similar spike trains are identified. The first group consists of 10 spike trains, these are: spike trains #6, #9, #3, #15, #7, #8, #11, #16, #12 and #14. In this group, the connection from spike train #6 to #9, spike train #15 to #9, spike train #8 to #9 and spike train #8 to #7 have the highest strength among other connections. Spike train #8 has 7 outgoing connections to spike trains #9, #3, #15, #7, #11, #16 and #12 and spike train #3 has 5 incoming connections from spike trains #9, #15, #7, #8 and #14. Thus from this group it can be concluded that spike train #8 acts as the most influential spike train. The second group consists of 2 spike trains: spike train #1 and spike train #10. The third group consists of 11 spike trains; these are: spike trains #19, #20, #24, #23, #27, #28, #26, #31, #25, #30 and #32. In this group, 4 connections (#19, #20), (#20, #19), (#30, #32) and (#32, #30) have strong strength compared to others. Spike train #32 has 8 outgoing connections to spike trains #19, #20, #24, #27, #26, #31, #25 and #30 and 5 incoming connections from spike trains #27, #28, #31, #25 and #30. This spike train can be considered as the most influential spike train for this group. Spike trains #2, #22, #13, #21, #18 and #17 do not form any group.

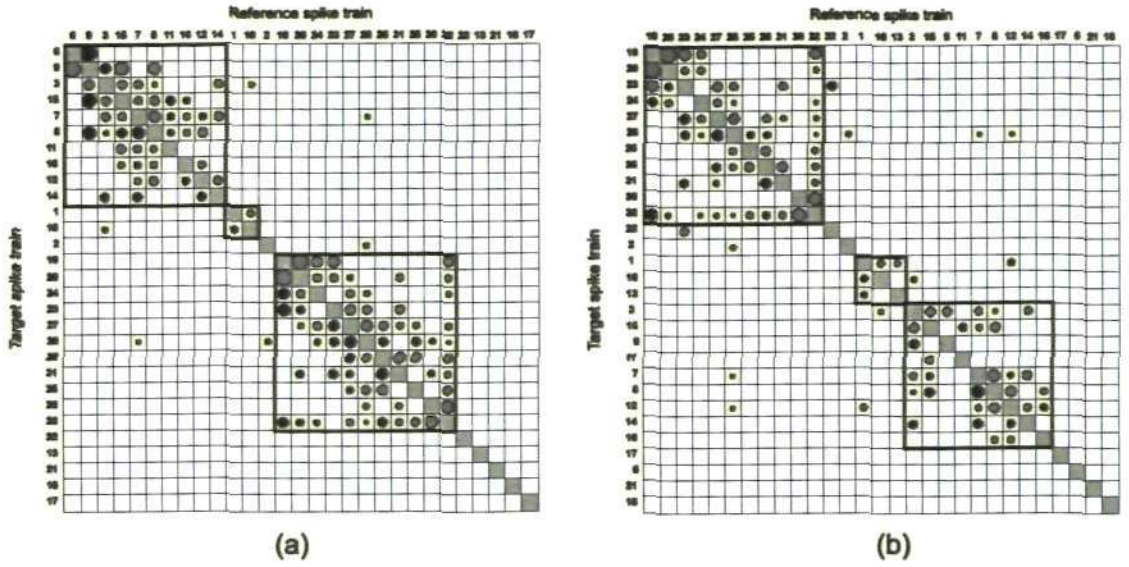


Figure B.5: Groups of similar spike trains revealed by the Cox metric of the 29 spike trains of stimulus 2. **(a)** Cox metric using pair-wise analysis. **(b)** Cox metric considering all spike trains at once.

Similarly, from Fig. B.5(b) three groups of similar spike trains are identified. The first group consists of 11 spike trains, these are: spike trains #19, #20, #23, #24, #27, #28, #25, #26, #31, #30 and #32. In this group, the connection from spike train #19 to #20, spike train #20 to #19, spike train #30 to #32 and spike train #32 to #30 have the highest strength among other connections. Spike train #32 has the 8 outgoing connections to spike trains #19, #20, #24, #27, #25, #26, #31 and #30 and 5 incoming connections from spike trains #27, #28, #25, #31 and #30. Thus from this group it can be concluded that spike train #32 is the most influential spike train. The second group consists of 3 spike trains, these are: spike trains #1, #10 and #13. The third group consists of 9 spike trains; these are: spike trains #3, #15, #9, #11, #7, #8, #12, #14 and #16. In this group, spike train #8 has 5 outgoing connections to spike trains #3, #15, #7, #12 and #16. As these spike trains have the highest outgoing connections, these spike trains can be considered as the most influential spike trains for this group. Spike trains #17, #6, #21 and #18 do not follow any group.

Investigation from Fig. B.5(a)-(b) reveal that the Cox metric identified by the pair-wise analysis and considering all spike trains at once show a good agreement. For example, the third group of figure B.5(a) consists 11 spike trains which are the same as the first group of figure B.5(b). Similarly, the other two groups of both figures consist same spike trains except a few. From this analysis it can be concluded that application of the Cox metric to experimental data using pair-wise analysis and considering all spike trains at once enables to create similar result.

B.1.4 Motif analysis

To find the pattern of interconnections among the 29 spike trains, a structural motif analysis is conducted using triplet-wise analysis of Cox method. A total of 3654 triplets are analysed. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\ ms$ and $\tau_s = 10\ ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval using (5.6). Functional connectivity of each triplet spike trains is used to identify the structural motif. Structural motif count is obtained by analysing all 3654 triplets of spike trains which is shown in Fig. B.6.

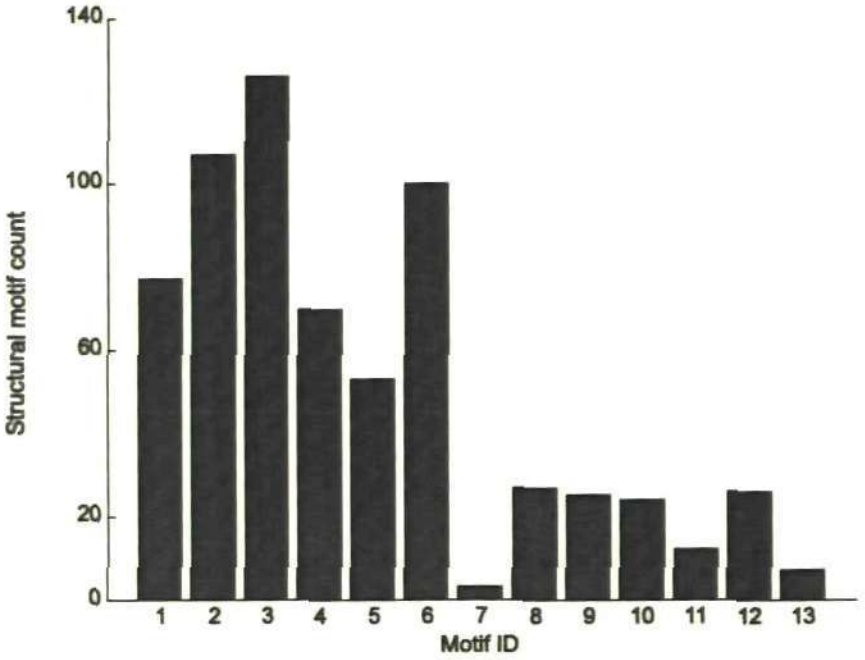


Figure B.6: Structural motif count of all 3654 triplets of the 29 spike trains of stimulus 2.

Out of 3654 triplets, 657 triplets have different structural motifs ID. Among the 657 triplets, 126 triplets have motif ID 3 which is the highest among others motif ID's. Only 3 triplets have motif ID 7 which is the lowest. Motif ID's 2 and 6 have similar number of triplets; motif ID's 1, 4 and 5 have similar number of triplets; motif ID's 8, 9, 10 and 12 have similar number of triplets. A total of 85 triplets have connected motifs (connected motifs are motif ID 7, 9, 10, 12 and 13). On the other hand, a total of 572 triplets have unconnected motifs. Thus there are low proportions of connected motifs (14.86%) in the groups of triplet spike trains which indicate that in the group, connection from every spike trains to every other spike trains is very low.

B.2 Analysis of functional connectivity of stimulus 3

The raster plot of 32 spike trains (Fig. B.7) shows that like stimuli 1 and 2, spike trains #4, #5 and #29 have high spiking rates compared to other spike trains. Therefore, these three spike trains (#4, #5 and #29) are considered to be outliers and are not considered in the analysis of functional connectivity. All other 29 spike trains have similar spiking

pattern. Spiking rates of these 29 spike trains is high over the time interval [43000 ms, 90000 ms]. These 29 spike trains are analysed to identify functional connectivity between them keeping the original numeration of the spike trains.

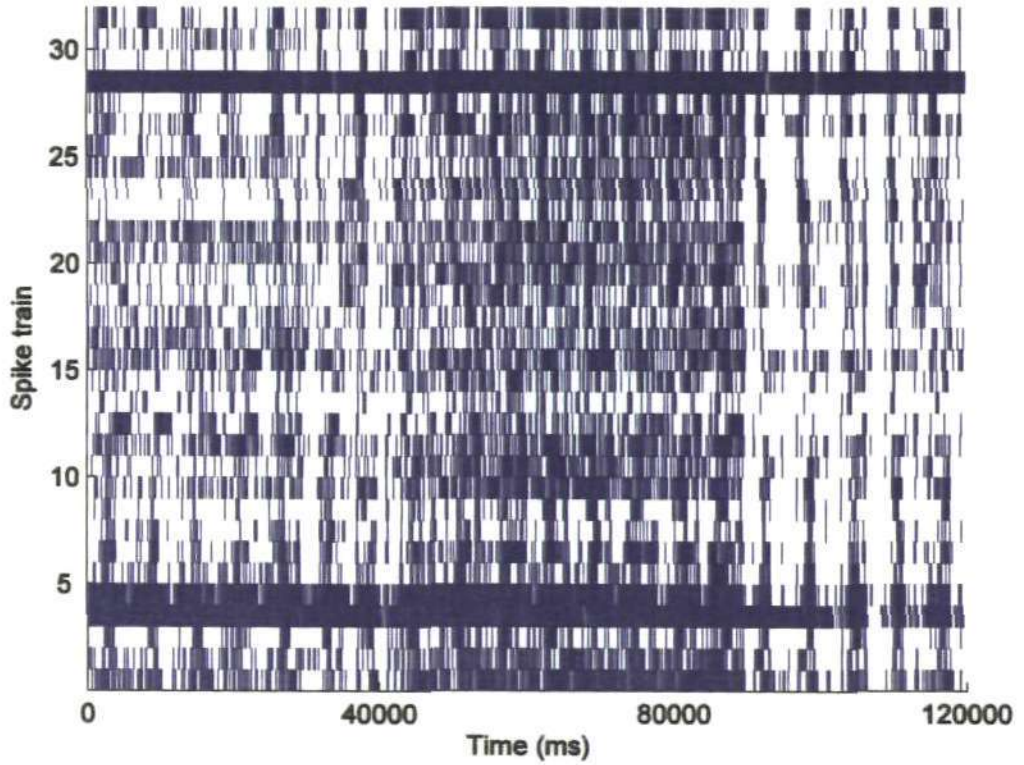


Figure B.7: Raster plot of 32 spike trains of stimulus 3. Spike trains #4, #5 and #29 have high spiking rates and are not considered for analysing functional connectivity.

B.2.1 MCG method

All pair-wise CCF are calculated with a bin size of 1 ms and a correlation window of 100 ms for the 29 spike trains. To test the independence of two spike trains the significance level $\alpha = 0.05$ is used with the Bonferroni correction. A total of 89 significant connections are found for 29 spike trains. These significant connections are shown in a matrix format in Fig. B.8(a) where the connections are indicated by circles. The direction of connection is considered from the reference spike train to the target spike train. Among the 29 spike trains, spike train #32 has 12 outgoing connections to

other spike trains which is the highest among 29 spike trains and similar 13 incoming connections from other spike trains.

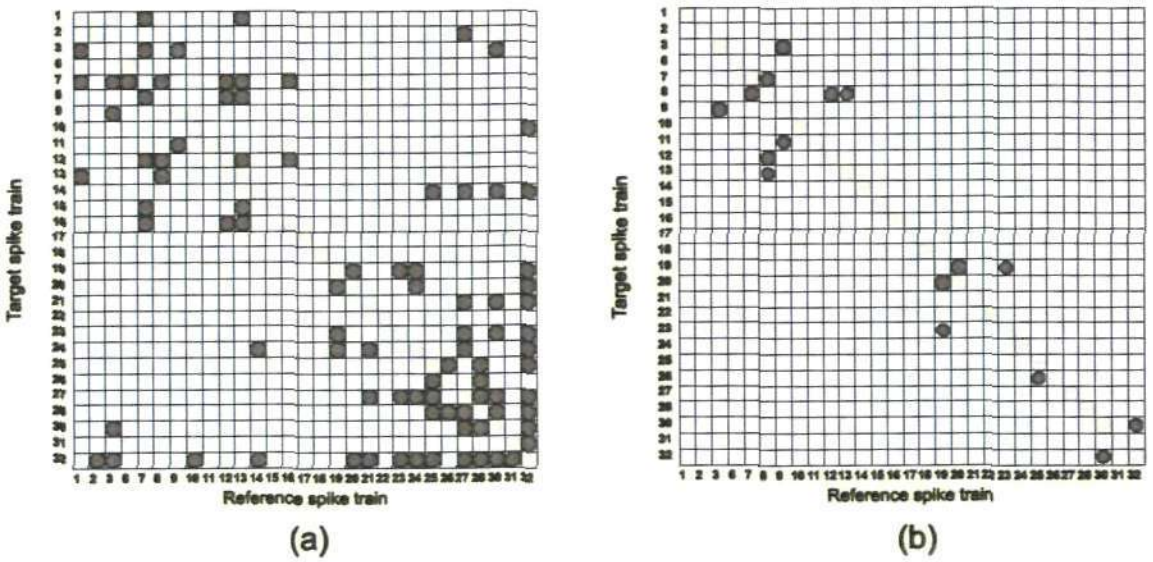


Figure B.8: (a) Significant connections obtained from pair-wise CCF analysis of the 29 spike trains of stimulus 3. **(b)** Direct connections obtained from the connections in (a) after the clustering algorithm. The radius of the circle indicates the strength of the connections.

A clustering algorithm is applied to the 89 significant connections to identify the direct connections and a total of 16 connections are identified as direct connections (Fig. B.8(b)). The radius of the circle indicates the strength of connection. All the 16 connections have strong strength. Spike train #8 has 3 outgoing connections to spike trains #7, #12 and #13; and 3 incoming connections from spike trains #7, #12 and #13, which are the highest among 29 spike trains. There are 7 pairs of connections where both spike trains have functional connectivity to each other. These pairs are: (#3, #9), (#7, #8), (#8, #12), (#8, #13), (#19, #20), (#19, #23) and (#30, #32).

B.2.2 Cox method

Application of Cox method to the 29 spike trains requires the identification of reference spike trains, target spike train and the influence function. The inter spike interval (ISI) histogram of three spike trains, spike train #15, #19 and #20 are given in (Fig. B.9).

These histograms have high count for the short ISI and the ISI count decreases with increase of the ISI length. That suggests that the influence function should be specified by the formula (5.3) with parameters $\tau_r = 0.1\ ms$, $\tau_s = 10\ ms$. Another parameter, the time lag Δ is specified from pair-wise CCF analysis. Thus, the influence functions are defined and the Cox coefficients and the corresponding confidence intervals are calculated using formulas (5.5) and (5.6). This procedure is repeated 29 times to obtain the full functional connectivity of the 29 spike trains. The confidence intervals are calculated using the significance level $\alpha = 0.05$ with Bonferroni correction.

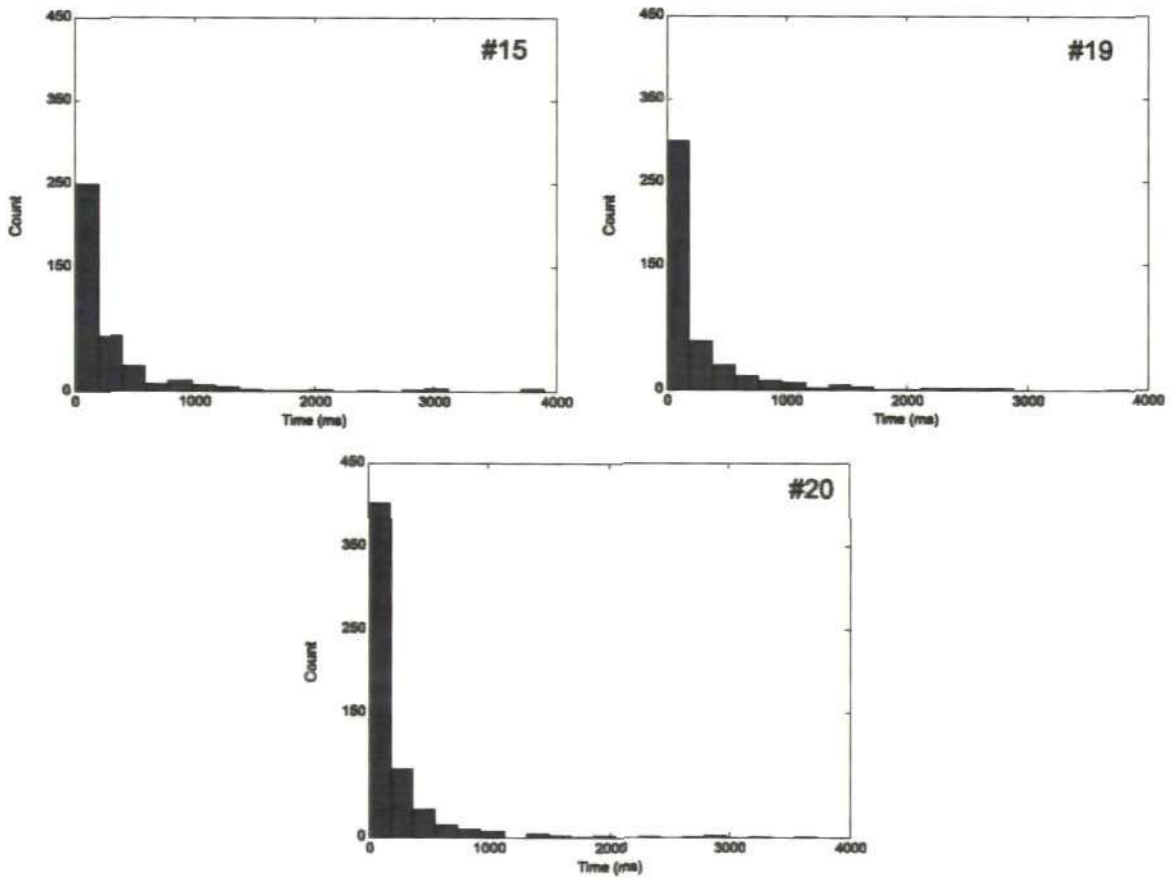
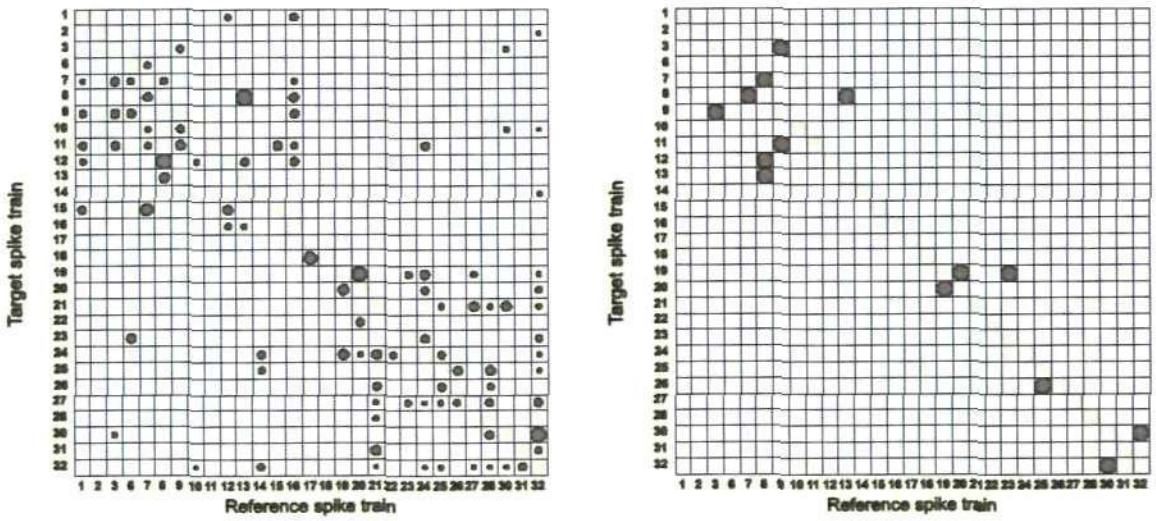


Figure B.9: Inter spike interval histogram of the spike trains #15, #19 and #20 of stimulus 3.

A total of 95 connections are identified by the Cox method which are shown by circles in Fig. B.10(a). The radius of the circle indicates the strength of functional connection. The direction of functional connection is from the reference spike train to the target spike train. Among the 95 connections, the 6 connections have stronger strength

compared to others. These connections are: (#7, #15), (#8, #12), (#13, #8), (#17, #18), (#20, #19) and (#32, #30). 6 connections have small strength compared to others. These connections are: (#10, #32), (#21, #32), (#24, #27), (#24, #32), (#32, #2) and (#32, #10). All other connections have a medium strength of connection. Spike train #32 has 12 outgoing connections and 9 incoming connections, which are the highest among 29 spike trains. There are 21 pairs of connections where both spike trains have functional connectivity to each other.



(a) (b)

Figure B.10: **(a)** Functional connectivity of the 29 spike trains identified by the Cox method of stimulus 3. Radius of the circle indicates strength of connection. **(b)** The connections that are identified both by the MCG method and the Cox method.

Functional connectivity obtained by the MCG method and the Cox method show good agreement between them (Fig. B.10(b)). There are 14 connections which are common in both techniques. Among the common connections spike train #8 has 3 outgoing connections to spike trains #7, #12 and #13. There are 5 pairs of connections where both spike trains have functional connectivity to each other. These connections are: (#3, #9), (#7, #8), (#8, #13), (#19, #20) and (#30, #32).

B.2.3 Cox metric

All 812 possible pairs of spike trains are analysed by pair-wise Cox method where one spike train is taken as a target and the other spike train as a reference. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\ ms$ and $\tau_s = 10\ ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval by (5.6). The Cox metric is applied to the significant connections to reveal the groups of mutually coupled spike trains. The result of Cox metric is shown in Fig. B.11(a) where the grey circles indicate the significant connections obtained by the pair-wise Cox method The black circles indicate symmetric of the grey circles but not identified by the pair-wise analysis of Cox method. Similarly, the Cox metric is applied to the functional connections identified by the Cox method considering all spike trains at once. This functional connection is shown in Fig. B.10(a) and the result of Cox metric is shown in Fig. B.11(b). In the figure the grey circles indicate the significant connections obtained by the Cox method considering all spike trains at once. The black circles indicate symmetric of the grey circles but not identified by the Cox method considering all spike trains at once.

From Fig. B.11(a) four groups of similar spike trains are identified. The first group consists of 12 spike trains, these are: spike trains #19, #20, #23, #14, #21, #24, #27, #30, #32, #25, #28 and #26. In this group, the connection from spike train #19 to #20, spike train #20 to #19, spike train #30 to #32 and spike train #32 to #30 have the highest strength among other connections. Spike train #32 has 9 outgoing connections to spike trains #19, #20, #23, #14, #21, #24. #27, #30 and #25 and spike train #27 has 9 incoming connections from spike trains #20, #23, #21, #24, #30, #32, #25, #28 and #26.

Thus from this group it can be concluded that spike train #32 is the most influential spike train. The second group consists of 7 spike trains; these are: spike trains #1, #16, #7, #15, #3, #9 and #11. In this group connection from spike train #7 to spike train #15 has stronger strength compared to others. Spike train #16 has connections to all other spike trains except spike train #3. Spike train ##11 has 5 incoming connections from all other spike trains. Spike train #16 can be considered as the most influential spike train for this group. The third group consists of 3 spike trains: #8, #12 and #13. All the connections have strong strength. Spike train #8 has connections to the spike trains #12 and #13. This spike train can be considered as the most influential spike train for this group. The fourth group consists only 2 spike trains: #17 and #18, where there is a connection from spike train #17 to spike train #18. Spike trains #31, #22, #2, #10 and #6 do not form any group.

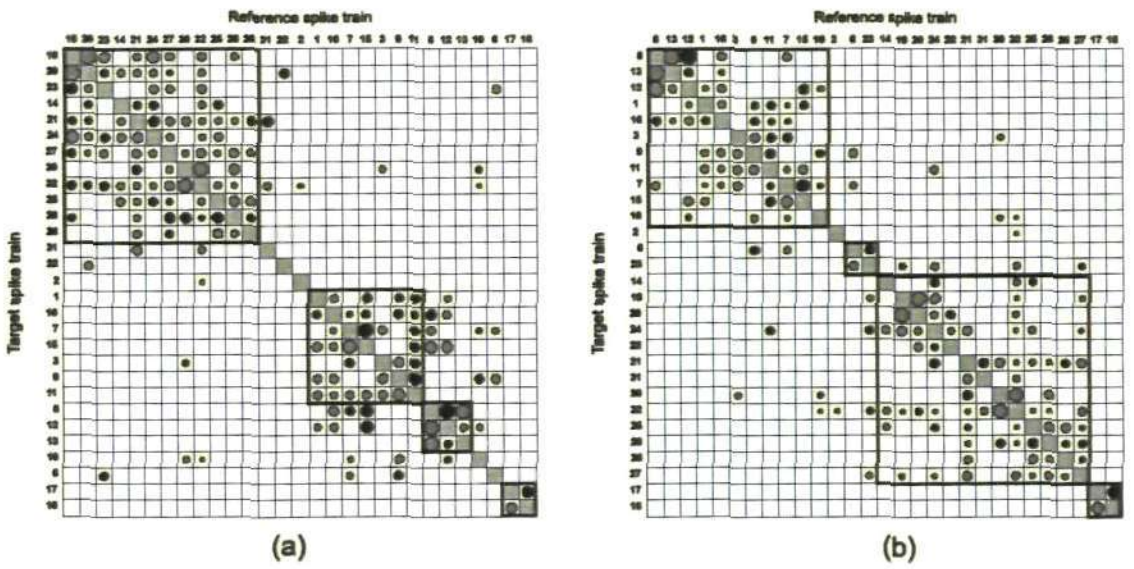


Figure B.11: Groups of similar spike trains revealed by the Cox metric of the 29 spike trains of stimulus 3. **(a)** Cox metric using pair-wise analysis. **(b)** Cox metric all spike trains at once.

Similarly, from Fig. B.11(b) four groups of similar spike trains are identified. The first group consists of 11 spike trains, these are: spike trains #8, #13, #12, #1, #16, #3, #9, #11, #7, #15 and #10. In this group, the connection from spike train #8 to #13 and spike

train #13 to #8 have the highest strength among other connections. Spike train #16 has the 6 outgoing connections to spike trains #8, #12, #1, #9, #11 and #7. As this spike train has the highest outgoing connections, this spike train is considered as the most influential spike train for this group. The second group consists of 2 spike trains, these are: spike trains #6 and #23. The third group consists of 13 spike trains; these are: spike trains #14, #19, #20, #24, #22, #21, #31, #30, #32, #25, #28, #26 and #27. In this group, four connections (#19, #20), (#20, #19), (#30, #32) and (#32, #30) have big strength compared to others. Spike train #32 has 9 outgoing connections to other spike trains of this group. As this spike train has the highest outgoing connections, this spike train can be considered as the most influential spike train for this group. The forth group consists only 2 spike trains, these are: spike trains #17 and #18. Spike trains #2 does not follow any group.

Investigation from Fig. B.11(a)-(b) reveal that the Cox metric identified by the pair-wise analysis and considering all spike trains at once show a good agreement. For example, the first group of Fig. B.11(a) and the third group of Fig. B.11(b) consist the same 11 spike trains (#19, #20, #14, #21, #24, #27, #30, #32, #25, #28 and #26) except one (#23) in Fig. B.11(a) and two (#22, #31) in Fig. B.11(b). Similarly, the other three groups of both figures consist same spike trains except a few. From this analysis it can be concluded that application of Cox metric to experimental data using pair-wise analysis and considering all spike trains at once enable to create similar result.

B.2.4 Motif analysis

To find the pattern of interconnections a structural motif analysis is conducted using triplet-wise Cox method to the 3654 triplets. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise

CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval using (5.6). Functional connectivity of each triplet spike trains is used to identify the structural motif among the triplets of spike trains. Structural motif count is obtained by analysing all 3654 triplets of spike trains which is shown in Fig. B.12.

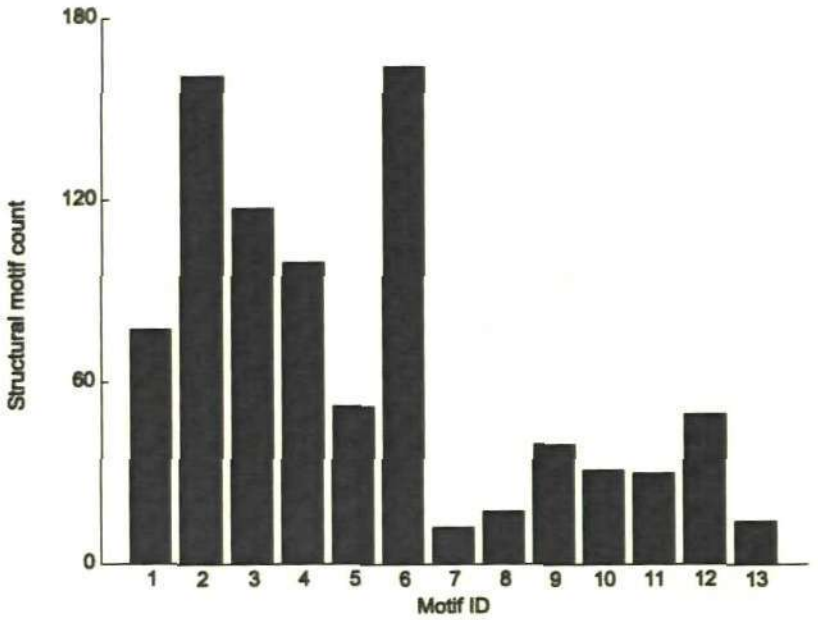


Figure B.12: Structural motif count of all 3654 triplets of the 29 spike trains of stimulus 3.

Out of 3654 triplets, 862 triplets have different structural motifs ID. Among the 862 triplets, 164 triplets have motif ID 6 which is the highest among other motif ID's. Only 12 triplets have motif ID 7 which is the lowest. 161 triplets have Motif ID 2 which is almost same as the motif ID 6. Motif ID's 1, 3 and 4 have similar number of triplets; motif ID's 9, 10, 11 and 12 have similar number of triplets. A total of 145 triplets have connected motifs (connected motifs are motif ID 7, 9, 10, 12 and 13). On the other hand, a total of 717 triplets have unconnected motifs. Thus there are low proportions of connected motifs (20.22%) in the groups of triplet spike trains which indicate that in the group, connection from every spike trains to every other spike trains is very low.

B.3 Analysis of functional connectivity of stimulus 4

Similar to stimuli 1, 2 and 3, spike trains #4, #5 and #29 have high spiking rates compared to other spike trains (Fig. B.13). Therefore, these three spike trains are considered to be outliers and they are excluded from analysis. All other 29 spike trains have similar spiking pattern. Spiking rates of these 29 spike trains are high over time interval [50000 ms, 83000 ms]. 29 spike trains are analysed for identifying functional connectivity between them keeping the original numeration of the spike trains.

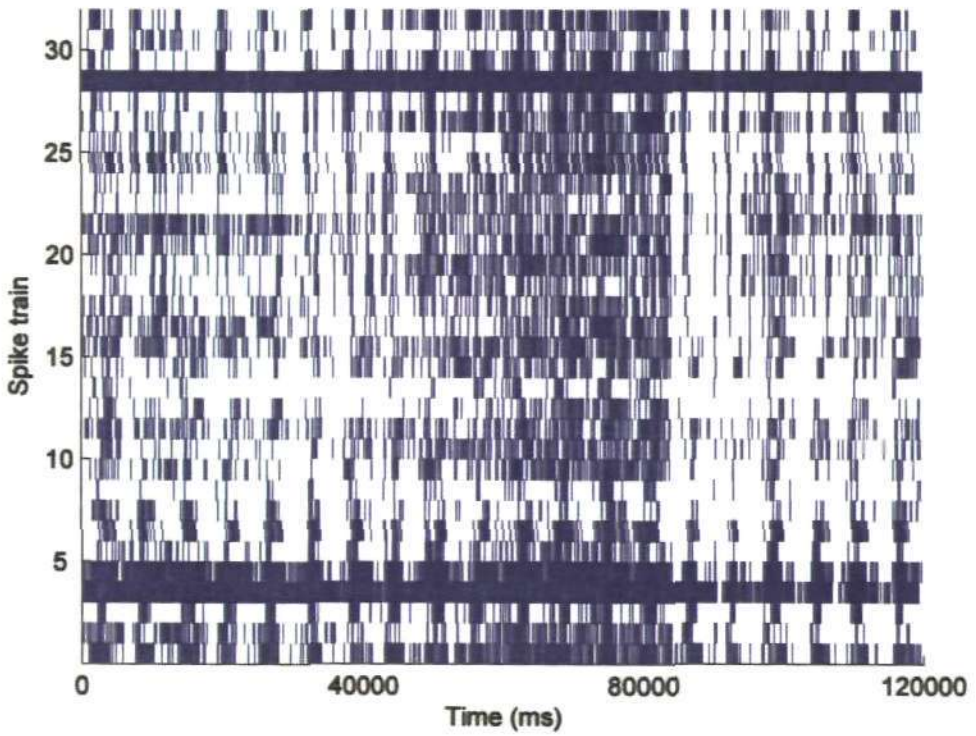


Figure B.13: Raster plot of 32 spike trains of stimulus 4. Spike trains #4, #5 and #29 have high spiking rates and are not considered for analysing functional connectivity.

B.3.1 MCG method

All pair-wise CCF are calculated with a bin size of 1 ms and a correlation window of 100 ms for the 29 spike trains. To test the independence of two spike trains the significance level $\alpha = 0.05$ is used with the Bonferroni correction. A connection is considered to be significant if a peak of the CCF exceeds the upper boundary of the 'confidence interval'.

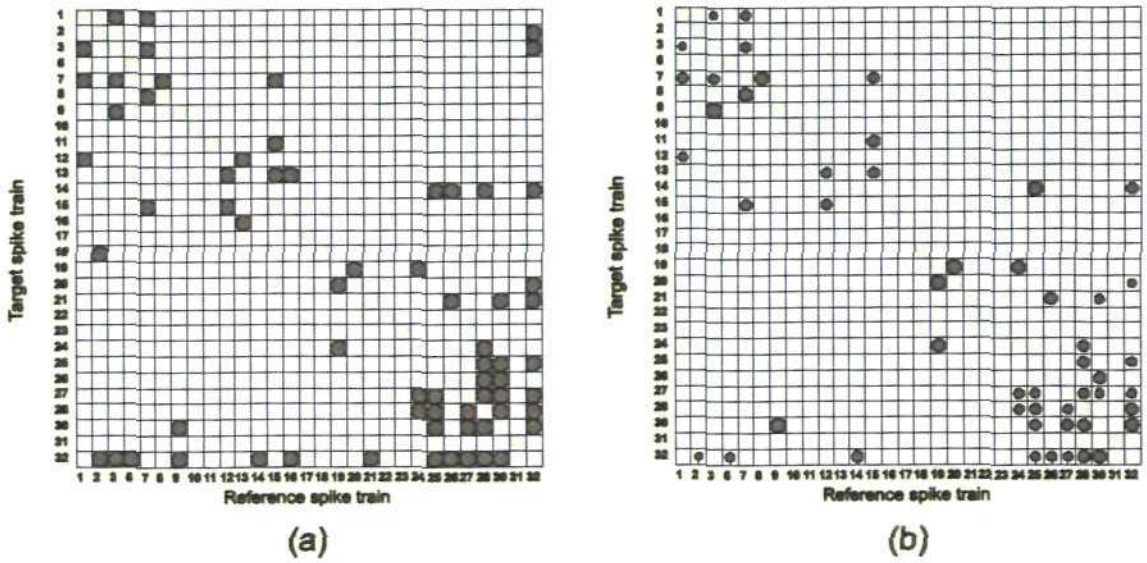


Figure B.14: (a) Significant connections obtained from pair-wise CCF analysis of the 29 spike trains of stimulus 4. **(b)** Direct connections obtained from the connections in (a) after the clustering algorithm. The radius of the circle indicates the strength of the connections.

A total of 67 significant connections are found for 29 spike trains. These significant connections are shown in a matrix format in Fig. B.14(a)) where the connections are indicated by circles. The direction of connection is considered from the reference spike train to the target spike train. Among the 29 spike trains, spike train #32 has 12 outgoing connections to other spike trains which is the highest among 29 spike trains and similar 12 incoming connections from other spike trains.

Application of clustering algorithm to the 67 significant connections reveals a total of 51 direct connections (Fig. B.14(b)). The radius of the circle indicates the strength of connection. Among 51 connections, 10 connections have strong strength. These connections are: (#3, #9), (#7, #8), (#8, #7), (#9, #30), (#19, #20), (#19, #24), (#20, #19), (#24, #19), (#25, #14), (#30, #32) and (#32, #30). Spike train #32 has 6 outgoing connections to spike trains #14, #20, #25, #27, #28 and #30; and 8 incoming connections from spike trains #2, #6, #14, #25, #26, #27, #28 and #30, which are the highest among 29 spike trains. There are 16 pairs of connections where both spike trains

have functional connectivity to each other. These pairs are: (#1, #3), (#1, #7), (#3, #7), (#7, #8), (#7, #15), (#14, #32), (#19, #20), (#19, #24), (#24, #28), (#25, #28), (#25, #32), (#27, #28), (#27, #30), (#27, #32), (#28, #32) and (#30, #32).

B.3.2 Cox method

To identify functional connectivity by the Cox method, the target spike train, reference spike trains and the influence function which determines how the reference spike train influences the target spike train should be specified. The inter spike interval (ISI) histogram of three spike trains, spike train #12, #19 and #20 are given in (Fig. B.15). These histograms have high count for the short ISI and the ISI count decreases with increase of the ISI length. That suggests that the influence function should be specified by the formula (5.3) with parameters $\tau_r = 0.1\ ms$, $\tau_s = 10\ ms$. Another parameter, the time lag Δ is specified by the pair-wise CCF. Thus, the influence functions are defined and the Cox coefficients and the corresponding confidence intervals are calculated using formula (5.5) and (5.6). This procedure is repeated 29 times to obtain the full functional connectivity of the 29 spike trains. The confidence intervals are calculated using the level of significance $\alpha = 0.05$ with Bonferroni correction.

The 71 connections, identified by the Cox method, are shown by circles in Fig. B.16(a). The radius of the circle indicates the strength of functional connection. The direction of functional connection is from the reference spike train to the target spike train. Among the 71 connections, the 5 connections have stronger strength compared to others. These connections are: (#10, #31), (#12, #9), (#20, #19), (#26, #31) and (#32, #30). All other connections have a medium strength of connection. Spike train #32 has 8 outgoing connections and 6 incoming connections, which are the highest among 29 spike trains. There are 19 pairs of connections where both spike trains have functional connectivity to each other.

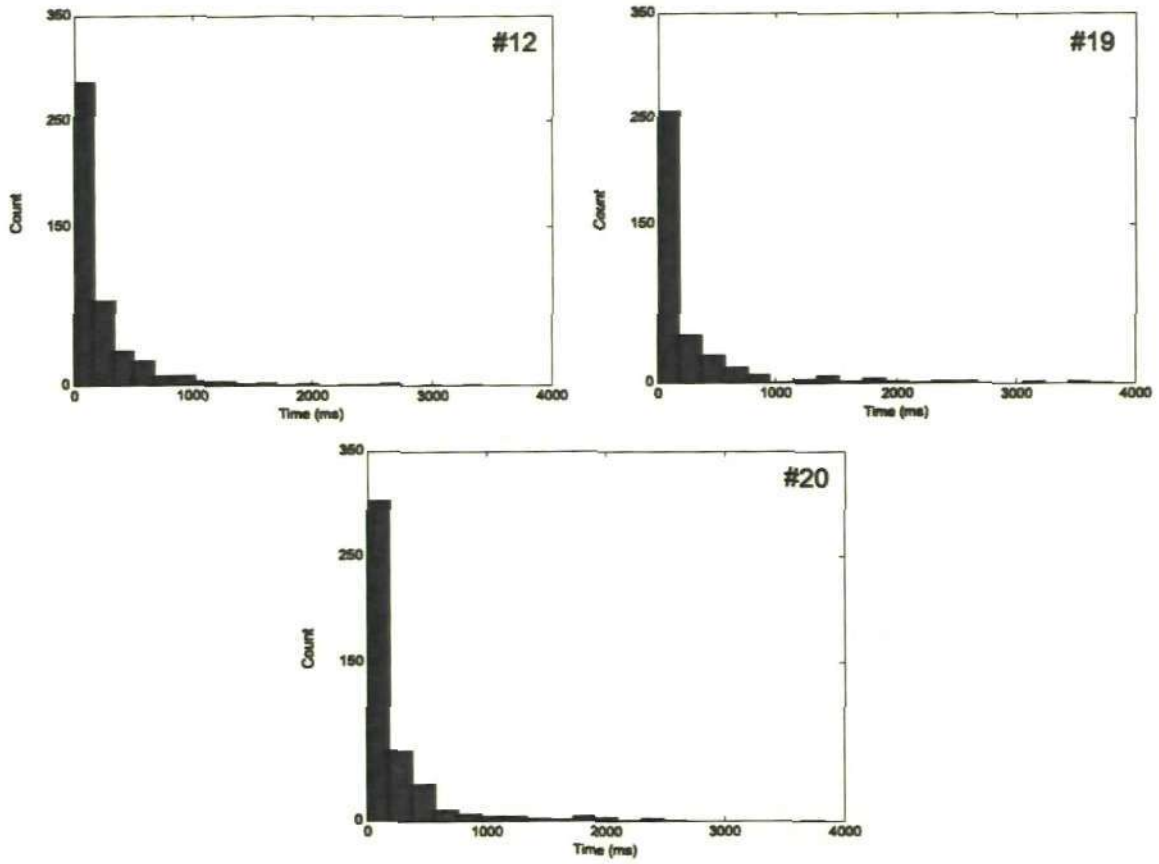


Figure B.15: Inter spike interval histogram of the spike trains #12, #19 and #20 of stimulus 4.

Functional connectivity obtained by the MCG method and the Cox method show a good agreement between them (Fig. B.16(b)). There are 31 connections which are common in both techniques. Among the common connections spike train #32 has 4 outgoing connections to spike trains #20, #25, #27 and #30 and 5 incoming connections from spike trains #2, #25, #26, #28 and #30. There are 6 pairs of connections where both spike trains have functional connectivity to each other. These connections are: (#3, #7), (#9, #20), (#25, #28), (#25, #32), (#27, #30) and (#30, #32).

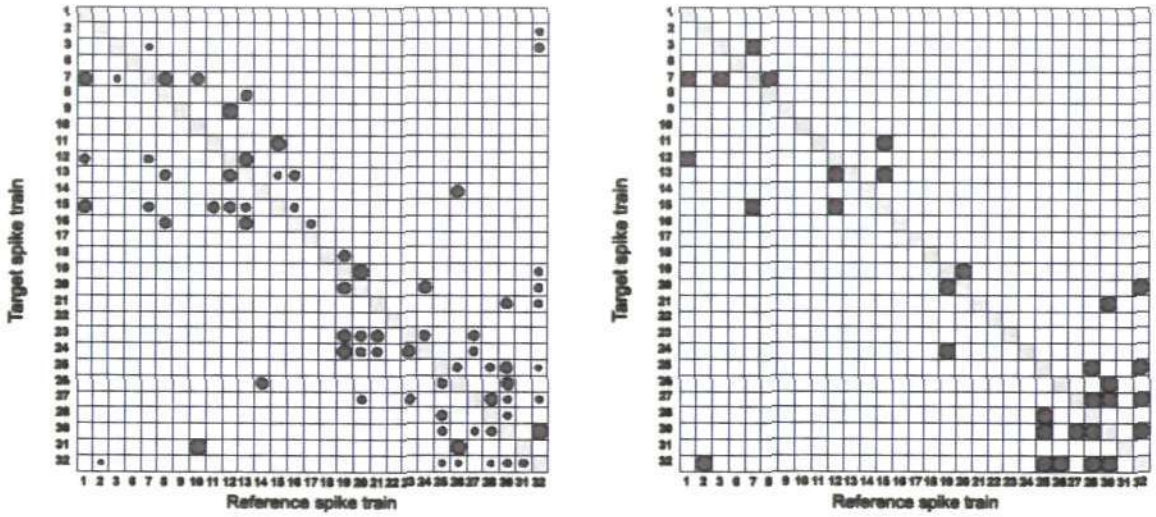


Figure B.16: (a) Functional connectivity of the 29 spike trains identified by the Cox method of stimulus 4. Radius of the circle indicates strength of connection. **(b)** The connections that are identified both by the MCG method and the Cox method.

B.3.3 Cox metric

To apply Cox metric to the 29 spike trains all possible pairs of spike trains are analysed by pair-wise Cox method where one spike train is taken as a target and the other spike train as a reference. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval by (5.6). The Cox metric is applied to the significant connections to reveal the groups of similar spike trains. The result of Cox metric is shown in Fig. B.17(a) where the grey circles indicate the significant connections obtained by the pair-wise Cox method. The black circles indicate symmetric of the grey circles but not identified by the pair-wise analysis of Cox method. Similarly, the Cox metric is applied to the functional connections identified by the Cox method considering all spike trains at once. This functional connection is shown in Fig. B.16(a) and the result of Cox metric is shown in Fig. B.17(b). In the

figure the grey circles indicate the significant connections obtained by the Cox method considering all spike trains at once. The black circles indicate symmetric of the grey circles but not identified by the Cox method considering all spike trains at once.

From Fig. B.17(a) five groups of similar spike trains are identified. The first group consists of 8 spike trains, these are: spike trains #30, #32, #21, #25, #27, #28, #14 and #26. In this group, the connection from spike train #30 to #32, spike train #32 to #30, spike train #14 to #26 and spike train #26 to #14 have the highest strength among other connections. Spike trains #30, #25 and #28 have 6 outgoing connections to other spike trains. Spike train #30 has 6 incoming connections from all 6 spike trains except spike train #14. Thus from this group it can be concluded that spike train #30 is the most influential spike train. The second group consists of 4 spike trains; these are: spike trains #19, #20, #23 and #24. In this group, all the connections have strong strength, which means that this group is strongly connected group. All the spike trains have connections to each other, which mean there is no most influential spike train in this group. The third group consists of only two spike trains: #10 and #31, where there is a connection from spike train #10 to spike train #31. The fourth group consists of 5 spike trains: #1, #7, #8, #11 and #15. All the connections in this group have a medium strength. The fifth group consists of 4 spike trains: #9, #12, #13 and #16. Connection from #12 to #9 has strong strength and spike train #12 can be considered as the most influential spike train of this group. Spike trains #16, #2, #3, #17, #22 and #6 do not form any group.

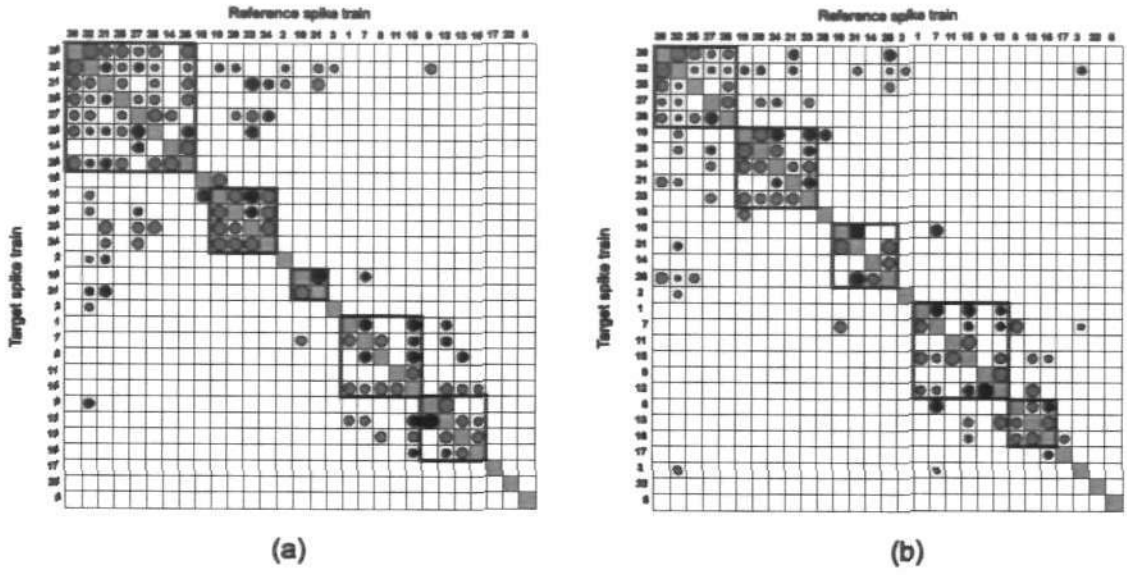


Figure B.17: Groups of similar spike trains revealed by the Cox metric of the 29 spike trains of stimulus 4. **(a)** Cox metric using pair-wise analysis. **(b)** Cox metric considering all spike trains at once.

Similarly, from Fig. B.17(b) five groups of similar spike trains are identified. The first group consists of 5 spike trains, these are: spike trains #30, #32, #25, #27 and #28. In this group, the connection from spike train #30 to #32 and spike train #32 to #30 have the highest strength among other connections. Spike trains #30 and #28 have outgoing connections to all other spike trains. Thus from this group it can be concluded that spike trains #30 and #28 are the most influential spike trains. The second group consists of 5 spike trains; these are: spike trains #19, #20, #24, #21 and #23. In this group, all the connections have strong strength, which means that this group is strongly connected group. In this group spike trains #19 and #20 have 3 outgoing connections to other spike trains and are considered as the most influential spike trains in this group. The third group consists of 4 spike trains: #10, #31, #14 and #26. Here spike train #26 has two outgoing connections to spike trains #31 and #14 and is considered as the most influential spike train. The fourth group consists of 6 spike trains: #1, #7, #11, #15, #9 and #12. There is no influential spike train in this group. The fifth group consists of 3

spike trains: #8, #13 and #16. Spike trains #16, #2, #3, #17, #22 and #6 do not form any group.

Investigation from Fig. B.17(a)-(b) reveal that the Cox metric identified by the pair-wise analysis and considering all spike trains at once show a good agreement. For example, the second group of both figures consists the same four spike trains (#19, #20, #23 and #24) except spike train #21 in Fig. B.17(b). Similarly, the other four groups of both figures consist same spike trains except a few. From this analysis it can be concluded that application of Cox metric to experimental data using pair-wise analysis and considering all spike trains at once enables to create similar result.

B.3.4 Motif analysis

A motif analysis is conducted to the 29 spike trains to find the pattern of interconnections. For 29 spike trains a total of 3654 triplets are analysed. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\ ms$ and $\tau_s = 10\ ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval using (5.6). Functional connectivity of each triplet spike trains is used for identifying the structural motif among the triplets. The structural motif count is obtained by analysing all 3654 triplets of spike trains which is shown in Fig. B.18.

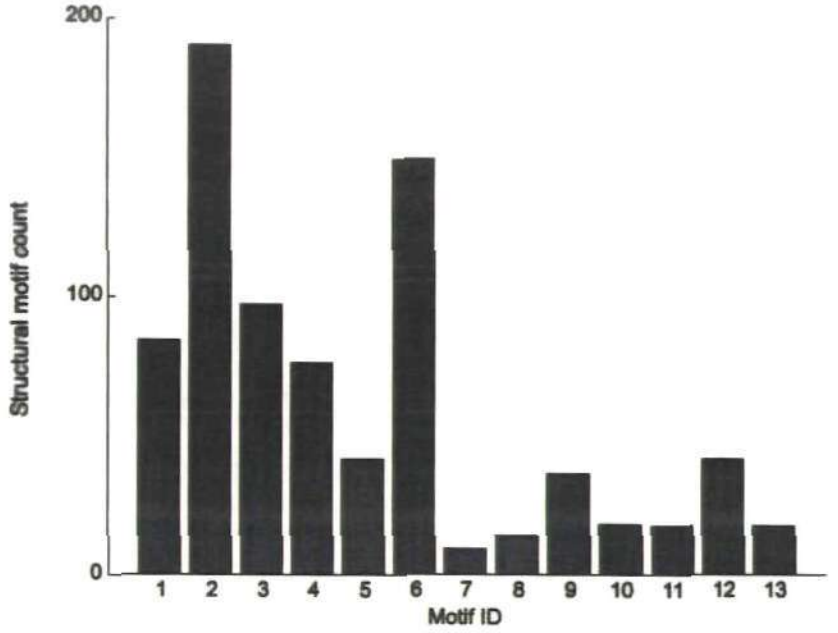


Figure B.18: Structural motif count of all 3654 triplets of the 29 spike trains of stimulus 4.

Out of 3654 triplets, 789 triplets have different structural motifs ID's. Among the 789 triplets, 190 triplets have motif ID 2 which is the highest among others motif ID's. Only 9 triplets have motif ID 7 which is the lowest. 149 triplets have Motif ID 6 which is the second highest among other motif ID's. Motif ID's 1, 3 and 4 have similar number of triplets; motif ID's 10, 11 and 13 have similar number of triplets and motif ID's 5, 9 and 12 have similar number of triplets. A total of 121 triplets have connected motifs (connected motifs are motif ID 7, 9, 10, 12 and 13). On the other hand, a total of 668 triplets have unconnected motifs. Thus there are low proportions of connected motifs (18.11%) in the groups of triplet spike trains which indicate that in the triplet, connection from every spike trains to every other spike trains is very low.

B.4 Analysis of functional connectivity of stimulus 5

The raster plot of 32 spike trains is shown in Fig. B.19. Similar to other stimuli, spike trains #4, #5 and #29 have high spiking rates compared to all other spike trains. Therefore, these three spike trains are considered to be outliers and they are excluded from analysis. All the 29 spike trains have similar spiking pattern. Spiking rates of these

29 spike trains are high over time interval [37000 ms, 100000 ms]. These 29 spike trains are analysed for identifying functional connectivity between them keeping the original numeration of the spike trains.

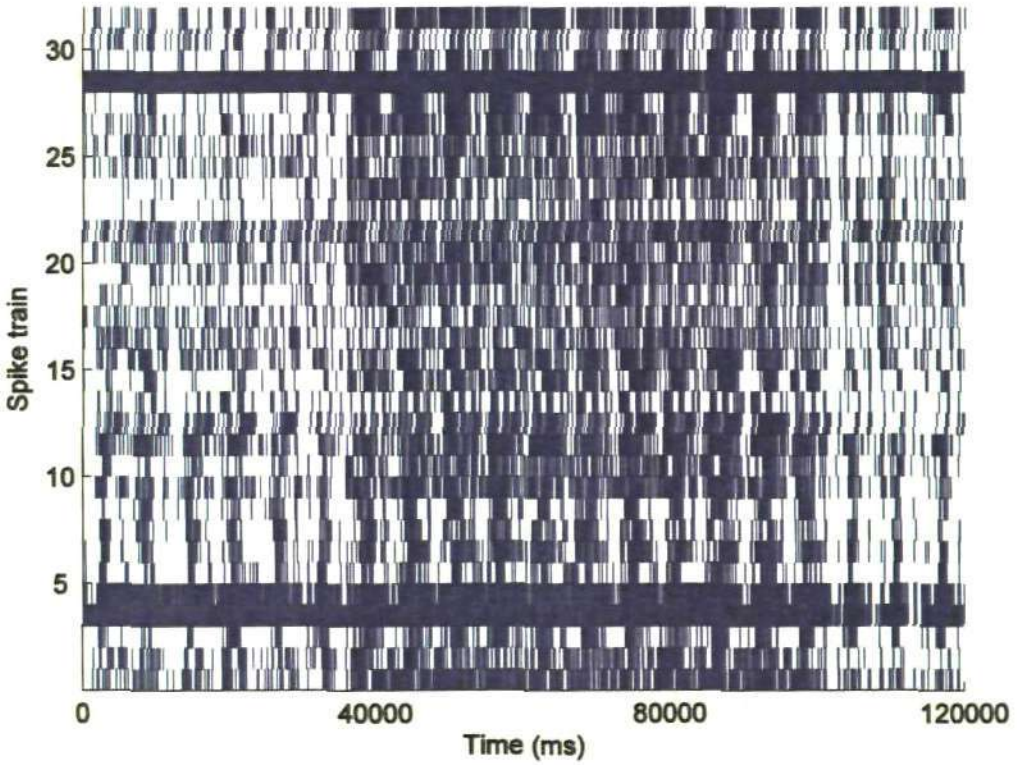


Figure B.19: Raster plot of 32 spike trains of stimulus 5. Spike trains #4, #5 and #29 have high spiking rates and are not considered for analysing functional connectivity.

B.4.1 MCG method

All possible pair-wise CCF are calculated with a bin size of 1 ms and a correlation window of 100 ms for the 29 spike trains. To test the independence of two spike trains the significance level $\alpha = 0.05$ is used with the Bonferroni correction. A connection is considered significant if a peak of the CCF exceeds the upper boundary of the confidence interval. A total of 175 significant connections are found for 29 spike trains. These significant connections are shown in a matrix format in Fig. B.20(a) where the connections are indicated by circles. The direction of connection is considered from the reference spike train to the target spike train. Among the 29 spike trains, spike trains

#30 and #32 have 17 outgoing connections to other spike trains and 16 incoming connections from other spike trains, which are the highest among 29 spike trains.

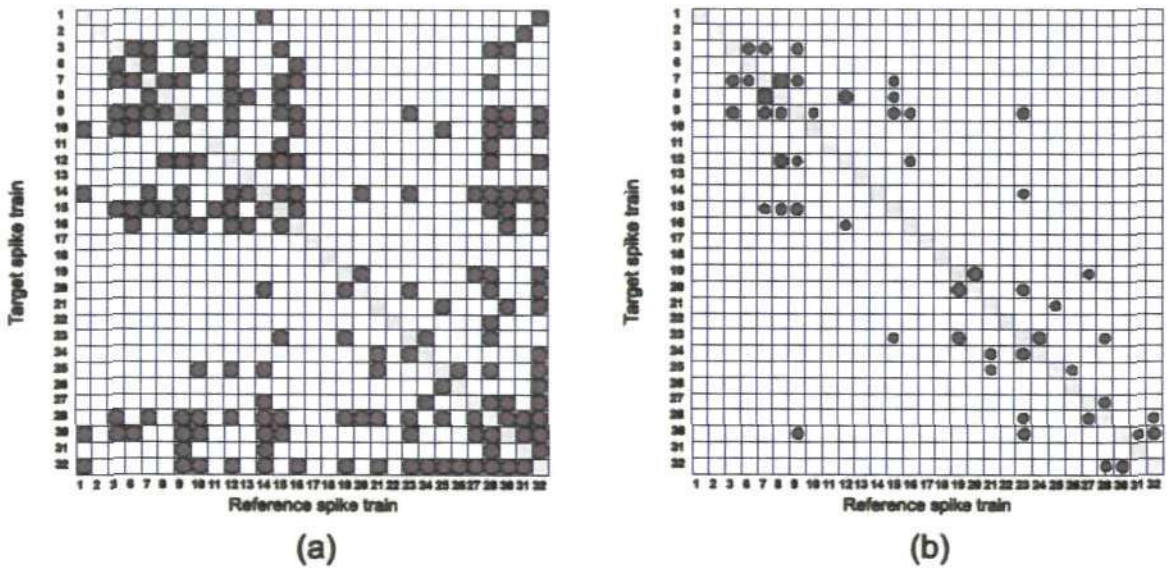


Figure B.20: (a) Significant connections obtained from pair-wise CCF analysis of the 29 spike trains of stimulus 5. **(b)** Direct connections obtained from the connections in (a) after the clustering algorithm. The radius of the circle indicates the strength of the connections.

A clustering algorithm is applied to the 175 significant connections for distinguishing direct connections from spurious ones (i.e., indirect connections and common source). After clustering a total of 49 connections are identified as direct connections (Fig. B.20(b)). The radius of the circle indicates the strength of connection. Among 49 connections, 9 connections have strong strength. These connections are: (#7, #8), (#8, #7), (#8, #12), (#12, #8), (#19, #20), (#19, #23), (#20, #19), (#23, #24) and (#24, #23). Spike train #23 has 6 outgoing connections to spike trains #9, #14, #20, #24, #28 and #30; and spike train 9 has 7 incoming connections from spike trains #3, #7, #8, #10, #15, #16 and #23, which are the highest among 29 spike trains. There are 16 pairs of connections where both spike trains have functional connectivity to each other. These pairs are: (#3, #7), (#3, #9), (#7, #8), (#7, #9), (#7, #15), (#8, #12), (#8, #15), (#9, #15),

(#12, #16), (#19, #20), (#21, #25), (#23, #24), (#23, #28), (#27, #28), (#28, #32) and (#30, #32).

B.4.2 Cox method

Cox method is applied to the 29 spike trains to analyse functional connectivity, considering one spike train as target spike train and other 28 spike trains as reference spike trains. The influence function which determines how the reference spike train influences the target spike train should be specified with parameters. All the influence functions are assumed identical. The inter spike interval (ISI) histogram of three spike trains, spike train #17, #18 and #22 are given in (Fig. B.21). These histograms have high count for the short ISI and the ISI count decreases with increase of the ISI length. That suggests that the influence function should be specified by the formula (5.3). The parameters of the influence function (5.3) are $\tau_r = 0.1\,ms$, $\tau_s = 10\,ms$. Another parameter, the time lag Δ is specified from pair-wise CCF analysis. Thus, the influence functions are defined and the Cox coefficients and the corresponding confidence intervals are calculated using formulas (5.5) and (5.6). This procedure is repeated 29 times to obtain the full functional connectivity of the 29 spike trains. The confidence intervals are calculated using the significance level $\alpha = 0.05$ with Bonferroni correction.

A total of 116 connections are identified by the Cox method which are shown by circles in Fig. B.22(a). The radius of the circle indicates the strength of functional connection. The direction of functional connection is from the reference spike train to the target spike train.

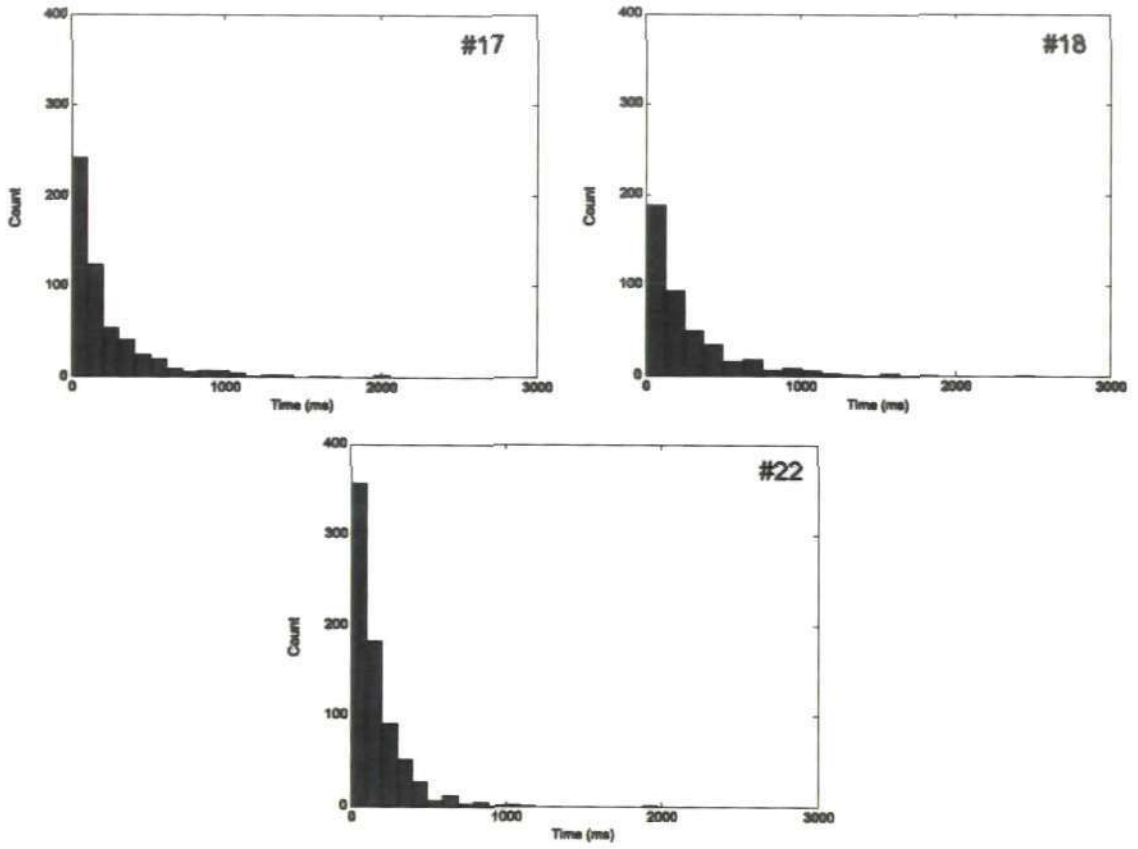


Figure B.21: Inter spike interval histogram of the spike trains #17, #18 and #22 of stimulus 5.

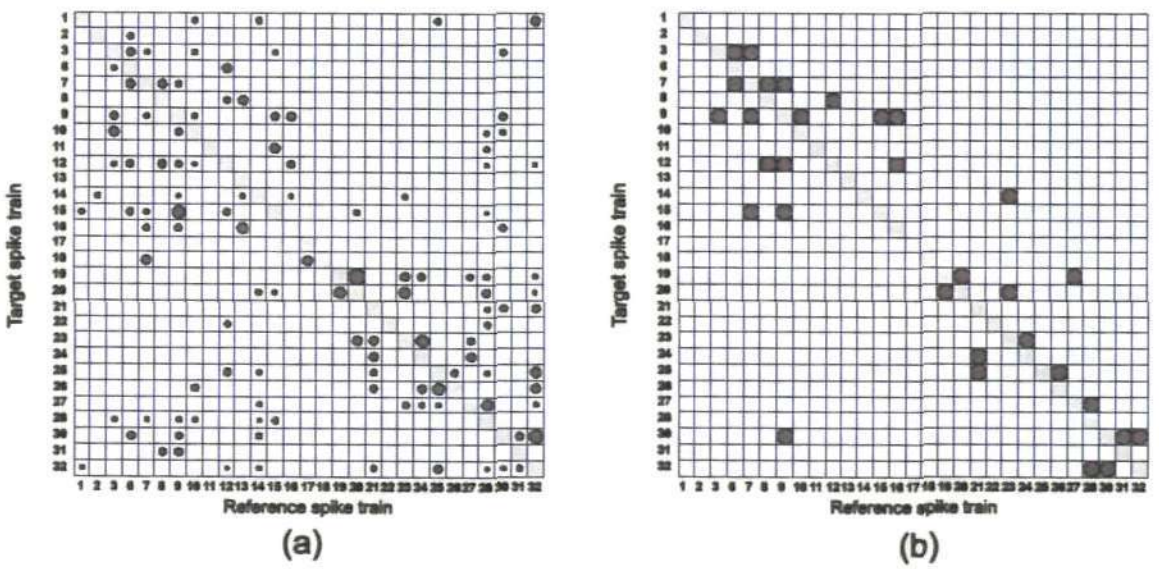


Figure B.22: **(a)** Functional connectivity of the 29 spike trains identified by the Cox method of stimulus 5. Radius of the circle indicates strength of connection. **(a)** The connections that are identified both by the MCG method and the Cox method.

Among 116 connections, 8 connections have strong strength compare to others. These connections are: (#9, #15), (#13, #16), (#19, #20), (#20, #19), (#24, #23), (#25, #26), (#28, #27) and (#32, #30). 5 connections have weak strength compared to others. These connections are: (#9, #14), (#12, #32), (#14, #28), (#28, #15) and (#28, #32). All other connections have a medium strength of connection. Spike train #28 has 11 outgoing connections and spike trains #12 and #32 have 8 incoming connections, which are the highest among 29 spike trains. There are 22 pairs of connections where both spike trains have functional connectivity to each other.

Functional connectivity obtained by the MCG method and the Cox method show a good agreement between them (Fig. B.22(b)). There are 31 connections which are common in both techniques. Among the common connections spike train #9 has 4 outgoing connections to spike trains #7, #12, #15 and #30 and 5 incoming connections from spike trains #3, #7, #10, #15 and #16. There are 5 pairs of connections where both spike trains have functional connectivity to each other. These connections are: (#7, #9), (#8, #12), (#9, #15), (#19, #20) and (#30, #32).

B.4.3 Cox metric

All possible pairs of spike trains are analysed by pair-wise Cox method taking one spike train as a target and the other spike train as a reference. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval using (5.6). The Cox metric is applied to the significant connections to reveal the groups of similar spike trains. The result of Cox metric is shown in Fig. B.23(a) where the grey circles indicate the significant connections obtained by the pair-

wise Cox method. The black circles indicate symmetric of the grey circles but not identified by the pair-wise analysis of Cox method. Similarly, the Cox metric is applied to the functional connections identified by the Cox method considering all spike trains at once. This functional connection is shown in Fig. B.22(a) and the result of the Cox metric is shown in Fig. B.23(b). In the figure the grey circles indicate the significant connections obtained by the Cox method considering all spike trains at once. The black circles indicate symmetric of the grey circles but not identified by the Cox method considering all spike trains at once.

From Fig. B.23(a) five groups of similar spike trains are identified. The First group consists of 7 spike trains, these are: spike trains #19, #20, #21, #23, #24, #27 and #28. In this group, the connection from spike train #19 to #20 and spike train #20 to #19 have the highest strength among other connections. All other connections have a medium strength. Spike train #28 has connections to all other spike trains and similar spike train #23 has incoming connections from all other spike trains. Thus from this group it can be concluded that spike train #28 is the most influential spike train. The second group consists of 6 spike trains; these are: spike trains #1, #14, #25, #26, #30 and #32. In this group, all the connections have medium strength. Spike train #32 has 4 outgoing connections to spike trains #1, #25, #26 and #30 and 5 incoming connections from all the spike trains. This spike train can be considered as the most influential spike train for this group. The third group also consists of 6 spike trains; these are: spike trains #3, #10, #6, #12, #9 and #15. In this group, spike train #3 has 4 out going connections to all other spike trains except spike train #15. Similarly, spike train #12 has 4 incoming connections from all other spike trains except spike train #15. In this group, spike train #3 is considered as the most influential spike train. The fourth group consists of 3 spike trains; spike trains #7, #18 and #8. The last group consists of 2 spike trains; spike trains #13 and #16. Spike trains #31, #11, #22, #17 and #2 do not form any group.

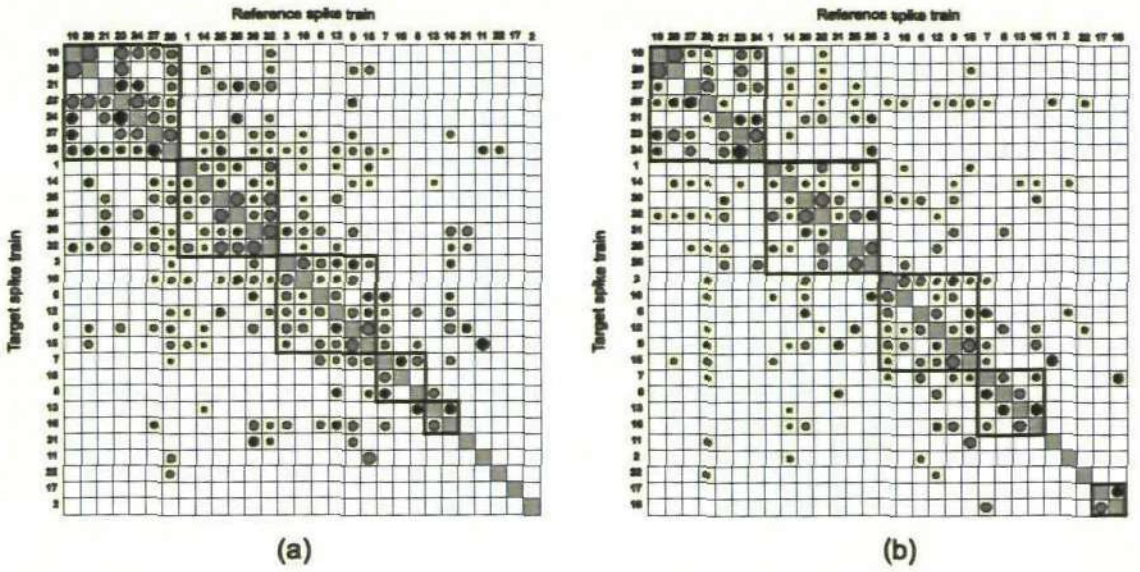


Figure B.23: *Groups of similar spike trains revealed by the Cox metric of the 29 spike trains of stimulus 5.* **(a)** Cox metric using pair-wise analysis. **(b)** Cox metric considering all spike trains at once.

Similarly, from Fig. B.23(b) five groups of similar spike trains are identified. The First group consists of 7 spike trains, these are: spike trains #19, #20, #27, #28, #21, #23 and #24. In this group, the connection from spike train #19 to #20 and spike train #20 to #19 have the highest strength among other connections. All other connections have a medium strength. Spike train #28 has connections to spike trains #19, #20, #27 and #21. From this group it can be concluded that spike train #28 is the most influential spike train. The second group consists of 7 spike trains; these are: spike trains #1, #14, #30, #32, #31, #25 and #26. In this group, the connection from spike train #30 to #32 and spike train #32 to #30 have the highest strength among other connections. Spike train #32 has 4 outgoing connections to spike trains #1, #30, #25 and #26 and 5 incoming connections from all the spike trains. This spike train can be considered as the most influential spike train for this group. The third group also consists of 6 spike trains; these are: spike trains #3, #10, #6, #12, #9 and #15. In this group, spike train #3 has 4 out going connections to all other spike trains except spike train #15. Spike train #3 is considered as the most influential spike train. The fourth group consists of 4 spike trains;

spike trains #7, #8, #13 and #16. The last group consists of 2 spike trains; spike trains #17 and #18. Spike trains #11, #2 and #22 do not form any group.

Investigation from Fig. B.23(a)-(b) reveal that the Cox metric identified by the pair-wise analysis and considering all spike trains at once show a good agreement. For example, the first and third group of both figures consists the same spike trains. Similarly, the other three groups of both figures consist same spike trains except a few. From this analysis it can be concluded that application of Cox metric to experimental data using pair-wise analysis and considering all spike trains at once enables to create similar result.

B.4.4 Motif analysis

To find the pattern of interconnections among the spike trains a structural motif analysis is conducted using triplet-wise analysis of Cox method. A total of 3654 triplets are analysed. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval using (5.6). Functional connectivity of each triplet spike trains is used for identifying the structural motif. The structural motif count is obtained by analysing all 3654 triplets of spike trains which is shown in Fig. B.24.

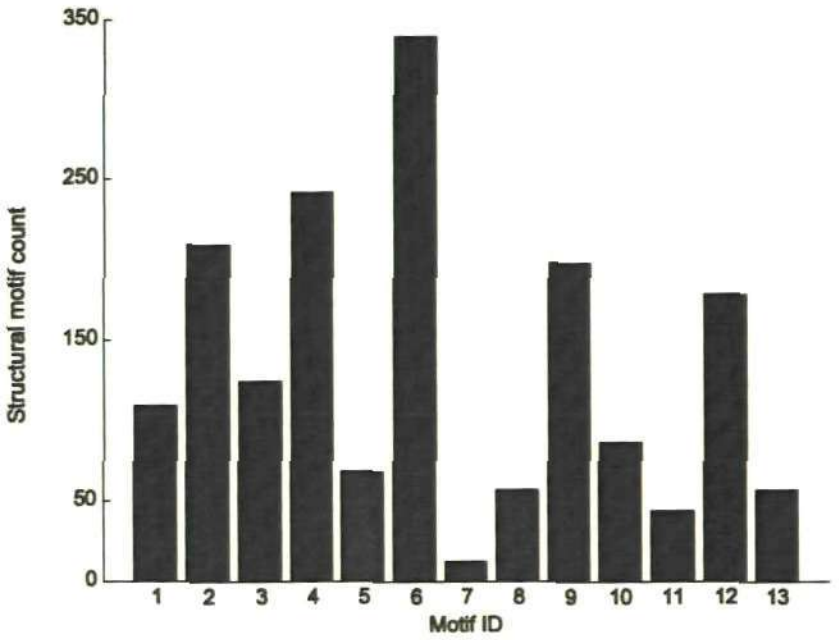


Figure B.24: Structural motif count of all 3654 triplets of the 29 spike trains of stimulus 5.

Out of 3654 triplets, 1724 groups have different structural motifs ID's. Among the 1724 triplets, 339 triplets have motif ID 6 which is the highest among others motif ID's. Only 12 triplets have motif ID 7 which is the lowest. Motif ID's 2, 4, 9 and 12 have similar number of triplets; motif ID's 1 and 3 have similar number of triplets. A total of 532 triplets have connected motifs (connected motifs are motif ID 7, 9, 10, 12 and 13). On the other hand, a total of 1938 triplets have unconnected motifs. Thus there are low proportions of connected motifs (27.45%) in the groups of triplet spike trains which indicate that connection from every spike train to every other spike trains is low.

B.5 Analysis of functional connectivity of stimulus 6

The raster plot of 32 spike trains is shown in Fig. B.25. Similar to all the 5 stimuli spike trains #4, #5 and #29 have high spiking rates compared to all other spike trains. Therefore, these three spike trains are considered to be outliers and they are excluded from analysis. All the 29 spike trains have similar spiking pattern. Spiking rates of these 29 spike trains are high over time interval [55000 ms, 95000 ms]. Like stimulus 2, there are no spikes over time interval [95000 ms, 101000 ms] of the 29 the spike trains except

spike train #1. These 29 spike trains are analysed for identifying functional connectivity between them keeping the original numeration of the spike trains.

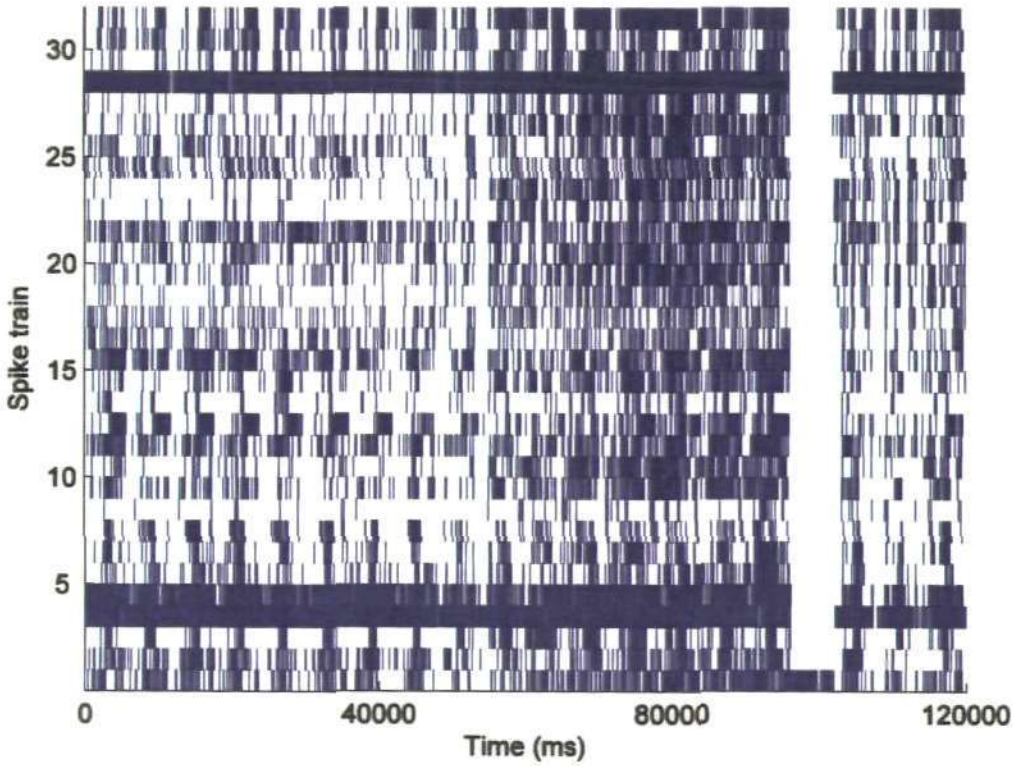


Figure B.25: Raster plot of 32 spike trains of stimulus 6. Spike trains #4, #5 and #29 have high spiking rates and are not considered for analysing functional connectivity.

B.5.1 MCG method

All possible pair-wise CCF are calculated with a bin size of 1 ms and a correlation window of 100 ms for the 29 spike trains. To test the independence of two spike trains the significance leve $\alpha = 0.05$ is used with the Bonferroni correction. A connection is considered significant if a peak in the cross-correlation function exceeds the upper boundary of the confidence interval. A total of 214 significant connections are identified for 29 spike trains. These significant connections are shown in a matrix format in Fig. B.26(a) where the connections are indicated by circles. The direction of connection is considered from the reference spike train to the target spike train. Among the 29 spike trains, spike train #32 have 17 outgoing connections to other spike trains and 18

incoming connections from other spike trains, which are the highest among 29 spike trains.

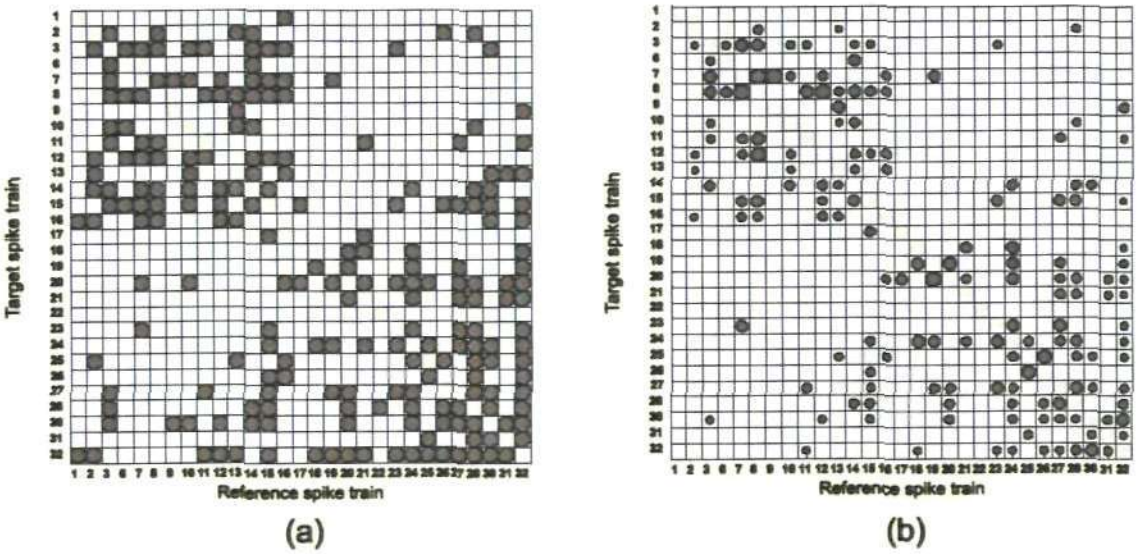


Figure B.26: (a) Significant connections obtained from pair-wise CCF analysis of the 29 spike trains of stimulus 6. **(b)** Direct connections obtained from the connections in (a) after the clustering algorithm. The radius of the circle indicates the strength of the connections.

Application of clustering algorithm to the 214 significant connections identified 154 direct connections (Fig. B.26(b)). The radius of the circle indicates the strength of connection. Among 154 connections, 6 connections have strong strength compared to others. These connections are: (#7, #8), (#8, #7), (#8, #12), (#9, #7), (#12, #8) and (#19, #20). Spike train #32 has 14 outgoing connections and spike train #30 has 10 incoming connections, which are the highest among 29 spike trains. There are 48 pairs of connections where both spike trains have functional connectivity to each other.

B.5.2 Cox method

Application of Cox method to the 29 spike trains requires the identification of target spike train, reference spike trains and the influence function. The inter spike interval (ISI) histogram of three spike trains, spike train #17, #18 and #19 are given in Fig. B.27. These histograms have high count for the short ISI and the ISI count decreases with

increase of the ISI length. That suggests that the influence function should be specified by the formula (5.3). The parameters of the influence function (5.3) are $\tau_r = 0.1\ ms$, $\tau_s = 10\ ms$. Another parameter, the time lag Δ is specified from pair-wise CCF analysis. Thus, the influence functions are defined and the Cox coefficients and the corresponding confidence intervals are calculated using formulas (5.5) and (5.6). This procedure is repeated 29 times to obtain the full functional connectivity of the 29 spike trains. The confidence intervals are calculated using the significance level $\alpha = 0.05$ with Bonferroni correction.

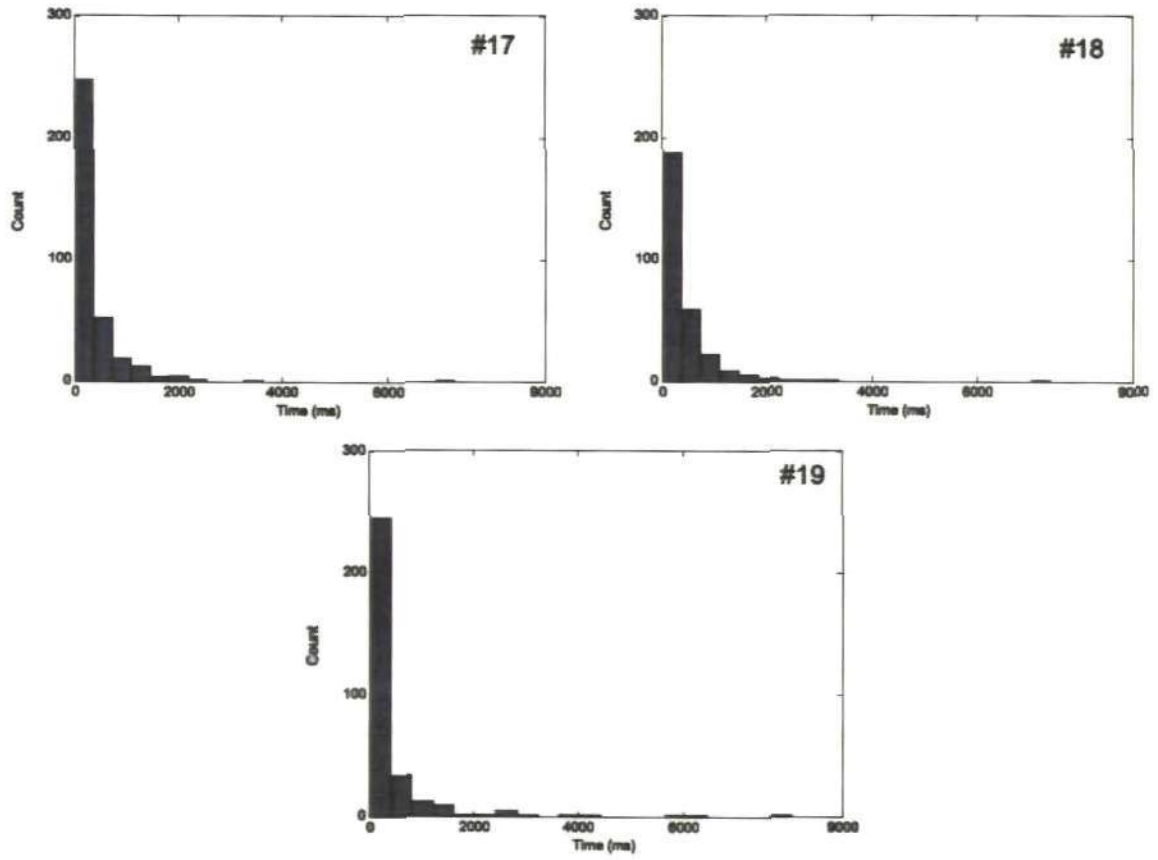


Figure B.27: Inter spike interval histogram of the spike trains #17, #18 and #19 of stimulus 6.

The 76 connections, identified by the Cox method, are shown by circles in Fig. B.28(a). The radius of the circle indicates the strength of functional connection. The direction of functional connection is from the reference spike train to the target spike train. Among 76 connections, 3 connections have strong strength compared to others. These

connections are: (#1, #9), (#14, #8) and (#32, #30). 4 connections have weak strength compared to others. These connections are: (#7, #3), (#16, #13), (#32, #13) and (#32, #28). All other connections have a medium strength of connection. Spike train #32 has 9 outgoing connections and spike trains #14 and #15 have 6 incoming connections, which are the highest among 29 spike trains. There are 17 pairs of connections where both spike trains have functional connectivity to each other.

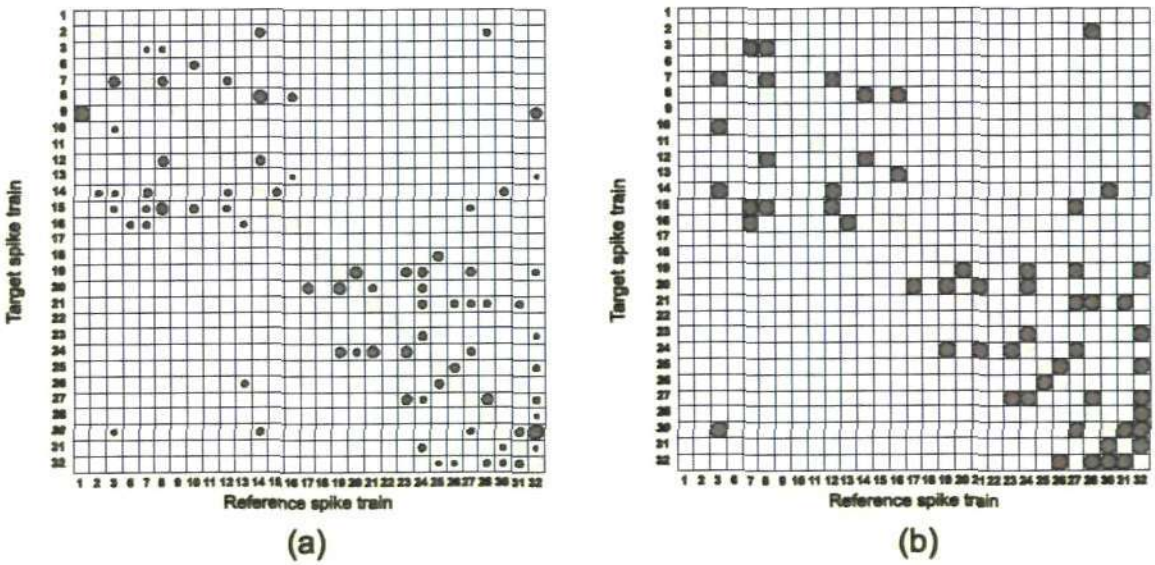


Figure B.28: (a) Functional connectivity of the 29 spike trains identified by the Cox method of stimulus 6. Radius of the circle indicates strength of connection. **(b)** The connections that are identified both by the MCG method and the Cox method.

Functional connectivity obtained by the MCG method and the Cox method show a good agreement between them (Fig. B.28(b)). There are 57 connections which are common in both techniques. Among the common connections spike train #32 has 8 outgoing connections to spike trains #9, #19, #23, #25, #27, #28, #30 and #31. Spike trains #15, #19, #20, #24, #27, #30 and #32 have 4 incoming connections from other spike trains. There are 12 pairs of connections where both spike trains have functional connectivity to each other.

B.5.3 Cox metric

Application of Cox metric to the 29 spike trains requires the analysis of 812 possible pairs of spike trains by pair-wise Cox method where one spike train is taken as a target and the other spike train as a reference. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\,ms$ and $\tau_s = 10\,ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval using (5.6). The Cox metric is applied to the significant connections to reveal the groups of similar spike trains. The result of Cox metric is shown in Fig. B.29(a) where the grey circles indicate the significant connections obtained by the pair-wise Cox method. The black circles indicate symmetric of the grey circles but not identified by the pair-wise analysis of Cox method. Similarly, the Cox metric is applied to the functional connections identified by the Cox method considering all spike trains at once. This functional connection is shown in Fig. B.28(a) and the result of Cox metric is shown in Fig. B.29(b). In the figure the grey circles indicate the significant connections obtained by the Cox method considering all spike trains at once. The black circles indicate symmetric of the grey circles but not identified by the Cox method considering all spike trains at once.

From Fig. B.29(a) four groups of similar spike trains are identified. The first group consists of 2 spike trains, spike trains #1 and #9, where there is a strong connection from #1 to #9. The second group consists of 7 spike trains; these are: spike trains #3, #7, #8, #15, #14, #12 and #10. In this group, all the connections have a medium strength. Spike train #3 has 4 outgoing connections to spike trains #7, #15, #14 and #10 and spike train #15 has 5 incoming connections from all other spike trains except spike train #14.

Spike train #3 can be considered as the most influential spike train for this group. The third group consists of 13 spike trains; these are: spike trains #18, #25, #26, #30, #32, #31, #19, #20, #21, #24, #27, #28 and #23. In this group, there are 6 strong strength of connections (#30, #32), (#32, #30), (#19, #20), (#20, #19), (#24, #19), (#19, #24) and (#23, #19). All other connections have a medium strength. Spike train #32 has outgoing connections to all other spike trains except spike train #26. This spike train can be considered as the most influential spike train. The fourth group consists of 2 spike trains, spike trains #13 and #16, where there is a connection from spike train #13 to #16. Spike trains #2, #17, #11, #6 and #22 do not form any group.

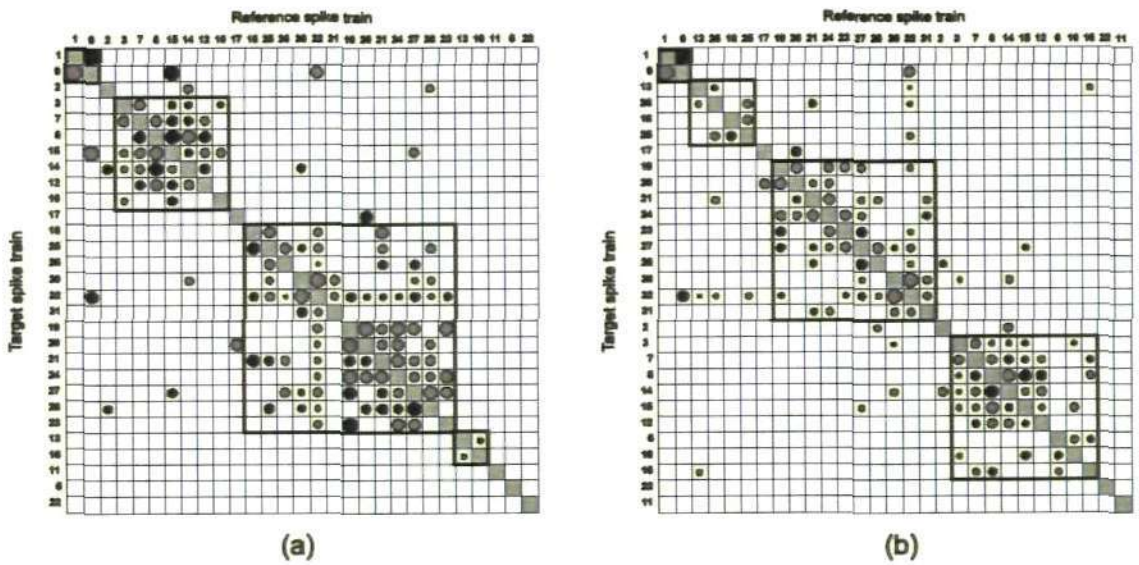


Figure B.29: Groups of similar spike trains revealed by the Cox metric of the 29 spike trains of stimulus 6. **(a)** Cox metric using pair-wise analysis. **(b)** Cox metric considering all spike trains at once.

Similarly, from Fig. B.29(b) four groups of similar spike trains are identified. The first group consists of 2 spike trains, spike trains #1 and #9, where there is a strong connection from #1 to #9. The second group consists of 4 spike trains; these are: spike trains #13, #26, #18 and #25. In this group, all the connections have a medium strength. The third group consists of 10 spike trains; these are: spike trains #19, #20, #21, #24, #23, #27, #28, #30, #32 and #31. In this group, there are 2 strong strength of

connections: (#30, #32) and (#32, #30). All other connections have a medium strength. Spike train #24 and #32 have 6 outgoing connections and can be considered as the most influential spike trains. The fourth group consists of 9 spike trains; spike trains #3, #7, #8, #14, #15, #12, #6, #10 and #16. In this group spike trains #3, #7 and #8 have 4 outgoing connections to other spike trains. Therefore, these spike trains are considered as the most influential spike trains for this group. Spike trains #17, #2, #22 and #11 do not form any group.

Investigation from Fig. B.29(a)-(b) reveal that the Cox metric identified by the pair-wise analysis and considering all spike trains at once show a good agreement. For example, the first group of both figures consists the same two spike trains (#1 and #9). Similarly, the other three groups of both figures consist same spike trains except a few. From this analysis it can be concluded that application of the Cox metric to experimental data using pair-wise analysis and considering all spike trains at once enables to create similar result.

B.5.4 Motif analysis

To find the pattern of interconnections among the 29 spike trains, a structural motif analysis is done using triplet-wise analysis of Cox method. For 29 spike trains a total of 3654 triplets are analysed. All the influence functions are considered identical and specified by (5.3) with the parameters $\tau_r = 0.1\ ms$ and $\tau_s = 10\ ms$. Another parameter of the influence function, the time lag Δ is obtained by the pair-wise CCF analysis. Using the parameter values, the influence functions are determined and the Cox coefficients are estimated using (5.5) with corresponding confidence interval using (5.6). Functional connectivity of each triplet spike trains is used to identify the structural motif. The structural motif count is obtained by analysing all 3654 triplets of spike trains which is shown in Fig. B.30.

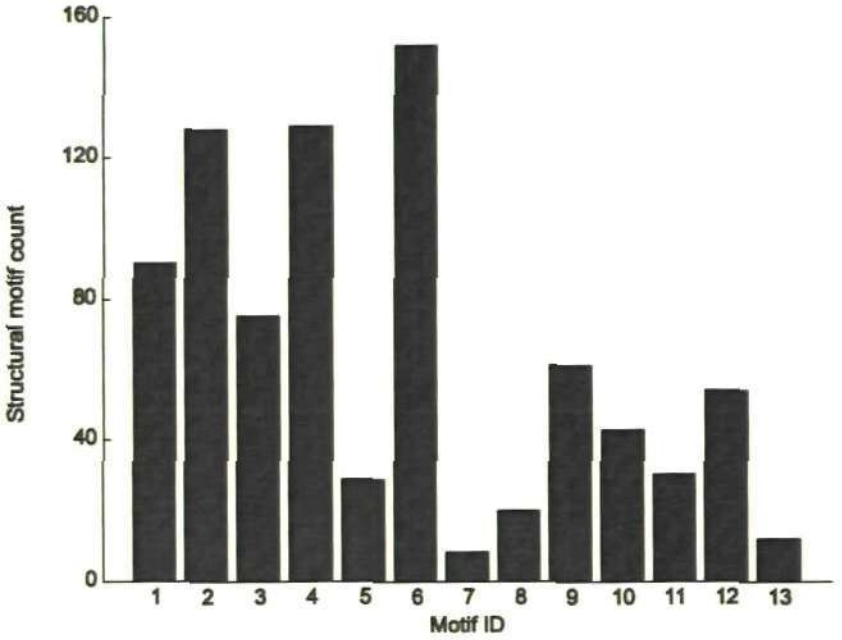


Figure B.30: Structural motif count of all 3654 triplets of the 29 spike trains of stimulus 6.

Out of 3654 triplets, 831 triplets have different structural motifs ID's. Among the 831 triplets, 152 triplets have motif ID 6 which is the highest among others motif ID's. Only 8 triplets have motif ID 7 which is the lowest. Motif ID's (2, 4); (1, 3); (9, 12) and (5, 11) have similar number of triplets. A total of 178 triplets have connected motifs (connected motifs are motif ID 7, 9, 10, 12 and 13). On the other hand, a total of 653 triplets have unconnected motifs. Thus there are low proportions of connected motifs (27.26%) in the groups of triplet spike trains which indicate that connection from every spike train to every other spike trains is low.

Appendix C

In chapter 7, some graph theoretical methods are applied to analyse the connectivity matrix of 29 spike trains. The results of analysing the first stimulus are presented in chapter 7. In this appendix the results of analysing the five stimuli are presented.

C.1 Analysis of connectivity of stimulus 2

The connection matrix of the 29 spike trains is shown in Fig. C.1(a). There are 62 connections of the connection matrix and the connection matrix contains a low density (0.0764). This indicates that in stimulus 2 the spike trains are not densely connected like stimulus 1. The sum of indegree and outdegree of the spike trains is shown in Fig. C.1(b). The degree of the spike trains varies widely from 0 to 13 showing the same number of degrees for certain spike trains. There are 4 spike trains that do not have any degree (#6, #17, #18 and #21). Some spike trains have high degree (#32, #28, #3 and #27); whereas some spike trains have low degree (#2, #9, #11, #13 and #22).

The characteristic path length of the connection matrix (2.7747) is greater than the characteristic path length obtained from a random network (2.6669). This characteristic path length indicates that, on average, to pass information from one spike train to another spike train, it takes approximately 3 edges like stimulus 1. Similarly, the global efficiency of the connection matrix (0.1871) (random: 0.2336) indicates that pairs of spike trains, on average, have long communication distances.

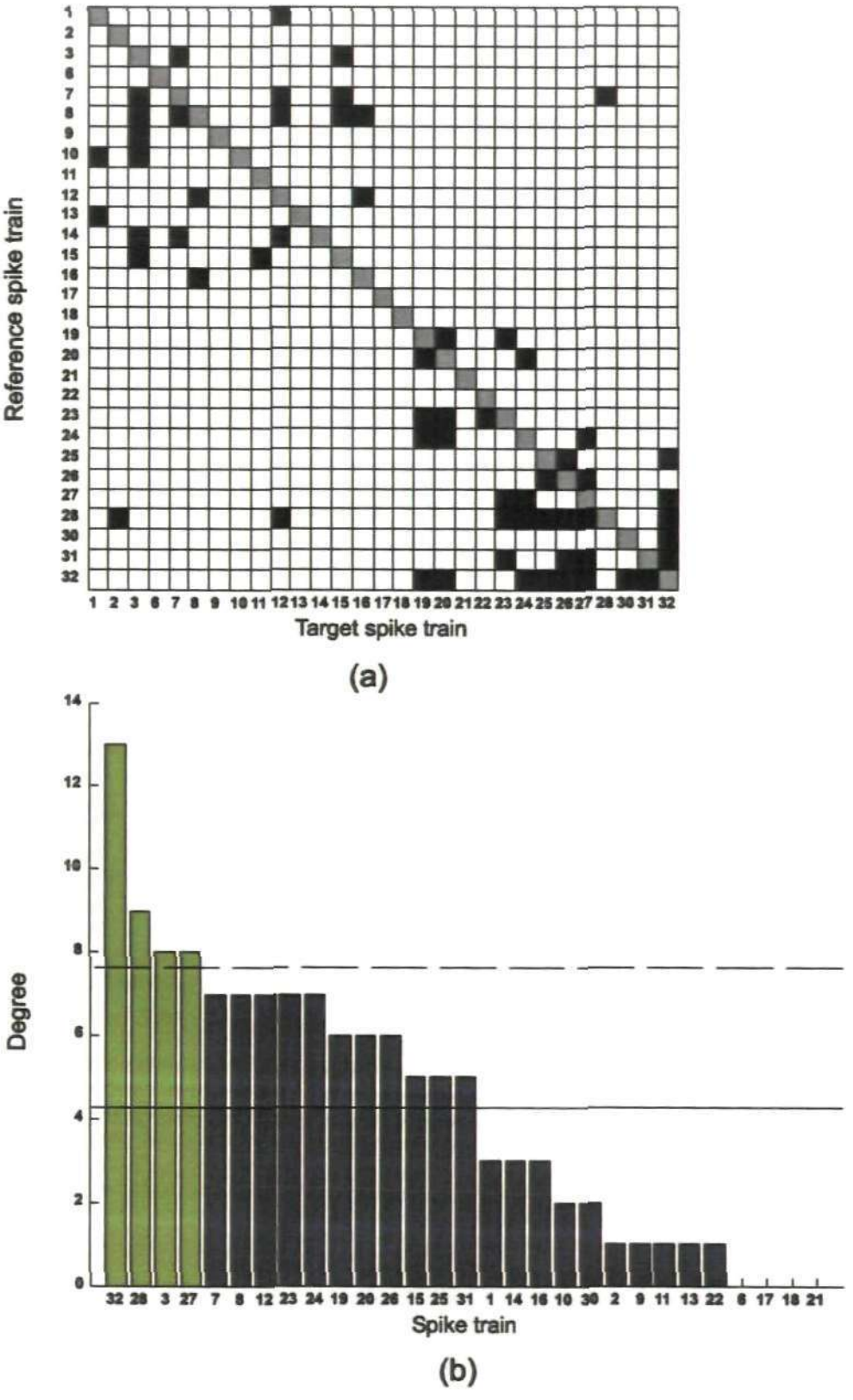


Figure C.1: **(a)** Connection matrix of the 29 spike trains of stimulus 2. Connection patterns are represented by the presence of connection (black square) and absence of connection (white square). Main diagonals are indicated in grey and self-connections are excluded. **(b)** Degree of the spike trains displayed in descending order. The solid horizontal line indicates the mean degree of the spike trains and the dashed horizontal line indicates the mean plus one standard deviation. High-degree spike trains are displayed as green.

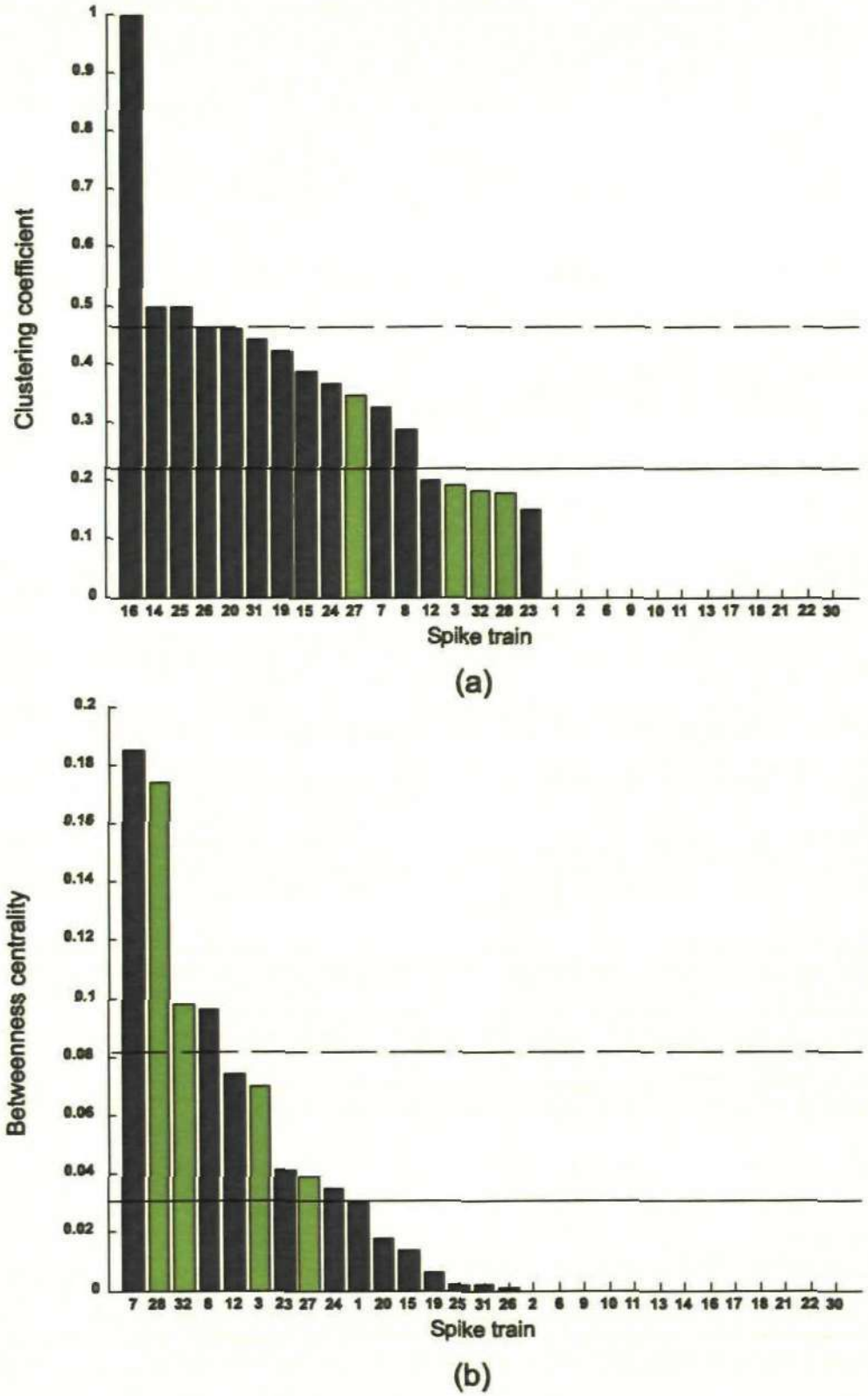
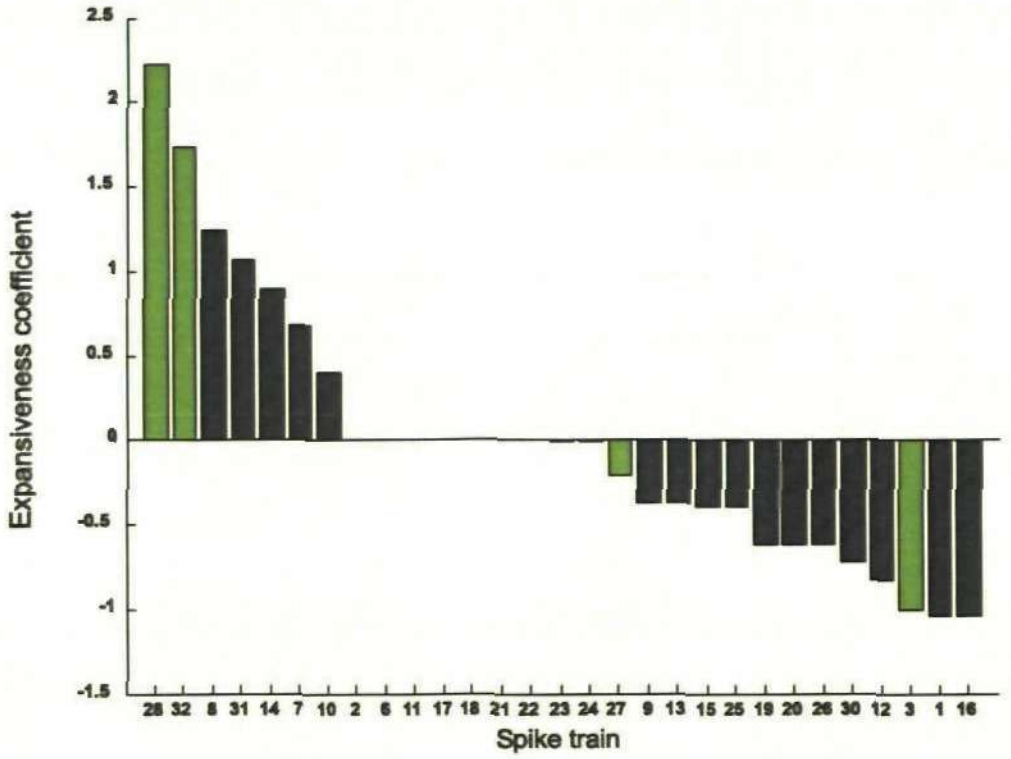
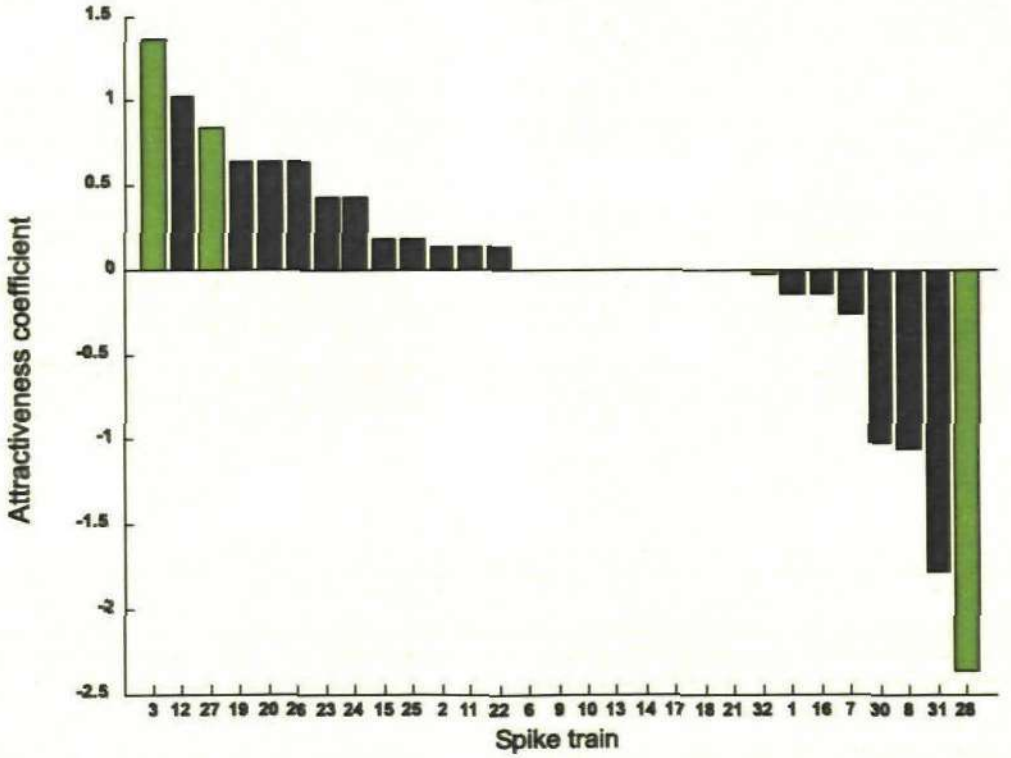


Figure C.2: (a) Clustering coefficient and betweenness centrality of the 29 spike trains of stimulus 2. The solid horizontal line indicates the mean and the dashed horizontal line indicates the mean plus one standard deviation. High-degree spike trains are displayed as green. **(a)** Clustering coefficient of 29 spike trains is displayed in descending order. **(b)** Betweenness centrality of the 29 spike trains is displayed in descending order.

Clustering coefficient of the spike trains is shown in Fig. C.2(a). Clustering coefficient of spike train #16 is 1 which means that this spike train completely makes a cluster with its neighbour spike trains. Investigation reveals (connection matrix) that spike train #16 has two neighbour spike trains (#12 and #8) and they are connected to each other. The clustering coefficient of the remaining spike trains ranges from 0 to 0.5. Some spike trains have high clustering coefficient (#14 and #25) compare to other spike trains indicating that the neighbours of these spike trains are also neighbours of each other. There are some spike trains which have low clustering coefficient (#12, #3, #32, #28 and #23), in fact below the mean of all the spike trains. Among the low clustering coefficient, three spike trains (#3, #32 and #28) have the high degree of all the spike trains which indicates that these spike trains communicate to other neighbour spike trains but the neighbour spike trains are not connected to each other. There are 12 spike trains that do not form any cluster to their neighbour spike trains. The clustering coefficient of these spike trains is zero. The global clustering coefficient (0.2213) (random: 0.1078) also indicates that many spike trains do not have neighbours which are connected to each other. Fig. C.2(b) shows the betweenness centrality of the spike trains. There are some central spike trains (#7, #28, #32 and #8) which transfer most of the information to the other spike trains. Among the central spike trains, spike trains #28 and #32 have the highest degree. This means that these two spike trains communicate to other spike trains through incoming and outgoing connections. There are 13 spike trains which do not pass any information to other spike trains. That means the betweenness centrality of these spike trains is zero.



(a)



(b)

Figure C.3: Expansiveness and attractiveness coefficient of the P1 model of the 29 spike trains of stimulus 2. High-degree spike trains are displayed as green. **(a)** Expansiveness coefficient displayed in descending order. **(b)** Attractiveness coefficient displayed in descending order.

There are two influential spike trains (#28 and #32) among the 29 spike trains (Fig. C.3(a)). Investigation from the connection matrix reveals that these spike trains have the high outdegree (8 each) and they have also high degree. There are 7 spike trains that do not show any expansiveness. Spike trains #1 and #16 have the most negative expansiveness coefficient. Both the spike trains have 2 indegree and 1 outdegree. Among the high degree, spike trains #27 (5 indegree, 3 outdegree) and #3 (6 indegree, 2 outdegree) show negative expansiveness coefficient. Spike trains #3, #12 and #27 are the most attractive spike trains (Fig. C.3(b)) as they have high indegree (6, 5 and 5 respectively) and two of them have the high degree (#3 and #27). Similar to expansiveness coefficient 8 spike trains do not show any attractiveness as their indegree are zero. Spike train #28 has the most negative attractiveness coefficient (8 outdegree and 1 indegree) and this spike train has the high degree.

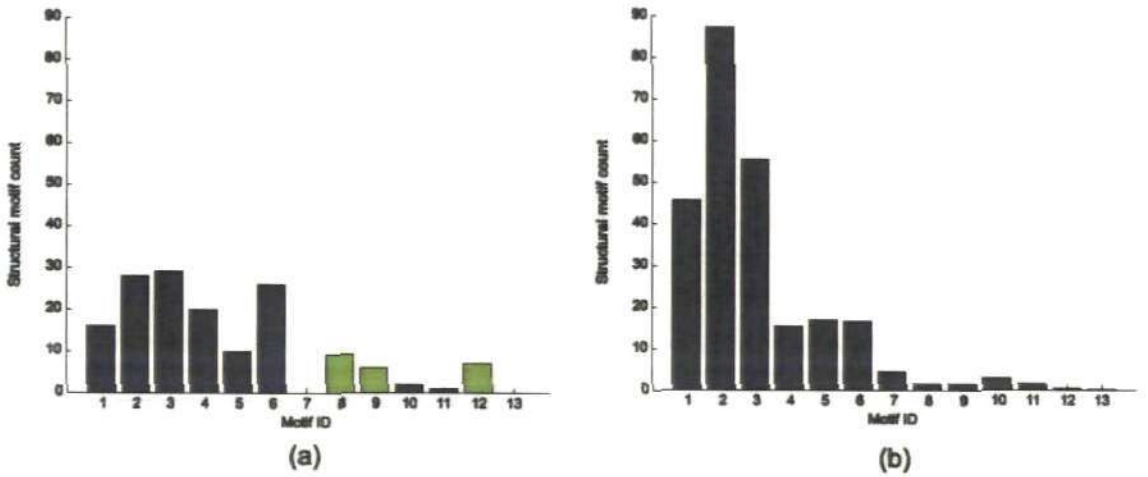


Figure C.4: a) Structural motif count of size m = 3 of the 29 spike trains of stimulus 2. Significant motif ID's are displayed as green. **(b)** Structural motif count of size m = 3 for the randomized diagram.

Fig. C.4(a) shows the motif count for structural motifs of size m = 3 found in the connection matrix of 29 spike trains. Motif ID 3 appears 29 times which is the highest among the motif ID's. Motif ID's 7 and 13 have no appearance in the connection matrix. To find the significant motif, 1000 random networks are generated keeping the same indegree and outdegree of the spike trains. The structural motif count of size m =

3 for the random network is shown in Fig. C.4(b). The motif ID's 8, 9 and 12 appear more than the random network. The Z-score of these motif ID's ($Z_8 \approx 5.98,\ p < .0001; Z_9 = 2.93,\ p \approx .005; Z_{12} = 9.02,\ p < .0001$) indicate that they are significant. There are a very low proportion of connected motifs (10.79%) in the connection matrix indicating that the spike trains are weakly connected.

C.2 Analysis of connectivity of stimulus 3

The connection matrix of the 29 spike trains is shown in Fig. C.5(a). There are 95 connections of the connection matrix and it contains a low density (0.1170) which means that in stimulus 3, the spike trains are not densely connected. The sum of indegree and outdegree of the spike trains is shown in Fig. C.5(b). The degree of the spike trains varies widely from 1 to 21 showing the same number of degrees for certain spike trains. Some spike trains have high degree (#32, #24 and #21); whereas some spike trains have low degree (#2, #17 and #18).

The characteristic path length of the connection matrix (2.8137) is greater than the characteristic path length obtained from a random network (2.5013). This characteristic path length indicates that, on average, to pass information from one spike train to another spike train, it takes approximately 3 edges. Similarly, the global efficiency of the connection matrix (0.3428) (random: 0.3955) indicates that pairs of spike trains, on average, have long communication distances.

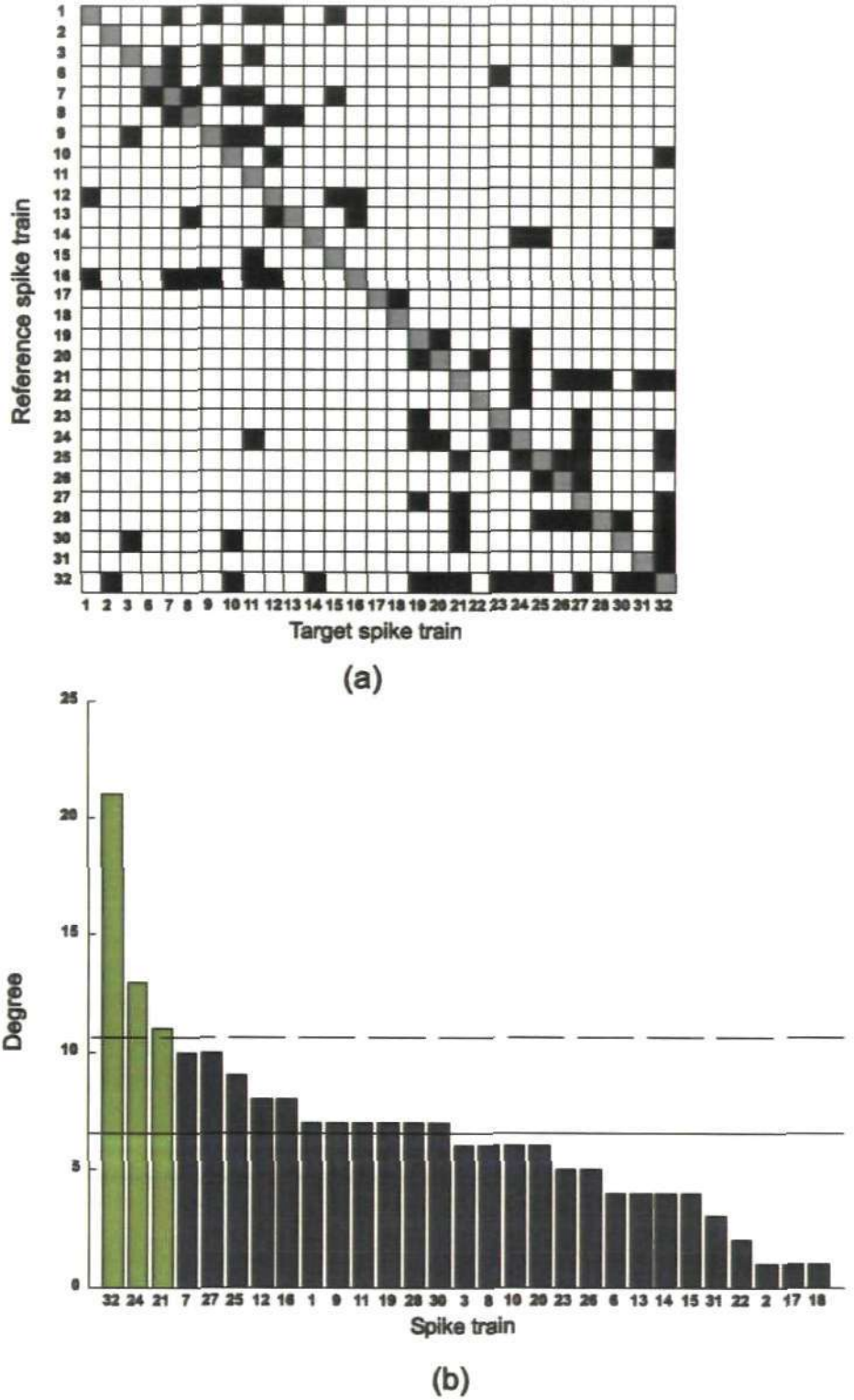


Figure C.5: **(a)** Connection matrix of the 29 spike trains of stimulus 3. Connection patterns are represented by the presence of connection (black square) and absence of connection (white square). Main diagonals are indicated in grey and self-connections are excluded. **(b)** Degree of the spike trains is displayed in descending order. The solid horizontal line indicates the mean degree of the spike trains and the dashed horizontal line indicates the mean plus one standard deviation of the spike trains. High-degree spike trains are displayed as green.

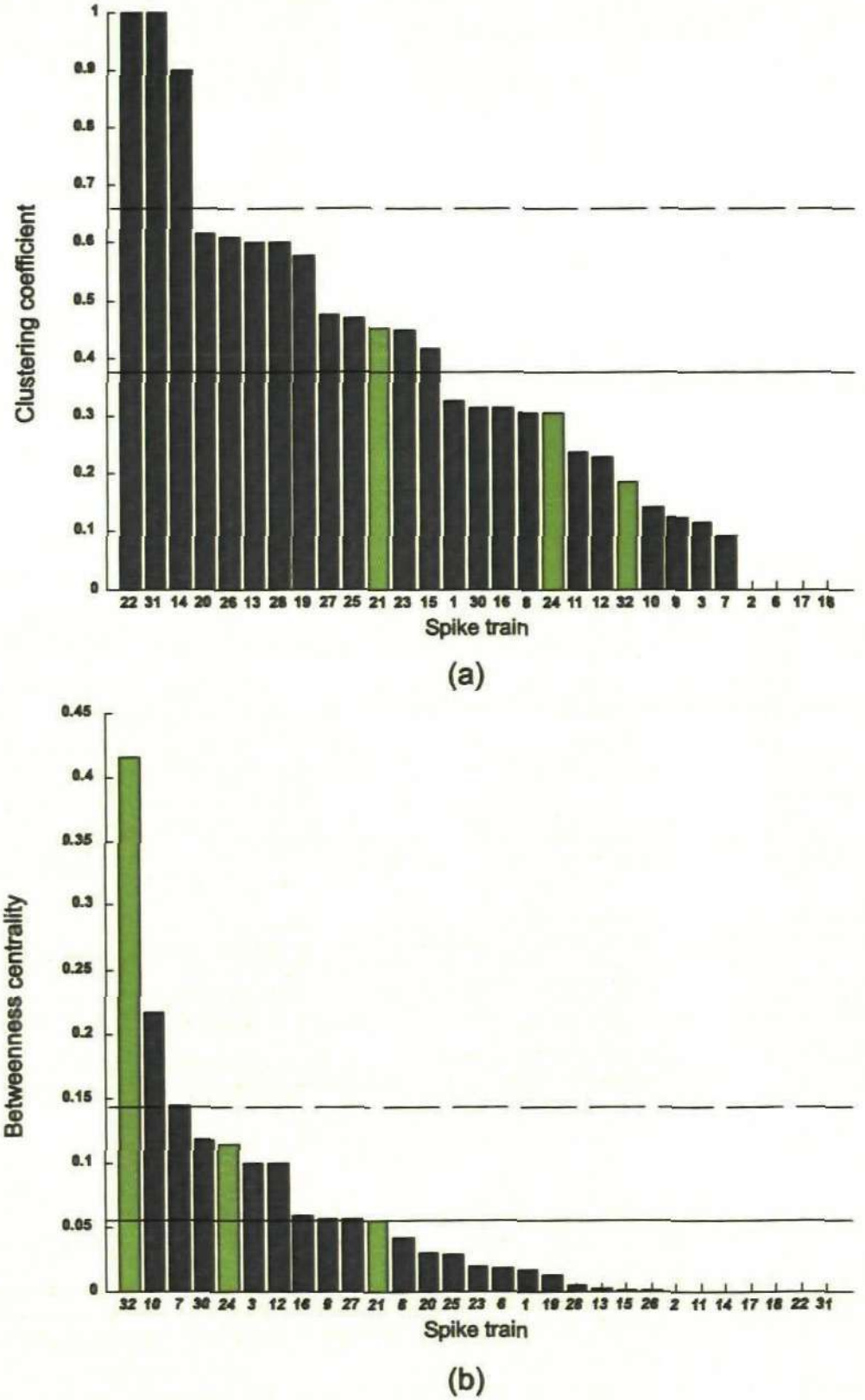


Figure C.6: Clustering coefficient and betweenness centrality of the 29 spike trains of stimulus 3. The solid horizontal line indicates the mean and the dashed horizontal line indicates the mean plus one standard deviation. High-degree spike trains are displayed as green. **(a)** Clustering coefficient of 29 spike trains is displayed in descending order. **(b)** Betweenness centrality of the 29 spike trains is displayed in descending order.

Clustering coefficient of the spike trains is shown in Fig. C.6(a). Clustering coefficient of spike trains #22 and #31 are 1 which means that these spike trains completely make a cluster with their neighbour spike trains. Investigation from connection matrix reveals that spike train #22 has two neighbour spike trains (#20 and #24) and they are connected to each other. Similarly, spike train #31 has two neighbour spike trains (#21 and #32) and they are connected to each other. The clustering coefficient of the remaining spike trains ranges from 0 to 0.9. Spike train #14 has three neighbour spike trains (#24, #25 and #32) and the neighbours are strongly connected to each other. Thus the clustering coefficient of this spike train is high which is 0.9. There are 12 low clustering spike trains which have the clustering coefficient below the mean of all the spike trains. Among them, two spike trains (#24 and #32) have the high degree of all the spike trains which indicate that these spike trains communicate to other neighbour spike trains but the neighbours are not connected to each other. There are 4 spike trains (#2, #6, #17 and #18) that do not form any cluster to their neighbour spike trains. The clustering coefficient of these spike trains is zero. The global clustering coefficient (0.3749) (random: 0.1701) also indicates that many spike trains do not have neighbours which are connected to each other. Fig. C.6(b) shows the betweenness centrality of the spike trains. There are some central spike trains (#32, #10 and #7) which transfer most of the information to the other spike trains. Among the central spike trains, spike train #32 has the highest degree. This means that these two spike trains communicate to other spike trains through incoming and outgoing connections. There are 7 spike trains which do not pass any information to other spike trains. That means the betweenness centrality of these spike trains is zero.

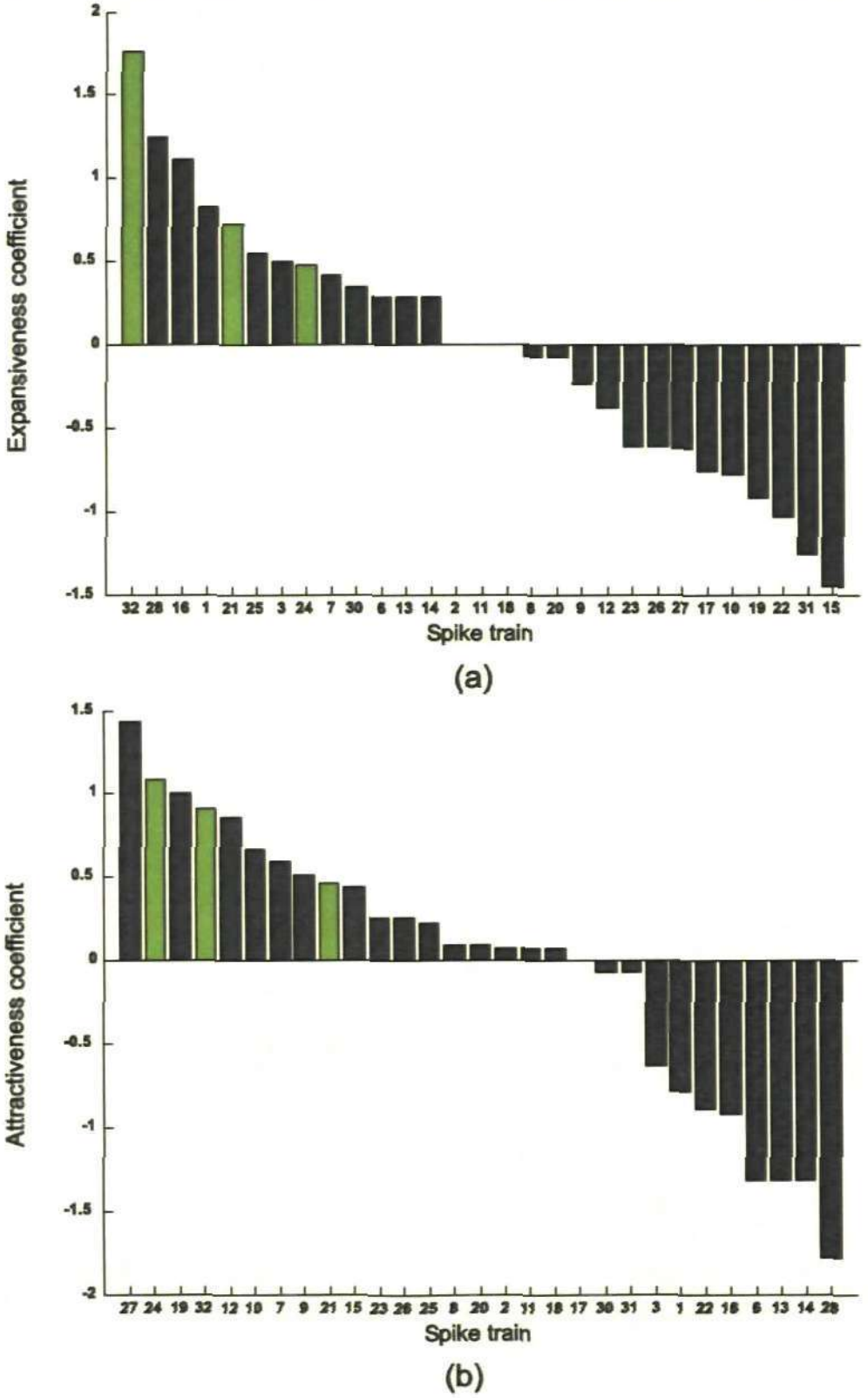


Figure C.7: Expansiveness and attractiveness coefficient of the P1 model of the 29 *spike trains of stimulus* 3. *High-degree spike trains are displayed as green.* **(a)** Expansiveness coefficient displayed in descending order. **(b)** Attractiveness coefficient displayed in descending order.

The most influential spike trains are shown in Fig. C.7(a). There are three (#32, #28 and #16) spike trains which are the most influential. Investigation from the connection matrix reveals that these spike trains have the high outdegree (12, 6 and 6 respectively). Among these high outdegree spike trains, spike train #32 has the highest degree also. There are 3 spike trains (#2, #11 and #18) that do not show any expansiveness. The outdegree of these spike trains is zero. Spike trains #31 and #15 have the most negative expansiveness coefficient. Both the spike trains have 1 outdegree each and 2 and 3 indegree respectively. Spike trains #27, #24, #19 and #32 are the most attractive spike trains (Fig. C.7(b)) as they have high indegree (7, 7, 5 and 9 respectively) and two of them have the high degree (#24 and #32). Similar to expansiveness coefficient, spike train #17 does not show any attractiveness as the indegree is zero. Spike train #28 has the most negative attractiveness coefficient (6 outdegree and 1 indegree).

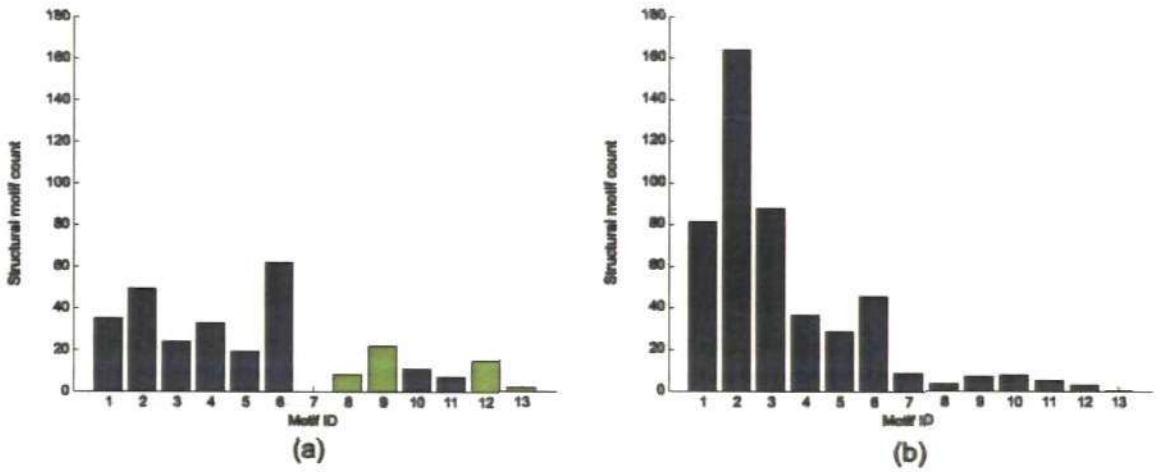


Figure C.8: (a) Structural motif count of size m = 3 of the 29 spike trains of stimulus 3. Significant motif ID's are displayed as green. **(b)** Structural motif count of size m = 3 for the randomized diagram.

Fig. C.8(a) shows the motif count for structural motifs of size m = 3 found in the connection matrix of 29 spike trains. Motif ID 6 appears 62 times which is the highest among the motif IDs. Motif ID 7 has no appearance in the connection matrix. To find the significant motif, 1000 random networks are generated keeping the same indegree and outdegree of the spike trains. The motif count for structural motifs of size m = 3 for

the random network is shown in Fig. C.8(b). The motif ID's 8, 9, 12 and 13 appear more than the random network. The Z-score of these motifs ($Z_8 = 2.56,\ p = .014$; $Z_9 = 3.50,\ p = .0009$; $Z_{12} = 6.764,\ p < .0001$; $Z_{13} = 4.95,\ p < .0001$) indicate that they are significant. There are a low proportion of connected motifs (21.10%) in the connection matrix indicating that the spike trains are weakly connected.

C.3 Analysis of connectivity of stimulus 4

The connection matrix (Fig. C.9(a)) of 29 spike trains shows 71 connections. The connection matrix contains a low density (0.0874) which means that in stimulus 4, the spike trains are not densely connected. The degree (Fig. C.9(b)) of the spike trains varies widely from 0 to 14 showing the same number of degrees for certain spike trains. Spike trains #6 and #22 have no indegree and outdegree, so their degrees are zero. Spike trains #32 and #30 have the high degree (14 and 10) compared to other spike trains. On the other hand, spike trains #9, #17 and #18 have low degree (1 each).

The characteristic path length of the connection matrix (3.1204) is greater than the characteristic path length obtained from a random network (2.6491). This characteristic path length indicates that, on average, to pass information from one spike train to another spike train, it takes approximately 3 edges. Similarly, the global efficiency of the connection matrix (0.2278) (random: 0.3055) indicates that pairs of spike trains, on average, have long communication distances.

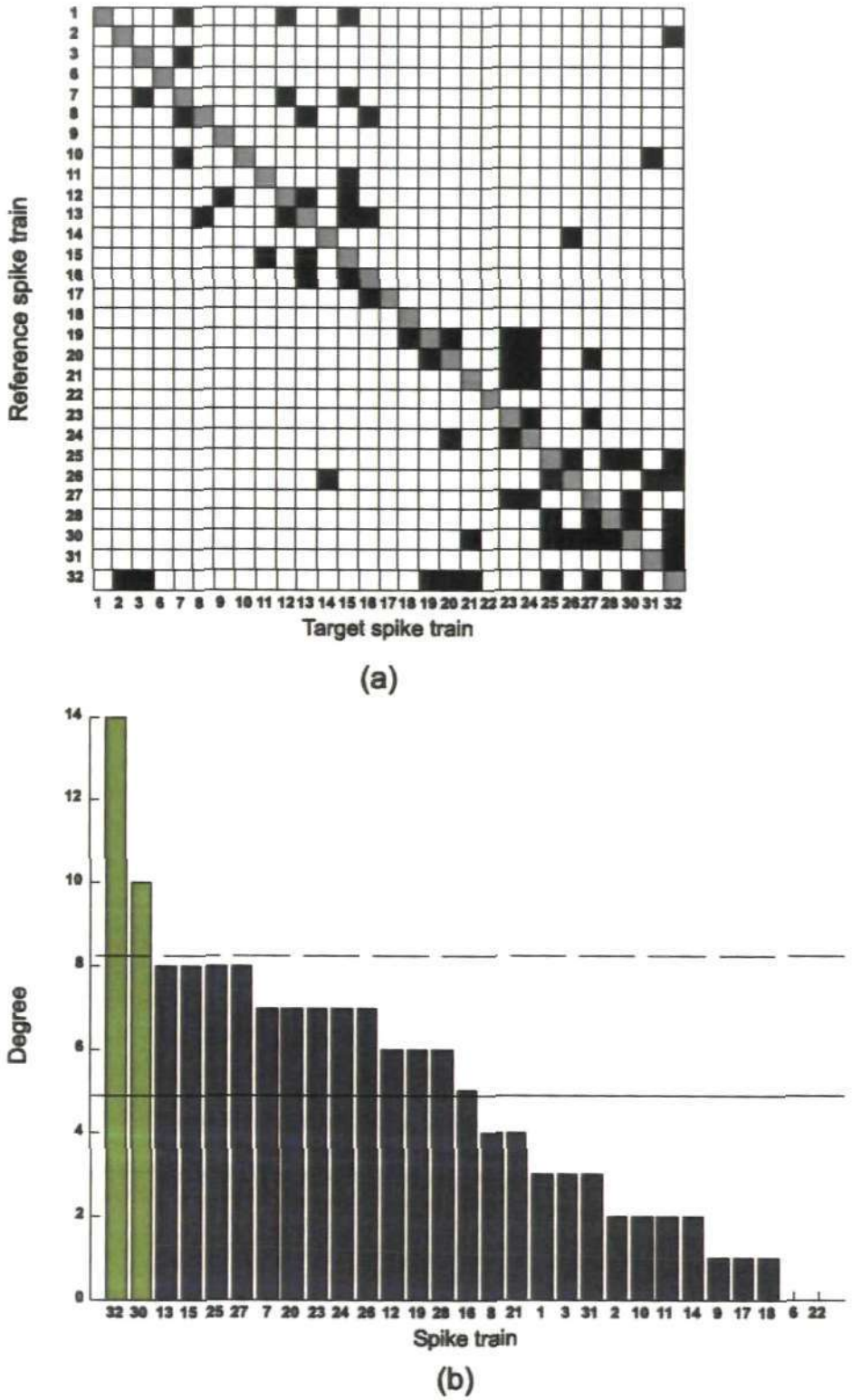


Figure C.9: **(a)** Connection matrix of the 29 spike trains of stimulus 4. Connection patterns are represented by the presence of connection (black square) and absence of connection (white square). Main diagonals are indicated in grey and self-connections are excluded. **(b)** Degree of the spike trains is displayed in descending order. The solid horizontal line indicates the mean degree of the spike trains and the dashed horizontal line indicates the mean plus one standard deviation of the spike trains. High-degree spike trains are displayed as green.

The cluster coefficient of the spike trains (Fig. C.10(a)) widely varies from 0 to 0.807. Spike train #28 has the highest clustering coefficient (0.807). This spike train has 4 neighbour spike trains (#25, #27, #30 and #32) and the neighbour spike trains are strongly connected to each other. Some spike trains have high clustering coefficient (#25, #1 and #24) compared to other spike trains indicating that the neighbours of these spike trains are also neighbours of each other. There are some spike trains which have low clustering coefficient (#32, #31 and #7), in fact below the mean of all the spike trains. Among the low clustering coefficient, spike train #32 has the highest degree of all the spike trains. This spike train communicates to other neighbour spike trains but the neighbours are not connected to each other. There are 10 spike trains that do not form any cluster to their neighbour spike trains. The clustering coefficient of these spike trains is zero. The overall clustering coefficient (0.2408) (random: 0.1220) also indicates that many spike trains do not have neighbours which are connected to each other.

The betweenness centrality of the spike trains is shown in Fig. C.10(b). Spike train #32 is the most central spike train which transfer most of the information to the other spike trains. There are some other central spike trains (#7, #3, #30 and #27) which pass most of the information to other spike trains. Among the central spike trains, spike trains #30 and #32 have the highest degree. This means that these two spike trains communicate to other spike trains through incoming and outgoing connections. There are 10 spike trains which do not pass any information to other spike trains. That means the betweenness centrality of these spike trains is zero.

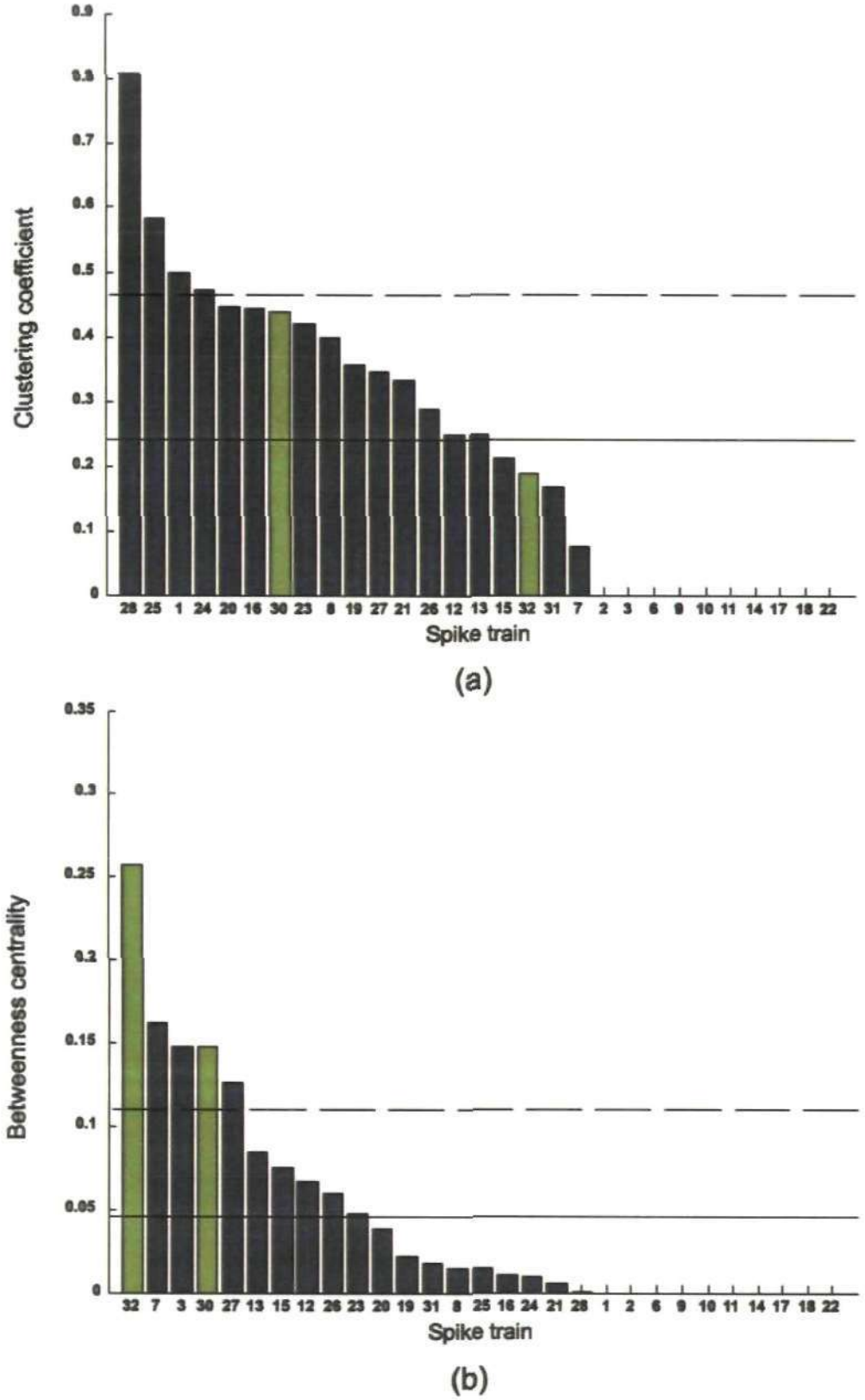


Figure C.10: Clustering coefficient and betweenness centrality of the 29 spike trains of stimulus 4. The solid horizontal line indicates the mean and the dashed horizontal line indicates the mean plus one standard deviation. High-degree spike trains are displayed as green. **(a)** Clustering coefficient of 29 spike trains is displayed in descending order. **(b)** Betweenness centrality of the 29 spike trains is displayed in descending order.

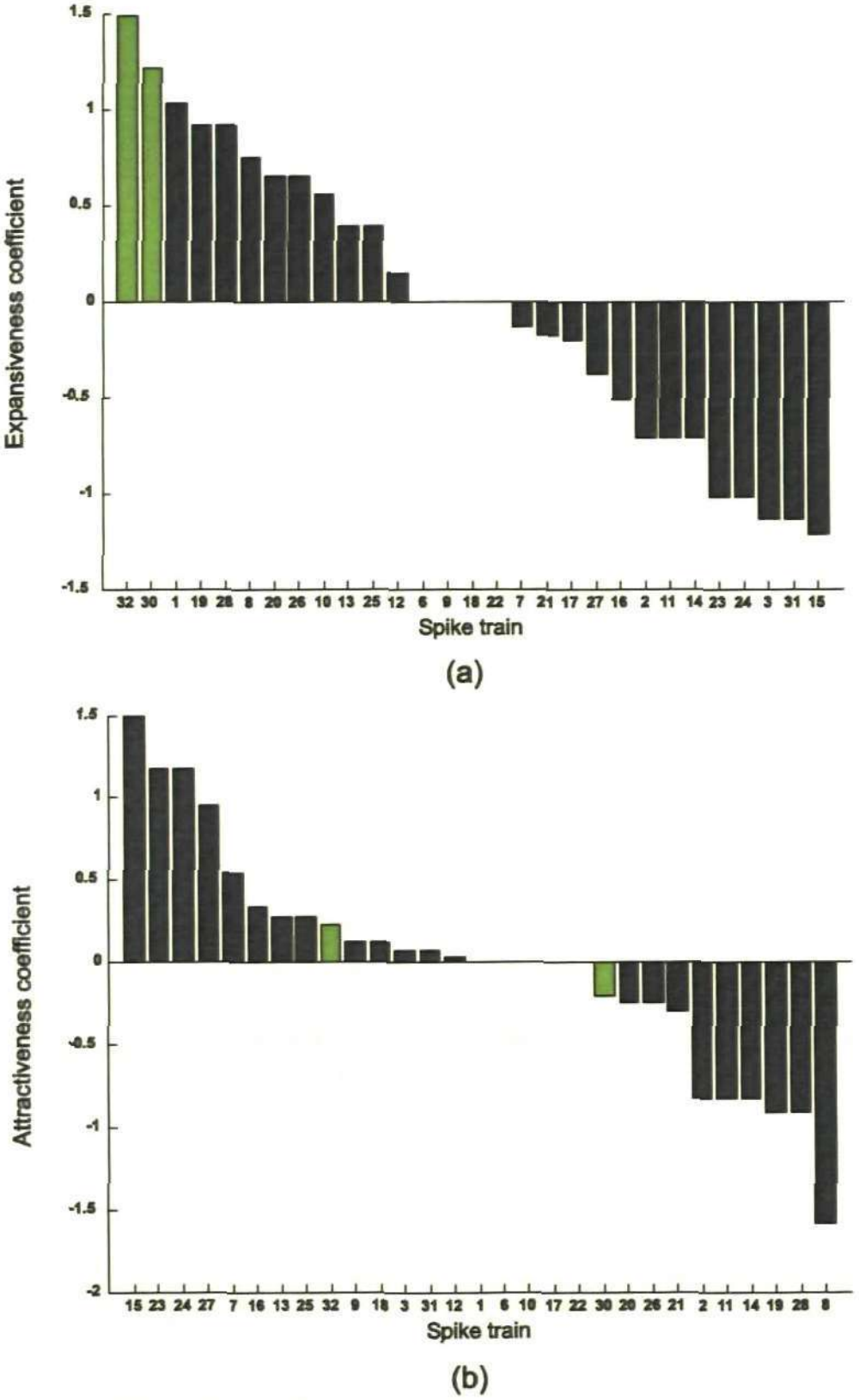


Figure C.11: Expansiveness and attractiveness coefficient of the P1 model of the 29 spike trains of stimulus 4. High-degree spike trains are displayed as green. **(a)** Expansiveness coefficient displayed in descending order. **(b)** Attractiveness coefficient displayed in descending order.

The most influential spike trains are shown in Fig. C.11(a). There are two (#32 and #30) spike trains which are the most influential. Investigation from the connection matrix reveals that these spike trains have the high outdegree (8 and 6). Both these spike trains have the highest degree also. There are 4 spike trains (#6, #9, #18 and #22) that do not show any expansiveness. The outdegree of these spike trains is zero. Spike train #15 has the most negative expansiveness coefficient. This spike train has 2 outdegree and 6 indegree connections. Spike trains #15, #23 and #24 are the most attractive spike trains (Fig. C.11(b)) as they have high indegree (6, 5 and 5 respectively). Similar to expansiveness coefficient, 4 spike trains (#6, #10, #17 and #22) do not show any attractiveness as the indegree of these spike trains are zero. Spike train #8 has the most negative attractiveness coefficient (3 outdegree and 1 indegree).

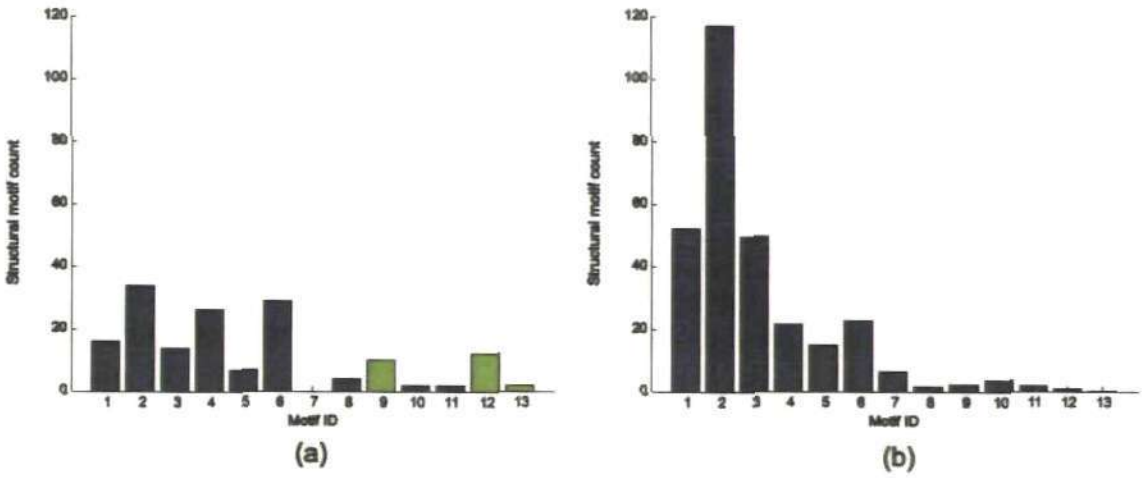


Figure C.12: (a) Structural motif count of size m = 3 of the 29 spike trains of stimulus 4. Significant motif ID's are displayed as green. **(b)** Structural motif count of size m = 3 for the randomized diagram.

Fig. C.12(a) shows the motif count for structural motifs of size m = 3 found in the connection matrix of 29 spike trains. Motif ID 2 appears 34 times which is the highest among the motif IDs. Motif ID 7 has no appearance in the connection matrix. To find the significant motif, 1000 random networks are generated keeping the same indegree and outdegree of the spike trains. The motif count for structural motifs of size m = 3 for

the random network is shown in Fig. C.12(b). Motif ID's 9, 12 and 13 appear more than the random network. The Z-score of these motif ID's ($Z_9 = 3.72,\ p = .0004$; $Z_{12} = 12.07,\ p < .0001$; $Z_{13} = 15.80,\ p < .0001$) indicate that they are significant. There are a low proportion of connected motifs (19.70%) in the connection matrix indicates that the spike trains are weakly connected.

C.4 Analysis of connectivity of stimulus 5

The connection matrix of 29 spike trains is shown in Fig. C.13(a). There are 116 connections in this stimulus which is higher than the previous stimuli. The connection matrix contains a low density (0.1429) which indicates that in stimulus 5 the spike trains are not densely connected. The degree (Fig. C.13(b)) of the spike trains varies widely from 0 to 17 showing the same number of degrees for certain spike trains. Some spike trains have the high degree (#28, #32, #9 and #12) whereas some spike trains have low degree (#2, #11, #18, #22 and #17). Spike train #28 and #32 have highest 17 indegree and outdegree connection whereas spike train 17 has 1 connection.

The characteristic path length of the connection matrix (2.3377) is greater than the characteristic path length obtained from a random network (2.2100). This characteristic path length indicates that, on average, to pass information from one spike train to another spike train, it takes approximately 3 edges. Similarly, the global efficiency of the connection matrix (0.4041) (random: 0.4257) indicates that pairs of spike trains, on average, have long communication distances. Compare to other stimuli this stimulus has short communication distances.

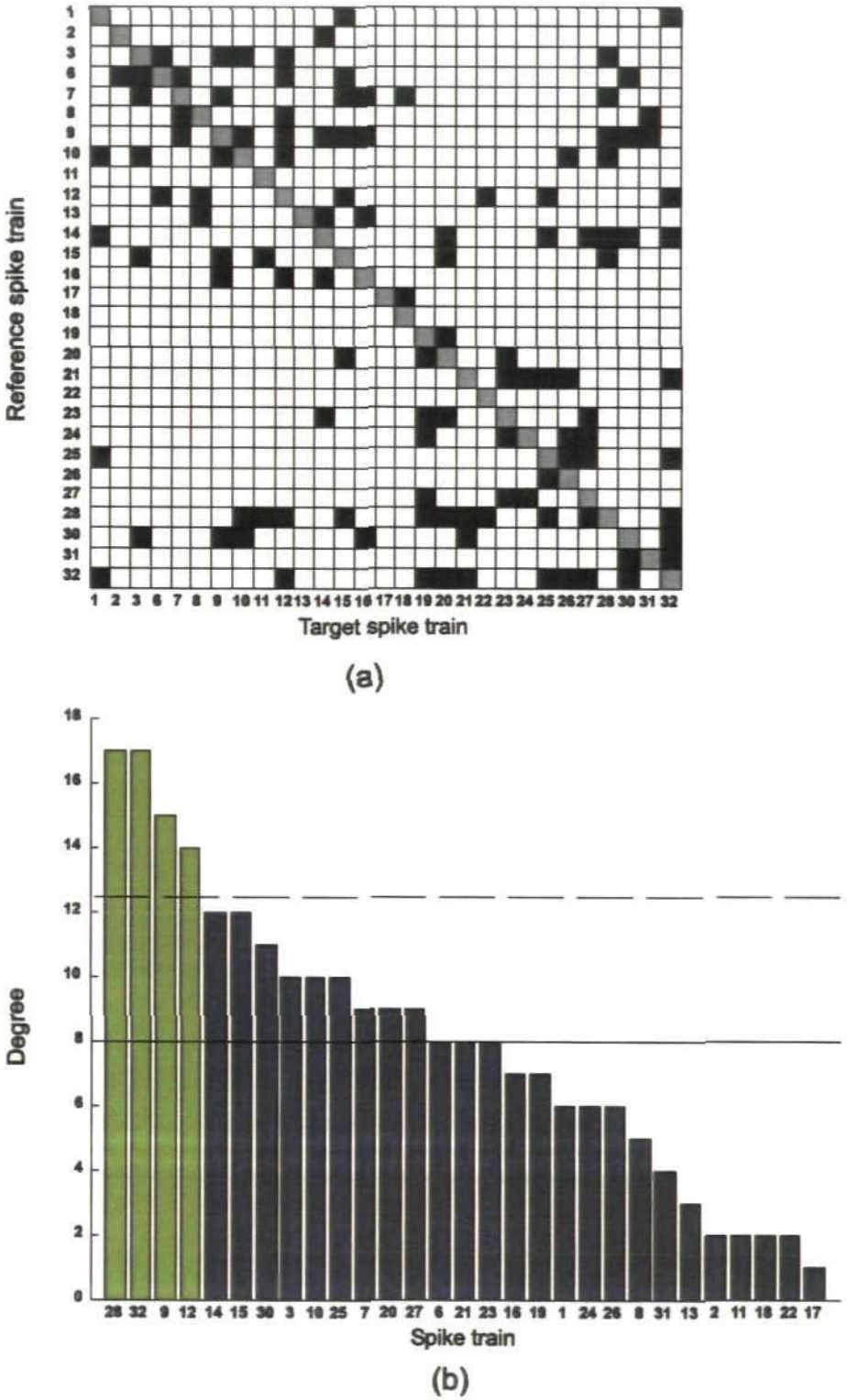


Figure C.13: (a) Connection matrix of the 29 spike trains of stimulus 5. Connection patterns are represented by the presence of connection (black square) and absence of connection (white square). Main diagonals are indicated in grey and self-connections are excluded. **(b)** Degree of the spike trains is displayed in descending order. The solid horizontal line indicates the mean degree of the spike trains and the dashed horizontal line indicates the mean plus one standard deviation of the spike trains. High-degree spike trains are displayed as green.

The clustering coefficient of the spike trains (Fig. C.14(a)) widely varies from 0 to 1. Spike train #11 has two neighbour spike trains (#15 and #28) and they are connected to each other, so the clustering coefficient of this spike train is 1. Spike train #22 has two neighbour spike trains (#12 and #28) and there is a connection from spike train #28 to spike train #12, so the clustering coefficient of this spike train is 0.5. All the high degree spike trains (#9, #28, #32 and #12) have the clustering coefficient below the mean of all the spike trains. These spike trains communicate to other neighbour spike trains but the neighbours are not connected to each other. There are 4 spike trains (#2, #8, #17 and #18) that do not form any cluster to their neighbour spike trains. The clustering coefficient of these spike trains is zero. The global clustering coefficient (0.2715) (random: 0.2080) also indicates that many spike trains do not have neighbours that are connected to each other.

The betweenness centrality of the spike trains is shown in Fig. C.14(b). Spike train #32 is the most central spike train which transfer most of the information to the other spike trains. There are some other central spike trains (#14, #15, #12, #9 and #28) which pass most of the information to other spike trains. Importantly, all the high degree spike trains have high betweenness centrality among the spike trains and most of the information go through these spike trains. There are 5 spike trains (#11, #13, #17, #18 and #22) which do not pass any information to other spike trains. That means the betweenness centrality of these spike trains is zero.

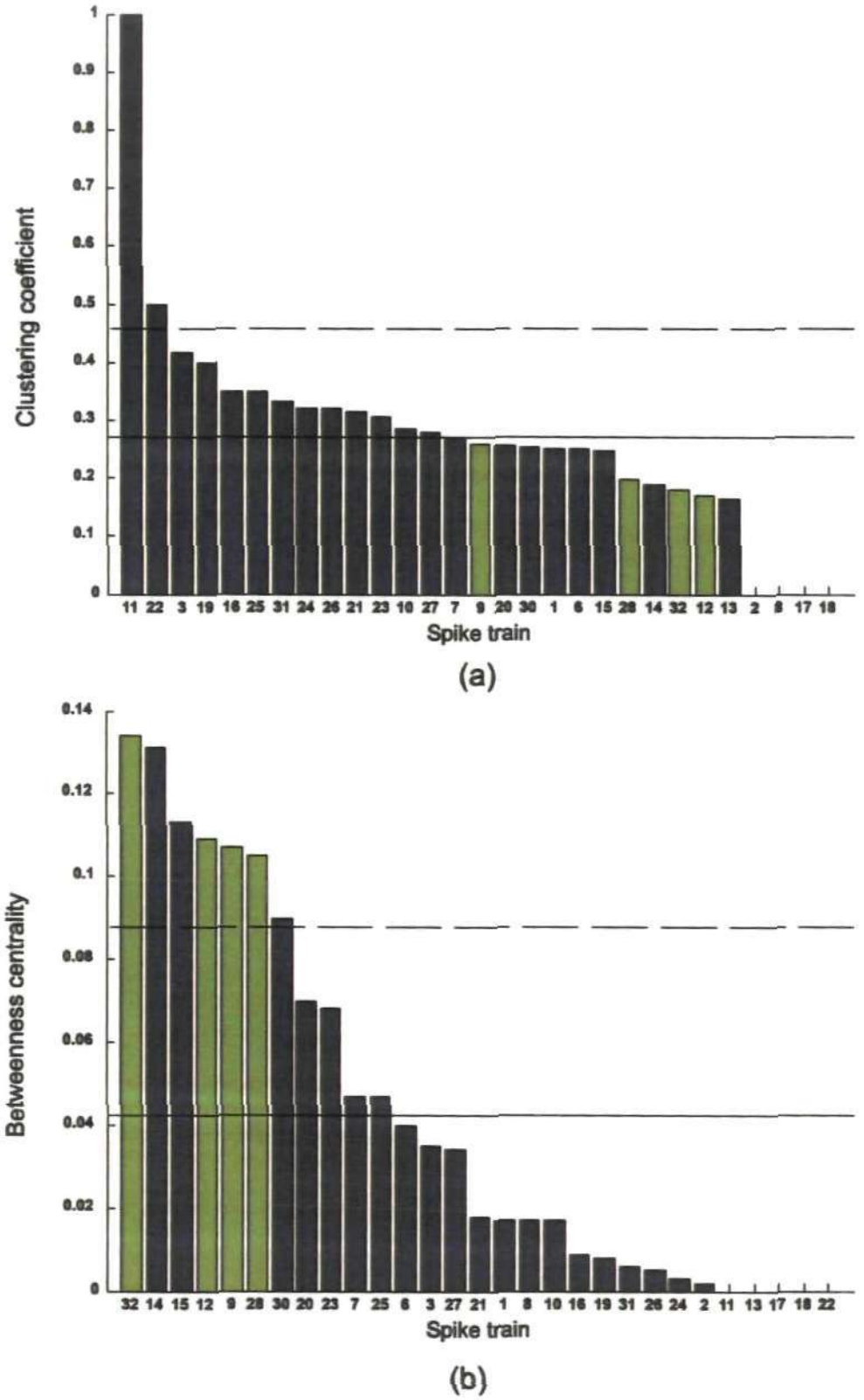


Figure C.14: Clustering coefficient and betweenness centrality of the 29 spike trains of stimulus 5. The solid horizontal line indicates the mean and the dashed horizontal line indicates the mean plus one standard deviation. High-degree spike trains are displayed as green. **(a)** Clustering coefficient of 29 spike trains is displayed in descending order. **(b)** Betweenness centrality of the 29 spike trains is displayed in descending order.

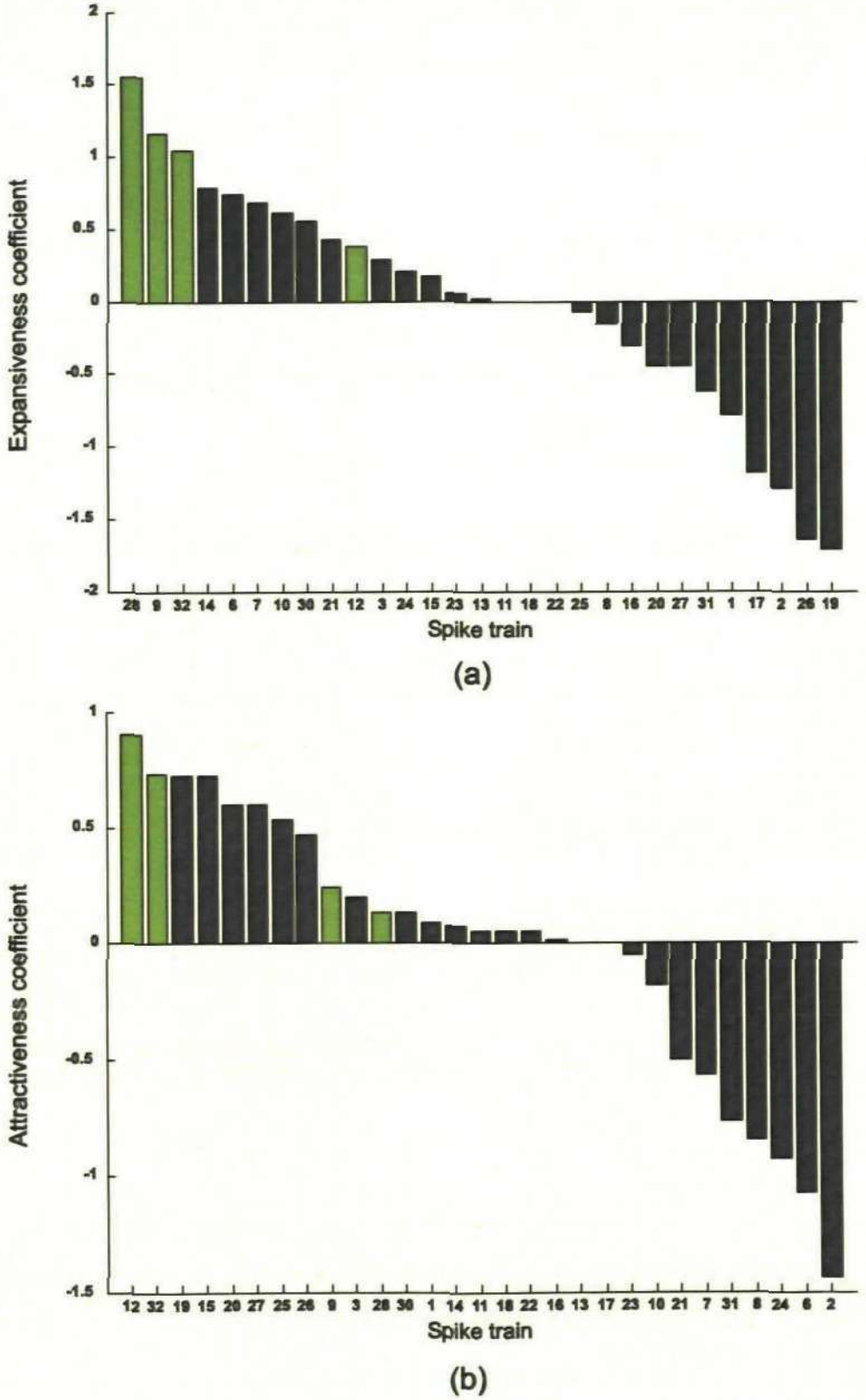


Figure C.15: Expansiveness and attractiveness coefficient of the P1 model of the 29 spike trains of stimulus 5. High-degree spike trains are displayed as green. **(a)** Expansiveness coefficient displayed in descending order. **(b)** Attractiveness coefficient displayed in descending order.

The most influential and attractive spike trains are shown in Fig. C.15. There are 3 (#28, #9 and #32) spike trains which are the most influential (Fig. C.15(a)). Investigation from the connection matrix reveals that these spike trains have the high outdegree (11, 9 and 9 respectively). All these spike trains have high degree also. There are 3 spike trains (#11, #18, and #22) that not show any expansiveness. The outdegree of these spike trains is zero. Spike train #19 has the most negative expansiveness coefficient. This spike train has 1 outdegree and 6 indegree connections. Spike trains #12, #32, #19 and #15 are the most attractive spike trains (Fig. C.15(b)) as they have high indegree (8, 8, 7 and 6 respectively). Among the attractive spike trains two (#12 and #32) have high degree also. Similar to expansiveness coefficient two spike trains (#13 and #17) do not show any attractiveness as the indegree of these spike trains are zero. Spike train #2 has the most negative attractiveness coefficient.

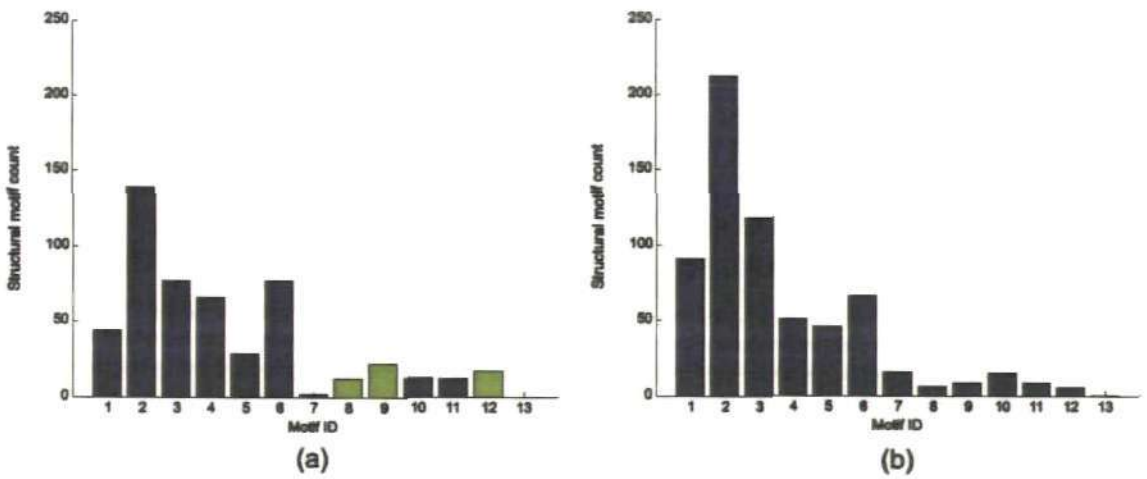


Figure C.16: (a) Structural motif count of size m = 3 of the 29 spike trains of stimulus 5. Significant motif ID's are displayed as green. **(b)** Structural motif count of size m = 3 for the randomized diagram.

Fig. C.16(a) shows the motif count for structural motifs of size m = 3 found in the connection matrix of 29 spike trains. Motif ID 2 appears 139 times which is the highest among the motif IDs. Motif ID 13 has no appearance in the connection matrix. To find the significant motif, 1000 random networks are generated keeping the same indegree

and outdegree of the spike trains. The motif count for structural motifs of size m = 3 for the random network is shown in Fig. C.16(b). The motif ID's 8, 9 and 12 appear more than the random network. The Z-score of these motifs ($Z_8 = 2.23,\ p = .0325$; $Z_9 = 3.12,\ p = .003$; $Z_{12} = 4.44,\ p < .0001$) indicate that they are significant. There are a very low proportion of connected motifs (11.84%) in the connection matrix indicating that the spike trains are weakly connected.

C.5 Analysis of connectivity of stimulus 6

The connection matrix of 29 spike trains is shown in Fig. C.17(a). There are 76 connections in this stimulus which is similar as stimuli 1, 2 and 4. The connection matrix contains a low density (0.0936) which indicates that in stimulus 6, the spike trains are not densely connected. The degree (Fig. C.17(b)) of the spike trains varies from 0 to 14 showing the same number of degrees for certain spike trains. Some spike trains have the high degree (#32, #24, #14 and #27) whereas some spike trains have low degree (#1, #17 and #18). Spike trains #11 and #22 have no indegree and outdegree, so the degree of these spike trains is zero.

The characteristic path length of the connection matrix is (3.0560) greater than the characteristic path length obtained from random network (2.6545). This characteristic path length indicates that, on average, to pass information from one spike train to another spike train, it takes approximately 3 edges. Similarly, the global efficiency of the connection matrix (0.2939) (random: 0.3242) indicates that pairs of spike trains, on average, have long communication distances.

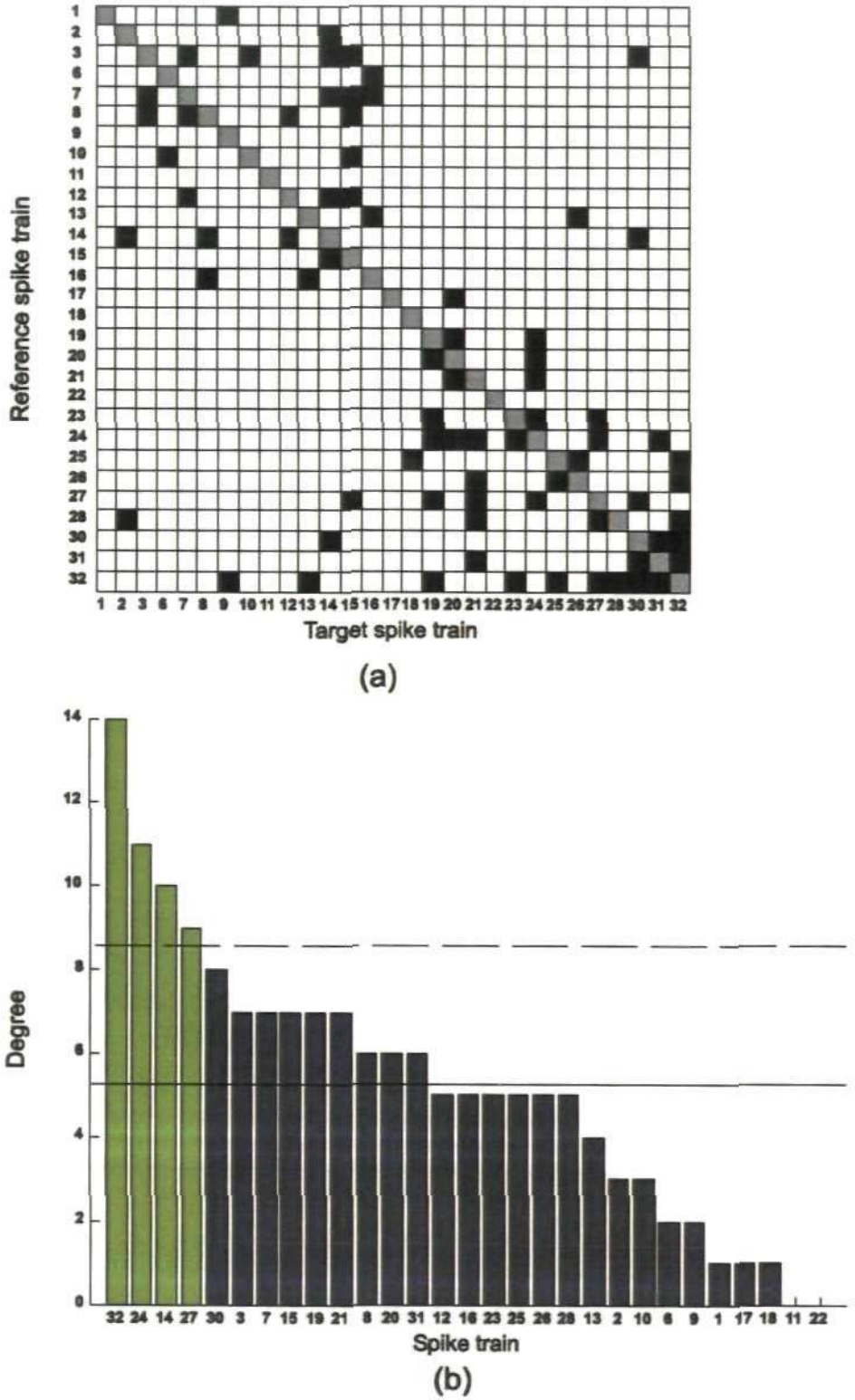


Figure C.17: (a) Connection matrix of the 29 spike trains of stimulus 6. Connection patterns are represented by the presence of connection (black square) and absence of connection (white square). Main diagonals are indicated in grey and self-connections are excluded. **(b)** Degree of the spike trains is displayed in descending order. The solid horizontal line indicates the mean degree of the spike trains and the dashed horizontal line indicates the mean plus one standard deviation of the spike trains. High-degree spike trains are displayed as green.

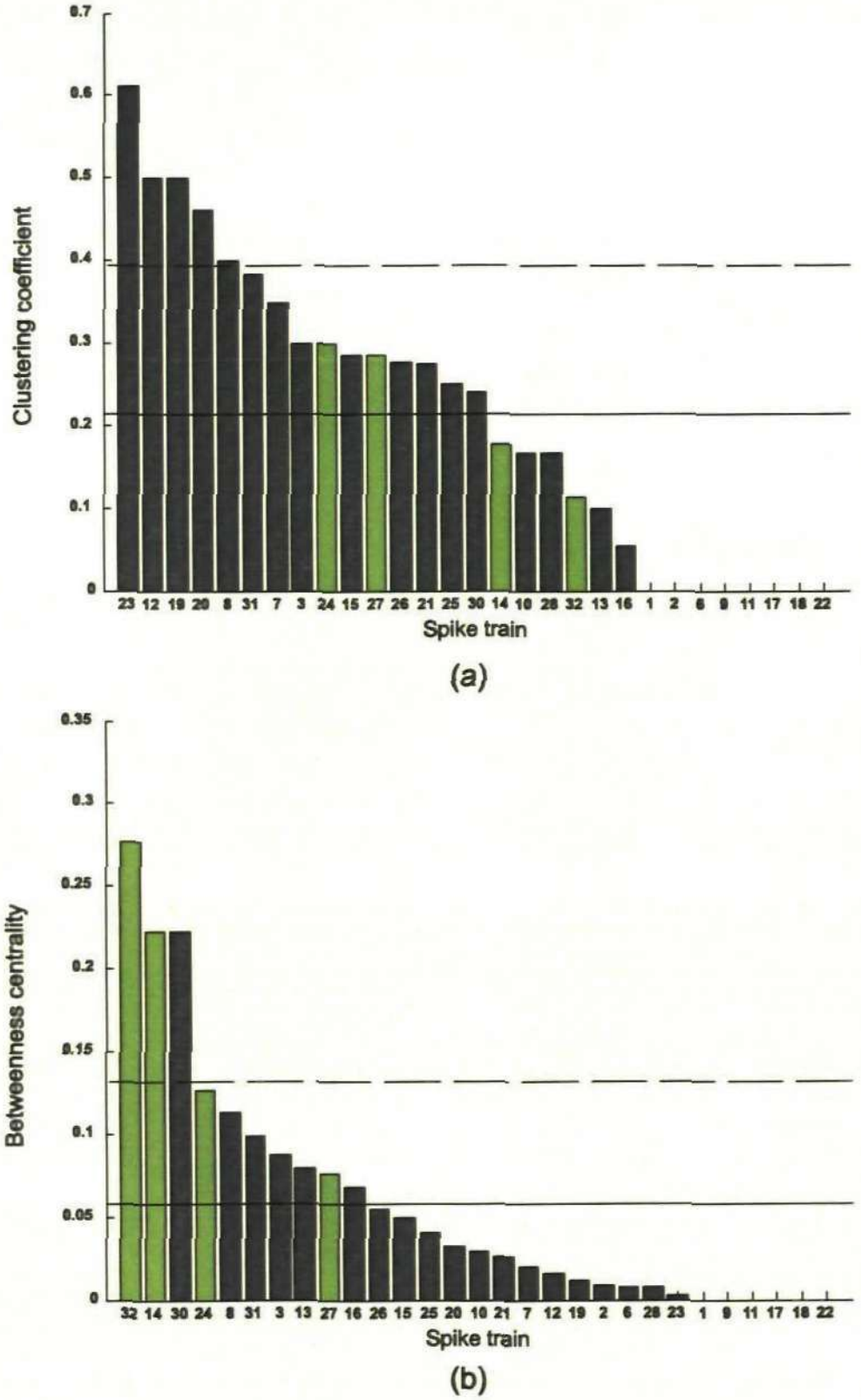
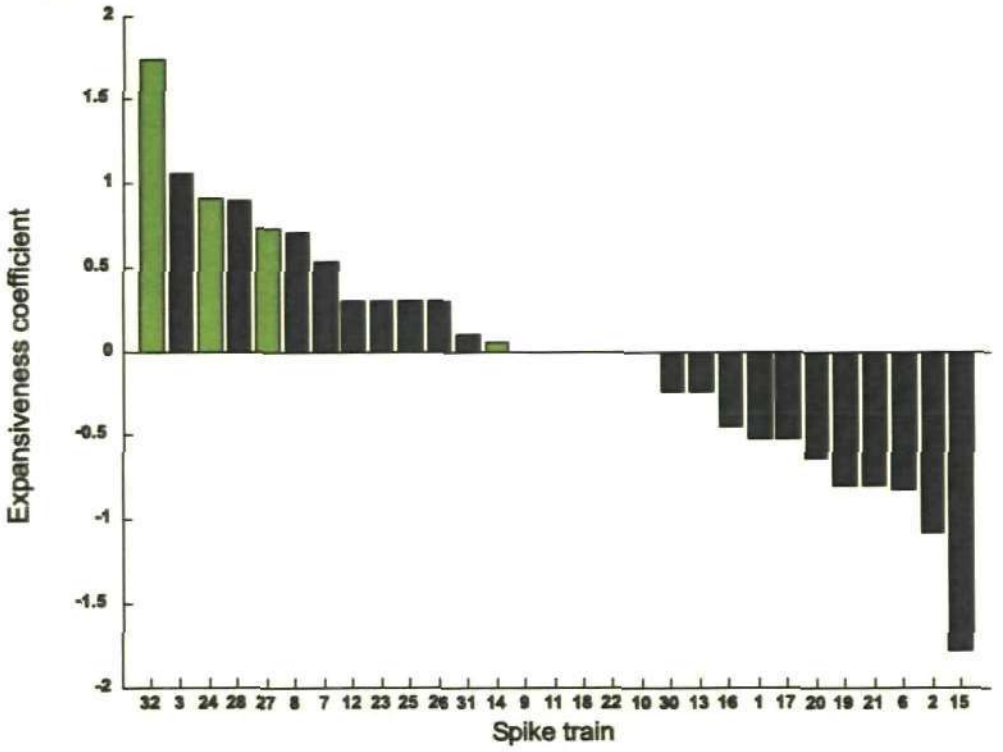
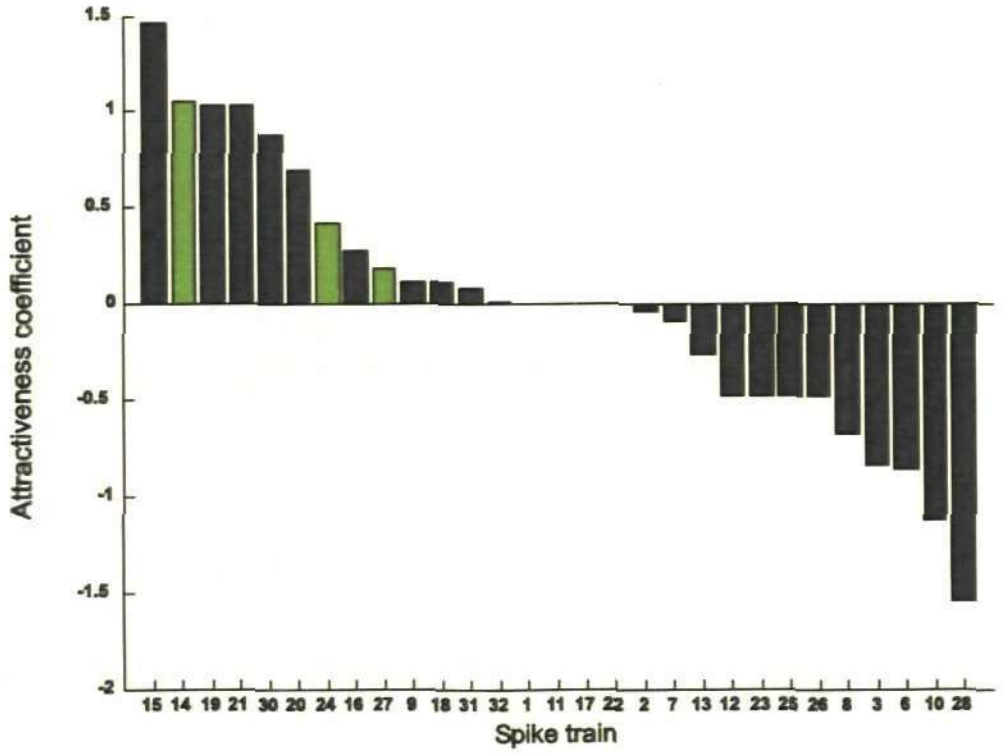


Figure C.18: Clustering coefficient and betweenness centrality of the 29 spike trains of stimulus 6. The solid horizontal line indicates the mean and the dashed horizontal line indicates the mean plus one standard deviation. High-degree spike trains are displayed as green. **(a)** Clustering coefficient of 29 spike trains is displayed in descending order. **(b)** Betweenness centrality of the 29 spike trains is displayed in descending order.

The clustering coefficient of the spike trains (Fig. C.18(a)) varies from 0 to 0.61. Spike train #23 has the highest clustering coefficient which is 0.61. This spike train has 4 neighbour spike trains (#19, #24, #27 and #32) and they have 7 connections to each other. That means the neighbour spike trains are strongly connected to each other. Spike train #12 has also 4 neighbour spike trains (#7, #8, #14 and #15) and they have 6 connections to each other. That means the neighbour spike trains are moderately connected to each other. Similarly, the neighbours of spike train #19 are moderately connected to each other as the clustering coefficient of this spike train is 0.5. All the high degree spike trains (#24, #27, #14 and #32) have the low clustering coefficient which means that these spike trains communicate to other neighbour spike trains but the neighbours are not connected to each other. There are 8 spike trains that do not form any cluster to their neighbour spike trains. The clustering coefficient of these spike trains is zero. The global clustering coefficient (0.2139) (random: 0.1222) also indicates that many spike trains do not have neighbours that are connected to each other. The betweenness centrality of the spike trains is shown in Fig. C.18(b). Spike train #32 is the most central spike train which transfer most of the information to the other spike trains. There are 2 other central spike trains (#14 and #30) which pass most of the information to other spike trains. There are 6 spike trains which do not pass any information to other spike trains. That means the betweenness centrality of these spike trains is zero.



(a)



(b)

Figure C.19: Expansiveness and attractiveness coefficient of the P1 model of the 29 spike trains of stimulus 6. High-degree spike trains are displayed as green. **(a)** Expansiveness coefficient displayed in descending order. **(b)** Attractiveness coefficient displayed in descending order.

The most influential and attractive spike trains are shown in Fig. C.19. There are 2 (#32 and #3) spike trains which are the most influential (Fig. C.19(a)). Investigation from the connection matrix reveals that these spike trains have the high outdegree (9 and 5). There are 4 spike trains (#9, #11, #18 and #22) that not show any expansiveness. The outdegree of these spike trains is zero. Spike train #15 has the most negative expansiveness coefficient. This spike train has 1 outdegree and 6 indegree connections. Spike trains #15, #14, #19 and #21 are the most attractive spike trains (Fig. C.19(b)) as they have high indegree (6, 6, 5 and 5 respectively). Similar to expansiveness coefficient four spike trains (#1, #11, #17 and #22) do not show any attractiveness as the indegree of these spike trains are zero. Spike train #28 has the most negative attractiveness coefficient (1 indegree and 4 outdegree).

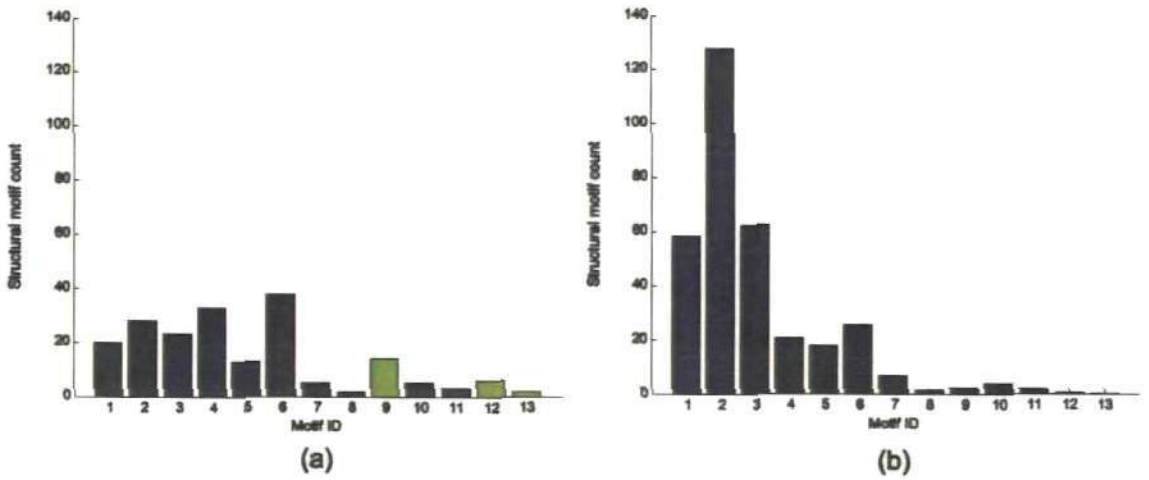


Figure C.20: (a) Structural motif count of size m = 3 of the 29 spike trains of stimulus 6. Significant motif ID's are displayed as green. **(b)** Structural motif count of size m = 3 for the randomized diagram.

Fig. C.20(a) shows the motif count for structural motifs of size m = 3 found in the connection matrix of 29 spike trains. Motif ID 6 appears 38 times which is the highest among the motif IDs. Motif ID's 8 and 13 appear twice in the connection matrix which is the lowest. To find the significant motif, 1000 random networks are generated keeping the same indegree and outdegree of the spike trains. The motif count for

structural motifs of size m = 3 for the random network is shown in Fig. C.20(b). The motif ID's 9, 12 and 13 appear more than the random network. The Z-score of these motifs ($Z_9 = 5.73,\ p < .0001;\quad Z_{12} = 5.47,\ p < .0001;\quad Z_{13} = 13.16,\ p < .0001$) indicate that they are significant. There are a low proportion of connected motifs (20%) in the connection matrix indicating that the spike trains are weakly connected.

List of references

Achard, S. and Bullmore, E. "Efficiency and cost of economical brain functional networks," *PloS Computaional Biology* 3(2): e17, 2007.

Achard, S., Salvador, R., Whitcher, B., Suckling, J. and Bullmore, E. "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *Journal of Neuroscience* 26, 2006, pp. 63-72.

Acuna, E. and Rodriguez, C. A. "Meta analysis study of outlier detection methods in classification", Technical report, Department of Mathematics, University of Puerto Rico at Mayaguez, 2004.

Aertsen, A., Vaadia, E., Abeles, M., Ahissar, E., Bergman, H., Karmon, B., Lavner, Y., Margalit, E., Nelken, I. and Rotter, S. "Neural interactions in the frontal-cortex of a behaving monkey – signs of dependence on stimulus context and behavioral state," *Journal of Hirnforschung* 32, 1991, pp. 735-743.

Aertsen, A. M. H. J., Bonhoeffer, T. and Kruger, J. "Physics of cognitive processes", *in* Caianiello, E. R., ed.,, World Sciences Publishers, Singapore, 1987, pp. 1-34.

Aertsen, A. M. H. J. and Gerstein, G. L. "Evaluation of neuronal connectivity: sensitivity of cross correlation," *Brain Research* 340, 1985, pp. 341-354.

Aertsen, A. M. H. J., Gerstein, G. L., Habib, M. K. and Palm, G. "Dynamics of neuronal firing correlation: Modulation of effective connectivity," *Journal of Neurophysiology* 61, 1989, pp. 900-917.

Arata, A., Hernandez, Y. M., Lindsey, B. G., Morris, K. F. and Shannon, R. "Transient configurations of baroresponsive respiratory-related brainstem neuronal assemblies in the cat," *Journal of Physiology* 525(2), 2000, pp. 509-530.

Aronov, D., Reich, D. S., Mechler, F. and Victor, J. D. "Neural coding of spatial phase in V1 of the macaque monkey," *Journal of Neurophysiology* 89, 2003, pp. 3304-3327.

Baker, S. N. and Gerstein, G. L. "Improvements to the sensitivity of gravitational clustering for multiple neuron recordings," *Neural Computation* 12, 2000, pp. 2597-2620.

Barabasi, A. L. and Oltavi, Z. N. "Network biology: understanding the cell's functional organization," *Nature Reviews Genetics* 5, 2004, pp. 101-111.

Bartlett, M. S. *An introduction to stochastic processes (2nd ed.)*, Cambridge: Cambridge University Press, 1966.

Bartolomei, F., Bosma, I., Klein, M., Baayen, J. C., Reijneveld, J. C., Postma, T. J., Heimans, J. J., van Dijk, B. W., de Munck, J. C., de Jongh, A., Cover, K. S. and Stam, C. J. "Disturbed functional connectivity in brain tumour patients: evaluation by graph analysis of synchronization matrices," *Clinical Neurophysiology* 117, 2006, pp. 2039-2049.

Bassett, D. S., Meyer-Linderberg, A., Achard, S., Duke, T. and Bullmore, E. "Adaptive reconfiguration of fractal small-world human brain functional networks," *Proceedings of the National Academy of Sciences* 103, 2006, pp. 19518-19523.

Bishop, P. O., Levick, W. R. and Williams, W. O. "Statistical analysis of the dark discharge of lateral geniculate neurones," *Journal of Physiology* 170, 1964, pp. 598-612.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. and Hwang, D. U. "Complex networks: Structure and dynamics," *Physics Reports* 424, 2006, pp. 175-308.

Borisyuk, G. N., Borisyuk, R. M., Kirillov, A. B., Kovalenko, E. I. and Kryukov, V. I. "A new statistical method for identifying interconnections between neuronal network elements," *Biological Cybernetics* 52, 1985, pp. 301-306.

Borisyuk, R. "Oscillatory activity in the neural networks of spiking elements," *BioSystems* 67, 2002, pp. 3-16.

Boven, K.-H., Fejtl, M., Möller, A., Nisch, W. and Stett, A. "Advances in network electrophysiology using multi-electrode arrays", *in* Baudry, M. and Taketani, M., ed.,, New York: Springer Press, 2006, pp. 24-37.

Brillinger, D. R. "Measuring the association of point processes: a case history," *The American Mathematical Monthly* 83(1), 1976c, pp. 16-22.

Brillinger, D. R. "Estimation of the second-order intensities of a bivariate stationary point process," *Journal of the Royal Statistical Society. Series B* 38(1), 1976b, pp. 60-66.

Brillinger, D. R. "Identification of synaptic interactions," *Biological Cybernetics* 22, 1976a, pp. 213-228.

Brillinger, D. R. "The identification of point process systems," *The Annals of Probability* 3:6, 1975b, pp. 909-929.

Brillinger, D. R. "Estimation of product densities", *in* Frane, J. W., ed.,'Computer Science and Statistics: 8th Annual Symposium', 1975a, pp. 431-438.

Brillinger, D. R. "Nerve cell spike train data analysis: a progression of technique," *Journal of the American Statistical Association* 87:418, 1992, pp. 260-271.

Brillinger, D. R. "Maximum likelihood analysis of spike trains of interacting nerve cells," *Biological Cybernetics* 59, 1988, pp. 189-200.

Brody, C. D. "Correlations without synchrony," *Neural Computation* 11, 1999, pp. 1537–1551.

Brown, E. N., Kass, R. E. and Mitra, P. P. "Multiple neural spike train data analysis: state-of-the-art and future challenges," *Nature Neuroscience* 7:5, 2004, pp. 456-461.

Buchel, C. and Friston, K. "Assessing interactions among neuronal systems using functional neuroimaging," *Neural Network* 13, 2000, pp. 871-882.

Bullmore, E. and Sporns, O. "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience* 10, 2009, pp. 186-198.

Chialvo, D. R. "Critical brain networks," *Physica A: Statistical Mechanics and its Applications* 340, 2004, pp. 756-765.

Chichilnisky, E. J. and Rieke, F. "Detection sensitivity and temporal resolution of visual signals near absolute threshold in the salamander retina," *Journal of Neuroscience* 25, 2005, pp. 318-330.

Chornoboy, E. S., Schramm, L. P. and Karr, A. F. "Maximum likelihood identification of neural point process systems," *Biological Cybernetics* 59, 1988, pp. 265-275.

Correia, M. J. and Landolt, J. P. "A point process analysis of the spontaneous activity of anterior semicircular canal units in the anesthetized pigeon," *Biological Cybernetics* 27, 1977, pp. 199-213.

Cox, D. R. "The statistical analysis of dependencies in point processes:", *in* Lewis, P. A., ed.,, New York: Willey, 1972, pp. 55-66.

Cox, D. R. and Lewis, P. A. W. "Multivariate point processes"'Proceedings Sixth Berkeley Symposium on Probability and Mathematical Statistics', 1972, pp. 401-448.

Czanner, G., Eden, U. T., Wirth, S., Yanike, M., Suzuki, W. and Brown, E. "Analysis of between-trail and within-trial neural spiking dynamics," *Journal of Neurophysiology* 99, 2008, pp. 2672-2693.

Dahlhaus, R., Eichler, M. and Sandkuhler, J. "Identification of synaptic connections in neural ensembles by graphical models," *Journal of Neuroscience Methods* 77, 1997, pp. 93-107.

Daley, D. and Vere-Jones, D., J. Gani, C.C. Heyde, T. K., (eds.) *An introduction to the theory of point processes: volume I: elementary theory and methods, second edition*, Springer, 2003.

Dayan, P. and Abbott, L. F. *Theoretical neuroscience: computational and mathematical modeling of neural systems*, Cambridge: The MIT Press, 2001.

Di Lorenzo, P. M. and Victor, J. D. "Taste response variability and temporal coding in the nucleus of the solitary tract of the rat," *Journal of Neurophysiology* 90, 2003, pp. 1418-1431.

Dodel, S., Hermann, J. M. and Geisel, T. "Functional connectivity by cross-correlation clustering," *Neurocomputing* 44-46, 2002, pp. 1065-1070.

Eden, U. T. "Point process models for neural spike trains", Technical report, Department of Mathematics and Statistics, Boston University, Boston, Massachusetts, 2008.

Eggermont, J. J. "Neural interaction in cat primary auditory cortex II. Effects of sound stimulation," *Journal of Neurophysiology* 71, 1994, pp. 246-270.

Eggermont, J. J. "Neuronal pair and triplet interactions in the auditory midbrain of the leopard frog," *Journal of Neurophysiology* 66:5, 1991, pp. 1549-1563.

Eguiluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M. and Apkarian, A. V. "Scale free brain functional networks," *Physical Review Letters* 94:018102, 2005.

Eichler, M., Dahlhaus, R. and Sandkuhler, J. "Partial correlation analysis for the identification of synaptic connections," *Biological Cybernetics* 89, 2003, pp. 289-302.

Eldawlatly, S., Jin, R. and Oweiss, K. G. "Identifying functional connectivity in large-scale neural ensemble recordings: a multiscale data mining approach," *Neural Computation* 21, 2009, pp. 450-477.

Espinosa, I. E. and Gerstein, G. L. "Cortical auditory neuron interactions during presentation of 3-tone sequences: effective connectivity," *Brain Research* 450(1-2), 1988, pp. 39-50.

Euler, L. "Solutio problematis ad geometriam situs pertinentis," *Commentarii Academiae Scientiarum Imperialis Petropolitanae* 8, 1736, pp. 128-140.

Fagiolo, G. "Clustering in complex directed networks," *Journal of Physical Review E, statistical, nonlinear and soft matter physics* 77, 2007, 026107.

Fee, M. S., Mitra, P. P. and Kleinfeld, D. "Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability," *Journal of Neuroscience Methods* 69, 1996, pp. 175-188.

Ferri, R., Rundo, F., Brunt, O., Terzano, M. G. and Stam, C. J. "Small-world network organization of functional connectivity of EEG slowwave activity during sleep," *Clinical Neurophysiology* 118, 2007, pp. 449-456.

Freeman, L. C. "Centrality in social networks: conceptual clarification," Social Network 1, 1978, pp. 215-239.

Friston, K. J. "Functional and effective connectivity in neuroimaging: A synthesis," *Human Brain Mapping* 2, 1994, pp. 56-78.

Friston, K. J., Frith, C. D., Liddle, P. F. and Frackowiak, R. S. J. "Functional connectivity: The principal-component analysis of large (PET) data sets. J Cereb Blood Flow Metab 1 3 : 5-14.," *Journal of Cerebral Blood Flow and Metabolosm* 13, 1993, pp. 5-14.

Gabbiani, F. and Koch, C. "Methods in Neuronal Modeling: From Ions to Networks", *in* Koch, C. and Segev, I., ed., Cambridge MA: MIT, 1998, pp. 313-360.

Gerstein, G. L. and Aertsen, A. M. "Representation of cooperative firing activity among simultaneously recorded neurons," *Journal of Neurophysiology* 54:6, 1985, pp. 1513-1528.

Gerstein, G. L., Bloom, M. J. and Maldonado, P. E. "Organization and perturbation of neuronal assemblies", *in* J, J. B. and Poon, P., ed., Plenum, New York, 1988.

Gerstein, G. L. and Perkel, D. H. "Simultaneously recorded trains of action potentials: analysis and functional interpretation," *Science* 164:3881, 1969, pp. 828-830.

Gerstein, G. L., Perkel, D. H. and Dayhoff, J. E. "Cooperative firing activity in simultaneously recorded populations of neurons: detection and measurement," *The Journal of Neuroscience* 5:4, 1985, pp. 881-889.

Gerstner, W. and Kistler, W. M. *Spiking neuron models: single neurons, populations, plasticity*, Cambridge University Press, 2002.

Gerwinn, S., Macke, J., Seeger, M. and Bethge, M. "Bayesian inference for spiking neuron models with a sparsity prior."'Advances in neural information processing systems 20', 2007, pp. 529-536.

Gochin, P., Miller, E., Gross, C. and Gerstein, G. "Functional interactions among neurons in inferior temporal cortex of the awake macaque," *Experimental Brain Research* 84, 1991, pp. 505-516.

Gochin, P. M., Gerstein, G. L. and Kaltenbach, J. A. "Dynamics of temporal properties of effective connections in rat dorsal cochlear nucleus," *Brain Research* 510, 1990, pp. 195-202.

Grammont, F. and Riehle, A. "Spike synchronization and firing rate in a population of motor cortical neurons in relation to movement direction and reaction time," *Biological Cybernetics* 88(5), 2003, pp. 360-373.

Grammont, F. and Riehle, A. "Precise spike synchronization in monkey motor cortex involved in preparation for movement," *Experimental Brain Research* 128, 1999, pp. 118-122.

Grewe, J., Kretzberg, J., Warzecha, A. K. and Egelhaaf, M. "Impact of photon noise on the reliability of a motion-sensitive neuron in the fly's visual system," *Journal of Neuroscience* 23, 2003, pp. 10776-10783.

Gross, G. W. "Simultaneous single unit recording in vitro with a photoetched laser deinsulated gold multi-microelectrode surface," *IEEE Transactions on Biomedical Engineering* BME-26, 1979, pp. 273-279.

Grun, S., Diesmann, M. and Aertsen, A. "Unitary events in multiple single-neuron spiking activity: II. nonstationary data," *Neural Computation* 14(1), 2002b, pp. 81-119.

Grun, S., Diesmann, M. and Aertsen, A. "Unitary Events in multiple single-neuron spiking activity. I. Detection and significance," *Neural Computation* 14(1), 2002a, pp. 43-80.

Grun, S., Riehle, A. and Diesmann, M. "Effect of cross-trial nonstationarity on joint-spike events," *Biological Cybernetics* 88(5), 2003, pp. 335-351.

Harary, F. and Palmer, E. M. *Graphical enumeration*, New York: Academic Press. p. 124., 1973.

Harris, K., Csicsvari, J., Hirase, H., Dragoi, G. and Buzsaki, G. "Organization of cell assemblies in the hippocampus," *Nature* 424, 2003, pp. 552-556.

Harris, K. D., Henze, D. A., Csicsvari, J., Hirase, H. and Buzsaki, G. "Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements," *Journal of Neurophysiology* 84, 2000, pp. 401-414.

Hatsopoulos, N. G., Ojakangas, C. L., Paninski, L. and Donoghue, J. P. "Information about movement direction obtained from synchronous activity of motor cortical neurons," *Proceedings of the National Academy of Sciences of USA* 95, 1998, pp. 15706-15711.

Hilgetag, C. C., Kotter, R., Stephan, K. E. and Sporns, O. "Computational neuroanatomy: principles and methods", *in* Giorgo, A., ed., Humana Press, Totowa, NJ., 2002, pp. 295-335.

Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., Branner, A., Chen, D., Penn, R. D. and Donoghue, J. P. "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature* 442, 2006, pp. 164-171.

Holland, P. W. and Leinhardt, S. "An exponential family of probability distributions for directed graphs," *Journal of the American Statistical Association* 76:373, 1981, pp. 33-50.

Houghton, C. "Studying spike trains using a van Rossum metric with a synapse-like filter," *Journal of Computational Neuroscience* 26, 2009, pp. 149-155.

Houghton, C. and Sen, K. "A new multineuron spike train metric," *Neural Computation* 20(6), 2008, pp. 1495-1511.

Hunter, J. D. and Milton, J. G. "Amplitude and frequency dependence of spike timing: Implications for dynamic regulation," *Journal of Neurophysiology* 90, 2003, pp. 387-394.

Iglewicz, B. and Hoaglin, D. *How to detect and handle outliers*, ASQC Quality Press, 1993.

Iyengar, S. and Liao, Q. "Modeling neural activity using the generalized inverse Gaussian distribution," *Biological Cybernetics* 77, 1997, pp. 289-295.

Jermakowicz, W. J., Chen, X., Khaytin, I., Bonds, A. B. and Casagrande, V. A. "Relationship between spontaneous and evoked spike-time correlations in primate visual cortex," *Journal of Neurophysiology* 101, 2009, pp. 2279-2289.

Kaiser, M. and Hilgetag, C. C. "Edge vulnerability in neural and metabolic networks," *Biological Cybernetics* 90, 2004, pp. 311-317.

Kandel, E. R. *Principles of neural science*, McGraw-Hill, 2000.

Keat, J., Reinagel, P., Reid, R. C. and Meister, M. "Predicting every spike: a model for the responses of visual neurons," *Neuron* 30, 2001, pp. 803-817.

Kilavik, B., Roux, S., Ponce-Alvarez, A., Confais, J., S, S. G. and Riehle, A. "Long-term modifications in motor cortical dynamics induced by intensive practice," *Journal of Neuroscience* 29(40), 2009, pp. 12653-12663.

Konig, P., Engel, A., Roelfsema, P. and Singer, W. "How precise is neuronal synchronization?" *Neural Computation* 7(3), 1995, pp. 469-485.

Kotter, R. "Neuroscience databases: tools for exploring brain-structure-function relationships," *Philosophical Transactions of the Royal Society London* B 356, 2001, pp. 1111-1120.

Kreiman, G., Krahe, R., Metzner, W., Koch, C. and Gabbiani, F. "Robustness and variability of neuronal coding by amplitude-sensitive afferents in the weakly electric fish eigenmannia," *Journal of Neurophysiology* 84, 2000, pp. 189-204.

Kreuz, T., Haas, J. S., Morelli, A., Abarbanel, H. D. and Politi, A. "Measuring spike train synchrony," *Journal of Neuroscience Methods* 165, 2007, pp. 151-161.

Kulkarni, J. and Paninski, L. "Common-input models for multiple neural spike-train data," *Network: Computation in Neural Systems* 18, 2007, pp. 375-407.

Lansky, P. and Ditlevsen, S. "A review of the methods for signal estimation in stochastic diffusion leaky integrate-and-fire neuronal models," *Biological Cybernetics* 99, 2008, pp. 253-262.

Latora, V. and Marchiori, M. "Efficient behavior of small-world networks," *Physical Review Letters* 87, 2001, pp. 198701.

Lee, A. K. and Wilson, M. A. "A combinatorial method for analyzing sequential firing patterns involving an arbitrary number of neurons based on relative time order," *Journal of Neurophysiology* 92, 2004, pp. 2555–2573.

Lee, A. K. and Wilson, M. A. "Memory of sequential experience in the hippocampus during slow wave sleep," *Neuron* 36, 2002, pp. 1183-1194.

Lewicki, M. S. "A review of methods for spike sorting: the detection and classification of neural action potentials," *Network: Computation in Neural Systems* 9, 1998, pp. 53-78.

Li, Z., Morris, K. F., Baekey, D. M., Shannon, R. and Lindsey, B. G. "Multimodal medullary neurons and correlational linkages of the respiratory network," *Journal of Neurophysiology* 82, 1999, pp. 188-201.

Lindsey, B. G. "How is the respiratory central pattern generator configured and reconfigured?" *Advances in Experimental Medicine and Biology* 499, 2001, pp. 179-184.

Lindsey, B. G. and Gerstein, G. L. "Two enhancements of the gravity algorithm for multiple spike train analysis," *Journal of Neuroscience Methods* 150, 2006, pp. 116-127.

Lindsey, B. G., Hernandez, Y. M., Morris, K., Shannon, R. and Gerstein, G. L. "Dynamic reconfiguration of brain stem neural assemblies: respiratory phase-dependent synchrony vs. modulation of firing rates," *Journal of Neurophysiology* 67, 1992b, pp. 923-930.

Lindsey, B. G., Hernandez, Y. M., Morris, K., Shannon, R. and Gerstein, G. L. "Respiratory related neural assemblies in the brain stem midline," *Journal of Neurophysiology* 67, 1992a, pp. 905-922.

Lindsey, B. G., Hernandez, Y. M., Morris, K. F. and Shannon, R. "Functional connectivity between brain stem midline neurons with respiratory-modulated firing rates," *Journal of Neurophysiology* 67:4, 1992c, pp. 890-904.

Lindsey, B. G., Morris, K. F., Shannon, R. and Gerstein, G. L. "Repeated patterns of distributed synchrony in neuronal assemblies," *Journal of Neurophysiology* 78, 1997, pp. 1714-1719.

Lindsey, B. G., Shannon, R. and Gerstein, G. L. "Gravitational representation of simultaneously recorded brain stem respiratory neuron spike trains," *Brain Research* 483, 1989, pp. 373-378.

Louie, K. and Wilson, M. A. "Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep," *Neuron* 29, 2001, pp. 145-156.

Machens, C., Prinz, P., Stemmler, M., Ronacher, B. and Herz, A. "Discrimination of behaviorally relevant signals by auditory receptor neurons," *Neurocomputing* 38, 2001, pp. 263-268.

MacLeod, K., Backer, A. and Laurent, G. "Who reads temporal information contained across synchronized and oscillatory spike trains?" *Nature* 395, 1998, pp. 693-698.

Makarov, V. A., Panetsosa, F. and de Feob, O. "A method for determining neural connectivity and inferring the underlying network dynamics using extracellular spike recordings," *Journal of Neuroscience Methods* 144, 2005, pp. 265-279.

Maldonado, P., Babul, C., Singer, W., Rodriguez, E., Berger, D. and Grun, S. "Synchronization of neuronal responses in primary visual cortex of monkeys viewing natural images," *Journal of Neurophysiology* 100, 2008, pp. 1523-1532.

Maldonado, P. E. and Gerstein, G. L. "Neuronal assembly dynamics in the rat auditory cortex during reorganization induced by intracortical microstimulation," *Experimental Brain Research* 112, 1996, pp. 431-441.

Martignon, L., Deco, G., Laskey, K., Diamond, M., Freiwald, W. and Vaadia, E. "Neural coding: higher-order temporal patterns in the neurostatistics of cell assemblies," *Neural Computation* 12, 2000, pp. 2621-2653.

Martignon, L., von Hasseln, H., Grun, S., Aertsen, A. and Palm, G. "Detecting higher-order interactions among the spiking events in a group of neurons," *Biological Cybernetics* 73, 1995, pp. 69-81.

Martin, P. D. "Locomotion towards a goal alters the synchronous firing of neurons recorded simultaneously in the subiculum and nucleus accumbens of rats," *Behavioural Brain Research* 124, 2001, pp. 19-28.

Mechler, F., Victor, J. D., Purpura, K. P. and Shapley, R. "Robust temporal coding of contrast by V1 neurons for transient but not for steady-state stimuli," *Journal of Neuroscience* 18, 1998, pp. 6583-6598.

Micheloyannis, S., Pachou, E., Stam, C. J., Breakspear, M., Bitsios, P., Vourkas, M., Erimaki, S. and Zervakis, M. "Small-world networks and disturbed functional connectivity in schizophrenia," *Schizophrenia Research* 87, 2006b, pp. 60-66.

Micheloyannis, S., Pachou, E., Stam, C. J., Vourkas, M., Erimaki, S. and Tsirka, V. "Using graph theoretical analysis of multi channel EEG to evaluate the neural efficiency hypothesis," *Neuroscience Letter* 402, 2006a, pp. 273-277.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. and Alon, U. "Network motifs: simple building blocks of complex networks," *Science* 298, 2002, pp. 824-827.

Montez, T., Linkenkaer-Hansen, K., van Dijk, B. W. and Stam, C. J. "Synchronization likelihood with explicit time-frequency priors. Neuroimage," *Neuroimage* 33, 2006, pp. 1117-1125.

Morris, K. F., Baekey, D. M., Nuding, S. C., Segers, L. S., Shannon, R. and Lindsey, B. G. "Neural network plasticity in respiratory control," *Journal of Applied Physiology* 94, 2003, pp. 1242-1252.

Nakahama, H., Suzuki, H., Yamamoto, M., Aikawa, S. and Nishioka, S. "A statistical analysis of spontaneous activity of central single neurons," *Physiology and Behavior* 3, 1968, pp. 745-752.

Nedungadi, A. G., Rangarajan, G., Jain, N. and Ding, M. "Analyzing multiple spike trains with nonparametric granger causality," *Journal of Computational Neuroscience* 27, 2009, pp. 55-64.

Nicolelis, M. A. L., Dimitrov, D., Carmena, J. M., Crist, R., Lehew, G., Kralik, J. D. and Wise, S. P. "Chronic, multisite, multielectrode recordings in macaque monkeys," *PNAS* 100:19, 2003, pp. 11041–11046.

Nikolic, D. "Non-parametric detection of temporal order across pairwise measurements of time delays," *Journal of Computational Neuroscience* 22:1, 2007, pp. 5-19.

Nykamp, D. Q. "Exploiting history-dependent effects to infer network connectivity," *SIAM Journal of Applied Mathematics* 68:2, 2007, pp. 354-391.

Nykamp, D. Q. "Revealing pairwise coupling in linear-nonlinear networks," *SIAM J. APPL. MATH* 65:6, 2005, pp. 2005-2032.

Okatan, M., A.Wilson, M. and Brown, E. N. "Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity," *Neural Computation* 17, 2005, pp. 1927-1961.

Paiva, A. R. C., Park, I. and Principe, J. C. "A comparison of binless spike train measures," *Neural Computation and Application* 19, 2009, pp. 405-419.

Paninski, L. "Maximum likelihood estimation of cascade point-process neural encoding models," *Network: Computation in Neural Systems* 15, 2004, pp. 243-262.

Paninski, L., Pillow, J. and Lewi, J. "Statistical models for neural encoding, decoding, and optimal stimulus design," *Progress Brain Research* 165, 2007, pp. 493-507.

Park, I., Paiva, A. R. C., DeMarse, T. B. and Principe, J. C. "An efficient algorithm for continuous time cross correlogram of spike trains," *Journal of Neuroscience Methods* 168, 2008, pp. 514-523.

Perkel, D. H., Gerstein, G. L. and Moore, G. P. "Neuronal spike trains and stochastic point processes II. simultaneous spike trains," *Biophysical Journal* 7, 1967, pp. 419-440.

Pillow, J. "Likelihood-based approaches to modeling the neural code", *in* Doya, K., Ishii, S., Pouget, A. and Rao, R., ed., MIT press, 2007, pp. 53-70.

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J. and Simoncelli, E. P. "Spatio-temporal correlations and visual signalling in a complete neuronal population," *Nature* 454, 2008, pp. 995-999.

Ponten, S. C., Bartolomei, F. and Stam, C. J. "Small-world networks and epilepsy: graph theoretical analysis of intracranially recorded mesial temporal lobe seizures," *Clinical Neurophysiology* 118, 2007, pp. 918-927.

Posthuma, D., de Geus, E. J. C., Mulder, E. J. C. M., Smit, D. J. A., Boomsma, D. I. and Stam, C. J. "Genetic components of functional connectivity in the brain: the heritability of synchronization likelihood," *Human Brain Mapping* 26, 2005, pp. 191-198.

Potter, S. M. "Distributed processing in cultured neuronal networks," *Progress Brain Research* 130, 2001, pp. 49-62.

Quiroga, R. Q., Nadasdy, Z. and Ben-Shaul, Y. "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural Computation* 16, 2004, pp. 1661–1687.

Reed, J. L. and Kaas, J. H. "Statistical analysis of large-scale neuronal recording data," *Neural Networks* 23, 2010, pp. 673-684.

Reich, D., Mechler, F. and Victor, J. "Temporal coding of contrast in primary visual cortex: when, what, and why?" *Journal of Neurophysiology* 85, 2001, pp. 1039-1050.

Reich, D. S., Victor, J. D. and Knight, B. W. "The power ratio and the interval map: spiking models and extracellular recordings," *Journal of Neuroscience* 18, 1998, pp. 10090-10104.

Reijneveld, J. C., Ponten, S. C., Berendse, H. W. and Stam, C. J. "The application of graph theoretical analysis to complex networks in the brain," *Clinical Neurophysiology* 118, 2007, pp. 2317-2331.

Reinagel, P. and Reid, R. C. "Precise firing events are conserved across neurons," *Journal of Neuroscience* 22, 2002, pp. 6837-6841.

Riehle, A., Grammont, F., Diesmann, M. and Grun, S. "Dynamical changes and temporal precision of synchronized spiking activity in monkey motor cortex during movement preparation," *Journal of Physiology* 94, 2000, pp. 569-582.

Riehle, A., Grun, S., Diesmann, M. and Aertsen, A. "Spike synchronization and rate modulation differentially involved in motor cortical function," *Science* 278, 1997, pp. 1950-1953.

Rieke, F., Warland, D., de Ruyter van Steveninck, R. R. and Bialek, W. *Spikes: Exploring the Neural Code*, Cambridge, MA: MIT, 1997.

Rigat, F., de Gunsty, M. and van Peltz, J. "Bayesian modelling and analysis of spatio-temporal neuronal networks," *Bayesian Analysis* 1:4, 2006, pp. 733-764.

Rodieck, R. W., Kiang, N. Y.-S. and Gerstein, G. L. "Some quantitative methods for the study of spontaneous activity of single neurons," *Biophysical Journal* 2, 1962, pp. 351-368.

van Rossum, M. C. W. "A novel spike distance," *Neural Computation* 13, 2001, pp. 751-763.

Rubinov, M. and Sporns, O. "Complex network measures of brain connectivity : uses and interpretations," *Neroimage* 52, 2010, pp. 1059-1069.

Salvador, R., Suckling, J., Coleman, M. R., Pickard, J. D., Menon, D. and Bullmore, E. "Neurophysiological architecture of functional magnetic resonance images of human brain," *Cerebral Cortex* 15, 2005a, pp. 1332-1342.

Salvador, R., Suckling, J., Schwarzbauer, C. and Bullmore, E. "Undirected graphs of frequency-dependent functional connectivity in whole brain networks," *Philosophical Transactions of the Royal Society of London. Series B , Biological Sciences* 360, 2005b, pp. 937-946.

Samonds, J. M., Allison, J. D., Brown, H. A. and Bonds, A. B. "Cooperation between area 17 neuron pairs enhances fine discrimination of orientation," *Journal of Neuroscience* 23, 2003, pp. 2416-2425.

Samonds, J. M. and Bonds, A. B. "From another angle: differences in cortical coding between fine and coarse discrimination of orientation," *Journal of Neurophysiology* 91, 2004, pp. 1193-1202.

Schneider, G., Havenith, M. N. and Nikolic, D. "Spatiotemporal structure in large neuronal networks detected from cross-correlation," *Neural Computation* 18, 2006, pp. 2387-2413.

Schreiber, S., Fellous, J., Whitmer, D., Tiesinga, P. and Sejnowski, T. "A new correlation-based measure of spike timing reliability," *Neurocomputing* 52-54, 2003, pp. 925-931.

Seo, S. "A review and comparison of methods for detecting outliers in univariate data sets", *Masters thesis,* University of Pittsburgh, 2006.

Shadlen, M. N. and Newsome, W. T. "The variable discharge of cortical neurons: implications for connectivity, computation, and information coding," *Journal of Neuroscience* 18, 1998, pp. 3870-3896.

Shahid, S. and Smith, L. S. "A novel technique for spike detection in extracellular neurophysiological recordings using cepstrum of bispectrum"'Proceedings of European Signal Processing Conference', 2008.

Shahid, S., Walker, J. and Smith, L. S. "A new spike detection algorithm for extracellular neural recordings," *Journal of IEEE Transactions on Biomedical Engineering* 57(4), 2010, pp. 853-866.

Shannon, R., Baekey, D. M., Morris, K. F., Li, Z. and Lindsey, B. G. "Functional connectivity among ventrolateral medullary respiratory neurones and responses during fictive cough in the cat," *Journal of Physiology* 525.1, 2000, pp. 207-224.

Shao, X. M. and Tsau, Y. "Measure and statistical test for cross-correlation between paired neuronal spike trains with small sample size," *Journal of Neuroscience Method* 70, 1996, pp. 141-152.

Simoncelli, E., Paninski, L., Pillow, J. and Schwartz, O. "Characterization of neural responses with stochastic stimuli", *in* Gazzaniga, M., ed., MIT press, 2004.

Skaggs, W. E. and McNaughton, B. L. "Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience," *Science* 271:5257, 1996, pp. 1870-1873.

Smit, D. J. A., Stam, C. J., Boomsma, D. I., Posthyma, D. and de Geus, E. J. C. "Heritability of "small-world" networks in the brain: a graph theoretical analysis of resting-state EEG functional connectivity," *Human Brain Mapping* 29, 2008, pp. 1368-1378.

Sporns, O. "Networks of the brain: quantitative analysis and modeling", Technical report, Department of Psychological and Brain Sciences, Indiana University, Bloomington, Indiana, 2010.

Sporns, O. "Brain connectivity," *Scholarpedia* 2(10):4695, 2007.

Sporns, O. "Neuroscience databases: a practical guide", *in* Kotter, R., ed., Klüwer, Boston, MA, 2002, pp. 171–186.

Sporns, O., Christopher, J. H. and Kotter, R. "Identification and classification of hubs in brain networks," *PLoS ONE* 2(10): e1049, 2007.

Sporns, O. and Kotter, R. "Motifs in brain networks," *PLOS Biology* 2, 2004, pp. 1910-1918.

Sporns, O. and Zwi, J. D. "The small-world of the cerebral cortex," *Neuroinformatics* 2, 2004, pp. 145-162.

Stam, C. J. "Functional connectivity patterns of human magnetoencephalographic recordings: a "small-world" network?" *Neuroscience Letter* 355, 2004, pp. 25-28.

Stam, C. J. and van Dijk, B. W. "Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets," *Physica D: Nonlinear Phenomena* 163, 2002, pp. 236-241.

Stam, C. J., Jones, B. F., Nolte, G., Breakspear, M. and Scheltens, P. "Small world networks and functional connectivity in Alzheimer's disease," *Cerebral Cortex* 17, 2007, pp. 92-99.

Stark, E., Drori, R. and Abeles, M. "Partial cross-correlation analysis resolves ambiguity in the encoding of multiple movement features," *Journal of Neurophysiology* 95, 2006, pp. 1966-1975.

Staude, B., Grun, S. and Rotter, S. "Higher order correlations and cumulants", *in* Grun, S. and Rotter, S., ed., Springer Series in Computational Neuroscience, Springer Verlag, 2010b, pp. 253-82.

Staude, B., Rotter, S. and Grün, S. "CuBIC: cumulant based inference of higher-order correlations in massively parallel spike trains," *Journal of Computational Neuroscience* 29(1-2), 2010a, pp. 327-350.

Stevenson, I. H., Rebesco, J. M., Hatsopoulos, N. G., Haga, Z., Miller, L. E. and Körding, K. P. "Bayesian inference of functional connectivity and network structure from spikes," *IEEE Transsactions on Neural Systems and Rehabilitation Engineering* 17(3):3, 2009, pp. 203-213.

Stevenson, I. H., Rebesco, J. M., Miller, L. E. and Kording, K. P. "Inferring functional connections between neurons," *Current Opinion in Neurobiology* 18, 2008, pp. 582-588.

Stuart, L., Walter, M. and Borisyuk, R. "The correlation grid: analysis of synchronous spiking in multi-dimensional spike train data and identification of feasible connection architectures," *BioSystems* 79, 2005, pp. 223-233.

Stuart, L., Walter, M. and Borisyuk, R. "Visualisation of synchronous firing in multi-dimensional spike trains," *BioSystems* 67, 2002, pp. 265-279.

Thomas, C. A., Springer, P. A., Loeb, G. E., Berwald-Netter, Y. and Okun, L. M. "A miniature microelectrode array to monitor the bioelectric activity of cultured cells," *Experimental Cell Research* 74, 1972, pp. 61-66.

Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P. and Brown, E. N. "A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects," *Journal of Neurophysiology* 93, 2005, pp. 1074-1089.

Truccolo, W., Friehs, G. M., Donoghue, J. P. and Hochberg, L. R. "Primary motor cortex tuning to intended movement kinematics in humans with tetraplegia," *Journal of Neuroscience* 28(5), 2008a, pp. 1163-1178.

Truccolo, W., Hochberg, L. R. and Donoghue, J. P. "Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes," *Nature Neuroscience* 13(1), 2010, pp. 105-111.

Truccolo, W., Hochberg, L. R., Eskandar, E., Cole, A. and Cash, S. S. "Multielectrode array recordings of single unit activity in humans with epilepsy"'Neural interfaces conference', 2008b.

Tuckwell, H. *Introduction to theoretical neurobiology*, New York: Cambridge University Press, 1988.

Utikal, K. J. "A new method for detecting neural interconnectivity," *Biological Cybernetics* 76, 1997, pp. 459-470.

Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H. and Aertsen, A. "Dynamics of neural interactions in monkey cortex in relation to behavioural events," *Nature* 373, 1995, pp. 515-518.

Vaadia, E., Kurata, K. and Wise, S. P. "Neuronal activity preceding directional and nondirectinal cues in the premotor cortex of rhesus monkeys," *Somatosensory and Motor Research* 6(2), 1988, pp. 207-230.

Vaknin, R., Goldberg, D. H., Victor, J. D., Gardner, E. P., Debowy, D. J., Babu, K. S. and Gardner, D. "Metric space analysis of responses in parietal cortex during prehension [abstract]," *Society for Neuroscience* 31:984.20., 2005.

Victor, J. D. and Purpura, K. P. "Metric-space analysis of spike trains: theory, algorithms and application," *Network: Computational Neural System* 8, 1997, pp. 127-164.

Victor, J. D. and Purpura, K. P. "Nature and precision of temporal coding in visual cortex: a metric-space analysis," *Journal of Neurophysiology* 76(2), 1996, pp. 1310-1326.

Watts, D. J. and Strogatz, S. H. "Collective dynamics of "small-world" networks," *Nature* 393, 1998, pp. 440-442.

Wilk, M. B. and Gnanadesikan, R. "Probability plotting methods for the analysis of data," *Biometrica* 55, 1968, pp. 1-17.

Wilson, M. A. and McNaughton, B. L. "Reactivation of hippocampal ensemble memories during sleep," *Science* 265, 1994, pp. 676-679.

Wilson, M. and McNaughton, B. "Dynamics of the hippocampal ensemble code for space," *Science* 261:5124, 1993, pp. 1055-1058.

Wise, K. D., Angell, J. B. and Starr, A. "An integrated-circuit approach to extracellular microelectrodes," *IEEE Transactions on Biomedical Engineering* BME-17, 1970, pp. 238-247.

Wu, H., Li, X. and Guan, X. "Networking property during epileptic seizure with multi-channel EEG recordings," *Lecture Notes in Computer Science* 3976, 2006, pp. 573-578.

Young, M. P. "The organization of neural systems in the primate cerebral cortex," *Proceedings of the Royal Society London* B 252, 1993, pp. 13-18.

Bound copy of published paper

Journal of Neuroscience Methods 196 (2011) 201–219

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Statistical technique for analysing functional connectivity of multiple spike trains

Mohammad Shahed Masud[a,*], Roman Borisyuk[a,b]

[a] School of Computing and Mathematics, University of Plymouth, A222, Portland Square, Plymouth, PL4 8AA, UK
[b] Institute of Mathematical Problems in Biology of the Russian Academy of Sciences, Pushchino, Moscow Region, 142290, Russia

ARTICLE INFO

Article history:
Received 24 May 2010
Received in revised form 5 January 2011
Accepted 6 January 2011

Keywords:
Hazard function
Modulated renewal process
Cox method
Conditional likelihood

ABSTRACT

A new statistical technique, the Cox method, used for analysing functional connectivity of simultaneously recorded multiple spike trains is presented. This method is based on the theory of modulated renewal processes and it estimates a vector of influence strengths from multiple spike trains (called reference trains) to the selected (target) spike train. Selecting another target spike train and repeating the calculation of the influence strengths from the reference spike trains enables researchers to find all functional connections among multiple spike trains. In order to study functional connectivity an "influence function" is identified. This function recognises the specificity of neuronal interactions and reflects the dynamics of postsynaptic potential. In comparison to existing techniques, the Cox method has the following advantages: it does not use bins (binless method); it is applicable to cases where the sample size is small; it is sufficiently sensitive such that it estimates weak influences; it supports the simultaneous analysis of multiple influences; it is able to identify a correct connectivity scheme in difficult cases of "common source" or "indirect" connectivity. The Cox method has been thoroughly tested using multiple sets of data generated by the neural network model of the leaky integrate and fire neurons with a prescribed architecture of connections. The results suggest that this method is highly successful for analysing functional connectivity of simultaneously recorded multiple spike trains.

1. Introduction

Multi-electrode array (MEA) enables a simultaneous recording of electrical activities of many neurons (Boven et al., 2006). From the application of a spike sorting technique (e.g. Quiroga et al., 2004) to such recordings, it is possible to extract multiple spike trains associated with different neurons. Simultaneously recorded spike trains are used to study how groups of neurons process information and how they interact with each other. Developing a new statistical method for analysing multiple spike trains and, in particular, estimating the functional connectivity between spike trains, is a challenging problem that has resulted in substantial research (Brown et al., 2004; Reed and Kaas, 2010).

The standard approach to analysing functional connectivity is based on calculation of the cross-correlation function (CCF) (Perkel et al., 1967). However, there are other methods to characterise the pair-wise dependencies between spike trains. These include the cross-intensity function (Cox and Lewis, 1972; Brillinger, 1976a, 1992), product densities, cumulant densities, cumulant spectra, method of moments (Bartlett, 1966; Brillinger, 1975a, 1975b), calculation of the coherence (Brillinger, 1976b, 1992), and the joint peristimulus time histogram (JPSTH) (Gerstein and Perkel, 1969, 1972; Aertsen et al., 1989). All these methods are focused on a pair of spike trains but they fail to consider all possible influences from other simultaneously recorded spike trains. For this reason, these pair-wise estimates of the functional connectivity sometimes can lead to inaccuracies. Furthermore, there are some well known limitations of statistical methods based on CCF: (1) Most of the techniques based on the CCF require the spike trains to be stationary, (2) The number of spikes should be large enough to ensure reliable estimation and (3) CCF based methods are linear and only consider a linear component of interconnectivity.

Alternatively, a different approach to analysing functional connectivity is based on the maximum likelihood (ML) method which estimates the probability of a spike as a result of multiple influences from other spike trains. Using the ML function, the algorithm calculates the regression parameters, which characterise the strength of the influences.

Recently, a new technique called the generalised linear models (GLM) was introduced to the study of neuronal interactions (Stevenson et al., 2008). The GLM assumes that the neuron's spike is influenced by such factors as the neuron's own recent spiking activity, the recent spiking activity of other neurons and the

* Corresponding author. Tel.: +44 01752 84920; fax: +44 01752 586300.
E-mail addresses: mohammad.masud@plymouth.ac.uk (M.S. Masud), rborisyuk@plymouth.ac.uk (R. Borisyuk).

0165-0270/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2011.01.003

activity of some external variables (stimuli). The conditional intensity function is the exponential of the linear combination of the factors being analysed. Brillinger (1988) developed such a GLM for a spike train based on the influence from other neurons within the group. Chornoboy et al. (1988) also use the GLM for the simultaneous analysis of multiple pair-wise interactions among neurons. In their model, the probability of a spike response depends upon the neuron's own spiking activity as well as the activity of other neurons in the population. Okatan et al. (2005) introduced a different method to estimate the functional connectivity of stochastic neural networks, based on a discrete time version of the approach developed by Chornoboy et al. (1988). Truccolo et al. (2005) used the GLM to estimate influences to the target spike train from both other spike trains of the neural population and the external inputs to the population. A good performance of this technique has been demonstrated in a special case of simulated activities of six neurons and in case of experimental recordings from motor neurons. A number of recent studies (Pillow, 2007; Pillow et al., 2008; Paninski et al., 2007) use the GLM to investigate the influence of sensory stimuli to spiking activity of the neural population (neuronal code). The difficulties associated with the study of neuronal coding are substantial but the evidence shows that the GLM based approach is useful.

However, there are some limitations of GLM approach. For example, the result of the analysis depends upon the size of the testing window (bin) as it is used to find estimates of parameters describing the influences to the spike train (Eldawlatly et al., 2009). It is clear that estimated values of parameter depend on the selected size of the bin. Additionally, GLM contains many parameters (approximately hundred). The result of this is that the optimization problem (finding the maximum of the likelihood function) could have a non-unique solution (see e.g. Stevenson et al., 2008; Chornoboy et al., 1988). A standard approach to resolve this difficulty is to incorporate prior knowledge about the nature of the influences. There are a variety of techniques to deal with this problem including the regularization method, Bayesian approach, and calculation of the maximum a posteriori (MAP) estimate (see Paninski, 2004; Rigat et al., 2006; Gerwinn et al., 2007; Stevenson et al., 2009).

In this work, study of the functional connectivity of neurons is based on the modulated renewal process (MRP) (Cox, 1972 and Borisyuk et al., 1985)[1]. The MRP is considered in terms of spike generation under multiple influences from other spike trains and estimates of the strength of each influence using the Cox method, which itself is based on conditional likelihood method (Cox, 1972). The MRP model describes the hazard function of spike appearance at the MRP and it includes a modulation which is the exponential of the linear combination of influence functions. In fact, this model is similar to the regression model and the set of influence strengths is similar to the regression coefficients. The definition of the influence function is based on some neurobiological details of spike generation and propagation. This function reflects the dynamics of postsynaptic potential under bombardment by spikes from other neurons.

When the original paper on the application of the Cox method to neuroscience data was published, back in 1985 (Borisyuk et al., 1985), its use was limited by the availability of computation power. At that time, only pairs and triplets of spike trains were considered. However, in this paper, the Cox method is further developed to support simultaneous consideration of any possible set of multiple spike trains. The corresponding formulas for the calculation of estimates of the influence strengths and their confidence intervals have been derived. Thus, this new development of the Cox method enables researchers to simultaneously analyse any number (n) of spike trains (where $n = 3, 4, 5, \ldots$). A numerical method and software application have been developed for the identification of the functional connectivity from the simultaneously recorded multiple spike trains. Testing results have shown that the Cox method is highly successful. Therefore, this method is recommended for the analysis of functional connectivity of neuronal circuits from multiple spike trains. The software is freely available from the authors on request.

Rigorous testing of the Cox method has shown that this statistical technique is not only highly efficient but it also overcomes some of the limitations of other classical methods. The Cox method has been tested on numerous sets of simultaneous spike trains; both artificially generated by different mathematical models as well as datasets recorded in experiments. These tests have shown that this method works well and provides reliable estimates of influence strengths. These influence strengths are used to define a diagram of functional connections. In particular, the Cox method can be used to analyse the functional connectivity of large groups (up to hundred) of spike trains. Thus, we conclude that this Cox method is a useful tool for analysing experimental data of multi-electrode recordings.

In conclusion, the main advantages of this method are: (1) It does not require the specification of a bin (see Paiva et al., 2009 for a review of binless techniques); (2) it supports the simultaneous analysis of multiple spike trains and provides statistical estimates of influence strengths and their confidence intervals (to test the hypothesis that the influence is zero); (3) it is applicable in situations where sample sizes are small; (4) it is sufficiently sensitive such that it estimates weak influences; and finally (5) it is able to identify a correct connectivity scheme in difficult cases of "common source" or "indirect" connectivity.

This paper describes the Cox method and demonstrates the application of the Cox method to both a small neural circuit of five spike trains and to a large circuit of twenty spike trains. In both cases, the Cox method is shown to be effective for analysing the functional connectivity. In these tests, an enhanced leaky integrate and fire model (ELIF) is used to generate the data with a prescribed scheme of connections. Naturally, this scheme is never used during the analysis phase. It is solely required for evaluation of the final results. Additionally, the Cox method is also compared to the CCF approach for analysing two and three spike trains. The results demonstrate that the Cox method has some significant advantages when compared to the pair-wise cross-correlation approach. Although in this paper the only excitatory connections are considered, the Cox method is applicable for analysing of both excitatory and inhibitory connection strengths.

Section 2 of this paper describes the Cox method and defines the influence function for use within neuroscience. Section 3 reports the results of applying the Cox method to a neural circuit of five and a neural circuit of twenty spike trains generated by ELIF model. In Section 4, a comparison of the Cox method with CCF is presented using pairs of spike trains generated by a probabilistic model. This probabilistic model is based on the modulated renewal process. Furthermore, it is a convenient technique for testing due to the common probabilistic basis of the MRP model and the Cox method. Subsequently, the analysis of a set of three spike trains is used to demonstrate that the Cox method can find a scheme of connections in a case of "common source" connection architectures. The method was similarly successful when used to analyse sets of three spike trains with "indirect connection" architecture. This is a problem for pair-wise methods to analyse these kinds of connectivity. A discussion of the work is presented in Section 5, which summarises the results and highlights the advantages of the method.

Appendix A presents the formulas for the calculation of the estimates of Cox coefficients. Appendix B presents the parameter values for the ELIF generator. Finally, in Appendix C, the flexibil-

[1] Also, in many cases a simplified approach based on the modulated Poisson process might be useful.

ity of the probabilistic model is shown and the parameters of the probabilistic model can be adjusted using an optimization procedure such that these parameter values can be subsequently used to generate spike train similar to the "integrate and fire" neuron model.

2. Statistical technique to analyse functional connectivity (Cox method)

In this section we revise the Cox method (Cox, 1972; Borisyuk et al., 1985) which provides statistical estimates (and their confidence intervals) for the strength of influence from one spike train to another. In fact, we test the hypothesis that the influence from one spike train to another is zero and in case if this hypothesis is rejected we use the value of the estimate as a measure of influence. This technique is based on the assumption of the renewal process and modulated renewal process. Considering the renewal process we assume that inter-spike intervals (ISIs) are independent with the probability density function $f(x)$. If $f(x)$ is the non-exponential density function then the process is called the renewal process. In fact, the assumption of the renewal processes is broadly used in neuroscience. Another approach to specification of the renewal process comes from survival analysis where spike rate density or rate of death (or failure) per time unit (hazard function) can be interpreted in terms of the density function $f(x)$. The hazard function is defined as the spike rate at time t conditional on survival time (without death) until time t or later:

$$\varphi(t) = \lim_{\Delta t \to 0} \frac{Pr(t \le X < t + \Delta t | t \le X)}{\Delta t} = \frac{f(t)}{1 - F(t)},$$

where X is inter-spike interval, $F(t)$ is the cumulative probability function of ISI.

The modulated renewal process allows introducing of dependencies (influences) between spike trains. Let us suppose that spike generation in spike train A depends on spikes of spike train B and the hazard function of process A is a product of two multipliers: one is the own hazard of process A without influence from B and another multiplier describes influence from spike train B. Thus, the hazard function at the moment t is:

$$\varphi(t) = \varphi_A(U_A(t)) \exp[\beta Z_B(t)], \quad (1)$$

where $\varphi_A(\cdot)$ is the hazard function of the renewal process A without influence from the point process B, $U_A(t)$ is the backward recurrence time of the process A at the moment t, $Z_B(t)$ is the influence function determining how the process B influences the process A, and β is the unknown parameter (Cox coefficient) describing the strength of the influence from process B to A. Therefore, given the influence function $Z_B(t)$, the goal is to estimate the parameter β. If $\beta = 0$ then there is no influence from spike train B to A. To test the hypothesis $H_0: \{\beta = 0\}$ we use statistical technique based on a conditional maximum likelihood principle (Cox, 1972). Application of the Cox method to analyse influences between two or three spike trains is described in (Borisyuk et al., 1985). Here we present a generalisation of the Cox method for simultaneous analysing of arbitrary number (n) of spike trains and demonstrate how this technique can be applied to study the functional connectivity of neural circuits.

The Cox method is applied to analyse a set of n simultaneously recorded spike trains. One spike train is selected to be considered as a target spike train and all other $(n-1)$ spike trains are considered to be the reference spike trains. The Cox method allows analysing of all n spike trains and estimating the $(n-1)$ dimensional vector β of regression coefficients under the assumption (1), where $Z_B(t)$ is $(n-1)$ dimensional vector-function of influences from reference spike trains to the target and $\beta Z_B(t)$ is the dot product. Application of the Cox method provides both the estimates of unknown parameter (Cox coefficients) $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_{n-1}$ and the corresponding confidence intervals of these estimates $\{[lb_i, ub_i], i = 1, 2, \ldots (n-1)\}$, where lb_i and ub_i are lower and upper boundaries respectively of the confidence interval for $\hat{\beta}_i$. The hypothesis $H_0^i: \beta_i = 0$ is accepted if the corresponding confidence interval contains zero ($0 \in [lb_i, ub_i]$) otherwise the hypothesis is rejected and the estimate $\hat{\beta}_i$ is considered as a measure of influence strength from the ith reference spike train to the target. To study the functional connectivity we apply this method consequently (n times) selecting the target and estimating the influence strengths from reference spike trains. The equations that are used to estimate the vector of regression coefficients and their confidence intervals are given in Appendix A.

2.1. The influence function

It is important for successful application of the Cox method to choose an appropriate influence function which takes into account some characteristic properties of neuronal interactions. We have studied (both analytically and numerically) different candidates and find that the alpha function which is used in neuroscience to describe synaptic connectivity between neurons (Gerstner and Kistler, 2002) is an appropriate description for modulation of the hazard function (see formula (1)). In fact, we find that the best influence function should describe the dynamics of postsynaptic potential: the influence function increases when spike arrives to the postsynaptic neuron and the probability of spike generation by postsynaptic neuron increases; after that the influence function decays to zero. Thus, the influence function is:

$$Z_B(t) = \frac{g_m}{(\tau_s - \tau_r)}(e^{-t/\tau_s} - e^{-t/\tau_r}) \quad (2)$$

where τ_s and τ_r are the characteristic times of decay and rise of postsynaptic potential respectively. Parameter g_m provides the normalization that the maximum of the influence function is one.

A simplified version of the influence function corresponding to the case $\tau_s = \tau_r$ is given by the following formula:

$$Z_B(t) = \frac{t}{\tau_s} e^{1 - t/\tau_s} \quad (3)$$

In formulas (2) and (3) we assume that the presynaptic neuron (spike train B) has generated an action potential at time zero and this spike arrives to the target neuron (spike train A) without any delay. Now we re-write formulas (2) and (3) taking into account times of spiking at the reference process B. We suppose the influence function depends on the backward recurrence time of the process B which we denote as $U_B(t)$ and substitute this variable as the argument to formula (2). It means that the influence described by formula (2) starts increasing from the last spike in B before time t. To take into account the time delay of spike propagation, the argument in formula (2) should be shifted by the time lag Δ. Thus the influence function is defined by the following formula:

$$Z_B(t) = \frac{g_m}{(\tau_s - \tau_r)}(e^{-U_B(t-\Delta)/\tau_s} - e^{-U_B(t-\Delta)/\tau_r}) \quad (4)$$

where $U_B(t)$ is the backward recurrence time of the process B; Δ is a time lag corresponding to delay of spike propagation from B to the target A. Similar to (3), a simplified version of the influence function with time lag Δ is given by the following formula:

$$Z_B(t) = \frac{1}{\tau_s} U_B(t - \Delta) e^{1 - U_B(t-\Delta)/\tau_s} \quad (5)$$

In formulas (4) and (5) we define the influence function taking into account only the last spike in B before the time $(t - \Delta)$. In some cases it is fruitful to consider accumulation of postsynaptic potential in time interval T considering a history of spiking in B over the time interval $(t - T, t)$. This type of the influence function is useful if

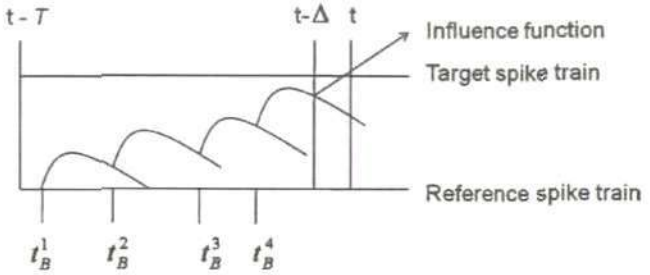


Fig. 1. Generalised influence function which accumulates influences from several spikes of the reference spike train in the history interval $(t-T, t)$ with propagation delay Δ.

decay time of the postsynaptic potential is relatively small in comparison with the mean interspike interval of the process B. Thus, a generalised influence function is:

$$Z_B(t) = \sum_{j=1}^{k} \frac{g_m}{(\tau_s - \tau_r)} \left(e^{-U_B^j(t-\Delta)/\tau_s} - e^{-U_B^j(t-\Delta)/\tau_r}\right) \quad (6)$$

here k is an index which denotes the highest order of the backward recurrence time in the history interval (the first order corresponds to a spike which is the most close to the moment $(t-\Delta)$ in backward time, the second order relates to the previous spike in the reward time, etc.). Fig. 1 shows the generalised influence function over the accumulation time T, thus, $U_B^k(t) < T$ and $U_B^{k+1}(t) > T$. A simplified version of the generalised influence function is given by the following formula:

$$Z_B(t) = \sum_{j=1}^{k} \frac{1}{\tau_s} U_B^j(t-\Delta) \exp\left(1 - \frac{1}{\tau_s} U_B^j(t-\Delta)\right) \quad (7)$$

In case of multiple reference spike trains $B = (B_1, B_2, \ldots, B_k)$, the influence function should be defined independently for each reference spike train and the hazard of the target spike train A is

$$\varphi(t) = \varphi_A(U_A(t)) \exp\{\beta_1 Z_{B_1}(t) + \beta_2 Z_{B_2}(t) + \cdots + \beta_k Z_{B_k}(t)\}, \quad (8)$$

where $\varphi_A(\cdot)$ is the hazard function of the renewal process A without influence from the reference processes B, $U_A(t)$ is the backward recurrence time of the process A at the moment t, $Z_{B_i}(t)$ is the influence function determining how the process B_i influences the process A, and β_i is the unknown parameter describing the strength of the influence from the process B_i to A $(i = 1,2,\ldots,k)$.

3. Functional connectivity of multiple spike trains

In this section we use the Cox method to identify functional connectivity of simulated multiple spike trains. For simulations we use the Enhanced Leaky Integrate and Fire model (Borisyuk, 2002) with a given scheme of coupling. Simulating a small neural circuit of five ELIF elements with prescribed connectivity we generate five spike trains and use the Cox method to identify the functional connectivity. We show that the method is effective and allows finding all connections and identifying their relative strengths. The Cox method is multivariate and enables to analyse all simultaneously recorded spike trains at ones. A relatively large circuit of twenty ELIF elements is used to generate twenty spike trains. Application of the Cox method to analyse twenty spike trains at ones can identify all functional connections; therefore, these results show that the Cox method is fruitful for study of functional connectivity of multiple spike trains.

3.1. Enhanced leaky integrate and fire model for data generation

We consider a neural network of enhanced leaky integrate and fire elements (Perkel, 1976; Borisyuk, 2002). A discrete-time version of the model neuron is used with the time increment equal to 1 ms. The state of each neuron at the moment t is characterised by a threshold and the total potential which is the sum of postsynaptic potentials and the noise. If the value of the total potential has reached the threshold, the neuron generates a spike. The spike propagates to other neurons with a time delay. The diagram of connection should be defined as well as connection strengths, time delays, and time decays of postsynaptic potentials. When the spike reaches another neuron, the postsynaptic potential jumps up or down depending on whether the spike is from an excitatory or inhibitory neuron, respectively. The value of the connection strength controls the jump height. The postsynaptic potential exponentially decays to the resting potential if there are no incoming spikes. After spike generation, the neuron is unable to generate a spike during an absolute refractory period. When this period expires, the threshold gets the highest value and then exponentially decays to the asymptotic threshold value. This decay is used to model a relative refractory period. To model a spontaneous background activity, the random noise is added to the membrane potential. The amplitude of the noise exponentially decays with time and a normally distributed random variable with zero mean and a fixed variance is added to the noise at each time step. The noise is independent random process for each element. If the amplitude of noise is large enough, then the element can be spontaneously active even without influences from other neurons. The formal description of the enhanced integrate-and-fire element and examples of parameter values are given in paper (Borisyuk, 2002) and a short description of parameters of the ELIF model is provided in Appendix B.

Table 1
Connection strengths, time delays of spike propagation and decay times of postsynaptic potential that are used for generating five spike trains.

Connection strength (w)	Time delay (Δ)	Decay time (τ_s)
$W_{4,1} = 10.786$	$\Delta_{41} = 12$	2.09
$W_{4,2} = 11.081$	$\Delta_{42} = 10$	1.63
$W_{4,3} = 8.973$	$\Delta_{43} = 10$	4.66
$W_{1,4} = 7.354$	$\Delta_{14} = 10$	4.35
$W_{3,5} = 6.901$	$\Delta_{35} = 6$	4.35

3.2. Analysis of functional connectivity of five spike trains

We generate the five spike trains using the connection architecture shown in Fig. 2A. The values of connection strength, time delay of spike propagation and time of decay of postsynaptic potential are given in Table 1. In this paper we do not consider self-coupling and therefore, we do not analyse self-influences. Fig. 2B shows the result of ELIF model simulation, i.e. the raster plot of spiking activity of these five neurons in time interval of 20 s. These spike trains are considered as a data set for analysing the functional connectivity. It is important to note that for analysing the functional connectivity we use only spike trains and we suppose that the scheme of connections is unknown. Also we assume that values of neuronal parameters and parameters characterising connections (connection strength, delay of spike propagation and time of decay of PSP) are also unknown. After completion the analysis, we compare the

Table 2
Statistics of the ISI of the neural circuit of five spike train.

Spike train	Number of spikes	Mean	Minimum	Maximum	Standard deviation
1	281	70.7	17	267	43.6
2	245	81.1	18	271	47.3
3	244	81.4	13	315	47.9
4	261	76.3	10	235	46.7
5	246	81.4	15	326	50.7

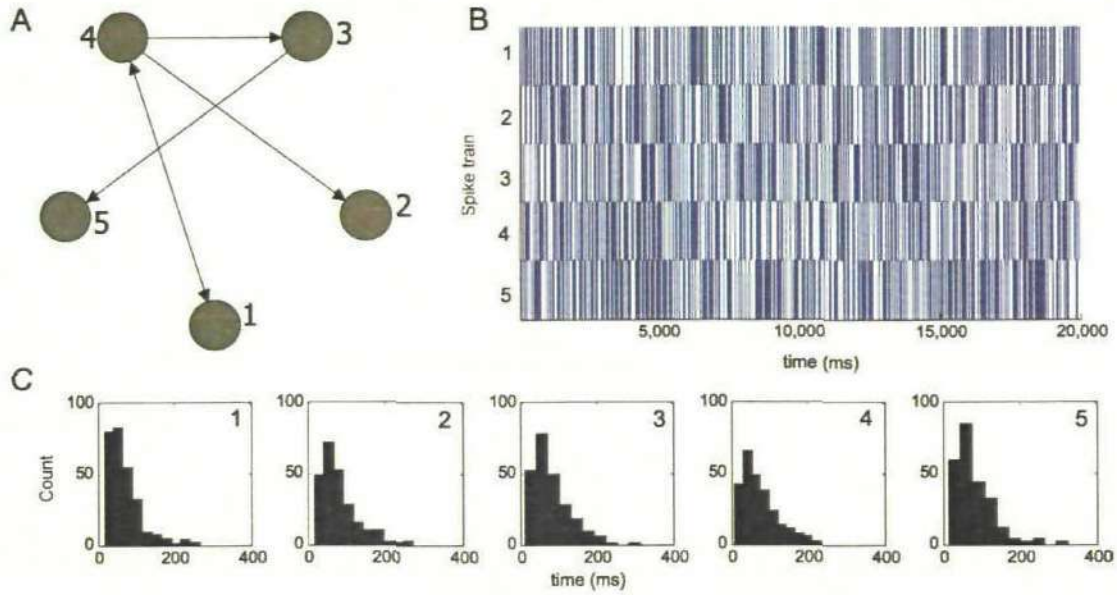


Fig. 2. (A) Connection scheme of the five spike train. There are five non zero connections which are shown by arrows. (B) Raster plot of five spike trains generated for the neural circuit (A) of the duration 20,000 ms. (C) ISI histograms of the generated five spike trains.

result of statistical analysis with parameter values which have been used for spike train generation. Fig. 2C shows the histogram of inter-spike intervals for each spike train and Table 2 provides the number of spikes, the mean ISI and the standard deviation of ISI for each spike train.

To apply the Cox method, one spike train should be selected as a target and other four are considered as the reference trains. Also, the influence functions should be specified to describe the influence from the reference spike train to the target. Here we assume that all influences are identical and the influence is specified by formula (4). This function includes three parameters (τ_s, τ_r, Δ) and their values should be defined for each reference spike train.

Characteristic times of rise and decay of the postsynaptic potential (PSP) are usually known from experimental recordings (e.g. they can be estimated from intra-cellular recordings). There are theoretical attempts to estimate characteristic decay of PSP from the histogram of inter-spike intervals (see e.g. Tuckwell and Richter, 1978; Lansky et al., 2006; Mullowney and Iyengar, 2008)[2]. The approach is based on consideration of a simple integrate and fire model with white noise for spike generation. Using some statistical technique (e.g. maximum likelihood), it is possible to estimate parameters of the model from the histogram of ISIs, including the time of decay of PSP. An approximation of recorded spike train by the integrate and fire model has a limited use due to over simplicity of this model and difficulties of numerical procedure.

In case of generated data, the values of PSP decay time are known (see column three in Table 1); however, we assume that these values are unknown and do not use them for analysis of connectivity. Instead, we estimate these values from the histogram of ISIs using technique from Tuckwell and Richter (1978) which shows that the PSP decay times are about 10 ms. Thus, we select $\tau_s = 10$ ms and $\tau_r = 0.1$ ms. It means that the PSP rises very fast and decays relatively slow. We have studied how the result of analysing the functional connectivity depends on chosen values of PSP characteristic times and found that the Cox method has low sensitivity to selected values of these parameters. Other words, there is no requirement to choose these times accurately. In fact, these characteristic times can be varied in a broad range and the results of analysing will be similar. Therefore, in this analysis we do not adjust characteristic times of postsynaptic potential for each neuron but use the same values of PSP characteristic times for ($\tau_s = 10$ ms and $\tau_r = 0.1$ ms) for all reference spike trains.

Also, the time lag Δ corresponding to the delay of spike propagation should be specified. To do this we consider a pair of spike trains: the target and the reference spike train and we analyse interdependences of two spike trains. A traditional approach is to apply the cross-correlation function which provides both a statistical estimate of dependency and a corresponding time lag.

The CCF is widely used in neuroscience to analyse dependencies and influences between pairs of simultaneously recorded spike trains. To measure the association between two spike trains A (target) and B (reference) over the time T the counting function $n_{AB}(u)$ is calculated. The function $n_{AB}(u)$ counts and accumulates the number of spikes of spike train A falling in a small interval of length h (bin) which is attached to a spike in B ($u = 0$) or shifted from a reference spike in B by u (left or right). A shift to right side means that with a high enough probability a spike in the reference trains causes a spike in the target spike train; therefore there is an influence from reference (B) to the target (A) spike train. This counting function is an estimate of the cross-product density p_{AB}. To test the independence of two spike trains Brillinger (1976c) considers the estimate $\hat{\rho}_{AB}(u) = \sqrt{\hat{p}_{AB}(u)/\hat{p}_A\hat{p}_B}$, where $\hat{p}_{AB}(u) = n_{AB}(u)/2hT$, $\hat{p}_A = n_A/T$ and $\hat{p}_B = n_B/T$ and normalises the counting function $n_{AB}(u)$ accordingly. Here n_A, n_B denote the number of spikes in A and B respectively. For a large sample size the random variables $\hat{\rho}_{AB}(u)$ are independent and the distribution of each of them is the normal with the mean $m = \sqrt{\hat{p}_{AB}(u)/\hat{p}_A\hat{p}_B}$ and the standard deviation $S = 1/(2\sqrt{2hT\hat{p}_A\hat{p}_B})$. Therefore, in case of two independent spike trains the mean of $\hat{\rho}_{AB}(u)$ equals to one (because in independent case $\hat{p}_{AB}(u) = \hat{p}_A(u)\hat{p}_B(u)$). To test the hypothesis H_0 that two spike trains are independent, the boundaries of the confidence interval with the significance level α are plotted by two horizontal lines at levels $1 \pm Q_{cr}^{\alpha}/1\sqrt{2hT\hat{p}_A\hat{p}_B}$, where Q_{cr}^{α} is a critical value of normal distribution corresponding to the significance level α. If H_0 is correct then all values of the cross-correlation function cor-

[2] All these over simplified models consider immediate jump of potential ($\tau_r = 0$) and decays with characteristic time τ_s.

Table 3
Time lags obtained from Fig. 3. These time lags are used to get the full functional connectivity of neural circuit of five spike train.

Target spike train	Reference spike train				
	1	2	3	4	5
1	0	2	2	12	0
2	0	0	9	10	0
3	21	0	0	10	0
4	10	0	0	0	41
5	0	0	6	0	0

responding to different bins fall inside the confidence interval and the estimated value of CCF ($\hat{\rho}_{AB}(u)$) is zero. If some value of CCF exceeds the upper boundary of the confidence interval then the hypothesis H_0 is rejected and we conclude that two spike trains are not independent. The highest value of CCF exceeding the upper boundary of the confidence interval can be considered as a measure of influence strength from one spike train to another and the corresponding time shift can be considered as a time delay in spike propagation (Stuart et al., 2005).

Fig. 3 shows CCFs of five spike trains. The highest peak outside the confidence interval is interpreted as an indicator of influence and the corresponding time shift of CCF is considered as a time lag corresponding to delay of spike propagation. Time lags are summarised in the Table 3 and these values are used for analysing the functional connectivity. For example, if the first spike train is selected as a target, then the first row of the Table 3 provides parameter values of time lags which are parameters of the respective influence functions: $\Delta_2 = 2$, $\Delta_3 = 2$, $\Delta_4 = 12$, $\Delta_5 = 0$

Remark: It is common that there are multiple peaks of CCF outside the confidence interval. They reflect a complex structure of interactive point processes. Here we analyse the excitatory connections only, therefore the drops of CCF below the lower boundary of significance interval are not considered.

Thus the general procedure of analysing the functional connectivity is: (1) select a target spike train and consider all others as the reference spike trains; (2) for each reference spike train specify the influence function using the values of three parameters; (3) apply the Cox method and calculate the estimates of Cox coefficients and their corresponding confidence intervals (with a prescribed confidence level).

Appendix A provides formulas (A.1) and (A.2) for calculation of estimates of Cox coefficients $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_p)$ p—is the number of reference spike trains as well as formulas (A.3) for calculation of confidence interval for each Cox coefficient.

Functional connections can be derived from these estimates and their confidence intervals. If the confidence interval contains zero, then we conclude that the corresponding Cox coefficient is not distinguishable from zero; therefore the functional connection from the reference spike train to the target is absent. If the confidence interval does not include zero, we conclude that there is a significant influence from the reference spike train to target and the value of the estimate characterises the strength of this functional connection. To define a complete diagram of functional connectivity we repeat the calculation of estimates and confidence intervals for each target spike train and define all functional connections incoming to the target. Thus, in case of n spike this procedure should be repeated n times.

Table 4 summarises the result of analysing spike trains by the Cox method. Each row of the table shows the Cox coefficients characterising the influence strength to the target spike trains. For example, the first row of Table 4 corresponds to the case that the first spike train is considered as a target and this row shows the estimates of Cox coefficients characterising influences to the target spike train (#1) from the reference spike trains (#2 to #5): $\hat{\beta}_{21} = 0.5$, $\hat{\beta}_{31} = 0.2$, $\hat{\beta}_{41} = 1.6$, $\hat{\beta}_{51} = 0.09$. Also the corresponding confidence intervals (the confidence level here is 0.95) are shown. These intervals are used to test the hypothesis that the Cox coefficient is zero: if the confidence interval includes zero then this hypothesis should be accepted and we conclude that there are no influence from reference spike train to the target (i.e. influence strength is zero). In the first row there is only one Cox coefficient that significantly differs from zero (shown in bold) which characterises the influence from spike train #4 to spike train #1. This non-zero influence strength is interpreted as strength of the functional connection from spike train #4 to spike train #1. All other Cox coefficients at the first row are not distinguishable from zero and the corresponding functional connections to the target spike train #1 are absent. This procedure of estimation of Cox coefficients is repeated for the target spike train #2 and the result is shown in row 2, etc. Thus, considering Table 4 we conclude that there are five Cox coefficients that significantly differ from zero; therefore there are five functional connections between spike trains. These functional connections are shown by circles in Fig. 4B and a radius of the circle is proportional to a relative strength of influence: a small radius corresponds to a relatively weak functional connection. Diagonal is shown by filled squares.

Comparison of the matrix of functional connectivity (Fig. 4B) with the matrix of connections (Fig. 4A)[3] which has been used for simulation of spike trains reveals a good correspondence between these two schemes of connections. Moreover, relative connection strengths have been correctly identified: circles with smaller radius correspond to weaker connections (see connection strengths in Table 1).

To emphasise how importance of this result, we note that the diagram of connectivity in Fig. 2A contains direct connections shown by arrows (e.g. from spike train #4 to spike train #3) and some "spurious" connections: connection due to a "common source"[4] and connection due to "indirect coupling". For example, there is no direct connection between spike trains #1 and #2; however, spike train #4 is a common source which delivers spikes to both spike trains (#1 and #2). Another example of spurious connection is "indirect coupling": there is no direct connection between spike trains #1 and #3, however, there is an indirect influence (coupling) from spike train #1 to spike train #3 via spike train #4.

Remarkably, the Cox method ignores all "spurious" connections and correctly finds the direct connections which have been used for data generation. Thus, it is shown that the Cox method can distinguish between "direct connection" and the connectivity due to a "common source" (or similar, to distinguish "direct "and "indirect" connections).

This problem is difficult for pair-wise methods. For example, the pair-wise CCF for spike trains #1 and #2 is shown in Fig. 3 (first row and second column). There is the significant peak of CCF corresponding to time lag of 2 ms (see Table 3) which can be erroneously interpreted as a connection from spike train #2 to spike train #1 with time lag of 2 ms. Similar, a spurious connection from #1 to #3 is correctly ignored by the Cox method but erroneously identified by the pair-wise method which shows a significant peak corresponding to 21 ms time lag (see the CCF in Fig. 3, third row and first column and the time lag value in Table 3). This advantage of the Cox method is due to a possibility to analyse all simultaneously recorded spike trains at once. Other advantages of the Cox method in comparison with CCF will be considered in the Section 4.

[3] In fact, this matrix is equivalent to the connection diagram in Fig. 2A.

[4] A common source is a neuron that simultaneously modulates the firing patterns of two or more other neurons. For example, in Fig. 2A the neuron #4 is a common source which modulates firing patterns of neurons #1 and #2.

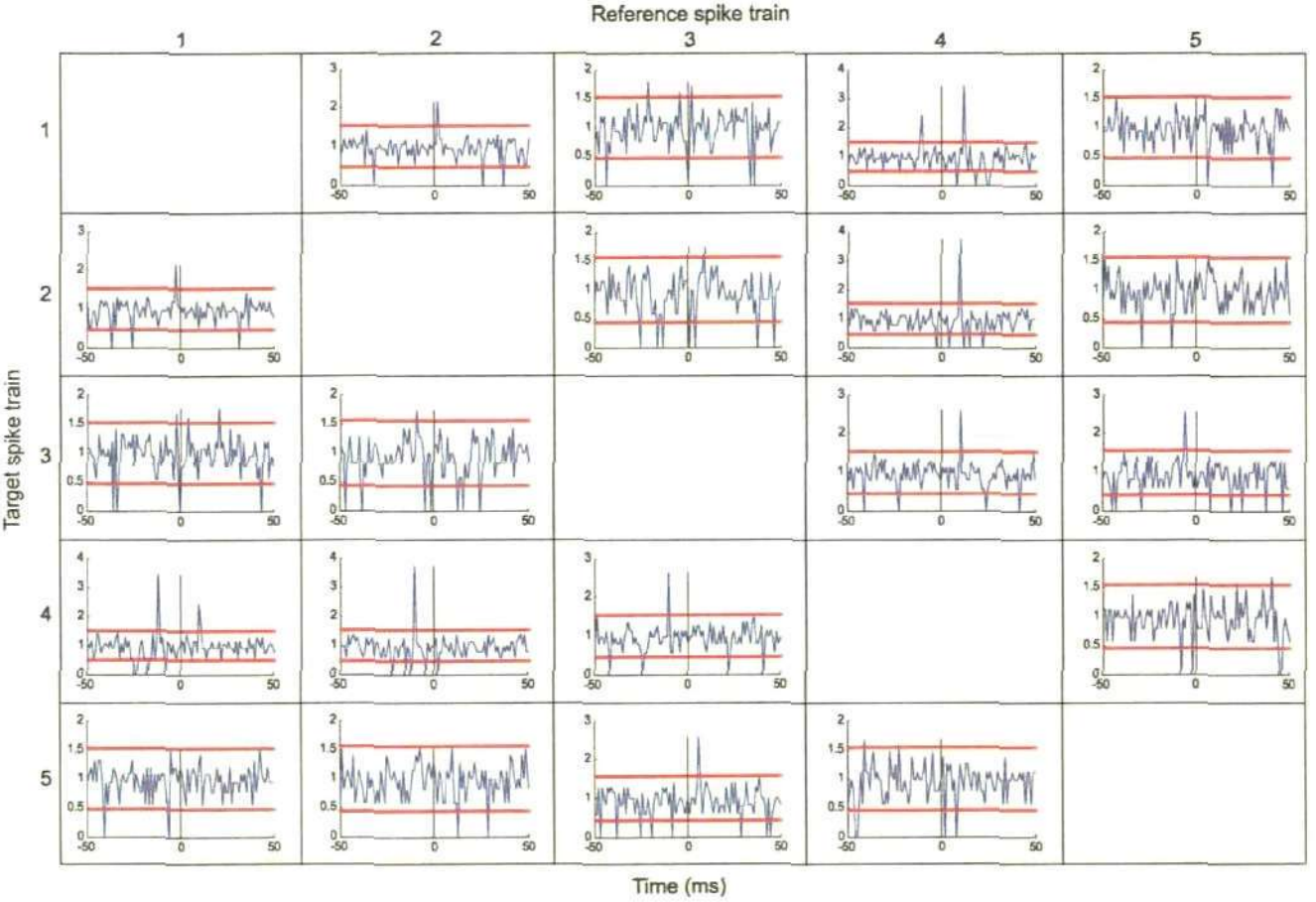


Fig. 3. Cross correlation function of the neural circuit of five spike trains.

3.3. Analysis of functional connectivity of twenty spike trains

In this section we analyse a relatively large set of twenty spike trains which are generated by the ELIF model with twenty elements and with forty two connections. A diagram of connections is shown in Fig. 5A. Parameter values of ELIF model are similar to the parameter values which have been used in the previous example of five spike trains. A simulation has been run in time interval of 50 s and Fig. 5C shows an initial part of the raster plot of twenty spike trains generated by the model (from 0 to 20 s). Fig. 5B shows an example of the four ISI histograms of spike trains #1 to #4.

The procedure for analysing the functional connectivity is the same as above. The target spike train is selected and other nineteen spike trains are considered as reference trains. The influence function is given by the formula (4) and characteristic times are the same as in case of five spike trains because spike trains have similar statistical characteristics of ISIs—we select $\tau_s = 10$ ms and $\tau_r = 0.1$ ms. To select a proper time lag for the influence function we calculate the pair-wise CCF between the reference spike train and the target spike train and identify the highest significant peak. The corresponding time shift of the CCF is used as the value of time lag Δ. If there are no significant peaks of CCF then $\Delta = 0$. Thus, the influence function is defined and using the Cox method we can calculate the estimates $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_{19}$ of Cox coefficients and their confidence intervals. Testing the hypothesis that the Cox coefficient is not distinguishable from zero, we identify "zero" Cox coefficients which are interpreted as absence of functional connection. All "non-zero" coefficients define functional connections and the value of estimate is considered as connection strength.

From statistical point of view, the repetitive application of the Cox method means that we carry out $M = 380$ independent statistical tests. In case of a large value of M, the probability of error is artificially inflated. Therefore, the significance level should be adjusted with taken into account the number of repetitions M. It is

Table 4
Result of analysis of five spike trains by the Cox method. The estimates of Cox coefficients and corresponding confidence intervals are shown. Cox coefficients which significantly differ from zero (i.e. the confidence interval does not include zero) are in bold.

Target spike train	Reference spike train				
	1	2	3	4	5
1	0	0.5 (−0.09, 1.1)	0.2 (−0.4, 0.9)	**1.6 (1.1, 2.1)**	0.09 (−0.6, 0.8)
2	−0.2 (−1.0, 0.4)	0	0.4 (−0.1, 1.1)	**1.9 (1.4, 2.4)**	0.3 (−0.3, 1.0)
3	0.5 (−0.06, 1.2)	−0.3 (−1.0, 0.4)	0	**1.1 (0.5, 1.7)**	−0.2 (−1.1, 0.5)
4	**1.2 (0.7, 1.8)**	−0.1 (−0.9, 0.6)	0.1 (−0.5, 0.8)	0	−0.2 (−1.0, 0.5)
5	−0.1 (−0.8, 0.6)	−0.2 (−1.1, 0.5)	**1.2 (0.7, 1.8)**	0.003 (−0.7, 0.7)	0

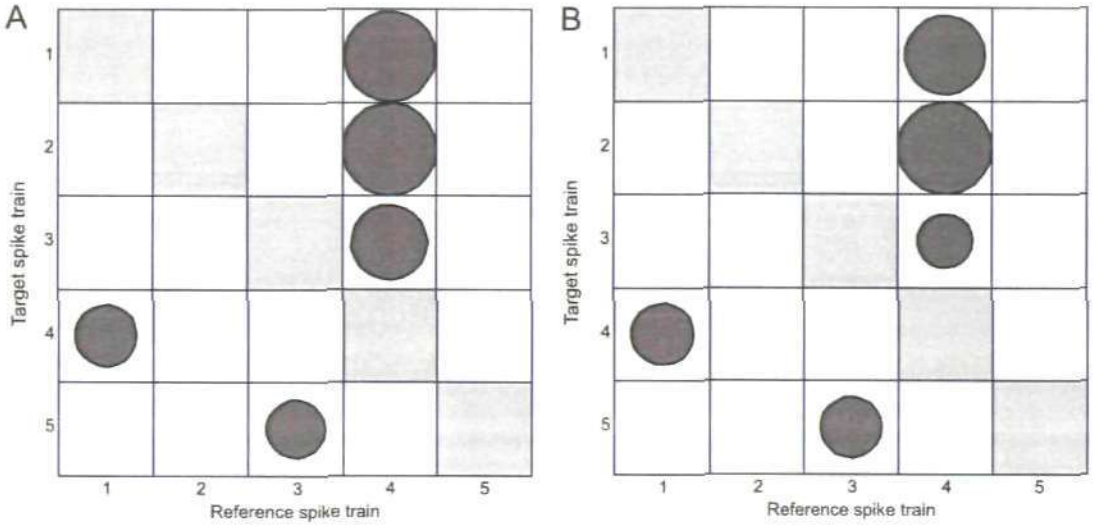


Fig. 4. (A) Connection scheme of five spike trains in matrix format (the same as the scheme shown in Fig. 2A in graph format). (B) A diagram of functional connections of five spike trains obtained by the Cox method.

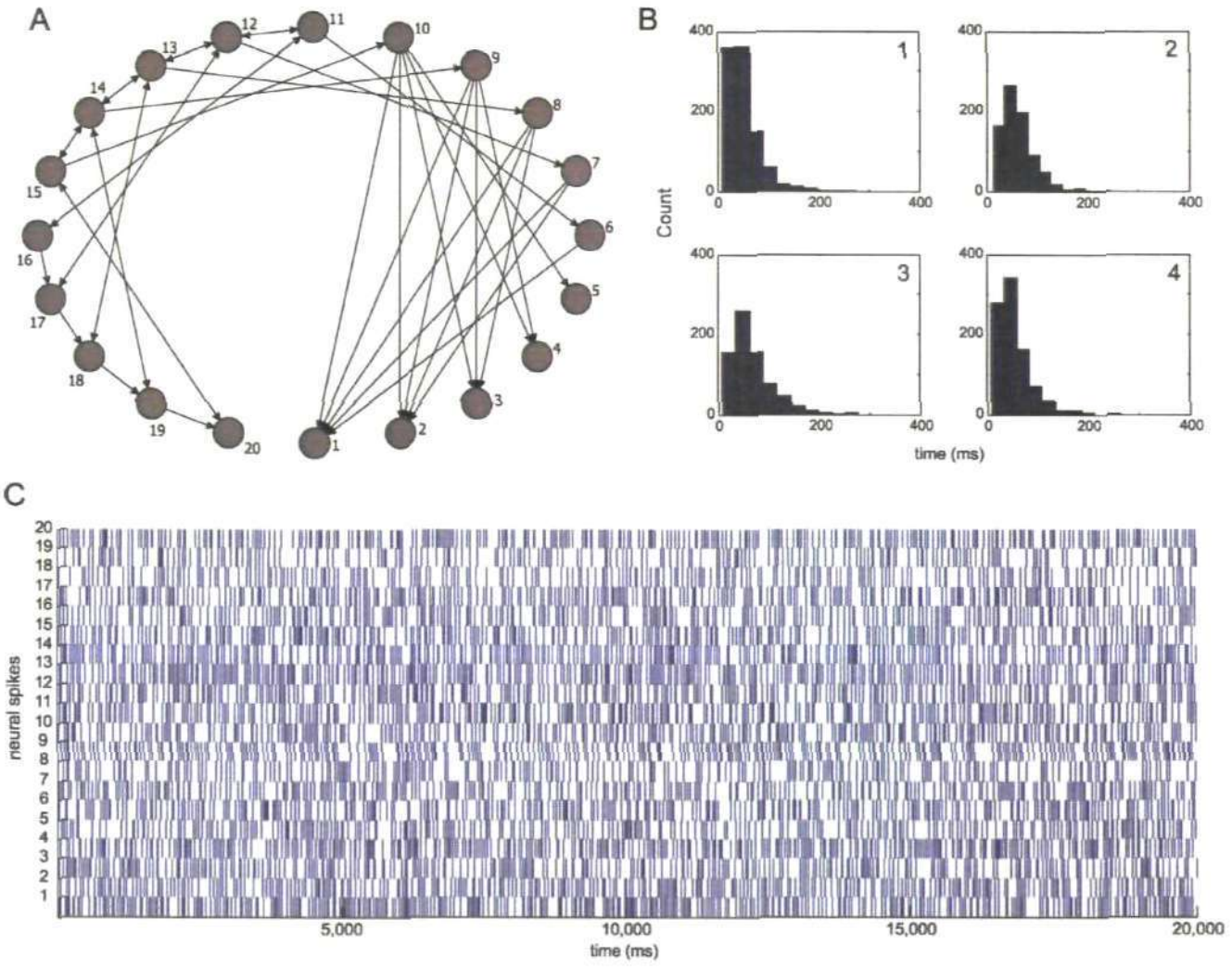


Fig. 5. (A) Connection scheme of the twenty spike trains. There are forty two non zero connections which are shown by arrows. (B) ISI histograms of the first four generated spike trains. (C) Raster plot of twenty spike trains generated for the neural circuit. This raster plot shows a portion of time (20,000 ms) of the duration 50,000 ms.

a common practice to employ a Bonferroni statistical adjustment to counteract the effect of multiple tests. Thus the significance level has been corrected: α/M. In the case of 20 spike trains we have used $\alpha \approx 0.05$ and $M = 380$ and the corrected significance level was 0.00013.

To simplify a comparison of the result of data analysis with a connection scheme used for data generation we show both connection schemes in matrix format. Fig. 6A shows connections of a neural network of twenty ELIF elements used for data generation (the same scheme is shown in Fig. 5A in a graph format). Fig. 6B shows a diagram of functional connections in matrix format which has been identified by the Cox method. A matrix of functional connectivity in Fig. 6B results from twenty repetitive applications of the Cox method. The first row of the matrix corresponds to the case when the first spike train is selected to be a target; the second row corresponds to the case when the second spike train is the target, etc. A circle indicates that there is a significant influence (functional connection) to the target spike train and the radius of the circle shows the relative strength of the influence. Comparison of connectivity matrix in Fig. 6A with matrix in Fig. 6B shows that the Cox method correctly identifies all forty two direct connections between spike trains. The connectivity matrix is derived from the repetitive testing of hypothesis that there are no dependencies between the target and reference spike trains using the Cox method. In the hypothesis testing, two null hypotheses of independence are incorrectly rejected. These two false positive connections are shown by green circles (Fig. 6B) and these erroneous connections are not present at the circuit for the spike train generation.

Fig. 6C shows the matrix of functional connectivity which has been constructed by using the pair-wise CCF technique. Comparison this matrix with the matrix of connections which have been used for data generation shows that all forty two null hypotheses of independency are correctly rejected. It means that all forty two non-zero connections have been correctly identified. However, in addition, fifteen null hypotheses of independency are incorrectly rejected. Thus, there are fifteen type I errors (false positives) and corresponding non-zero erroneous connections are shown by green circles (Fig. 6C). A radius of circles corresponding to these erroneous non-zero connections is relatively large, therefore, a strength of erroneous influence is also relatively large[5]. This comparison shows that the Cox method has some advantages over CCF technique.

4. The Cox method versus cross-correlation function

In this section we compare the Cox method with a traditional technique based on the cross-correlation function and show advantages of the Cox method especially in cases which are difficult for analyses. The CCF is a pair-wise method, therefore in this section we mainly analyse a connection of two spike trains. The main assumption of the Cox method is that the target process is the modulated renewal process with the hazard function described by formula (1). A probabilistic model has been developed to generate the MRP. It is expected that for this data the estimate $\hat{\beta}$ of the Cox coefficient equals to the influence strength β in formula (1). Of course, in a general case of data generation using the ELIF model, we do not expect that the target spike train is MRP. However, we demonstrate that the Cox method can be successfully applied to analyse functional connectivity and the estimate $\hat{\beta}$ monotonically increases with increase of connection strength in generated data. Also, in this section we study connectivity of three spike trains generated by ELIF model with "common source" connections. This connection scheme is very difficult for analysing by pair-wise methods and in particular by CCF. We show that the Cox method which can analyse three spike trains at once can be successfully applied to identify a functional connectivity. In a similar way we study another set of three spike trains with "indirect" connections.

4.1. Description of the probabilistic model

The probabilistic model generates two spike trains A and B. Spike train B is a renewal process with the gamma-distribution $\gamma(x; a, b)$ of interspike intervals, where parameters a and b are the shape and the scale parameters respectively. Spike train B influences spike train A and spikes of B modulate the probability of spike generation in the process A which is the modulated renewal process with the hazard defined by formula (1). The hazard function of A without the influence from B depends on the backward recurrence time of the process A. This assumption corresponds to a standard consideration of neuronal spiking (Daley and Vere-Jones, 2003; Truccolo et al., 2005). For example, after firing a spike, the process is less likely to fire again immediately afterward. Also, we suppose that the backward recurrence times have the Weibull distribution $W(x; c, d)$ (see Cox, 1972), where parameters c and d are the shape and the scale respectively. In fact, the types of the distribution of ISI of process B and the distribution of backward recurrence times of A can vary. Our choice of the gamma distribution and the Weibull distribution is motivated by the fact that both families include the exponential distribution which can be seen in many neuroscience data.

To generate spike trains, the influence strength β should be selected. The influence function describing an impact of the spike train B to the spike train A is given by formula (5) with the characteristic time τ_s and the time lag Δ. Thus, the hazard of the modulated renewal process A is completely defined and we use this function to generate spikes of A using a small time step and calculating the probability of spike at discrete times t_k.

The probabilistic model includes five parameters and we would like to adjust them in such a way that the ISI distribution of the MRP is similar to the ISI distribution of the spike train which is generated by the ELIF model. In Appendix C we briefly describe the cost function which enables us to find an optimal set of five parameters of the probabilistic model.

4.2. Analysis of two spike trains

In this section we consider two spike trains (A and B) which are generated by the probabilistic model. The probabilistic model generates a pair of spike trains A (target) and B (reference) with only one connection from B to A and the connection strength is β_{BA}. We use both the Cox method and CCF to estimate the connection strength.

The influence function of the Cox method is given by formula (5) with characteristic times $\tau_s = \tau_r = 5$ ms and zero time lag ($\Delta = 0$). The estimate $\hat{\beta}_{BA}$ of the Cox coefficient is calculated as well as the confidence interval. Also, the CCF has been calculated and the value of highest peak outside of the confidence interval $\hat{\rho}_{BA}$ is considered as an estimate of the connection strength. Of course, if there are no peaks outside of the upper bound of the confidence interval, the connection strength is considered to be zero and $\hat{\rho}_{BA} = 0$. We demonstrate some advantages of the Cox method both in case of short spike trains and in case of a weak coupling.

Moderate and high strength of influence: Varying the strength of influence β_{BA} in a range from moderate to high, we generate eight pairs of spike trains and for each pair we analyse connectivity. The average number of spikes in the reference spike train B is about 400 and target spike train A has a larger number of spikes. The blue line in Fig. 7A shows the estimated Cox coefficient $\hat{\beta}_{BA}$ (with corresponding confidence interval which is shown by black vertical bar)

[5] The Bonferroni correction is applied to test the hypothesis of independency for CCF method.

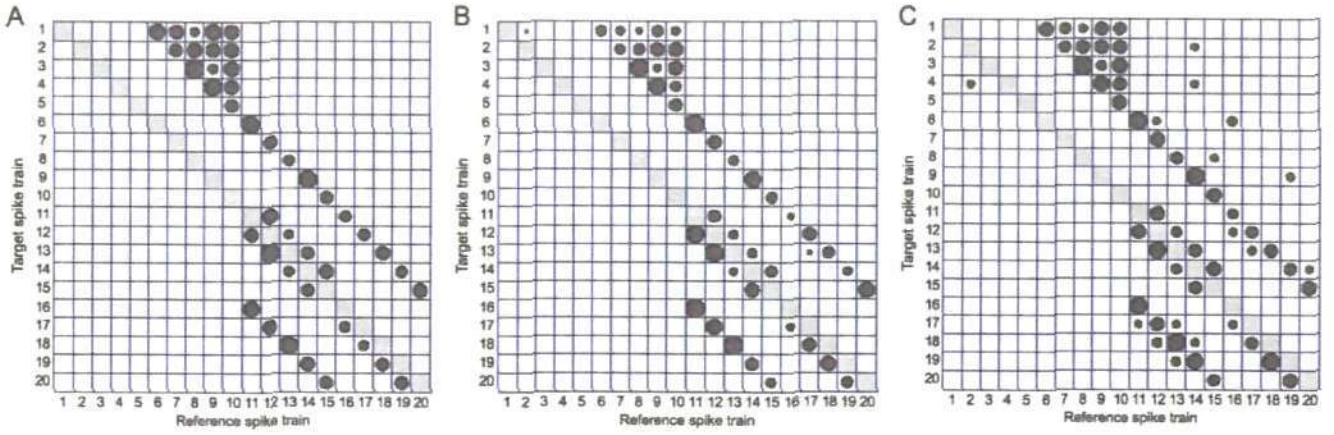


Fig. 6. (A) Connection scheme of the neural circuit of twenty spike trains in matrix format (the same as a scheme of connections in a graph format in Fig. 5A). (B) Functional connections identified by the Cox method. (C) Functional connections obtained by the CCF method.

versus the strength of influence β_{BA}. This plot shows that the estimated values are close to the values which have been used for data generation. The magenta line shows the independency measure of CCF $\hat{\rho}_{BA}$ versus the strength of influence β_{BA}. It is clear from Fig. 7A that the CCF also allows to identify this connection; however, the plot of CCF is not monotonically changing and fails to indicate the increase of influence.

Short spike train: To test sensitivity of the Cox method in a case of short spike trains we use the same pairs of spike trains as in the previous example but for the analysis we cut the epoch time and consider only a part where the independent spike train has about 70 spikes, thus a time epoch is about six times shorter than in the previous example. Fig. 7B shows the estimate of Cox coefficient $\hat{\beta}_{BA}$ versus the strength of influence β_{BA} for the moderate and high strength of influence (shown by the blue line with the black vertical bars of confidence intervals). This line shows that the estimated values are similar to the strengths which have been used for data generation with the only one exception: for $\beta_{BA} = 0.5$ the confidence interval contains zero and therefore, the Cox coefficient is not distinguishable from zero.

The magenta line in Fig. 7B shows that the CCF fails to identify functional connection for the moderate influences $\beta_{BA} = (0.5, 1)$. For the higher influences $\beta_{BA} = (1.5, 2, 2.5, 3, 3.5, 4)$ the CCF measure is nearly constant and fails to indicate the increase of influence. Therefore, the Cox method has some advantage in a case of short spike trains.

Weak influence: To test an efficiency of methods and to identify weak connection strength we use the same probabilistic model to generate another eight pairs of spike trains with weak influence: $\beta_{BA} = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8)$. In this case the time epoch should be long enough (about 1400 spikes in the reference spike train B) to allow a distinguishing of weak influence. Fig. 7C shows that both methods demonstrate a good result and identify the connection. The Cox coefficient increases with the connection strength increase, but the CCF measure is not monotonically increasing.

Sensitivity to the length of spike train: Here we study how a sensitivity of methods depends on a length of spike trains under a constant value of the influence strength. The conclusion is that for shorter trains, the Cox method identifies the connection but the CCF fails. The strength of influence $\beta_{BA} = 1.0$ is relatively small. We have fixed the value of influence and generated eight pairs of spike trains with the different number of spikes in the reference spike train B: n=50,60, ..., 120. Fig. 7D shows that the estimated Cox coefficient is almost constant ($\hat{\beta}_{BA} = 1$) and does not depend on the length of spike train. The CCF measure ($\hat{\rho}_{BA}$) shows the connection for the larger spike trains ($n = 90, 100, 110, 120$) but fails to identify a strength of influence for the shorter lengths of reference spike trains ($n = 50, 60, 70, 80$).

4.3. Analysis of three spike trains: common source and indirect connection

Here we show that the Cox method is very effective to analyse connections which are not direct such as "common source" circuit (Fig. 8A) and "indirect connection" circuit (Fig. 10A). Usually it is very difficult to analyse these types of connection using pair-wise CCF technique. The Cox method is multivariate and can analyse three spike trains at ones and it makes this method more sensitive than pair-wise CCF. For example, this advantage enables the Cox method to distinguish between "direct" connection and connection due to a "common source" in case of a moderate influence form the common source.

Common source: Three spike trains ({#1, #2, #3}) are generated using ELIF with the same parameters like in previous examples and a "common source" connections (Fig. 8A). The "common source" circuit includes two connections from spike train #1 to spike trains #2 and #3. Connection strengths are 12.6 and 10.6; delays of spike propagation are 11 ms 14 ms, respectively.

We analyse these three spike trains by the Cox method with the influence function given by formula (4), characteristic times are $\tau_s = 10$ ms and $\tau_r = 0.1$ ms. To prescribe the time lags we calculate the CCF function for all pairs of spike trains (Fig. 9): $\Delta_{12} = 11$ ms, $\Delta_{13} = 14$ ms, $\Delta_{23} = 3$ ms, and all other time lags are zero. We apply the Cox method three times (each of spike trains is selected to be a target spike trains).

Suppose that the target spike train is #k, ($k = 1,2,3$). The estimates $(\tilde{\beta}_{ik}, \tilde{\beta}_{jk})(i = 1, 2, 3; k = 1, 2, 3, i \neq j, i \neq k, j \neq k)$ of two Cox coefficients have been calculated using formulas (A.1) and (A.2) as well as a confidence region on plane (β_{ik}, β_{jk}) using formulas (A.4). The confidence region ($\alpha = 0.05$) has an elliptic shape and the centre of the confidence region is located at point $(\hat{\beta}_{ik}, \hat{\beta}_{jk})$.

We test the hypothesis that a pair of Cox coefficients $(\hat{\beta}_{ik}, \hat{\beta}_{jk})$ equals to zero (i.e. both component of the pair are zero). We accept the hypothesis (i.e. the data does not contradict to the hypothesis) if the origin is inside of the confidence region and conclude that both connections are absent (i.e. connection strength is zero). If the hypothesis is rejected then we test the hypothesis that one Cox coefficient equals to zero. This hypothesis is tested separately for each coefficient. We consider two projections of the elliptical confidence region to the coordinate axis: (β_{ik}) and (β_{jk}). If projection to the axis (β_{ik}) contains zero then the hypothesis is accepted and we

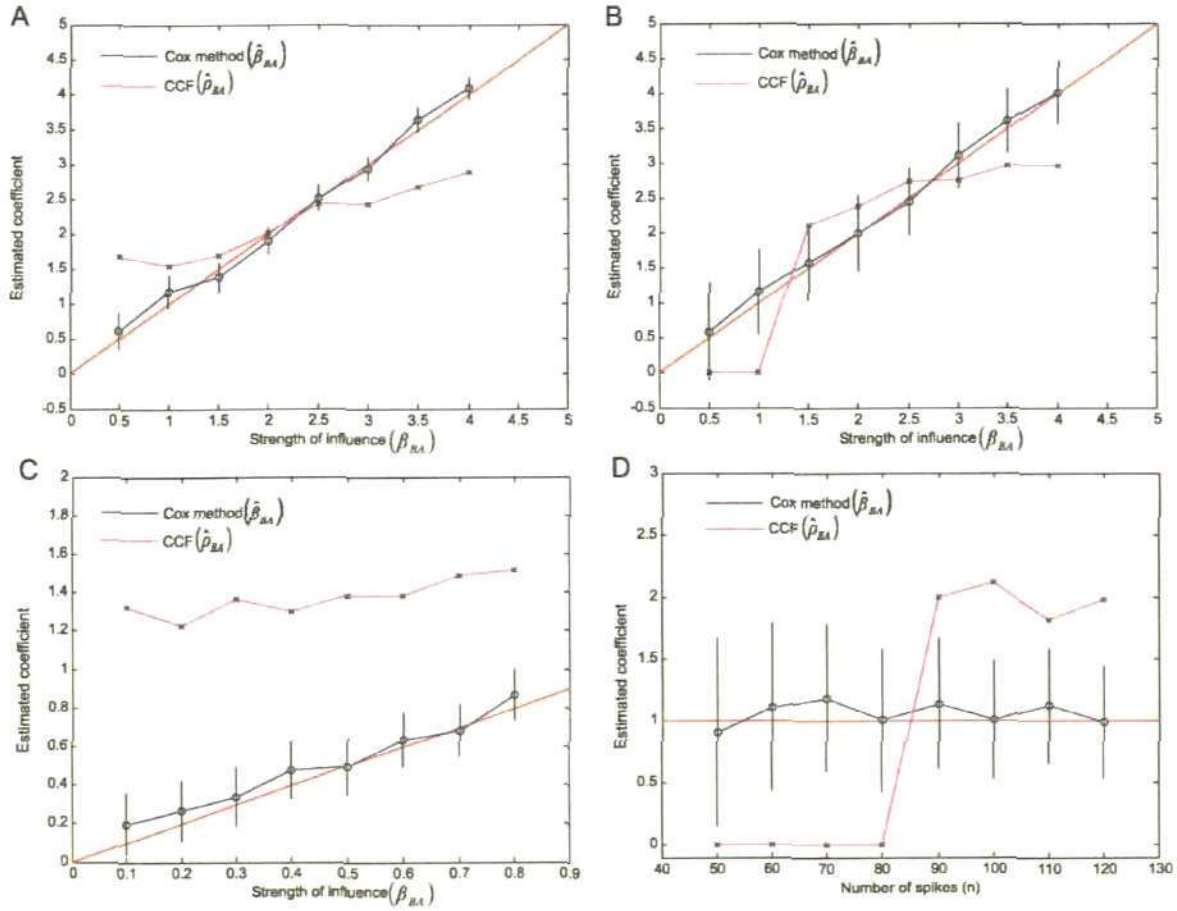


Fig. 7. Estimate of Cox coefficient and CCF measure for two spike trains. Estimated Cox coefficients are shown by black circles and the confidence interval of the estimates are shown by black vertical lines. Estimated measures of independency using CCF are shown by black cross sign. (A) Moderate and strong influence. Eight pairs of spike trains are generated using the probabilistic model taking the strength of influence in range from 0.5 to 4: $\beta_{BA} = \{0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4\}$. The average number of spikes in the reference spike train B is about 400. Estimated Cox coefficients $\hat{\beta}_{BA}$ identify accurately all these strength of influences (blue line with circle markers and vertical black bars for confidence intervals) and show monotonic increasing. The highest peaks $\hat{\rho}_{BA}$ of the CCF (independency measure) are shown by the magenta line (with cross markers) and they also can identify functional connectivity but do not show monotonic increase. (B) Short spike train. A short version of eight pairs of spike trains described in (A) are considered. The average number of spikes in the reference spike train B is about 70. The estimated Cox coefficients $\hat{\beta}_{BA}$ identify accurately all these strengths of influence β_{BA} except of one ($\beta_{BA} = 0.5$) and show monotonic increase. The independency measure of CCF ($\hat{\rho}_{BA}$) show connection for large strength but they fail to identify connection for $\beta_{BA} = \{0.5, 1\}$. Also these values do not show monotonic increase. (C) Weak influence. Eight pairs of spike trains are generated taking weak influence $\beta_{BA} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$. The number of spikes in the reference spike train B is about 1400. Estimated Cox coefficients ($\hat{\beta}_{BA}$) identify accurately all these strength of influence (β_{BA}) and show monotonic increasing. Independency measures of CCF ($\hat{\rho}_{BA}$) identify functional connectivity though they do not indicate an increase of influence. (D) Length of spike train. Eight pair of spike trains of a different length are generated keeping the same connection strength $\beta_{BA} = 1$. The length n of the reference spike train B increases: $n = 50, 60, \ldots, 120$. Estimated Cox coefficients ($\hat{\beta}_{BA}$) are almost constant for all lengths but independency measures of CCF ($\hat{\rho}_{BA}$) fail to identify strength of influence for shorter lengths of reference spike trains ($n = 50, 60, 70, 80$).

conclude that the connection is absent (i.e. the connection strength is zero), otherwise the hypothesis is rejected and a centre of the interval ($\hat{\beta}_{ik}$) is considered as strength of connection. Similar, if a projection to another axis (β_{jk}) contains zero then the hypothesis is accepted and we conclude that the connection is absent, otherwise the hypothesis is rejected and a centre of the interval ($\hat{\beta}_{jk}$) is considered as strength of connection.

Fig. 8B shows three confidence regions. A region on the left side corresponds to the target spike train #1; region in the middle corresponds to target spike train #2, region in right side corresponds to target spike train #3. It is shown in Fig. 8B that the region on the left side contains zero, therefore both connections to spike train #1 are absent. This result is shown in Fig. 8C by two dashed arrows pointing to #1. These dashed arrows mean an absence of both connections.

The region in the middle does not contain the origin and it means that the hypothesis that both Cox coefficients are zero should be rejected. The centre of the confidence region is shown by the cross and its coordinates are the estimates ($\hat{\beta}_{12}$, $\hat{\beta}_{32}$). The projection to the vertical axis β_{32} contains zero, therefore, we conclude that hypothsis is accepted and connection from #3 to #2 is absent. The projection to the horizontal axis β_{12} does not contain zero, therefore, we conclude that the hypothesis is rejected and the estimate $\hat{\beta}_{12}$ is strength of connection from #1 to #2. This result is shown in Fig. 8C by two arrows pointing to the #2: the dashed arrow means absence of connection and the solid lines means presence of connection from #1 to #2 and the value of connection strength is 2.6.

The region in right side can be interpreted in a similar way. The result is shown in Fig. 8C by two arrows pointing to #3: the dashed arrow means the absence of connection #2 to #3 and the solid arrow means the presence of connection from #1 to #3 and the value of connection strength is 1.7

The result of analysing is in a good agreement with the architecture of connections which has been used to generate these data

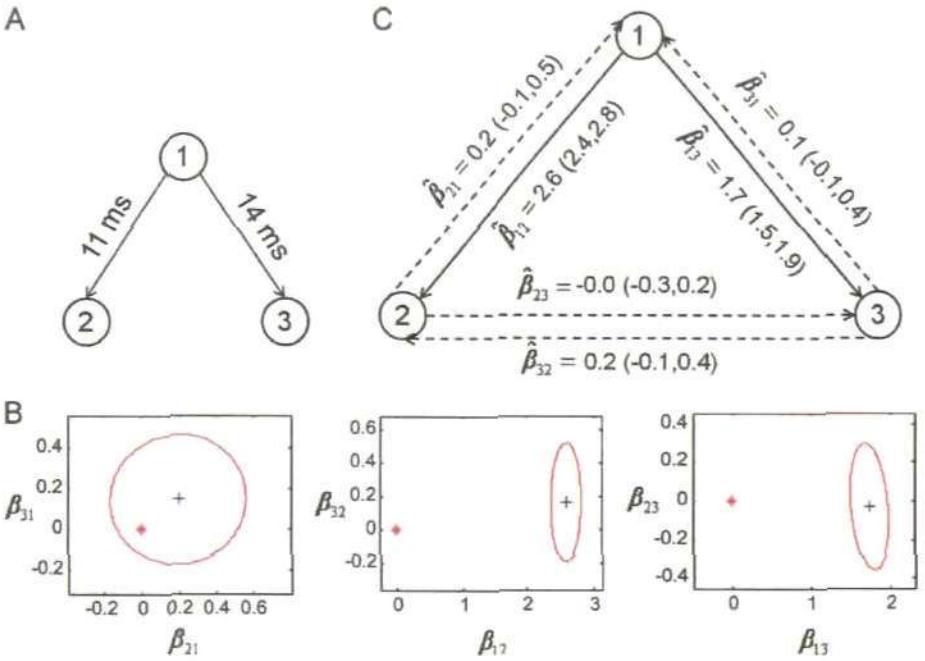


Fig. 8. (A) Connection scheme of three spike trains which have common source. Spike train #1 influences both spike train #2 and spike train #3 with time delays 11 ms and 14 ms, respectively. (B) Confidence regions of the estimated Cox coefficients in three cases: influences to spike train #1 (left), influences to spike train #2 (middle), influences to spike train #3 (right). (C) Estimated coefficients of Cox method with confidence intervals. Significant connections are indicated by solid lines.

(compare Fig. 8A with Fig. 8C). For example, for data generation the higher connection strength has been selected for connection from #1 to #2 and the estimated connection strength from #1 to #2 is also higher than estimated connection strength from #1 to #3.

Thus, the result in Fig. 8C indicates that there are two significant influences only (shown by solid lines, all others are shown by the dashed lines): from spike train #1 to spike train #2 and from spike train #1 to spike train #3 and the

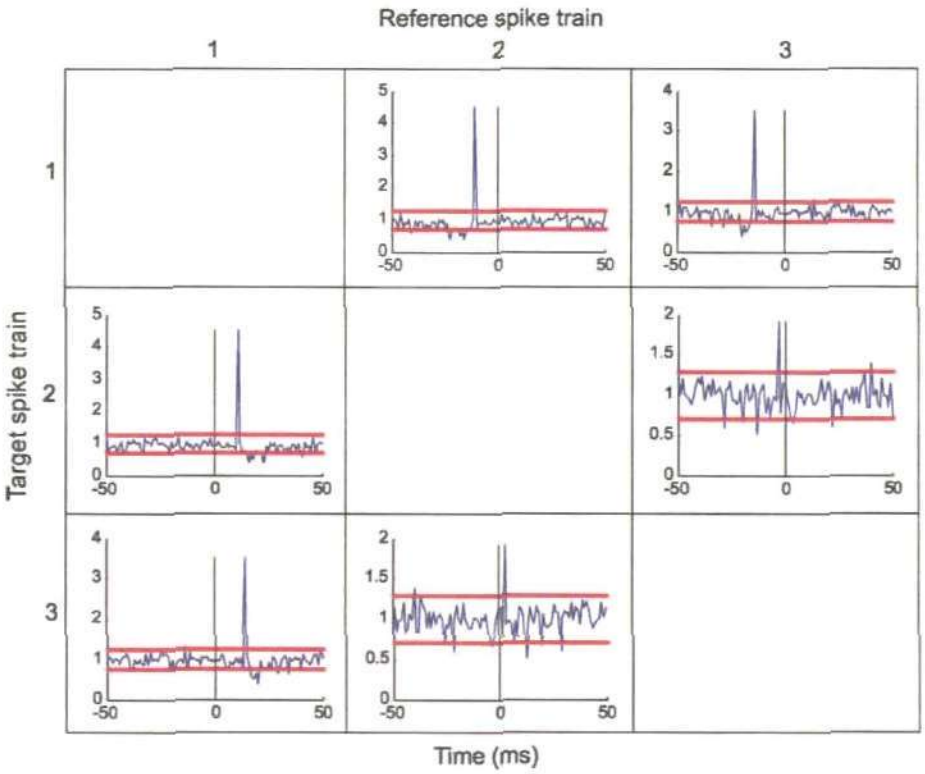


Fig. 9. Pair-wise cross correlation functions of three spike trains. Each CCF is shown for selected pair of spike trains (called target and reference). Diagram of connections (common source) is shown in Fig. 8A.

influence strengths are shown with their confidence intervals ($\alpha = 0.05$).

Fig. 9 shows the result of analysing of the same three spike trains by the pair-wise CCF. Each row of the figure shows two CCF corresponding to the selected target spike train – spike train #1 is the target for the first row, spike train #2 is the target for the second row, etc. The CCF analysis reveals three connections: from spike train #1 to spike train #2 (second row, first column); from spike train #1 to spike train #3 (third row, first column); from spike train #2 to spike train #3 (third row, second column). First two of these connections correspond to the diagram of connectivity (Fig. 8A) but the third one is erroneous and this connection appears due the common source to spike trains #2 and #3. Thus, the Cox method is able to distinguish the common source from the direct connections but the CCF fails.

Indirect connection: Similar to the previous example, we generate a set of three spike trains ($\{\#1, \#2, \#3\}$) using ELIF model with the same parameters as in the previous examples and indirect connections (Fig. 10A). The "indirect connection" circuit includes two direct influences: from spike train #1 to spike train #2 with the time lag 11 ms and from spike train #2 to spike train #3 with a time lag 12 ms. The connection strengths are 11.2 and 9.1, respectively.

To analyse functional connectivity by the Cox method with the influence function given by formula (4), we specify the characteristic times are $\tau_s = 10$ ms and $\tau_r = 0.1$ ms. To specify a time lag we calculate the CCF function for all pairs of spike trains (Fig. 11): $\Delta_{12} = 11$ ms, $\Delta_{23} = 12$ ms, $\Delta_{13} = 23$ ms, and all other lags are zero. We apply the Cox method three times (each of spike trains is selected to be a target spike trains).

Suppose that the target spike train is #k, ($k = 1,2,3$). The estimates $(\hat{\beta}_{ik}, \hat{\beta}_{jk})(i = 1, 2, 3; j = 1, 2, 3; k = 1, 2, 3, i \neq j, i \neq k, j \neq k)$ of two Cox coefficients have been calculated using formulas (A.1) and (A.2) as well as a confidence region on plane (β_{ik}, β_{jk}) using formulas (A.4). The confidence region ($\alpha = 0.05$) has an elliptic shape and the centre of the confidence region is located at point $(\hat{\beta}_{ik}, \hat{\beta}_{jk})$.

Fig. 10B shows three confidence regions. The region on the left side corresponds to the target spike train #1; region in the middle – target spike train #2, region in right side – target spike train #3. The region on left side contains zero, therefore both connections to spike train #1 are absent. This result is shown in Fig. 10C: two dashed arrows pointing to #1 mean the absence of both connections.

The region in the middle does not contain the origin and it means that the hypothesis that both Cox coefficients are zero should be rejected. The centre of the confidence region is shown by the cross and its coordinates are the estimates $(\hat{\beta}_{12}, \hat{\beta}_{32})$. The projection to the vertical axis β_{32} contains zero, therefore, we conclude that hypothsis is accepted and connection from #3 to #2 is absent. The projection to the horizontal axis β_{12} does not contain zero, therefore, we conclude that the hypothesis is rejected and the estimate $\hat{\beta}_{12}$ is strength of connection from #1 to #2. This result is shown in Fig. 10C by two arrows pointing to the #2: the dashed arrow means absence of connection and the solid lines means presence of connection from #1 to #2 and the value of connection strength is 2.3.

The region in right side can be interpreted in a similar way. The result is shown in Fig. 10C by two arrows pointing to #3: the dashed arrow means the absence of connection #1 to #3 and the solid arrow means the presence of connection from #2 to #3 and the value of connection strength is 1.5

A result of analysis is in a good agreement with the architecture of connections which has been used to generate these data (compare Fig. 10A with C). For example, for data generation the higher connection strength has been selected for connection from #1 to #2 and the estimated connection strength from #1 to #2 is also higher than estimated connection strength from #2 to #3.

Thus, the result in Fig. 10C indicates that there are two significant influences only (shown by solid lines, all others are shown by the dashed lines): from spike train #1 to spike train #2 and from spike train #2 to spike train #3 and the influence strengths are shown with their confidence intervals ($\alpha = 0.05$).

Fig. 11 shows the result of analysing of the same three spike trains by the pair-wise CCF. Each row of the figure shows two CCF corresponding to the selected target spike train – spike train #1 is the target for the first row, spike train #2 is the target for the second row, etc. The CCF analysis reveals three connections: from spike train #1 to spike train #2 (second row, first column); from spike train #2 to spike train #3 (third row, second column); from spike train #1 to spike train #3 (third row, first column). First two of these connections correspond to the diagram of connectivity (Fig. 10A) but the third one is spurious (from #1 to #3) and this connection appears due the "indirect" connection from trains #1 to train #3. Thus, the Cox method is able to distinguish the "indirect" connection from the direct connections but the CCF fails.

4.4. How to find a time lag of spike propagation by pair-wise Cox method

In this section we show that a time lag of spike propagation can be found by analysing a pair of spike trains with the Cox method. We suggest that there is a pair of simultaneous spike trains and we would like to study connection from spike train B (reference) to another spike train A (target). We generate a pair of spike trains by ELIF model with the same parameters as above and connection from B to A, the connection strength is 18.04, time delay of spike propagation is 11 ms. Of course, we suppose that all parameters which have been used for data generation are unknown and their values cannot be used when we analyse connections.

Let us suppose that two spikes of the reference spike train B appear at times t_B^1 and t_B^2. We assume that there is a time delay δ_0 of spike propagation from the reference spike train to the target spike train. It means that if there is a spike in train B at time moment t_B^2 then the probability of spike at train A at time moment $t_A = t_B^2 + \delta_0$ is very high. The influence function $Z_B(t)$ is described by formula (4) requires prescribing of a time lag Δ corresponding to the delay of spike propagation from the reference to target spike train. The time delay δ_0 is unknown, therefore we will repeat the calculation of the Cox coefficient for different values of time lag Δ. Fig. 12A shows estimates $\hat{\beta}_{BA}$ versus time lag Δ. The increment of time lag is 1 ms and the corresponding confidence intervals are shown by vertical bars. The estimate $\hat{\beta}_{BA}$ increases with increase of Δ and reaches its highest value at $\Delta = 10$ ms but for $\Delta = 11$ ms this coefficient drops down to a negative value (Fig. 12A). We calculate the Cox coefficient for values of the time lag in the interval [10.95, 11.03] (ms) with an increment of 0.01 ms. The result is shown in Fig. 12B. The Cox coefficient drops down from a high positive value to a negative value in a small interval [10.99, 11] (ms).

We conclude that a time delay of spike propagation is considered to be 11 ms ($\delta_0 = 11$) and an estimate of the Cox coefficient $\hat{\beta}_{BA} = 2.2$. To justify this interpretation of the data analysis, let us assume that a chosen time lag is smaller than the time delay of spike propagation: $\Delta_S < \delta_0$. According to formula (4), the backward recurrence time is calculated at the moment $(t_A - \Delta_S)$ and this backward recurrence time is smaller than the time delay of spike propagation: $U_B(t_A - \Delta_S) < \delta_0$. Therefore, the value of the influence function depends on the backward recurrence time $Z_B(U_B(t_A - \Delta_S))$, which is shown by the circle in Fig. 13. This value is less than the maximum of the influence function and if the time lag Δ_S increases, then the influence function also increases and tends to the maximum of the influence function if the time lag tends to δ_0. A described calculation of the backward recurrence time can be applied in a small vicinity of each spike of the train B under the main assumption that the time

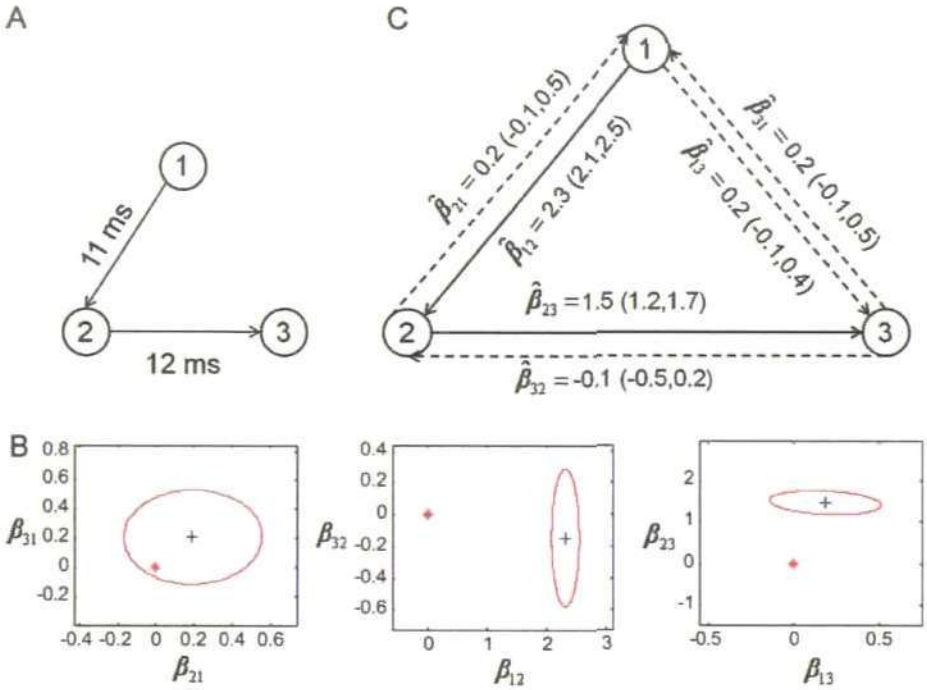


Fig. 10. (A) "Indirect connection" scheme of three spike trains. Spike train #1 influences spike train #2 which influences spike train #3 with time delays 11 ms and 12 ms, respectively. (B) Confidence regions of the estimated Cox coefficients in three cases: influences to spike train #1 (left), influences to spike train #2 (middle), influences to spike train #3 (right). (C) Estimated coefficients of Cox method with confidence intervals. Significant connections are indicated by solid lines.

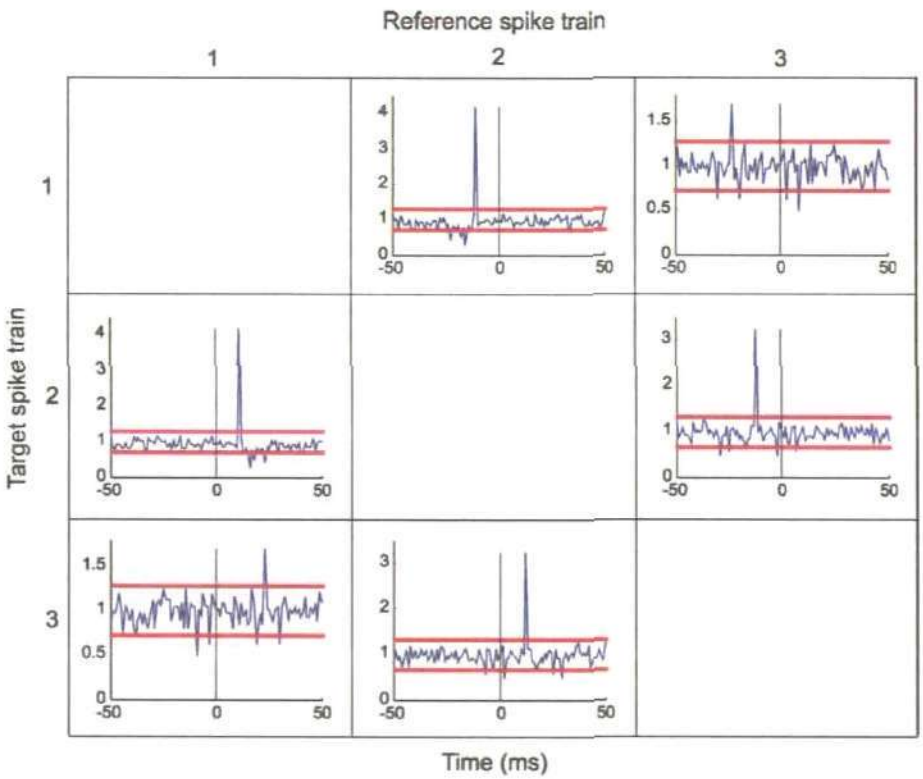


Fig. 11. Pair-wise cross correlation functions of three spike trains. Each CCF is shown for selected pair of spike trains (called target and reference). Diagram of connections (indirect connection) is shown in Fig. 10A.

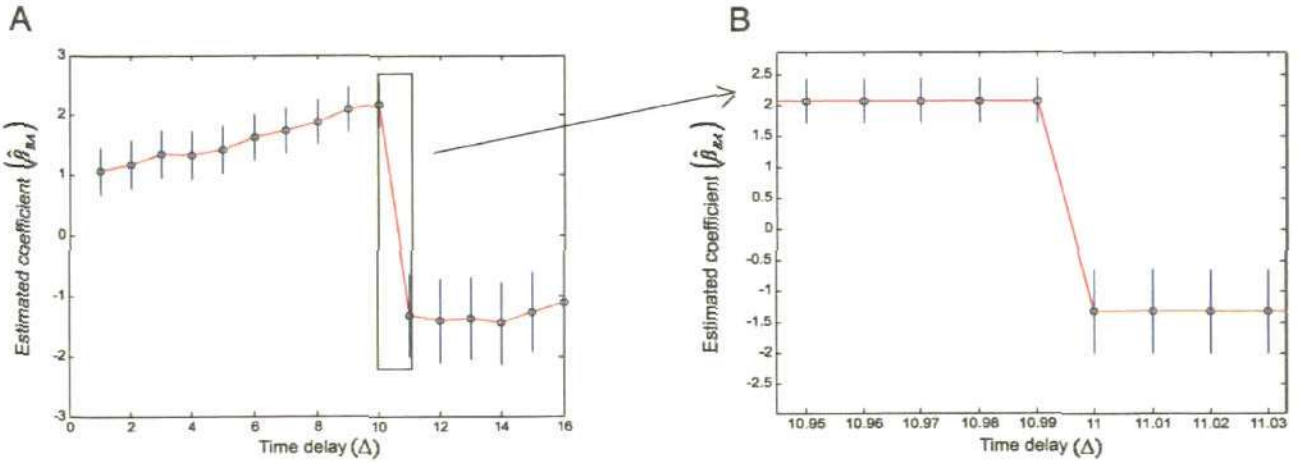


Fig. 12. Cox coefficient for the analysis of two spike trains. (A) Cox coefficient $\hat{\beta}_{BA}$ with different time delays (B) Cox coefficient $\hat{\beta}_{BA}$ in the interval 10 ms to 11 ms with a small step of time delays.

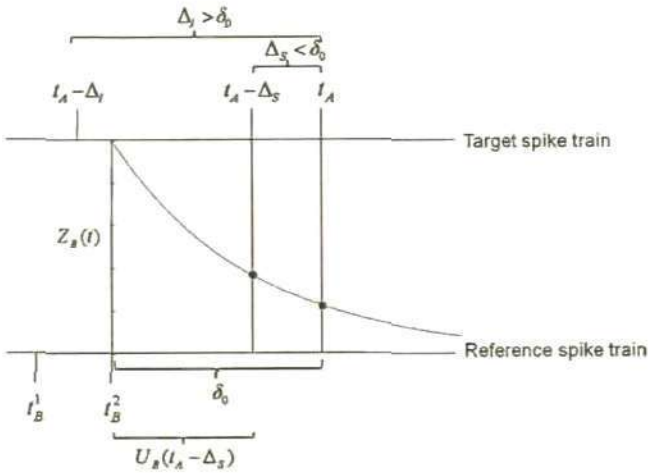


Fig. 13. Illustration for the calculation of influence function.

delay of spike propagation is δ_0. This consistency in calculation of the backward recurrence time is important for a reliable numerical procedure for calculation of estimate of the Cox coefficient.

Now we assume that a time lag $^{\Delta}I$ of the influence function is larger than the time delay of spike propagation $^{\Delta}I > \delta_0$ (see Fig. 13) and the calculation of the backward recurrence time will be based on the spike at time t_B^1, which is a previous spike according to the spike at moment t_B^2 (see Fig. 13). Thus, for the time lag $^{\Delta}I > \delta_0$ the backward recurrence time can get an arbitrary value. The backward recurrence time is calculated for each spike of train B, therefore, the estimate of the Cox coefficient is calculated using arbitrary (random) values of the influence function. Thus, the estimate will be very different from the correct value and it might be zero or a negative number.

5. Discussion

We have presented here a statistical technique (Cox, 1972) for analysing dependencies of point processes for application to neuroscience data. We find that the Cox method is an efficient tool to study functional connectivity. Comparison with the cross-correlation technique which is traditionally used in neuroscience shows significant advantages of the Cox method.

The Cox method is based on mathematical ideas of the modulated renewal process developed by Cox (1972). This approach provides a useful mathematical tool to study dependencies and mutual influences of point processes. The main advantage is that the Cox method can analyse all available simultaneously recorded spike trains at once. Another important advantage of the Cox method is that this method is binless. Also, the estimates of the Cox coefficients indicate a relative strength of influence from the reference to the target spike train. The Cox method enables us to select the influence function which takes into account the specificity of neural connections. We have found that the best candidate for the influence function reflects the dynamics of postsynaptic potential.

Recent progress in development of statistical methods for analysing multiple spike trains includes both techniques based on pair-wise spike coincidence analysis (e.g. Pipa et al., 2003; Stuart et al., 2005) and approaches which consider all (or several) spike trains at once (e.g. Staude et al., 2010a). For example, a procedure of using some appropriate surrogate data looks like a promising improvement of pair-wise correlation based method (see, Louis et al., 2010a,b). This method of surrogate data has been successfully applied to analyse data of multi-electrode recordings from the visual cortex of a cat (Berger et al., 2007, 2010) and, in particular, to decide on participation of a neuron in synchronous population activity.

There are several recent techniques which can analyse multiple spike trains at once. For example, the method based on the estimation of higher order correlations has been suggested in papers Martignon et al. (1995, 2000). This technique is aimed to estimate a huge amount of parameters and, therefore, requires very long recordings and can be applied to a relatively small number of spike trains (about ten spike trains). This approach has been further developed in Staude et al. (2010a) where a cumulant-based inference of higher-order correlations (CuBIC) method has been presented. This method estimates the low-order cumulants and is able to decide whether the high-order correlations are needed. Thus, both a direct calculation of higher-order correlations and a requirement of a large sample size might be avoided. In Staude et al. (2010b) a modified version of the CuBIC method has been reported. This version is based on a statistical model which includes the non-stationary compound Poisson process.

Another technique which is well known in neuroscience and has been applied for analysing multiple spike trains is the generalised linear model (GLM). Comparison of MRP and GLM assumptions shows that they are similar: both methods assume that the prob-

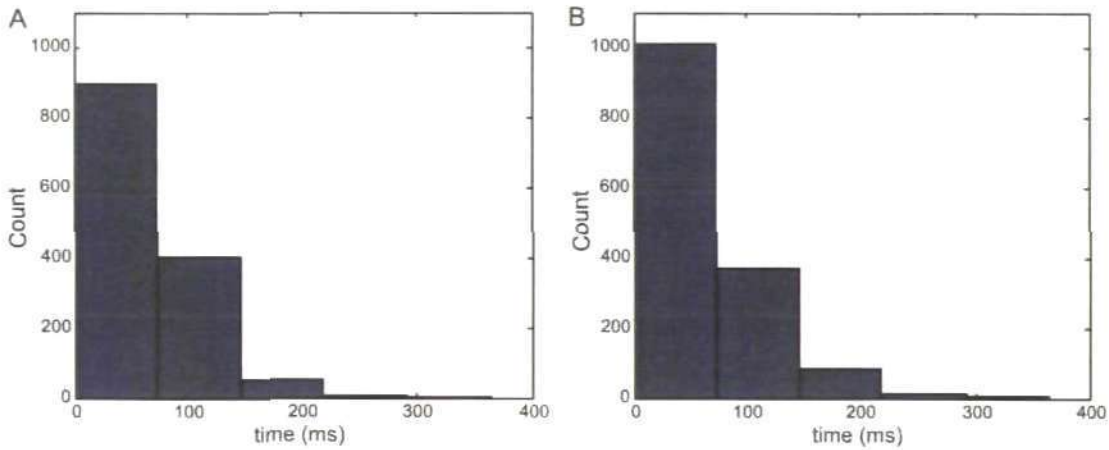


Fig. 14. (A) The ISI histogram for the data generated by the ELIF model. (B) The ISI histogram for the data generated by the probabilistic model with optimal values of parameters.

ability of spike generation at time t in the target train (under a given set of reference spike trains) depends on the intensity of the target train (without influences) modulated by a term which describes influences. To describe influences, the GLM based methods consider a number of time windows preceding time t and there is a prescribed parameter which indicates the strength of influence corresponding to this window (Okatan et al., 2005; Truccolo et al., 2005; Paninski et al., 2007). Thus, statistical procedures of GLM based methods count the number of spikes in reference spike trains falling into the window (bin) and use this information to calculate the strength of influence corresponding to the window. Therefore, the number of parameters equals to the number of windows and in some cases this number might be very large and it causes an over-fitting of data (Stevenson et al., 2008).

It is worth to apply all three methods (the Cox method, the higher order correlation technique and the GLM) to analyse the same data (both generated and experimental) and compare the results.

The Cox method has been applied to analyse functional connectivity of multiple spike trains generated by the enhanced leaky integrate and fire model (Borisyuk, 2002). To generate spike trains we prescribe the architecture of connections between ELIF elements. However, to analyse the generate data we suppose that connections are unknown. A diagram of functional connections is derived as a result of data analysis. We compare this diagram of identified functional connections with connections of ELIF model which have been used for data generation. Two examples have been considered. The result of analysing the five spike trains shows that the Cox method accurately identifies all functional connections. The result of analysing a large neural circuit of twenty spike trains also accurately identifies all forty two connections but also finds two erroneous connections which are relatively weak.

Comparison with the cross-correlation function shows that the Cox method has advantages over the CCF technique. In particular, the Cox method is more accurate in difficult situations such as a weak strength or short spike trains. One important advantage of the Cox method is that this method allows to analyse all simultaneously recorded spike trains. To demonstrate this advantage we apply the Cox method to analyse three spike trains coupled according to "common source" scheme and show that couplings can be correctly identified, but the pair-wise CCF fails to distinguish between the direct connection and the connection due to a common source. A similar example of three spike trains with "indirect connection" also demonstrates an advantage of the Cox method over the CCF.

For comparison with CCF we use a probabilistic model to generate data which satisfy with the assumption of the modulated renewal process. In this case the estimated Cox coefficient equals the prescribed strength of influence for data generation. A study of the probabilistic MRP model shows that this model can be fitted to a wide range of spike trains either generated by the integrate-and-fire model or experimentally recorded. In general the case we do not know whether the analysed spike trains satisfy with the assumption of MRP. However, our multiple tests show that the Cox method is robust and can be successfully used to find a functional connectivity for a wide range of point processes. In particular, we applied the Cox method for analysing connections in many cases of spike trains generated by the ELIF model. This nonlinear model is based on consideration of a threshold mechanism of spike generation and the postulates of this model are different from the MRP approach. Nevertheless, these tests show that the Cox method can be successfully applied to analyse data which are generated by the ELIF model.

Also, we have applied the Cox method to analyse functional connectivity of experimental recordings. Of course, in this case we cannot be sure that the result is correct. However, we can use some implicit tests to increase our confidence of the method. For example, we use the following methodology for testing the robustness of the statistical estimates. The epoch of recording is divided into two equal subintervals of time (left and right) and for each subinterval we apply the Cox method to analyse the functional connectivity of multiple spike trains. Thus, we have three sets of estimated coefficients: for the left interval, for the right interval, and for the whole epoch. If all three sets are identical or similar, then it means that the result is robust and the functional connections are the same for all three time intervals. A detailed report on the application of the Cox method to analyse experimental recordings is under preparation.

We have applied the Cox method to study the functional connectivity of many data sets of both generated and experimentally recorded, small size and large size (up to 100 spike trains), and we found that the Cox method is efficient, allows to identify the functional connectivity, and has some advantages in comparison with the method based on the cross-correlation function. Thus, we conclude that the Cox method can be successfully used to analyse functional connectivity of both generated and experimental data.

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

We would like to thank the anonymous reviewer for valuable comments. This work was supported by the EPSRC grant EP/E002331/1 (CARMEN project: www.carmen.org.uk).

Appendix A.

A.1. Derivation of formulas for the Cox method

A description of the method follows the paper by Cox (1972). To describe dependencies and influences among $(p+1)$ spike trains we select a target spike train A, other p spike trains are called reference spike trains and they are denoted by $B=(B_1, B_2, \ldots, B_p)$. The goal is to estimate the vector of unknown parameters $\beta=(\beta_1, \beta_2, \ldots, \beta_p)$ which describe the strengths of influences from reference trains to the target. Thus β_m represents the strength of influence from the spike train B_m to the target A $(m=1,2,\ldots,p)$. The main assumption is that the point process A is the modulated renewal process with the hazard function:

$$C_A(t) = \varphi(U_A(t)) \exp\left\{\sum\nolimits_{m=1}^{p} \beta_m Z_{B_m}(t)\right\},$$

where $C_A(t)$ is the hazard of the modulated renewal process A, $\varphi_A(\cdot)$ is the hazard function of the renewal process A without modulation (i.e. without influence from another point processes), $U_A(t)$ is the backward recurrence time of the process A at the moment t, $Z_{B_m}(t)$ is the influence function determining how the process B_m influence the process A, and β_m is the parameter describing the influence strength from the spike train B_m to the target A. To estimate the parameters $\beta=(\beta_1, \beta_2, \ldots, \beta_p)$ we use the method of conditional likelihood which eliminates the nuisance function $\varphi_A(\cdot)$ (Cox, 1972).

Let us suppose that the spike train A contains n interspike intervals $x_1, x_2, \ldots, x_n$. For simplicity we assume that all intervals $x_1, x_2, \ldots, x_n$ are of different length. If there are several identical intervals we use a randomization procedure and add a small normally distributed random number to the interval length. We arrange intervals in order of increasing size $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$. For $i>j$, let $x_{(i)}=x_k$ and $x_{(j)}=x_l$. We define $Z_{B_{mij}}$ to be the value of $Z_{B_m}(t)(m=1,2,\ldots,p)$ where time t is calculated in the following way: the interval x_l is allocated inside of the interval x_k and the left ends of both intervals coincide, time t corresponds to the right end of the interval x_l. Respectively, $Z_{B_{mii}}$ is the value of $Z_{B_m}(t)$ at the right end of the interval x_k.

We build up likelihood for the data conditionally on the magnitudes of the intervals, by considering contributions in order starting with the smallest interval. Thus the contribution from the first interval is

$$\frac{\exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{m11}}\right)}{\sum_{l=1}^{n} \exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{ml1}}\right)}$$

Conditionally on the first interval, the contribution from the second interval is

$$\frac{\exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{m22}}\right)}{\sum_{l=2}^{n} \exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{ml2}}\right)}$$

and so on.

The log likelihood is

$$L(\beta_1, \beta_2, \ldots\ldots, \beta_p) = \beta \sum\nolimits_{i=1}^{n} Z_{B_{mii}} - \sum\nolimits_{i=1}^{n} \log\left\{\sum\nolimits_{l=i}^{n} \exp\left(\sum\nolimits_{m=1}^{p} \beta_m Z_{B_{mli}}\right)\right\}$$

Now the first derivatives of the log likelihood is

$$\frac{\partial L}{\partial \beta_m} = \sum\nolimits_{i=1}^{n} Z_{B_{mii}} - \sum\nolimits_{i=1}^{n} \left[\frac{\sum_{l=i}^{n} Z_{B_{mli}} \exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{mli}}\right)}{\sum_{l=i}^{n} \exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{mli}}\right)}\right] ; (m = 1, 2, \ldots, p)$$

The estimate for β is obtained by setting the first derivatives at zero.

The second derivatives can be obtained by

$$\frac{\partial^2 L}{\partial \beta_r \partial \beta_s} = -\sum_{i=1}^{n} \left[\frac{\sum_{l=i}^{n} Z_{B_{rli}} Z_{B_{sli}} \exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{mli}}\right)}{\sum_{l=i}^{n} \exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{mli}}\right)}\right] + \sum_{i=1}^{n} \left[\frac{\sum_{l=i}^{n} Z_{B_{rli}} \exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{mli}}\right) \sum_{l=i}^{n} Z_{B_{sli}} \exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{mli}}\right)}{\left[\sum_{l=i}^{n} \exp\left(\sum_{m=1}^{p} \beta_m Z_{B_{mli}}\right)\right]^2}\right];$$

$(r, s = 1, 2, \ldots\ldots, p)$

Let $U(\beta) = \left[\partial L/\partial \beta_m\right]_{p\times 1}$ is the score vector and $I(\beta) = \left[-\partial^2 L/\partial \beta_r \partial \beta_s\right]_{p\times p}$ is the observed information matrix.

Now to obtain maximum likelihood estimate $\widehat{\beta}_q$ we have to find a numerical solution of the equation $U(\beta) = 0$. We use the Newton–Raphson iterative method starting from the initial guess $\beta^{(0)}$. Formula for iterations is:

$$\hat{\beta}_q = \hat{\beta}_{q-1} + [I(\hat{\beta}_{q-1})]^{-1} U(\hat{\beta}_{q-1}),\ q = 1, 2, \ldots, Q. \tag{A.1}$$

The iterations converge to the estimates:

$$\hat{\beta}_q \underset{q\to\infty}{\longrightarrow} \hat{\beta},\ \hat{\beta} = \left(\hat{\beta}_1, \hat{\beta}_2, \ldots\ldots, \hat{\beta}_p\right) \tag{A.2}$$

To obtain the confidence interval for β_m, we use the fact that $\partial L/\partial \beta_m$ has asymptotically normal distribution $N(0, I(\beta_m))$ where $I(\beta_m)$ is the mth diagonal element of $I(\beta)$. The confidence interval with the confidence level γ is

$$\left[\hat{\beta}_m - \sqrt{I\left(\hat{\beta}_m\right)} K_{(1-\gamma)/2},\ \hat{\beta}_m + \sqrt{I\left(\hat{\beta}_m\right)} K_{(1-\gamma)/2}\right] \tag{A.3}$$

where $K_{(1-\gamma)/2}$ is the upper $(1-\gamma)/2$ quantile of normal distribution. We accept the hypothesis that spike train $B_m; (m = 1, 2, \ldots\ldots, p)$ does not influence A if the confidence interval includes zero.

To obtain the 2D confidence region for β_r and β_s, we use the fact that $U^T(I_{rs})^{-1}U$ has χ^2 distribution with two degrees of freedom. Here $U = \left(\partial L/\partial \beta_r, \partial L/\partial \beta_s\right)$ and the matrix I_{rs} is

$$I_{rs} = \begin{bmatrix} -\dfrac{\partial L}{\partial \beta_r^2} & -\dfrac{\partial^2 L}{\partial \beta_r \partial \beta_s} \\ -\dfrac{\partial^2 L}{\partial \beta_s \partial \beta_r} & -\dfrac{\partial^2 L}{\partial \beta_s^2} \end{bmatrix}$$

The confidence region on the plane $\left(\beta_r, \beta_s\right)$ with the confidence level γ is defined by the following equation:

$$\left(\hat{\beta}_r - \beta_r, \hat{\beta}_s - \beta_s\right)' \left[I_{rs}\left(\hat{\beta}_r, \hat{\beta}_s\right)\right]^{-1} \times \left(\hat{\beta}_r - \beta_r, \hat{\beta}_s - \beta_s\right) \le \chi^2_{(1-\gamma,2)}, \tag{A.4}$$

where $\chi^2_{(1-\gamma,2)}$ is the upper $(1-\gamma)$ quantile of chi-square distribution with two degrees of freedom. We accept the hypothesis that Cox coefficients $\left(\hat{\beta}_r, \hat{\beta}_s\right) = 0$ if the confidence region includes zero.

Appendix B.

B.1. Parameter values for spike train generation by enhanced leaky integrate and fire model

An ELIF model can be simulated using software from the following web-site: http://www.tech.plymouth.ac.uk/infovis. To run the simulation, the parameters of ELIF neurons and their coupling should be specified.

Parameters of ELIF neuron:

(a) Maximum value of the threshold, (b) Threshold decay rate, (c) Asymptotic threshold value, (d) Amplitude of the noise (i.e. the standard deviation of the normally distributed random variable), (e) Noise decay rate, (f) Initial value of after spike hyperpolarisation, (g) Soma's membrane potential decay rate, (h) External input, (i) Absolute refractory period, and (j) Type of the neuron (0 – non-pacemaker, 1 – pacemaker).

Connection parameters describe synaptic coupling between elements:

(a) List of the numbers of those neurons which send their connections to the current neuron, (b) Connection strengths for these connections (positive for excitatory connection and negative for inhibitory), (c) Decay rates of postsynaptic potential for each connection respectively and (d) Time lag of spike propagation for each incoming connection (ms).

Below we show some examples of parameter values which have been used in simulations.

Neural parameters:

# of neuron	a	b	c	d	e	f	g	h	i	j
1	46.12	2.97	15.17	4.60	9.99	−28.94	20.76	0.73	6	0
2	43.07	2.87	14.33	5.38	9.91	−28.31	19.53	0.26	4	0

Connection parameters:

Neuron	The numbers of presynaptic neurons coupled with current neuron	Connection strength for these incoming connections	Decay rate	Time lag (ms) of spike propagation from pre to post-synaptic neuron
2	1 5	12.37 11.99	3.18 3.35	12 9

Appendix C.

C.1. Fitting the probabilistic model with the ELIF model

Probabilistic model includes several parameters. Here we describe how parameter values have been selected. We start from the ELIF model which generates two spike trains (A_1 and B_1) with connection from B_1 to A_1. An optimization procedure is used to find the parameters of the probabilistic model which provides the best fit to the spike trains generated by the ELIF model. In optimization procedure we use a cost function which depends on the difference of the histogram of ISI of the probabilistic model and the ELIF model. The minimum value of this cost function shows the minimum difference of the histogram of ISI of the probabilistic model and the ELIF model and suggests that both models generate nearly the same shape of ISI distribution.

First, we generate two spike trains using the enhanced leaky integrate and fire model with directed coupling from one neuron to another. The independent spike train B_1 simulates spikes generated by a neuron under some constant stimulation. This spike train influence the spike train A_1 emulating the synaptic connection from B_1 to A_1. Thus, we consider these two spike trains as the given data and would like to adjust parameters of the probabilistic model in such a way that this model will be able to generate spike trains which are similar to the given ones. To find the best values of parameters of the probabilistic model we use an optimization procedure and the cost function is calculated in the following way. Let select some parameter values of the probabilistic model (β, a, b, c, d), where β is the strength of influence from one spike train to another, (a,b) are the shape and scale parameters of the gamma distribution and (c,d) are the shape and scale parameters of the Weibull distribution. Using these parameter values we generate the independent renewal process B with the gamma distributed ISIs and another spike train A which is the modulated renewal process. The cost function Q takes into account a difference between the ISI distributions of the ELIF generated spike train A_1 and the MRP denoted by A:

$$Q = \sum_{i=1}^{k} \left(h_i^{A_1} - h_i^{A}\right)^2$$

here $h_i^{A_1}$ and h_i^{A} are the frequency of appearance of ISI in the bin i of the histogram for A_1 and A respectively and k is the number of bins. This bin number is arbitrary and we use $k = 5$.

The optimization procedure provides the optimal parameter values corresponding to the best fitting and the histograms are shown in Fig. 14. Fig. 14A shows the ISI histogram for the ELIF generated data and Fig. 14B shows the ISI histogram for the MRP data generated by the probabilistic model with the optimal values of parameters. These histograms are similar and the corresponding value of the cost function is small enough: Q is 4.9964×10^{-6}.

In the cost function we use the histogram of ISI which consider the shape of ISI distribution. This approach can be improved in such a way that the cost function includes both the histogram of ISI and the second order histogram of ISI (histogram of pairs of adjoin intervals) to reflect both the shape of ISI distribution and allocation of ISI in time. Of course, further improvement along this line leads to the cost function which includes both the histogram of ISI and the auto-correlation function. For simplicity we use the histogram of ISI in the cost function.

References

Aertsen AMHJ, Gerstein GL, Habib MK, Palm G. Dynamics of neuronal firing correlation: modulation of effective connectivity. J Neurophysiol 1989;61:900–17.

Bartlett MS. An introduction to stochastic processes. 2nd ed. Cambridge: Cambridge University Press; 1966.

Berger D, Warren D, Normann R, Arieli A, Grun S. Spatially organized spike correlation in cat visual cortex. Neurocomputing 2007;70:2112–6.

Berger D, Borgelt C, Louis S, Morrison A, Grun S. Efficient identification of assembly neurons within massively parallel spike trains. Comput Intell Neurosci 2010:1–18.

Borisyuk GN, Borisyuk RM, Kirillov AB, Kovalenko EI, Kryukov VI. A new statistical method for identifying interconnections between neuronal network elements. Biol Cybern 1985;52:301–6.

Borisyuk R. Oscillatory activity in the neural networks of spiking elements. BioSystems 2002;67:3–16.

Boven K-H, Fejtl M, Möller A, Nisch W, Stett A. On micro-electrode array revival. In: Advances in network electrophysiology using multi-electrode arrays. New York: Springer Press; 2006. p. 24–37.

Brillinger DR. Estimation of product densities. In: Frane JW, editor. Computer Science and Statistics: 8th Annual Symposium, Los Angeles; 1975a. p. 431–8.

Brillinger DR. The identification of point process systems. Ann Probab 1975b;3:909–29.

Brillinger DR. Estimation of the second-order intensities of a bivariate stationary point process. J R Stat Soc B 1976a;38:60–6.

Brillinger DR. Identification of synaptic interactions. Biol Cybern 1976b;22:213–28.

Brillinger DR. Measuring the association of point process: a case history. Am Math Monthly 1976c;83(1):16–22.

Brillinger DR. Maximum likelihood analysis of spike trains of interacting nerve cells. Biol Cybern 1988;59:189–200.

Brillinger DR. Nerve cell spike train data analysis: a progression of technique. J Am Stat Assoc 1992;87:260–71.

Brown EN, Kass RE, Mitra PP. Multiple neural spike train data analysis: state-of-the-art and future challenges. Nat Neurosci 2004;7:456–61.
Chornoboy ES, Schramm LP, Karr AF. Maximum likelihood identification of neuronal point process systems. Biol Cybern 1988;59:265–75.
Cox DR, Lewis PAW. Multivariate point processes. In: Proceedings Sixth Berkeley Symposium on Probability and Mathematical Statistics, vol. 3; 1972. p. 401–48.
Cox DR. The statistical analysis of dependencies in point processes. In: Lewis PA, editor. Stochastic point processes. New York: Willey; 1972. p. 55–66.
Daley D, Vere-Jones D. An introduction to the theory of point processes. New York: Springer-Verlag; 2003.
Eldawlatly S, Jin R, Oweiss KG. Identifying functional connectivity in large-scale neural ensemble recordings: a multiscale data mining approach. Neural Comput 2009;21:450–77.
Gerstein GL, Perkel DH. Simultaneously recorded trains of action potentials: analysis and functional interpretation. Science 1969;164:828–30.
Gerstein GL, Perkel DH. Mutual temporal relationships among neuronal spike trains: statistical techniques for display and analysis. Biophys J 1972;12:453–73.
Gerstner W, Kistler WM. Spiking neuron models: single neurons, populations, plasticity. Cambridge University Press; 2002.
Gerwinn S, Macke JH, Seeger M, Bethge M. Bayesian inference for spiking neuron models with a sparsity prior. Adv Neural Inf Process Syst 2007:20.
Lansky P, Sanda P, He J. The parameters of the stochastic leaky integrate-and-fire neuronal model. J Comput Neurosci 2006;21:211–23.
Louis S, Gerstein GL, Grun S, Diesmann M. Surrogate spike train generation through dithering in operational time. Front Comput Neurosci 2010a;4:127.
Louis S, Borgelt C, Grun S. Generation and selection of surrogate methods for correlation analysis. In: Grun S, Rotter S, editors. Analysis of parallel spike trains. Springer series in computational neuroscience, vol. 7. Springer Verlag; 2010b. p. 359–82.
Martignon L, von Hasseln H, Grun S, Aertsen A, Palm G. Detecting higher-order interactions among the spiking events in a group of neurons. Biol Cybern 1995;73:69–81.
Martignon L, Deco G, Laskey K, Diamond M, Freiwald W, Vaadia E. Neural coding: higher-order temporal patterns in the neurostatistics of cell assemblies. Neural Comput 2000;12:2621–53.
Mullowney P, Iyengar S. Parameter estimation for a leaky integrate-and-fire neuronal model from ISI data. J Comput Neurosci 2008;24:179–94.
Okatan M, Wilson MA, Brown EN. Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. Neural Comput 2005;17:1927–61.
Paiva ARC, Park I, Principe JC. A comparison of binless spike train measures. Neural Comput Appl 2009;19:405–19.
Paninski L, Pillow J, Lewi. Statistical models for neural encoding, decoding, and optimal stimulus design. Prog Brain Res 2007;165:493–508.
Paninski L. Maximum likelihood estimation of cascade point process neural encoding model. Network: Comput Neural Syst 2004;15:243–62.
Perkel DH. A computer program for simulating a network of interacting neurons. I. Organization and physiological assumptions. Comput Biomed Res 1976;9:31–43.
Perkel DH, Gerstein GL, Moore GP. Neuronal spike trains and stochastic point processes II. Simultaneous spike trains. Biophys J 1967;7:419–40.
Pillow J. Likelihood-based approaches to modeling the neural code. In: Kenji Doya, Ishii S, Alexandre Pouget, Rao Rajesh PN, editors. Bayesian brain probabilistic approaches to neural coding. MIT Press; 2007. p. 53–70.
Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP. Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature 2008;454:995–9.
Pipa G, Diesmann M, Grun S. Significance of joint-spike events based on trial-shuffling by efficient combinatorial methods. Complexity 2003;8(4):79–86.
Quiroga RQ, Nadasdy Z, Shaul YB. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput 2004;16:1661–87.
Reed JL, Kaas JH. Statistical analysis of large-scale neuronal recording data. Neural Networks 2010;23:673–84.
Rigat F, de Gunst M, van Pelt J. Bayesian modelling and analysis of spatio-temporal neuronal networks. Bayesian Anal 2006;1:733–64.
Staude B, Rotter S, Grun S. CuBIC: cumulant based inference of higher-order correlations in massively parallel spike trains. J Comput Neurosci 2010a;29(1–2):327–50.
Staude B, Grun S, Rotter S. Higher order correlations and cumulants. In: Grun S, Rotter S, editors. Analysis of parallel spike trains. Springer series in computational neuroscience, vol. 7. Springer Verlag; 2010b. p. 253–82.
Stevenson IH, Rebesco JM, Hatsopoulos NG, Haga Z, Miller LE, Körding KP. Inferring functional connections between neurons. Curr Opin Neurobiol 2008;18:582–8.
Stevenson IH, Rebesco JM, Hatsopoulos NG, Haga Z, Miller LE, Körding KP. Bayesian inference of functional connectivity and network structure from spikes. Brain Connect 2009;17(3):203–13.
Stuart L, Walter M, Borisyuk R. The correlation grid: analysis of synchronous spiking in multi-dimensional spike train data and identification of feasible connection architectures. BioSystems 2005;79:223–33.
Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. J Neurophysiol 2005;93:1074–89.
Tuckwell H, Richter W. Neuronal interspike time distributions and the estimation of neurophysiological and neuroanatomical parameters. J Theor Biol 1978;71:167–83.