UNIVERSIDADE DA CORUÑA

PhD Thesis

Nonparametric inference for neural synchrony under low firing activity

Author:

Aldana M. González Montoro

Supervisors:

Ricardo Cao Abad Jorge Mariño Alfonso

Departamento de Matemáticas
Facultade de Informática
Universidade da Coruña
SPAIN

February, 2013

UNIVERSIDADE DA CORUÑA

Los abajo firmantes hacen constar que son los directores de la Tesis Doctoral titulada "**Nonparametric inference for neural synchrony under low firing activity**", desarrollada por Aldana M. González Montoro en el ámbito del programa de doctorado de Estadística e Investigación Operativa ofertado por el Departamento de Matemáticas de la Universidade da Coruña, dando su consentimiento para que su autora, cuya firma también se incluye, proceda a su presentación y posterior defensa.

The undersigned hereby certify to be the supervisors of the Thesis entitled "**Nonparametric inference for neural synchrony under low firing activity**" developed by Aldana M. González Montoro, whose signature is also presented, under the PhD program Statistics and Operations Research at the Department of Mathematics of the University of A Coruña, consenting to its presentation and posterior defense.

18 de febrero, 2013

Directores:

Ricardo Cao Abad Jorge Mariño Alfonso

Doctoranda:

Aldana M. González Montoro

A mi familia

Dijo Tennyson que si pudiéramos comprender
una sola flor sabríamos quiénes somos
y qué es el mundo.

J.L. BORGES
El Aleph, 1949

Acknowledgments

One looks back with appreciation to the brilliant teachers,
but with gratitude to those who touched our human feelings.
The curriculum is so much necessary raw material,
but warmth is the vital element for the growing plant
and for the soul of the child.

Carl Jung

En primer lugar quiero agradecer a los directores de esta tesis por la oportunidad que me dieron al proponerme este problema y por la confianza que han tenido en mí. He aprendido muchísimo de ambos y este trabajo no hubiera sido posible sin su tan buena disposición todos estos años. Ha sido un placer trabajar junto a Ricardo desde el principio ya que, además de una mente brillante, es una gran persona y un gran profesor. Agradezco enormemente que siempre, incluso esta última temporada de tanto trabajo para él, haya sabido encontrar tiempo para nuestras reuniones de cuatro horas. Por otro lado, Xurxo ha sido una fuente de motivación constante, su pasión por la ciencia y su caracter comunicador han sido muy inspiradores para mí.

Quiero agradecer también a Nelson por estar siempre dispuesto a responder mis preguntas. A Germán AP. por sus palabras de ánimo, confianza y ayuda con la burocracia. A mis compañeros de máster, Javi y Miguel, que siempre han estado cerca y con nada más que palabras de cortesía para mí.

Al profesor Marcelo Montemurro por recibirme en su grupo en la Universidad de Manchester. A Christel Faes y Geert Molenberghs por permitirme realizar una estancia en su centro y por su colaboración en el capítulo cinco de esta tesis. Ha sido un gusto trabajar con ellos esos meses. A Emery Brown, por la oportunidad de realizar una estancia en su laboratorio en MIT y MGH. Emery es un gran científico y una maravillosa persona. Todo su grupo de me acogió de maravilla y me mostró lo que es trabajar en equipo. En particular, gracias a Aaron y a Sheri, quienes supieron solucionar todos

mis problemas. También a mi mate-mate con quién tuve pocas pero muy enriquecedoras charlas. Al profesor Juan Cuesta, por permitirme la visita a la UNICAN. Tengo que agradecerle la gran cantidad de horas que me dedicó a pesar de las *expropiaciones*. Ha sido un placer trabajar con él y espero poder seguir haciéndolo en el futuro.

Por hacer los días más amenos, a mis compañeros de laboratorio y cafés de todos estos años, María, Carmen, Marta, Álvaro, J. Germán, Paulas, Diego, Manu, Danis, Sasi, Brais, Javier, Migueles, Jesús, Raquel, Gonzalo, Moli, Xurxo y Porta. A estos últimos dos tengo que agradecerles por la compañía y *palabras* en este último tiempo de tanto trabajo.

Para la causante de que yo tenga la mejor panda del mundo no me alcanzan las palabras. Mi estadía estos años en Galicia y mi paso por este doctorado no hubieran sido lo mismo sin ella. Mis amigos gallegos han sido mi familia estos años cuando la de sangre ha estado tan lejos. María S., Tono, Diana, Óscar, María D., Bics y Sanchito, mil gracias por hacer las cosas tanto más fáciles y por esos fines de semana y viajecitos tan recargadores de energía. A los argentinos de siempre también porque, aunque lejos, siempre se hacen hacer sentir cerca, en particular a Any con sus mini mails de ánimo de este último tiempo!

A Eli por la buena onda, compañía, ayuda y tantas otras cosas desde que llegamos. Siempre has estado ahí para todo y has sido un pedacito de casa y familia a la vuelta de la esquina, gracias!

A Sancho por tanto que tuvo que aguantar estos años y por tanto cariño. Gracias por no dejarme caer en la desesperación tantas veces y por esa mirada tan esperanzadora de la vida!

A mi familia que, a pesar de estar tan lejos, siempre ha estado muy cerca. A mis hermanos por ser tan buena onda. Los chats intermitentes con Aye, los express con Né y los colgados con la Drú han sido una constante compañía y cable a casa a lo largo de estos años. Por último, a mis padres, por apoyarme siempre, sin importar cuán locas fueran mis decisiones y por intentar mostrarme siempre *lo que no veo en mí*.

Funding

La doctoranda ha sido beneficiaria de una beca FPI (BES-2009-017772) asociada al proyecto de investigacin MTM2008-00166 del antiguo Ministerio de Ciencia e Innovación de España. El desarrollo de esta tesis también ha sido parcialmente financiado por el proyecto de investigación MTM2011-22392, del Ministerio de Economía y Competitividad de España, y el proyecto INCITE09 137 272 PR, de la Xunta de Galicia.

Abstract

The aim of this thesis is to introduce statistical tools to study neural synchrony under spontaneous activity. The data analyzed comes from extracellular recordings of the primary visual cortex of anesthetized cats. The effect of the disruption of the normal spontaneous oscillations in neural synchrony is studied. Disruptions of the typical sleep-like patterns are achieved by means of electrical stimulations in the different neural nuclei that regulate the sleep-wake transitions, namely, the brainstem and the basal forebrain. Nonparametric methods are proposed to estimate the neural synchrony and bootstrap tests are proposed to test several hypotheses regarding the effects of stimulation in the associations between neurons at a pairwise and population level. Moreover, the relationship between the orientation selectivity of neurons and the synchrony among them is studied. Results indicate that the methods proposed in this thesis are succesfull in achieving the goals of the study. Significant decreases in synchrony due to the stimulations and differential effects of the stimuli are found. Moreover, at a populational level, our methods succeed on proving that functional affinity between neurons, regarding their orientations selectivity, affects the pairwise synchrony strength and dynamics.

Resumen

El objetivo de esta tesis es el de introducir herramientas estadísticas para el estudio de la sincronización neuronal bajo actividad espontánea. Los datos analizados provienen de registros extracelulares realizados en la corteza visual primaria de gatos anestesiados. Se estudia el efecto de la disrupción de la actividad oscilatoria espontánea en la sincronía entre pares de neuronas. La disrupción de los patrones típicos en estado de sueño se obtiene a partir de la micro estimulación eléctrica en los núcleos que regulan la transición del sueño a la vigilia, el tronco encefálico y el el área peribraqueal. Se proponen métodos noparamétricos para estimar la sincronía y contrastes bootstrap para contrastar hipótesis vinculadas con los efectos de la estimulación en la asociación entre neuronas tanto a nivel de pares de neuronas como a nivel poblacional. Más aún, se estudia la relación entre la selectividad a la orientación de las células y la sincronización entre ellas. Se encuentran disminuciones significativas en la sincronización debido a la estimulación y se encuentran efectos diferenciales entre las áreas estimuladas. Además, a nivel poblacional, nuestros métodos encuentran que la afinidad funcional entre neuronas, debido a la selectividad a la orientación, afecta la fuerza y la dinámica de la sincronización de pares de neuronas.

Resumo

O obxectivo desta tese é o de introducir ferramentas estatísticas para o estudo da sincronización neuronal baixo actividade espontánea. Os datos analizados proveñen de rexistros extracelulares realizados na cortiza visual primaria de gatos anestesiados. Estúdase o efecto da ruptura da actividade oscilatoria espontánea na sincronía entre pares de neuronas. A ruptura dos patróns típicos en estado de soño obtense a partir da microestimulación eléctrica nos núcleos que regulan a transición do soño á vixilia, o tronco encefálico e a área peribraqueal. Propóñense métodos non paramétricos para estimar a sincronía e métodos bootstrap para contrastar hipóteses vinculadas cos efectos da estimulación na asociación entre neuronas, tanto a nivel de pares de neuronas como a nivel poboacional. Máis aínda, estúdase a relación entre a selectividade á orientación das células e a sincronización entre elas. Atópanse disminucións significativas na sincronización debido á estimulación, así como efectos diferenciais entre as áreas estimuladas. Ademais, a nivel poboacional, os nosos métodos atopan que a afinidade funcional entre neuronas, debida á selectividade á orientación, afecta á forza e á dinámica da sincronización de pares de neuronas.

Contents

Chapter 1

Introduction

The aim of this thesis is to present statistical tools to deal with several methodological problems regarding the analysis of electrophysiological data; specifically, the estimation of synchrony dynamics between pairs of neurons under a regime of low firing activity. The methods are applied to real data and inferences about the estimated synchrony are made. The biological problem was proposed to the statistics group, MODES, of the Universidade da Coruña, by researchers of the neuroscience group, Neurocom, of the same university. Both, the questions and the data, resulted very challenging from a statistical point of view. This thesis describes the approaches to tackle the problem and the process to reach the objectives. This memoir can be found online at http://dm.udc.es/profesores/ricardo/PhDadvisorships.htm.

The experimental work that posited the statistical problems addressed in this work and that provided the neurophysiologycal data, aims to study brain functional connectivity through the synchronization dynamics between neurons of the primary visual cortex of anesthetized cats. We will present some methods to measure synchrony and to test several hypotheses regarding the effects of of stimulation in two precise areas of the brain. Finally, we will also introduce another factor regarding a particular characteristic of neurons of the visual cortex: the orientation selectivity. The objectives of the experimental study will be clearly stated in Chapter 2 together with a description of the experiment that led to the data. But, first of all, let us start introducing the area of study as well as the context of the problem.

1.1 Neuroscience context

Neuroscience is the field of knowledge that studies the structure and function of the nervous system, in particular, the human brain. It has numerous areas of study and brings together several disciplines such as medicine, psychology, biology and engineering among others. The electrophysiology is a branch of neuroscience that deals with the electrical properties and electrical activity of neurons. Technological advances have made possible the simultaneous electrophysiological recording of groups of neurons, generating large amounts of data that require specific methodological tools for their analysis. Areas like mathematics, physics, statistics and computational sciences are nowadays much involved in neuroscience, developing methods for data analysis to cope with the demand that electrophysiological problems generate.

1.1.1 Action potentials and spike trains

Neurons are specialized cells, which, along with glial cells, are the basic structural and functional units of the nervous system. These cells are organized in large and complex networks and they shape and connect the three main components of the nervous system: sensory, integration and motor. This is, they carry information from the sensory regions, analyze it, and then convey the responses to the corresponding regions. Neurons are formed by three functional parts: dendrites, soma and axon. The dendrites are ramifications that receive signals from other neurons or sensory organs and carry them into the soma. The soma is the central processing unit for the signals. Roughly speaking, it processes the information and generates a response signal which will travel through the axon to be delivered to other neurons or muscle cells. This information is carried as electrical impulses, which are called *action potentials* and often referred to as *spikes*.

Neurons are characterized by their capacity to propagate information very rapidly thorough very long distances. The generation of action potentials involves various specific morphological and physiological properties of the neurons. A simple way to describe it is the following. Every neuron is a kind of biological battery, with an electrical potential (Vm) between the inner and outer part of the cellular membrane (plasma membrane). The Vm is not fixed, but varies depending on the neuron inputs. When the electrical signal received at the soma surpasses a certain threshold, a sudden change in the cell's membrane potential is produced giving birth to an action potential, which will travel along the cell's membrane. These electrical pulses

are relatively easy to record, as they are abrupt changes in the Vm with a relatively high amplitude (∼100 mV). To do so, electrophysiologists place tiny electrodes inside or close to the neuron's soma or axon and record the electrical activity. These signals are referred to as intracellular or extracellular recordings basically depending on whether the electrode penetrates the cell or not. Figure 1.1 shows a neuron and three simulated recordings of its activity. We can see two intracellular recordings at the soma and axon of the neuron in which we can observe, not only the action potentials, but also the subthreshold activity of the cell. On the other hand, in the extracellular recording (center) we can only observe action potentials, as it is outside the cell and the electrode can only distinguish with clarity high amplitude changes of the extracellular potential.

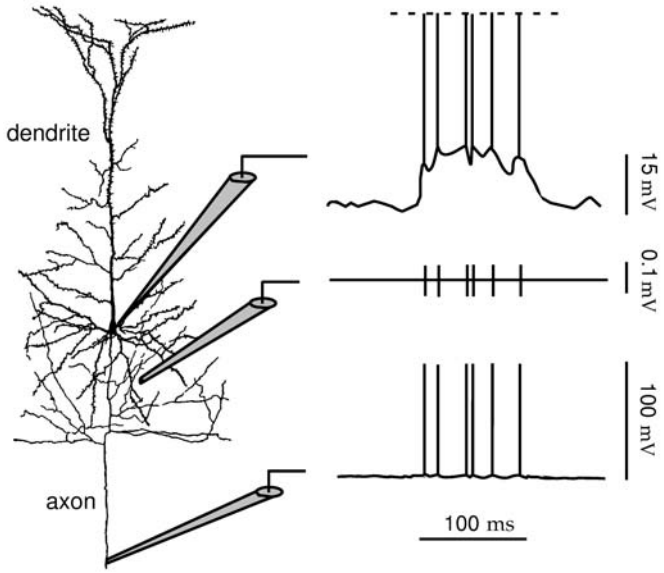

Figure 1.1: Representation of intracellular and extracellular recordings from Dayan and Abbott (2001).

Action potentials have an approximate amplitude of 100 mV and a duration of around 1 ms. The shape of these pulses remains practically constant while they travel through the axon. For this reason, it is believed that the information is carried by sequences of these spikes. The sequences of action potentials are called *spike trains* and they are the main object of study of this project. Figure 1.2 shows a typical plot to display spike trains, called *raster plot*. It corresponds to 50 seconds of spontaneous activity of a group of neurons which were recorded simultaneously. Action potentials are repre-

sented by vertical lines in the plot. Each row of each panel represents the spike activity of a neuron and each panel belongs to a different trial of the same group of neurons. As the principles of neural information processing are not well understood, the means by which spike trains carry information are a matter of debate and there exist several possible properties to be investigated. *Firing rates* and *exact time of firing* are two main views of possible neural codes (Shadlen and Movshon (1999), Singer (1999)), also associations and temporal correlations among neurons are key features in neural coding (Singer (1999)). For a more complete introduction to neuroscience in general, to spike trains in particular and mathematical modeling of neuronal data, please refer to, for example, Dayan and Abbott (2001) and Gerstner and Kristler (2002).

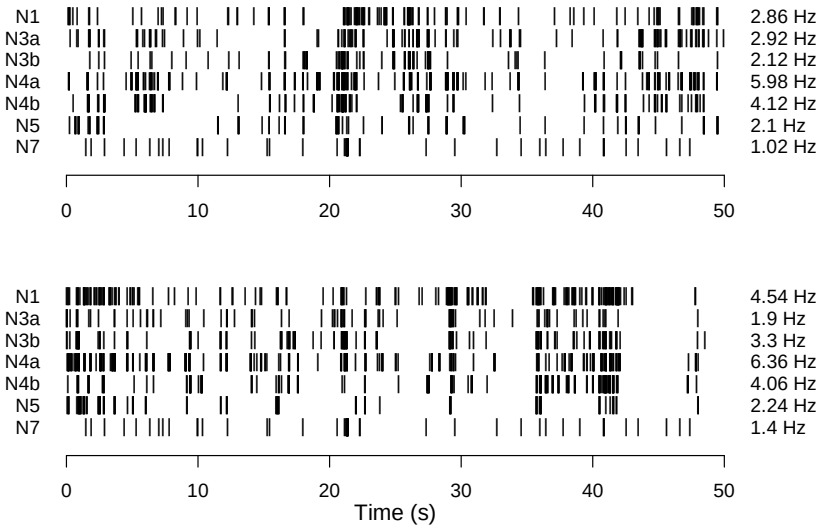

Figure 1.2: Raster plots of two trials (top and bottom panels) of a group of simultaneously recorded neurons (N1, N3a, N3b, N4a, N4b, N5 and N7) during 50 s of spontaneous activity. The average firing rate of each spike train is shown at the right of each row.

1.1.2 Sleep-like activity versus awake-like activity

The global brain activity observed during deep sleep is strikingly different to the one observed during the awake state. During the most profound phase of sleep, most neurons of the cerebral cortex display an oscillatory behavior, generating bursts of spikes with a dominant rhythm of about 1–4 Hz (1–0.25 s between bursts) which is called delta rhythm. This oscillation is highly synchronized between neurons, giving rise to the almost simultaneous generation of millions of action potentials by millions of neurons in the cortex and other brain regions. Because of this massive synchronization, the global electrical activity displays a high amplitude oscillation which is easily recorded even in the surface of the head: this is the delta oscillation of the electroencephalogram (EEG).

Under experimental conditions this global oscillatory activity can be induced by some anesthetics, giving rise to a sleep-like activity that is very useful to study the characteristic electrophysiological neuronal properties of this period. The left part of Figure 1.3 shows the EEG recording of an anesthetized cat, with a conspicuous delta oscillation, resulting from the synchronization of neuronal spikes that is taking place inside the brain.

During the awake state such global oscillatory synchronized activity does not exist, and neuronal spikes are not organized in repetitive bursts of activity, but follow what could be seen as a more random response, generating trains of spikes with different patterns and frequencies. This is called tonic activity, in contrast with the mentioned slow oscillatory activity. Of course, this tonic activity is not random, but is used to convey all kinds of information; and also, during this period, there are several types of high frequency (> 15 Hz) oscillatory patterns, but those are not global, but carried out by specific and small groups of neurons. Due to the fact that during this tonic activity, characteristic of the awake state, there is not a global neuronal synchronization, the EEG does not show changes of big amplitude, and remains fairly flat.

Interestingly, this tonic awake-like activity can also be reproduced experimentally and induced in the anesthetized cat by means of the electrical micro-stimulation of some regions or activating pathways. This is shown on the right part of Figure 1.3. At time 120 the global mode of operation of the cerebral cortex was artificially changed due to the delivery for a short period (usually 2 s) of micro-stimulation in the basal forebrain (one of the activating pathways). The result was the disruption of the global delta oscillation. It

can be seen that the previous high amplitude oscillation disappears, and for a period of several seconds the brain operates in an awake-like mode.

The example in Figure 1.3 shows a very interesting experimental model, in which we have, to the left, the spontaneous activity of a sleeping brain and, after the induced activation, some seconds of awake-like activity. After the stimulation, this awake-like activity last for up to 20 s and spontaneously returns to the sleep pattern. Hence, this experimental setup provides us with a model of sleep and awake activity, in which the same neurons can be studied while they are spontaneously interacting. The awake-like behavior is induced by a stimulation, but the subsequent activity can be thought of as spontaneous, as it last for a very long period (in electrophysiological terms) and spontaneously -and slowly- returns to the sleep-like pattern without any other intervention.

Most electrophysiological works rest upon the study of neuronal responses to some kind of stimulus. But we think that the study of the neuronal spontaneous electrical activity can also be of great help to unveil the underlying functional architecture. Because it has been much less studied, there are few mathematical and statistical tools to deal with the electrophysiological characteristics of the spontaneous spike activity. One of the main drawbacks is that, usually, the number of spikes is fairly small. Immediately after the application of an appropriate stimulus, neurons generate trains of spikes that can be easily analyzed using common statistics; but during spontaneous activity (either sleep-like or awake-like) the response of neurons is less robust and needs a specific statistical approach.

This is the case of the present work. It was developed to help in the analysis of the dynamic synchronization between pairs of neurons under two types of spontaneous activity: the anesthetic-induced sleep-like activity, and the electrically-induced awake-like activity. Throughout the present work we will use the term *stimulus* to refer to the micro-stimulation applied to activating pathways to induce the awake-like mode of operation, but, regarding the type of spike activity and the statistical problems under study, will consider both the sleep-like and the awake-like signals as periods of spontaneous activity. In any case, this is only a terminological issue that does not change at all the mathematical and statistical results and conclusions.

1.1.3 Spontaneous activity and activating ascending pathways

As it has just been explained, the data we will deal with along this work comes from a particular scenario in brain activity, referred to as *spontaneous activity*, which is the electrical activity that is observed in the absence of any discernible or controlled stimuli. It can be thought of as the brain activity at a resting state. The functional significance of this activity is not well understood and has not been widely studied, some views consider it as just *noise*, while others consider it a carrier of information. Rodieck et al. (1962) presented several descriptive methods to investigate these data. In recent years, some neurologist have focused their attention on a similar type of activity, coining the term "default mode network" or DMN, to refer to several brain regions implicated in the organization of neuronal activity while the brain is somehow idle. This is a term that has cognitive implications. On the other hand, in this work we study the spontaneous activity, which has a much more broader sense, and refers to any type of activity which is not directly driven by a controlled stimulus. The electrophysiological part of our work is not intended to dwell on the cognitive properties of the DMN, but to study some functional neuronal connections using the spontaneous electrical activity.

The spike activity analyzed in this work was obtained from an experimental model that combines a slow oscillatory mode (sleep-like) and a tonic mode (awake-like), using to switch between modes the transient stimulation of two activating pathways. Let us see this model with a bit more detail. As previously indicated, during the sleep state, the EEG is characterized by low frequency and high amplitude oscillations, while in the awake state, it presents a pattern of low amplitude and high frequency rhythms (Steriade et al. (1993)). There are several studies that indicate that these oscillations of the sleep state are originated in the thalamus and cerebral cortex and regulated by the brainstem (bs) modulatory systems (Steriade (1994)). Located in the upper brainstem, posterior hypothalamus and basal forebrain (bf), there are nerve paths, called the *ascending pathways*, that innervate the entire cortex and thalamus and release different neurotransmitters. These ascending pathways are responsible for the modulation and transition from the sleep state to the awake state. To induce this transition, the released neurotransmitters abolish the low-frequency rhythms in the thalamo-cortical network and promote the appearance of high-frequency oscillations, characteristic of the awake brain (Steriade et al. (1990)). For details on thalamic functions and the control systems of these nuclei, refer to Steriade et al.

(1997). The experimental electrical stimulation of these nuclei can change the EEG pattern from the typical sleep-like pattern to the one expected in an awake individual. This happens because stimulation suppresses slow waves (< 1 Hz), delta waves (1–4 Hz) and spindle-wave oscillations (7–14 Hz), and enhances gamma oscillations (30–100 Hz) and other high frequency patterns, thus introducing a tool to study the effects of the mechanisms that underlie the sleep-wake cycle (Hu et al. (1989); Burlet et al. (2002); Mariño and Cudeiro (2003)). Figure 1.3 shows a real EEG recording on the primary visual cortex of a cat. The cat was under deep anesthesia and, after 120 seconds of recording, an electrical stimulus was applied in its basal forebrain. A very noticeable change occurs in the EEG recording, as the salient high amplitude waves are lost for several seconds.

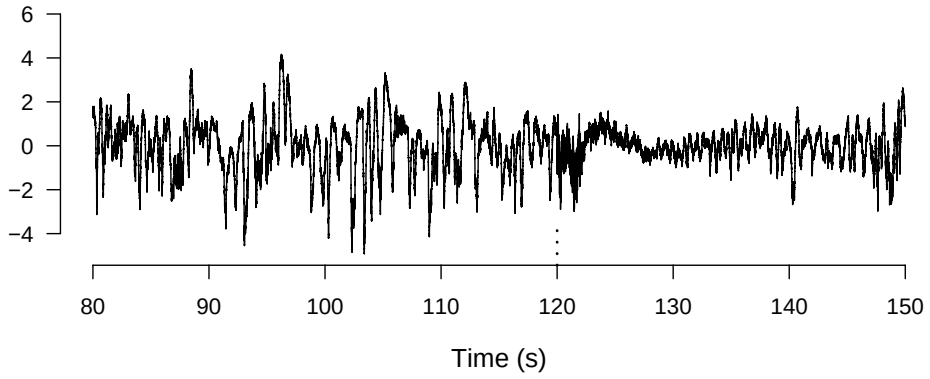

Figure 1.3: EEG recording of an anesthetized cat. Electrically stimulation wass applied at second 120 (dotted vertical line).

1.1.4 Neural synchrony

One of the key differences between the sleep and awake states is the synchronization between cortical neurons. This issue is the center in which we focus the present statistical work.

The oscillatory activity that is so clear in the EEG during the different states of the brain (different stages of sleep) reveals a massive neural synchronization. This synchronization can also be studied at the level of single

neurons and, in particular, of spike trains. Association measures are most commonly used to analyze synchronization of isolated single-neuron activities under certain conditions, like sensory stimulation or electrical activation of brain areas. Those approaches are essential to study the information coding and functional organization of the brain, but the spontaneous spike activity can also provide important clues to brain structure and function. Many methods have shown to be useful and competitive but most of them are designed to work with large spike densities (high firing rates) or plenty of repetitions (trials) of an experiment.

One of the most used tools to measure neural associations is the cross-correlation analysis. For example, the joint peristimulus time histogram (JPSTH) (Gerstein and Perkel (1969); Aertsen (1989)) displays the dynamics of correlation between neurons. This measure is the generalization for two neurons of the peristimulus time histogram (PSTH), which accumulates the spike times across trials of a single cell. The JPSTH is a two dimensional histogram of the joint firing count at time t of the first neuron and at time u of the second one. Its normalized version is just the Pearson correlation coefficient (computed across trials) of the firing counts of both neurons at two different time bins. This measure assumes that all the trials are indistinguishable and therefore it cannot take into account trial to trial variations. The cross-correlogram is the sum of the diagonals of the JPSTH and therefore shows the lagged firing-together of the two neurons. This is, the cross-correlogram is a histogram of the joint firing as a function of the time lag.

Other methods commonly used to capture synchrony are those based on 'unitary events' (Grün (1996); Grün et al. (2002); Grün (2009)). These methods rely on binned trains, where each bin will have a 1 if a spike occurs or a 0 otherwise. Unitary events refer to the occurrence of coincident spikes, or 1-1 matches, in the neurons under study. The unitary events analysis estimates the probabilities of joint-firing under the hypothesis of independence of the two spike trains. These probabilities are used to compute the expected numbers of joint spikes. Tests for synchrony are defined in terms of the difference between the expected frequencies and the observed ones. Faes et al. (2008) proposed a synchrony index, the Conditional Synchrony Measure, which is calculated, also, with binned trains. It is a flexible method, based on estimating the probability of joint-firing given that one of the neurons fired. However, as all the methods previously described, it was developed for the presence of many trials or, at least, a high firing rate. Quiroga et al. (2002) presented a nonlinear method based on relative timings of events in

time series, which can be applied not only to spike trains but also to EEGs or any other time series. Another interesting method is presented by Kruskal et al. (2007). Their method is based on the smoothing of spike trains themselves. They propose the use of kernel methods for this smoothing. There exist other approaches to neural associations which are defined in the frequency domain. For example, the correlation of the Fourier transform of two processes is called the *coherency*, and it is a function of the frequency. The squared modulus of the coherency is called *coherence* and it is an association method widely used in the literature (Brillinger (1992)). For general and state-of-the-art methods on spike train analysis, including correlation among neurons, see Brown et al. (2004) and Kass et al. (2005).

Usually, experiments are repeated several times under the same conditions, to be able to average across trials and, in this way, reduce the within-trial noise. On the other hand, averaging among trials has the natural drawback of possible between-trial noise that can be originated by uncontrollable differences in the experiment setting. In our particular case, the data for each neuron came from a small number of trials and also the overall spike activity was fairly low. The number of recorded trials is low because of several methodological restrictions, such as the long duration of the protocol applied to every group of simultaneously recorded neurons, and the methodological necessity to keep the number of electrical stimulations as low as possible. Regarding the low spike activity, it is characteristic (among other regions) of the spontaneous activity recorded in our area of study, the primary visual cortex. It is then necessary to develop specific statistical tools to analyze neural dynamic synchronization under these circumstances.

1.1.5 Primary visual cortex and orientation selectivity

Towards the last part of the present work some of the synchronization analysis between pairs of neurons are related to one more variable (together with the sleep vs awake mode and the two activating pathways studied): the orientation selectivity of visual cortex cells.

The visual cortex is the part of the cortex responsible of processing visual information. It is composed by the *primary visual cortex* (V1) and the extrastriate visual cortical areas, such as, V2, V3, V4 and V5. In mammals, including humans, V1 is located in the posterior pole of the occipital cortex and it is the one that receives the visual input from the retina, after being processed at thalamic level.

David Hubel and Tortsen Wiesel discovered in 1958 that cells in the visual cortex are selective to orientation. This means that neurons respond (fire more) if they detect local bars or edges at some particular angle in the processed image (Hubel and Wiesel (1962)). Figure 1.4 shows an example of this fact. In the left panel, it shows extracellular recordings of a neuron in the primary visual cortex of a monkey while oriented light bars where shown to the animal. It is clear how the amount of spikes depend on the orientation of the bar. In the right panel, the figure shows the average firing rate of a neuron in a cat V1 as a function of the angle of the bar presented.

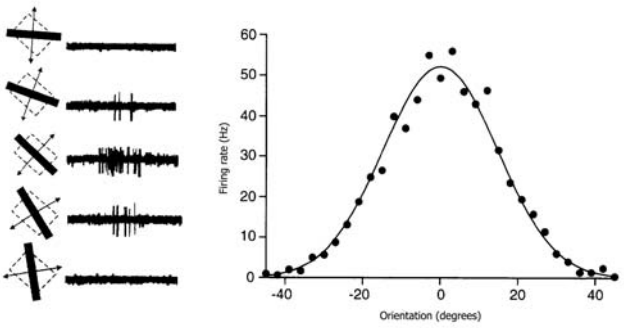

Figure 1.4: Left panel: response firing of a neuron in the primary visual cortex of a monkey. Right panel: average firing rate of a cat V1 neuron; the data is from Henry et al. (1974).

1.2 Nonparametric methods

In general, we will use nonparametric methods to do descriptive analysis of the data, develop association methods and make inference based on them. In particular, kernel smoothing methods will be repeatedly used. We will use these methods for the estimation of density and regression curves. In each chapter, the methodology used or proposed is basically explained and references are provided for more details. Anyhow, it is worth a general overview of kernel methods. We refer the reader to the book of Wasserman (2006) for general nonparametric statistics or the books of Wand and Jones (1995) and Simonoff (1996) for theory and details on kernel smoothing.

When it cannot be assumed that the distribution of the variable under study belongs to a given finite-dimensional parametric space, nonparametric methods are used. Parametric estimators are the most efficient while the

model assumed for the data is the correct one, otherwise these estimators can be inconsistent. On the other hand, nonparametric estimators are usually consistent although they tend to be less efficient for samples of small sizes when a parametric model holds. In the spike train data context, Poisson process models are often used to describe spike times. Nevertheless, there are several intrinsic characteristics of spike trains that make the Poisson model inadequate, such as refractory periods or burst activity, for example. Statistical models for spike data have been widely discussed (Gerstein and Madlebrot (1964); Tuckwell (1988); Smith and Smith (1965); Shadlen and Newsome (1998)) and other models have been proposed and discussed by Barbieri et al. (2001), Kass and Ventura (2001) and Reich et al. (1998), among others. In Chapter 3 we will show the Poisson model is not useful for the data under study in this thesis.

To estimate curves, such as density or regression functions, we will assume they are smooth and make use of kernel estimators, which were first introduced by Rosenblatt (1956) and Parzen (1962).

1.2.1 Kernel density estimation

Let X be a random variable with density function f and let $x_1, \ldots, x_n$ represent a random sample of size n coming from the density f. The histogram is probably the most simple and widely used nonparametric method to estimate f. Another estimator of the density function, is the *kernel density estimator*:

$$\hat{f}_{n,K}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right),$$

where K is a kernel function, typically non-negative and satisfying $\int K(u)\mathrm{du} = 1$, $\int uK(u)\mathrm{du} = 0$ and $\int u^2 K(u)\mathrm{du} = \sigma_K^2 > 0$. The positive number, h, is called the *bandwidth* and it is the parameter which controls the amount of smoothing. Some commonly used kernel functions are shown in Table 1.1.

Figure 1.5 shows the estimated density function of a small set of data using a histogram and a kernel density estimator. The data are the logarithms of the intersipke intervals which will be described in Chapter 3. The values of the sample are depicted underneath the density estimations and correspond to a real spike train. A Gaussian kernel has been used for this estimation. It can be observed that even though the histogram is informative, the kernel estimator has the advantage of being smooth and much more sensitive to local properties of the underlying density.

Table 1.1: Some common kernel functions.

Epanechnikov	$K(u) = \frac{3}{4}(1-u^2)\,\mathbb{I}\{u \in [-1,1]\}$
Biweight	$K(u) = \frac{15}{16}(1-u^2)^2\,\mathbb{I}\{u \in [-1,1]\}$
Triweight	$K(u) = \frac{35}{32}(1-u^2)^3\,\mathbb{I}\{u \in [-1,1]\}$
Gaussian	$K(u) = \frac{1}{\sqrt{2\pi}}\exp(-u^2/2)$
Uniform	$K(u) = \frac{1}{2}\,\mathbb{I}\{u \in [-1,1]\}$

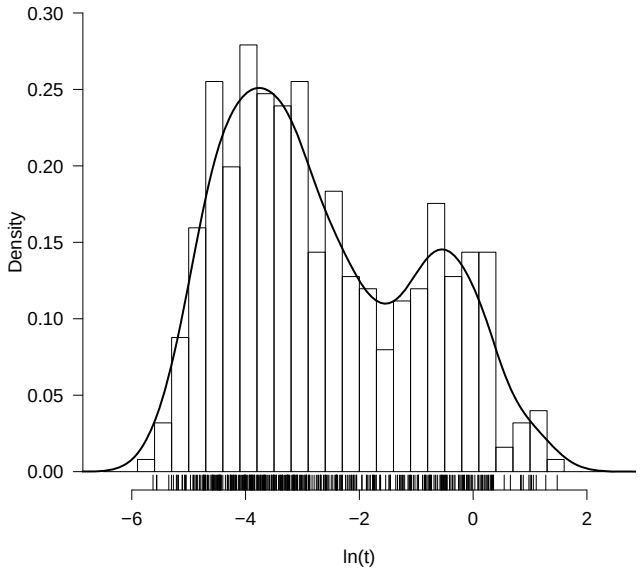

Figure 1.5: Histogram and (Gaussian) kernel density estimator for a sample of the logarithms of intersike intervals.

It is well known that the choice of the kernel function in not too important. When suitable bandwidths are chosen, the results obtained for different kernels will not be considerably different, see, for example, Wand and Jones (1995), pp. 28–31 or Simonoff (1996), pp. 41–44. This fact is also shown in Figure 1.6 where the Epanechnikov, the biweight and the uniform kernels have been used. The bandwidths have been chosen to obtain equivalent estimations to the one shown in Figure 1.5 (see Simonoff (1996), p. 45). Note that the results are comparable.

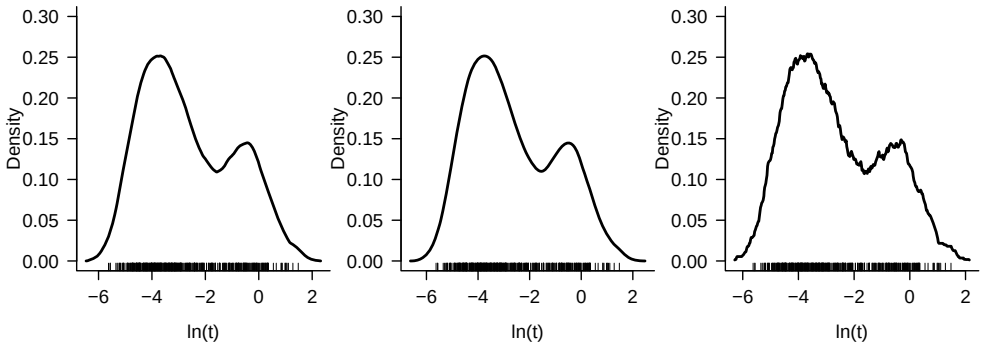

Figure 1.6: Kernel density estimation using Epanechnikov (left), biweight (center) and uniform (right) kernels, for a sample of the logarithms of intersike intervals.

On the other hand, the choice of the bandwidth, h, is really important. Small bandwidths give rough estimates while larger bandwidths give oversmoothed estimates. There exist several methods to conveniently choose this smoothing parameter. Let us discuss briefly some of them. The simplest method is the one that looks out for optimization of the mean integrated square error (MISE). The mean square error (MSE)

$$\mathrm{MSE}(x) = \mathbb{E}((\hat{f}_{n,K}(x) - f(x))^2)$$

is a measure often used to evaluate the error of the estimator $\hat{f}_{n,K}$. This quantity can be rewritten as the sum of the squared-bias and the variance of the estimator:

$$\mathrm{MSE}(x) = (\mathbb{E}(\hat{f}_{n,K}(x) - f(x)))^2 + \mathrm{Var}(\hat{f}_{n,K}(x)).$$

Under some regularity conditions (f'' absolutely continuous, and $f''' \in \mathrm{L}^2$), if $h \to 0$ with $nh \to \infty$ as $n \to \infty$, then, it can be proven using Taylor series expansions,

$$\mathbb{E}(\hat{f}_{n,K}(x) - f(x)) = \frac{h^2 \sigma_K^2 f''(x)}{2} + O(h^4)$$

and

$$\mathrm{Var}(\hat{f}_{n,K}(x)) = \frac{f(x)R(K)}{nh} + O(n^{-1}),$$

with $R(K) = \int K^2(u)\mathrm{du}$. Integrating the MSE we get the MISE, which is a global measure of accuracy. Its asymptotic value can be written as:

$$\mathrm{AMISE} = \frac{R(K)}{nh} + \frac{h^4 \sigma_K^4 R(f'')}{4}.$$

The value of h that minimizes AMISE is

$$h_{AMISE} = \left(\frac{R(K)}{\sigma_K^4 R(f'')n} \right)^{\frac{1}{5}}. \tag{1.1}$$

This gives a possible method to select h: choose a reference density function and substitute it into (1.1). For example, if f is a normal density with standard deviation σ then,

$$h_{AMISE} = \left(\frac{8\sqrt{\pi}R(K)}{3\sigma_K^2 n} \right)^{\frac{1}{5}} \sigma.$$

If the kernel is the Gaussian one, and replacing σ by a sample estimate, $\hat{\sigma}$, we get that the optimal h is

$$\hat{h}_{AMISE} = 1.059\hat{\sigma}n^{-1/5}.$$

We now consider a second method. Following the *plug-in* principle, the asymptotically optimal h in (1.1) can be estimated by

$$\hat{h} = \left(\frac{R(K)}{\sigma_K^4 \widehat{R(f'')}n} \right)^{\frac{1}{5}},$$

where $\widehat{R(f'')}$ is an estimate of $R(f'')$. Sheather and Jones (1991) proposed $\widehat{R(f'')} = R(\hat{f}'')$ where $\hat{f}$ is an estimate of f computed with a different pilot bandwidth.

The last approach we will describe is the *cross-validation* method. Instead of minimizing the MISE, the integrated square error (ISE) is considered here:

$$\begin{aligned} \text{ISE}(h) &= \int (\hat{f}_{n,K}(u) - f(u))^2 \text{du} = \\ &= \int \hat{f}_{n,K}^2(u)\text{du} - 2\int \hat{f}_{n,K}(u)f(u)\text{du} + \int f^2(u)\text{du}. \end{aligned}$$

The term $\int f^2(u)\text{du}$ clearly does not depend on h, therefore we can forget about that term. On the other hand,

$$\int \hat{f}_{n,K}(u)f(u)\text{du} = \mathbb{E}(\hat{f}_{n,K}(Y)|x_1, \cdots, x_n),$$

where Y is a random variable with density f, which needs to be estimated. Let $\hat{f}_{n,K}^{-i}(x_i)$ be an estimate of f without using the i-th element of the sample

and evaluated in x_i. The expected value of that estimate is $\mathbb{E}(\hat{f}_{n,K}(x))$ and therefore

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{f}_{n,K}^{-i}(x_i)\right) = \mathbb{E}(\hat{f}_{n,K}(x)) = \int \hat{f}_{n,K}(u)f(u)\mathrm{du}$$

So, the cross-validation method consists in choosing h such that minimizes

$$\mathrm{CV}(h) = \int \hat{f}_{n,K}^{2}(u)\mathrm{du} - \frac{2}{n}\sum_{i=1}^{n}\hat{f}_{n,K}^{-i}(x_i).$$

In Figure 1.5 the bandwidth was chosen by the Sheather and Jones plug-in method. The Sheather and Jones plug-in smoothing parameter turned out to be $h = 0.38$. To exemplify the importance of the bandwidth selection, in Figure 1.7 the same density is estimated again. In this case, the smoothing parameters were chosen to be $h = 0.095$ and $h = 0.85$. Large differences can be observed. On one side, when a small bandwidth is used, the density is noticeably undersmoothed. On the other hand, when a large bandwidth is used, the density results oversmoothed and important information could be lost such as, for example, in this case, the bimodality of this density function could have been disregarded.

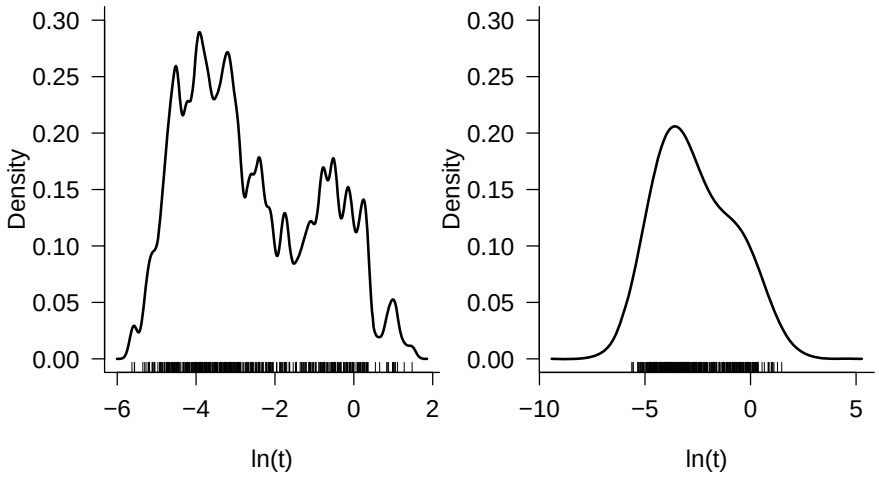

Figure 1.7: Kernel density estimation using a Gaussian kernel for a sample of the logarithms of interspike intervals. The bandwidth paramteters are $h = 0.095$ (left panel) and $h = 0.85$ (right panel).

1.2.2 Kernel regression estimation

In the regression context, the interest lies on analyzing the relation between a response variable, Y, and a covariable, X. This is, if $\{(x_i, y_i)\}_{i=1}^n$ is a sample of (X, Y), then, the regression model states:

$$y_i = m(x_i) + \epsilon_i.$$

By definition, $m(x)$ is the conditional expectation $\mathbb{E}(Y|X = x)$, with $\mathbb{E}(\epsilon|X = x) = 0$ and $\text{Var}(\epsilon|X = x) = \sigma^2(x)$. In the nonparametric case, $m(x)$ is just assumed to be a smooth general curve, contrary to, for example, the most commonly used model, the linear regression, where a specific parametric form is assumed: $m(x) = a + bx$.

The *kernel regression estimator* is defined as:

$$\hat{m}_{n,K}(x) = \sum_{i=1}^{n} w_i y_i \text{ with}$$

$$w_i = \frac{K(\frac{x-x_i}{h})}{\sum_{j=1}^{n} K\left(\frac{x-x_j}{h}\right)} = \frac{\frac{1}{nh}K(\frac{x-x_i}{h})}{\frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{x-x_j}{h}\right)},$$

K is a kernel function and h is the bandwidth as before. This estimator is called the *Nadaraya-Watson kernel estimator* (Nadaraya (1964) and Watson (1964)).

The Nadaraya-Watson estimator is a local weighted mean of the observations of Y with $\sum_{i=1}^{n} w_i = 1$ and it happens to be the solution to the weighted least squares problem, being the quantity that minimizes the function:

$$\Psi_0(\gamma_0) = \sum_{t=1}^{n} (y_i - \gamma_0)^2 K\left(\frac{x - x_i}{h}\right).$$

This idea suggests fitting a higher order polynomial and gives place to *local polynomial regression*. This method was introduced by Stone (1977) and Cleveland (1979) but it gained importance with the papers of Ruppert and Wand (1994) and Fan and Gijbels (1995). The local polynomial estimator is obtained by using weighted least squares to locally fit polynomials of a given degree previously specified. It is a method of easy computation, it adaptates to estimate derivatives and it also has other nice properties (see Wand and Jones (1995) or Fan and Gijbels (1996) for details).

Observe that, if m has $p+1$ continuous derivatives at the point x, it can be approximated (using a Taylor expansion) by

$$m(z) \approx m(x) + m'(x)(z-x) + \frac{m''(x)}{2!}(z-x)^2 + \ldots + \frac{m^{(p)}(x)}{p!}(z-x)^p.$$

This polynomial can be fitted by minimizing the function:

$$\Psi_p(\gamma) = \sum_{t=1}^{n} \left(y_i - \sum_{j=0}^{p} \gamma_j (x_i - x)^j \right)^2 K\left(\frac{x - x_i}{h} \right),$$

where $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_p)$. The local polynomial regression estimator of the j-th derivative of $m(x)$, $m^{(j)}(x)$, is $\hat{m}_j(x) = j!\hat{\gamma}_j$, $j = 1, \cdots, p$, where $\hat{\gamma} = (\hat{\gamma}_0, \hat{\gamma}_1, \cdots, \hat{\gamma}_p)$ is the minimizer of Ψ_p. In particular, $\hat{m}(x) = \hat{\gamma}_0$ is an estimator of $m(x)$. The Nadaraya-Watson estimator is the particular case of the local polynomial kernel estimator when $p = 0$.

As in the density function estimation case, the choice of the kernel function K is not as relevant as the proper choice of the smoothing parameter h. The methods described in Section 1.2.1 have a similar formulation for the regression problem. For example, the cross-validation method, using the *leave-one-out* procedure in this context is as follows. The cross-validation function is

$$\mathrm{CV}(h) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{m}^{-i}(x_i))^2$$

where $\hat{m}^{-i}$ is built as $\hat{m}$ but with all the data except the pair (x_i, y_i) and evaluated in x_i. The bandwidth selector obtained by this method is the minimizer of $\mathrm{CV}(h)$.

Instead of a constant bandwidth, a local variable bandwidth can be used. A variable bandwidth, $h(x)$, allows for different degrees of smoothing, giving flexibility to the fitting and helping to reduce the bias in rough regions while possibly reducing the variance in flat regions. For example, for the cross-validation method, the data can be separated in blocks or sliding windows and the parameter $h(x)$ estimated at each block\window. On the other hand, Fan and Gijbels (1996) give a theoretical optimal local bandwidth for $\hat{m}_j(x)$, the estimator of the j-th derivative of m. It is obtained by minimizing the conditional mean squared error (MSE):

$$MSE(x) = \left(\mathrm{Bias}\left(\hat{m}_j(x) | \{x-1, \cdots, x_n\} \right) \right)^2 + \mathrm{Var}\left(\hat{m}_j(x | \{x-1, \cdots, x_n\}) \right).$$

This theoretical optimal choice can be approximated by the value that minimizes the asymptotic conditional mean squared error (AMSE):

$$AMSE(x) = \left(h_n^{p+1-j} \frac{m^{(p+1)}(x) j!}{(p+1)!} \int uK(u)\mathrm{du} \right)^2 + \\ + \frac{h_n^{2j+1} \sigma^2(x) (j!)^2}{f(x)} \int K^2(u)\mathrm{du}$$

and results in

$$h_{AMSE}(x) = C_{j,p}(K) \left(\frac{\sigma^2(x)}{(m^{(p+1)}(x))^2 f(x)} \right)^{1/(2p+3)} n^{-1/(2p+3)},$$

where

$$C_{j,p}(K) = \left(\frac{(p+1)!^2 (2j+1) \int K^2(u)\mathrm{du}}{2(p+1-j)(\int u^{p+1} K(u)\mathrm{du})^2} \right)^{1/(2p+3)}.$$

Of course, there are still unknown quantities in the expression for the asymptotically optimal bandwidth, such as, the density f, the conditional variance σ^2 or the $(p+1)$-th derivative of the function m. These quantities can be estimated from the data and *pluged in* the formula. Thus, this is the plug-in method for the local bandwidth selection. Details of this and other methods to choose the bandwidth for the local polynomial kernel estimator can be found in Fan and Gijbels (1996).

Another important quantity to choose is the order of the polynomial, as high orders may decrease the bias while increasing the variability. Although there exist automatic procedures for this selection, in general, $p = j + 1$ is sufficient and an odd value p is advised (Fan and Gijbels (1996), Chapter 3).

1.2.3 Bootstrap

Along this work we make an extensive use of bootstrap methods to approximate the sampling distribution of some estimators. The bootstrap is a resampling technique that is relatively easy to use. It is an appealing alternative to parametric inference when this is impossible or requires complicated mathematical calculations. Bootstrap methods obtain estimates for estimators' properties by resampling with replacement from an original data sample. These techniques are most often used to make inference on estimators by approximating their variance or building confidence intervals but they can also be used to conduct hypothesis tests. The bootstrap was first introduced

by Efron (1979) to estimate the standard error of a statistic. This is probably the simplest and most commonly used application of the bootstrap. As before, assume we have a sample $\mathbf{x} = \{x_1, \cdots, x_n\}$ from an unknown probability distribution, F, and let θ by a parameter of that distribution, that needs to be estimated. To do this, we can compute the estimate $\hat{\theta} = s(\mathbf{x})$ from the sample $\mathbf{x}$. To have a notion of how accurate $\hat{\theta}$ is, we could estimate its standard error, $\text{se}(\hat{\theta})$. In this context, the bootstrap procedure is very simple. A *bootstrap resample*, $\mathbf{x}^* = \{x_1^*, \cdots, x_n^*\}$ of size n is defined as a random sample drawn with replacement from the empirical distribution, $\hat{F}$, that assigns probability $1/n$ to each point in the sample $\mathbf{x}$. The *bootstrap statistic* can be computed from $\mathbf{x}$: $\hat{\theta}^* = s(\mathbf{x}^*)$. So, the so called *uniform bootstrap* procedure is as follows:

1. Draw a random sample $\mathbf{x}^* = \{x_1^*, \cdots, x_n^*\}$ from $\hat{F}$.
2. Evaluate the bootstrap statistic $\hat{\theta}^* = s(\mathbf{x}^*)$.
3. Repeat 1 and 2, B times to obtain $\hat{\theta}^{*1}, \cdots, \hat{\theta}^{*B}$ and compute their sample standard error. This gives $\hat{\text{se}}(\hat{\theta}^*)$ which is an estimate of $\text{se}(\hat{\theta})$.

The general idea of the uniform bootstrap is repeated in almost every context where the bootstrap can be applied. Although resampling procedures may change, bootstrap samples are obtained by Monte Carlo from the original sample, and the desired characteristic of the estimator distribution approximated by the corresponding bootstrap analogue. We will briefly discuss some of the different bootstrap procedures.

If F is a continuous distribution with density function f, it can be estimated by the integral of the kernel density estimator, with bandwidth h, discussed in Section 1.2.1:

$$\hat{F}_h(x) = \int_{-\infty}^{x} \hat{f}_h(u)\text{du}.$$

Therefore, the random samples can be obtained by resampling from $\hat{F}_h$ instead of doing it from$\hat{F}$. This is called the *smooth bootstrap* (Efron and Tibshirani (1993) pp. 231; Silverman and Young (1987)). The Monte Carlo resampling results easy: let W be a random variable with density function $K(w)$ and X_0 another random variable with distribution $\hat{F}$, the empirical distribution of the data. It results that $\hat{F}_h$ is the distribution of $hW + X_0$ and therefore a bootstrap sample can be obtained by $x_i^* = hw_i + z_i$ with w_i a realization of W and z_i drawn with equiprobability from $x_1, x_2, \cdots, x_n$.

Instead of using a nonparametric estimator for F, a parametric estimation, $\tilde{F}$, is sometimes plausible. If this is the case, a *parametric bootstrap* can be used, resampling from $\tilde{F}$.

The bootstrap can be also used to make inference on regression models. Consider the linear model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i,$$

where inference on $\hat{\beta} = (\beta_0, \beta_1)^t$ is aimed. In this case, a possible procedure is to resample from the residuals of the fitted model. This is, consider the least squares estimator of β, $\hat{\beta}$, and let $\hat{\epsilon} = y_i - \hat{y}_i$, with $\hat{y}_i$ the fitted values of the model. Then, a bootstrap resample for this problem can be obtained as $y_i^* = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\epsilon}_i^*$, being $\{\hat{\epsilon}_1^*, \cdots, \hat{\epsilon}_n^*\}$ a bootstrap resample obtained from the empirical distribution of $\{\hat{\epsilon}_1, \cdots, \hat{\epsilon}_n\}$.

Of course, the smooth bootstrap and the parametric bootstrap described above can be adapted to this regression context as well. Also, there exist variants that take into account heteroscedasticity (wild bootstrap; Wu (1986)) or dependence in the data. If this last one is the case, just resampling from the estimated distribution of the residuals would not imitate the correlation of the data and therefore the method would fail. There exists methods to overcome this drawback which are widely used in time series analysis. The moving blocks bootstrap (Künsch (1989); Liu and Singh (1992)) or the stationary bootstrap (Politis and Romano (1994)) are examples of bootstrap methods designed for dependent data. In the moving block bootstrap, the resampling procedure is made on fixed-length blocks of data instead of single data points. The stationary bootstrap, allows for variable length and it improves the moving blocks procedure because, as it name indicates, it is stationary while the previous method is not. The stationary bootstrap procedure consists in choosing a real number $p \in [0, 1]$, drawing at random y_1^* from $\hat{F}$ and, once $y_i^* = y_j$ has been chosen (for $j \in \{1, \cdots, n-1\}$), defining y_{i+1}^* as y_{j+1} with probability $1 - p$ and drawing from $\hat{F}$ with probability p. In the case $j = n$, the observation y_{j+1} is replaced by y_1.

1.3 Summary of the following chapters

The content of the rest of the thesis is briefly summarized now. In Chapter 2 the objectives of the experimental study are outlined. Also, the experiment is described as the data used in this work are briefly presented. Chapter 3 is an introduction to spike trains. A mathematical definition is presented and

single spike train correlation measures are discussed. Chapter 4 introduces interaction between two spike trains. A method to measure associations based on a generalization of inter spike intervals for two neurons is proposed and a procedure to test for significance in the effect of each stimulus on these measures is presented. A second method to measure synchrony is presented in Chapter 5. This method is based on the cross-correlogram. Hypothesis tests are presented for the differential effect of stimulation. In Chapter 6 an alternative method to measure synchrony is presented with its respective hypothesis tests. The orientation selectivity of neurons is introduced in the analysis in Chapter 7 as the synchrony between neurons and the effects of stimulation of the activating neural pathways are analyzed at a group level. Finally, in Chapter 8 a brief analysis at a population level is made. The estimations of synchrony are gathered across many experiments and the averages of these curves analyzed. Chapter 9 gathers the general conclusions of the thesis.

Chapter 2

Objectives and experimental setting

In this chapter we present the objectives of the study, describe the experiments and briefly present the data that will be used throughout the thesis.

2.1 Objectives

Sleep is a fundamental part of our everyday life. Although it is a death-life activity, there are still many questions that remain to be answered. One of those questions is how is the sleep-wake cycle regulated by the neuronal networks. As already mentioned, an important characteristic of the sleep state is the highly synchronized activity that can be observed in the EEG. How is the spontaneous dynamics of synchronized cortical neurons? How is that synchronization disrupted by the ascending systems? How are the temporal patterns of synchronization during the awake state? How does they evolve into the sleep state? These are some of the questions that guide one of the research projects of the Neurocom group.

The main hypothesis of that project is that it is possible to extract information on cortical functional architecture from the spontaneous behavior of neurons, a behavior obtained from their spike activity and reflected in the synchronization strength between pairs of cells. In their current work, Neurocom researchers are interested in defining the synchronization dynamics of pairs of neurons in the sleep mode and also in the awake mode, using the already described experimental model: a model in which the switch from the sleep-like to the awake-like pattern is achieved by means of electrical microstimulation on specific locations of either the brainstem (bs) or the basal

forebrain (*bf*). But there is a methodological problem that makes it difficult to define statistical significances: the scarce number of action potentials (spike activity) which is typical of the spontaneous activity.

The present work is the result of our research to find appropriate statistical tools to define, under the above mentioned experimental conditions:

- the synchronization dynamics of pairs of neurons under spontaneous activity.
- the differences in synchronization strength between the sleep-like and the awake-like periods.
- the efficacy of *bs* and *bf* in generating the transition from the sleep to the awake mode, and the relative difference in such effect of *bs* versus *bf*.
- the synchronization dynamics of pairs of neurons in the above conditions regarding their orientation selectivity.

2.2 Experiment

The data analyzed in this thesis comes from experiments performed in adult cats. All the procedures were performed by the researchers of Neurocom group and according to national and international guidelines. The animals were anesthetized and paralyzed. Ketamine (15 mg/kg) and xylazine (3 mg/kg) were used to induce the anesthesia and isofluorane (0.5–1%) in nitrus oxide (70%) and oxygen (30%) to maintain a state of deep anesthesia and a stable pattern of delta (1–5 Hz) slow oscillatory activity. Paralysis was obtained with gallamine triethiodide. The cats were artificially ventilated and their body temperature was mantained at 37–38 °C. Four cranotomies were performed. One in the primary somatosensory cortex (S1) to record the electrocorticographic activity (ECoG), another one in V1 to perform the extracellular electrophysiological recordings, and the other two for electrical stimulation of ascending pathways located in the basal forebrain and brainstem. Cells in V1 were recorded using an eight-points multielectrode. Figure 2.1 shows a sketch of the experiment.

Once a group of neurons was selected for recording, visual stimulation was used to detect each neuron's preferred orientation. Visual stimulation was performed using a monitor situated at 57 cm from the cats eyes. The

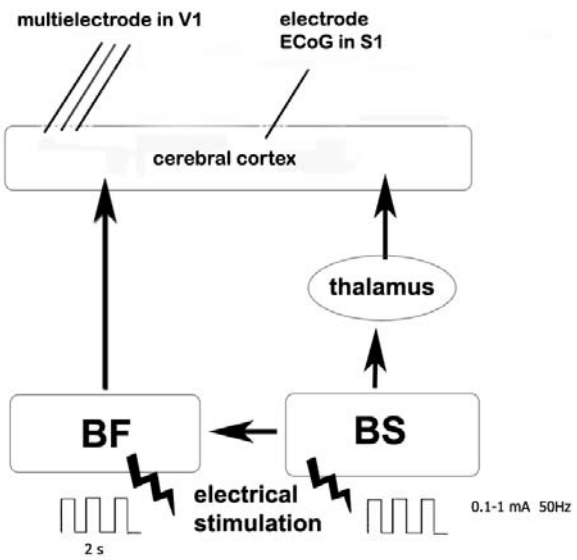

Figure 2.1: Sketch of the experimental setting described in Section 2.2.

stimuli used were oriented light bars. Figure 2.2 depicts the nature of the visual stimuli as well as the spiking activity of a real neuron provoked by such stimuli.

After the visual stimulation procedure took place, *bs* and *bf* were stimulated one at a time, three or four times each. The order of stimulation was randomly chosen. Stimuli were trains of rectangular cathodal shocks (0.05 ms, 0.1–1 mA) delivered at a frequency of 50 Hz for periods of 2 s through bipolar electrodes. Intervals of 8–10 minutes passed between stimulations to let the neurons recover. Each recorded trial lasted for approximately 600 s long, with the stimulus presented after around 120 s. Figure 2.3 shows the location of the brainstem and basal forebrain in a cat's brain as well as the approximate locationof the recording and stimulating electrodes.

2.3 Data

The data resulting from the experiment are spike trains. As already stated, these spike trains are temporal sequences of neuronal action potentials, which

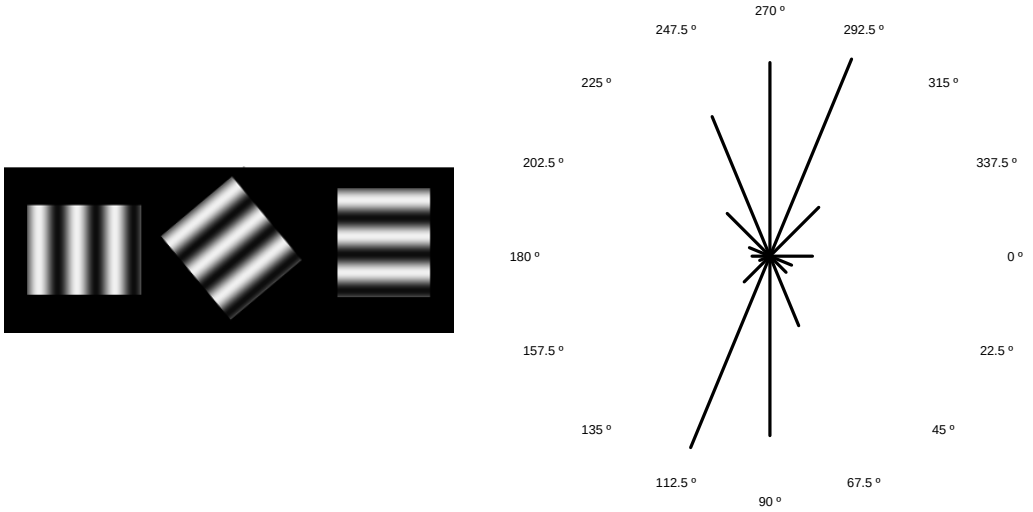

Figure 2.2: Light bars presented as visual stimuli to determine the favorite orientation of the neurons (left panel). Spiking activity of a real neuron provoked by the oriented light bars (right panel).

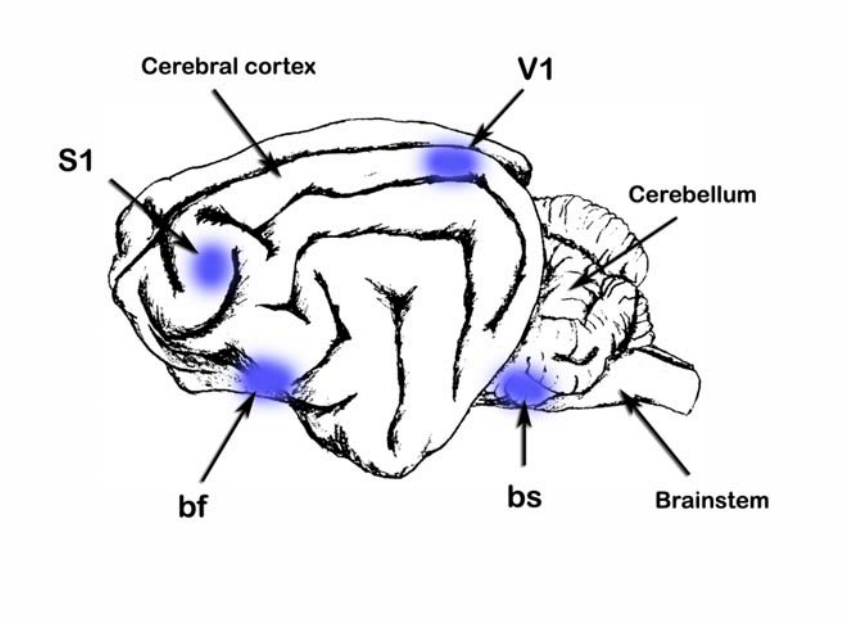

Figure 2.3: Approximate location of recording (S1, V1) and stimulating (*bs*, *bf*) electrodes in the cat brain.

can be easily visualized in Figure 1.2. In the following four chapters we will apply different statistical methods to a group of seven simultaneously

recorded neurons. We will denote these neurons with the names N1, N3a, N3b, N4a, N4b, N5 and N7 (as in the Neurocom database). As already mentioned, the neurons were recorded with an eight-points multielectrode. A given electrode can record, none, one or more than one neuron in a given trial. This is because the points of the electrode can be situated close to none, one or more that one neuron respectively. This is the reason for the names of the neurons we will work with. In this case, the first electrode recorded one neuron, the second electrode recorded none and the third electrode recorded two neurons at the same time, and so on. The activity of the two neurons recorded by the third electrode was sorted off-line by the researchers. Originally, the data is stored in a large data base containing all the trials of every recording of several experiments. As a first step we organized the data in smaller data frames containing the information for each recording. Also the data was aligned at the stimulation time. Table 2.1 shows the first eight spike times from the beginning of the recording of trial one of the previously discussed group of neurons. It also shows the eight first spike times after the onset of the *bs* stimulus (120 s).

In Chapter 3 we will discuss some of the characteristics of spike trains, such as their firing rates: frequencies at which the neurons fire. In the meanwhile, Figure 2.4 shows the raster plots and corresponding firing rates, estimated using kernel smoothers, forr 160 s of three spike trains. They belong to three neurons of the aforementioned group, namely, N1, N3a and N3b.

In Chapter 7 another group of neurons is used. The group chosen to exemplify the use of the methodology described is a group of 8 simultaneously recorded neurons. Also, four trials for each stimuli were performed in that case.

In Chapter 8 the recordings of nine experiments like the one described in Section 2.2 were combined together for the analyses.

2.4 Software

The `R` language was used for the statistical analyses. Although the major part of the functions and algorithms were programmed by ourselves, several `R` packages were often used. Those packages are: `stats`, `MASS`, `mgcv`, `fda`, `fda.usc`, `KernSmooth`.

Table 2.1: Raw spike times of one trial of a group of simultaneously recorded neurons.

N1	N3a	N3b	N4a	N4b	N5	N7
0.987150	1.797775	2.117425	1.319200	1.174950	1.091600	1.033425
1.030650	1.812725	3.171450	1.322850	1.181975	1.096325	1.818850
1.041150	1.992225	3.226000	1.325175	1.183875	3.169600	1.823025
1.065000	2.121600	3.232725	1.327375	1.185800	3.175000	2.068100
1.206925	2.474200	4.204950	1.329775	1.188300	3.183150	2.070325
1.402850	2.510000	4.208350	1.332400	1.605100	3.307175	2.111125
1.517475	3.230400	4.212950	1.954500	1.611050	3.309450	2.112825
1.592500	3.588325	5.809900	2.028775	1.612725	3.314000	2.115100
⋮	⋮	⋮	⋮	⋮	⋮	⋮
121.4606	121.4458	121.7348	120.1265	121.7576	121.4010	122.9900
122.1192	121.4866	121.9913	120.1305	121.8439	121.9499	123.0143
122.3384	121.7293	122.1219	120.1348	121.8606	122.1055	123.0178
122.5217	121.7900	122.2710	120.1369	122.0996	122.1103	123.0376
122.5441	122.0254	122.4586	120.1390	122.1465	122.1126	123.0396
122.5695	122.3074	122.4820	120.1415	122.1988	122.5300	123.4560
122.8900	122.3501	122.4841	121.4034	122.2008	122.5333	124.6872
122.9102	122.5324	122.4882	121.4182	122.2595	122.5352	124.6905

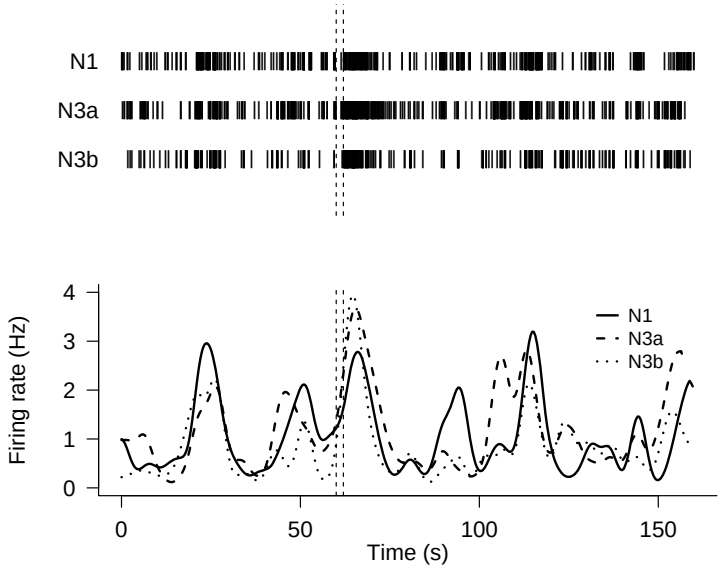

Figure 2.4: Raster plots of one trial (upper panel) of three simultaneously recorded neurons during 160 s of spontaneous activity. Firing rates profiles of the same three neurons (bottom panel) estimated using kernel smoothing. Electrical stimulation was applied at 60 s.

Chapter 3

Analysis of single spike trains

As a first approach to the study of spike trains, in this chapter we present a mathematical representation for them and we discuss some of their properties. First we consider the firing rates, estimate them using the kernel method and discuss how the selection of the smoothing parameter affects these estimations. Second, we present the *inter-spike intervals* (ISIs), which are the periods of time that elapse from one action potential to the following. We use kernel estimators to describe the density function of the logarithm of ISI and investigate its stationarity. Finally, we present autocorrelation measures for spike trains and use them to propose a bootstrap-based hypothesis test for independence of the ISIs. Part of the resutls of this chapter can be found in González-Montoro et a. (2011).

3.1 Point processes and rate functions

Spike trains can be described by means of point processes. A *point process* is a stochastic process that consists of a set of discrete events occurring in continuous time. A neural spike train is completely described by point process where the events are the time points, $0 < X_1 < X_2 < \ldots$, when the spikes occur. Apart from the event times, there are other ways to represent a point process and, in particular, a spike train. Let $S_1, S_2, \ldots$ be the set of random variables that describe the possible *waiting times* between consecutive occurrences of the previous point process. We will refer to these variables as *inter-spike intervals*, ISI. A third possible way to describe the spike trains is by the *counting process*, $N(t)$, defined as the number of events that occur in the time interval $(0, t]$, this is, $N(t) = \#\{X_i \leq t, i = 1, 2, \ldots\}$. Also, we can denote the amount of events in an arbitrary time interval $(t_1, t_2]$ as

$\Delta N_{(t_1,t_2)} = N(t_2) - N(t_1)$.

The three ways of characterizing a spike train discussed above are equivalent. All of them fully specify the point process and, any of them, completely determine the other two. The ISI can be computed as the difference between consecutive spike times, $S_1 = X_2 - X_1, S_2 = X_3 - X_2, \ldots$, and, knowing the ISI we can compute the spike times using the cumulative sum $X_n = \sum_{i=1}^{n} S_i$. On the other hand, $X_j = u$ if and only if $N(u) = j$ and $N(t) < j$ for all $t < u$. Both, the spike times and the inter-spike intervals, are variables with a discrete index which take values in $\mathbb{R}$, while the counting process takes integer values but its index is continuous, indicating a point in time. These three forms of representing the same point process could be useful to solve different problems and we will use the three of them throughout the text.

When working with real data, the point processes are observed in a time interval and, therefore, only a finite set of variables can be observed. Given an observational interval $(0, T]$, we will work with single realizations of the point process consisting on the observed spiking times in that interval. We will denote these observed realizations with lower case letters: $X_1 = x_1, X_2 = x_2, \ldots, X_J = x_J$. In a similar way, we will use $S_1 = s_1 = x_2 - x_1, \ldots, S_{J-1} = s_{J-1} = x_J - x_{J-1}$ for the ISI.

It is often found in the neuroscience literature that Poisson processes are used to model spike trains. Formally, a homogeneous Poisson process with rate λ is a point process, $N(t)$, which satisfies, a) $N(0) = 0$, b) for any interval $(t, t + \Delta t)$, $\Delta N_{(t,t+\Delta t)} \approx Poisson(\lambda \Delta t)$ and c) $N(t)$has independent increments, this is, for any two non-overlapping intervals, $(t, t + \Delta t]$ and $(s, s + \Delta s]$, $\Delta N_{(t,t+\Delta t)}$ and $\Delta N_{(s,s+\Delta s)}$ are independent random variables. A major drawback of using Poisson processes to model neuronal data is the fact the these processes assume no dependence on spiking history. Consider S_i, the inter spike interval between the $(i-1)$-th and i-th spike. The event $S_i > t$ occurs if and only if $\Delta N_{(X_{i-1},X_{i-1}+t)} = 0$, and therefore, $P(S_i > t) = P(\Delta N_{(X_{i-1},X_{i-1}+t)} = 0) = \exp(-\lambda t)$, by the definition of Poisson process. This means that the ISI follow a exponential distribution with mean $E(S_i) = \lambda^{-1}$, as $F_{S_i}(t) = 1 - \exp(-\lambda t)$.

In our particular problem, after the application of the stimulus to switch from the sleep-like to the awake-like mode, it is clear that the spike train can not be modeled with a homogeneous point process because of the non-stationarity provoked by the appearance and disappearance of the stimulus. But, what about the period before the stimulus? We performed a

Kolmogorov-Smirnov goodness of fit test for several trials of all the neurons obtaining p-values < 0.001 in every case. Figure 3.1 shows, a Q-Q plot of the ISIs of the recording before stimulation of one trial of neuron N1. The plot exhibits how the distribution of the ISIs diverges from the exponential distribution.

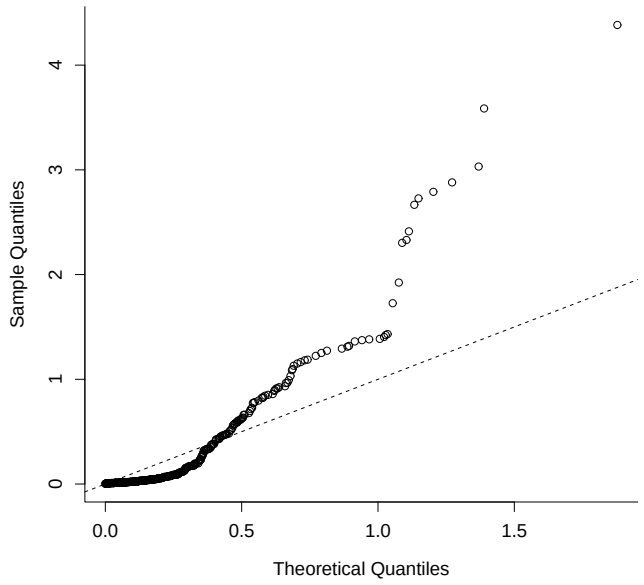

Figure 3.1: Q-Q plot for one trial of N1.

So far, we have considered processes that are stationary in time. This assumption is not realistic when working with spike trains in general and moreover when a sudden change (as the one induced by the stimulation in *bs* or *bf*) is being held. The *rate function* or *intensity function* can be defined as the instantaneous probability of observing an event at every time point per unit of time, this is:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(N(t + \Delta t) - N(t))}{\Delta t}$$

The rate function of spike trains has been widely studied since it is one of the features in which information can be encoded. The rate function of a point process is the instantaneous probability of finding an event at time t per unit of time and, of course, this probability has to be estimated. There exist several ways to estimate this quantity as, for example, the *global count*

rate defined as the total amount of action potentials divided by the length of the recording interval of time, $r_0 = \frac{J}{T}$, as shown in Figures 1.2 and 2.4, by the raster plots, and Table 3.1. This measure has no information whatsoever about the dynamics of the firing rate over time so, more sophisticated methods are needed.

A time varying firing rate can be defined to estimate the rate function using different type of estimators. For example a *histogram-like* estimator could be defined as follows. Given $t_0 = 0$ and a bin width h_0, let us define the firing rate as constant in the intervals $\{(t_0 + h_0 m, t_0 + h_0(m+1)] : m \in \mathbb{N}\}$. This is, in each interval $[t_m, t_{m+1})$ (where $t_m = t_0 + h_0 m$) define, for $t \in [t_m, t_{m+1})$,

$$\tilde{r}_{h_0}(t) = \frac{1}{h_0} \sum_{j=1}^{J} \mathbb{I}\{t_m < X_j < t_{m+1}\}. \tag{3.1}$$

Figure 3.2 shows the histogram firing rate estimations for N1 using three different bin widths.

Kernel estimators are also useful to estimate rate functions. Given a window length, h, and a kernel function, K, such that $\int_{-\infty}^{\infty} K(t)\mathrm{dt} = 1$ and generally non-negative, we can estimate the rate function as

$$\hat{r}_h(t) = \sum_{i=1}^{J} \frac{1}{h} K\left(\frac{t - X_i}{h}\right)$$

In the case of having several trials, say N, of the same neuron, some resolution can be gained using the mean of the firing rates. If this is the case, the firing rate would be defined as:

$$\hat{r}_{h,N}(t) = \frac{1}{N} \sum_{k=1}^{N} \hat{r}_h^{(k)}(t) = \frac{1}{N} \sum_{k=1}^{N} \sum_{j=1}^{J_k} \frac{1}{h} K\left(\frac{t - X_i^{(k)}}{h}\right) \tag{3.2}$$

where $\hat{r}_h^{(k)}(t)$ is the kernel estimator of the rate function, $X_i^{(k)}$ are the firing times and J_k are the amount of spikes of the k-th trial.

An open and interesting discussion is how to choose the smoothing parameter h. Nawrot et al. (1999) studied the influence of the shape of the chosen kernel function and the width of the smoothing parameter in the firing rate estimation. Figure 3.3 shows kernel firing rate estimations for several smoothing windows.

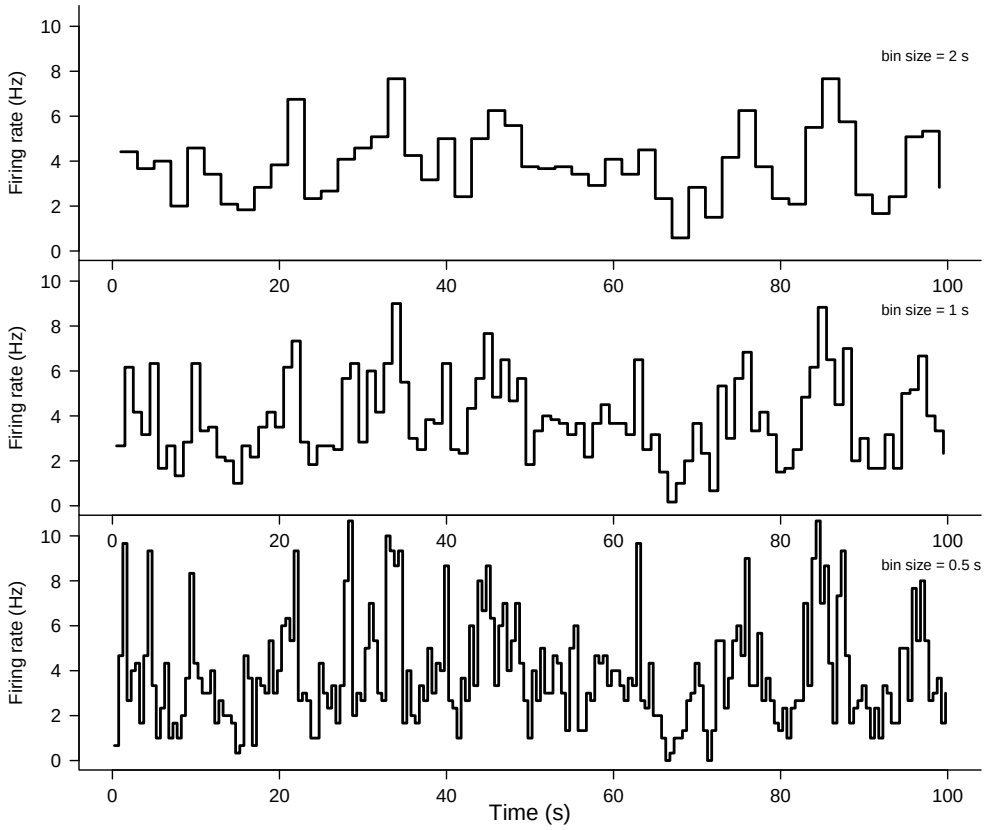

Figure 3.2: Firing rate for neuron N1 averaged over three trials estimated with the histogram-like method defined in (3.1). using three different choices of bin size.

3.2 Inter-spike intervals

In this section we will investigate the nature of the ISIs in more detail. At this point, it is important to note that the time to recover from a stimulus varies from neuron to neuron and the recovery is not sudden but a continuous process. Despite of this fact, and just for the present chapter, we use a partition of the time in three stages: *pre* (from the beginning of the recording to the *bs*/*bf* stimulus), *post* (20 s, starting at the end of the stimulus) and *final* (the rest). The choice of 20 s for the *post* part was made based on previous work (Mariño and Cudeiro (2003)) and was intended to capture the period in which the cerebral cortex is dominated by an awake-like pattern of activity. Table 3.1 shows the global firing rates for each neuron in each of the stages.

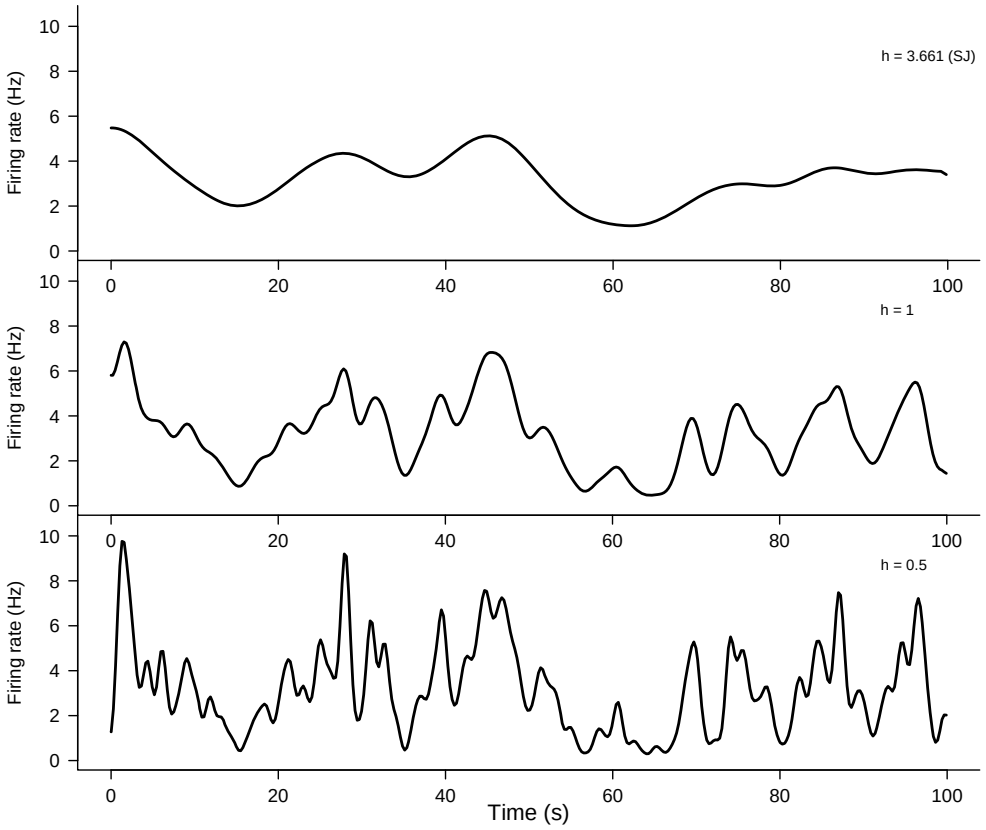

Figure 3.3: Firing rate for three trials of neuron N1 estimated with the kernel estimator defined in (3.1) using three different choices of smoothing parameter.

Figure 3.4 shows an example of the distribution of the ISIs for one trial of the *pre* and *post* part of N1. It can be observed that the distribution of this variable is highly concentrated for small values but has a very heavy tail to the right. In this particular example, before stimulation, more than 300 of the 418 ISIs are smaller than 0.25 s but there also exist some intervals up to 4 s long without any firing. After the stimulation, the large ISIs dissapear almost completely for a period of time to reappear gradually later on. Consequently, due to the nature of the data, we will work with the natural logarithm of the ISIs to make results more easy to interpret.

We estimate the density functions of the logarithms of the ISIs and show some examples to study stationarity. For the estimation of the densities we use kernel estimators and, as before, we use a Gaussian kernel function. The

Table 3.1: Average number of spikes per second in each stage.

	N1		N3a		N3b		N4a	
	bs	*bf*	*bs*	*bf*	*bs*	*bf*	*bs*	*bf*
pre	3.459	4.534	3.319	2.932	2.815	3.617	5.036	6.591
post	3.950	3.800	5.900	2.450	4.700	2.900	21.450	4.250
final	3.844	4.081	3.136	3.867	2.324	3.964	9.506	4.419

	N4b		N5		N7	
	bs	*bf*	*bs*	*bf*	*bs*	*bf*
pre	4.260	4.443	2.237	2.122	1.271	1.453
post	5.600	4.950	3.750	2.900	1.150	0.050
final	3.39	4.083	2.336	2.404	1.386	0.731

smoothing parameter is chosen by the automatic plug-in window selection described by Sheather and Jones (1991). These estimates are shown in Figures 3.5 to 3.8, where the left and right panels correspond to trials of the *bs* and *bf* stimulations respectively.

An important fact that arises is that each neuron has a particular density; some of them present two or three modes in the *pre* stage. In general, all densities present a mode in very small values and another one around zero, which corresponds to $e^0 \approx 1\,\mathrm{s}$. Also, it can be observed that the densities change considerably in the *post* stage to recover in the *final* stage, in most of the cases, a shape very similar to the original. This difference and recuperation is very noticeable for neuron N1 in Figure 3.5 and neuron N3b (Fig. 3.7) for the *bf* stimulation. This is also the case for Neuron N4b (Fig. 3.8) in the *bs* panel although, the *bf* stimulation affects very little the density of the ISIs, where the only considerable change can be observed in a reduction of the principal mode. For neuron N3a (Fig. 3.6) the original density does not seem to be recovered in the *final* stage, specially for the *bf* stimulation. An interesting point is that, although the computations have been made using the three trials of each condition altogether, these characteristics can be observed for each of the isolated trials, as shown in Figure 3.9.

In order to be able to asses that the distribution of the ISIs is recovered after the effect of the stimulus, we need to know whether the density during the sleep-like *pre* period is stationary. To do this, we use sliding windows in the *pre* part to estimate the density of the ISIs in each window. Figure 3.10

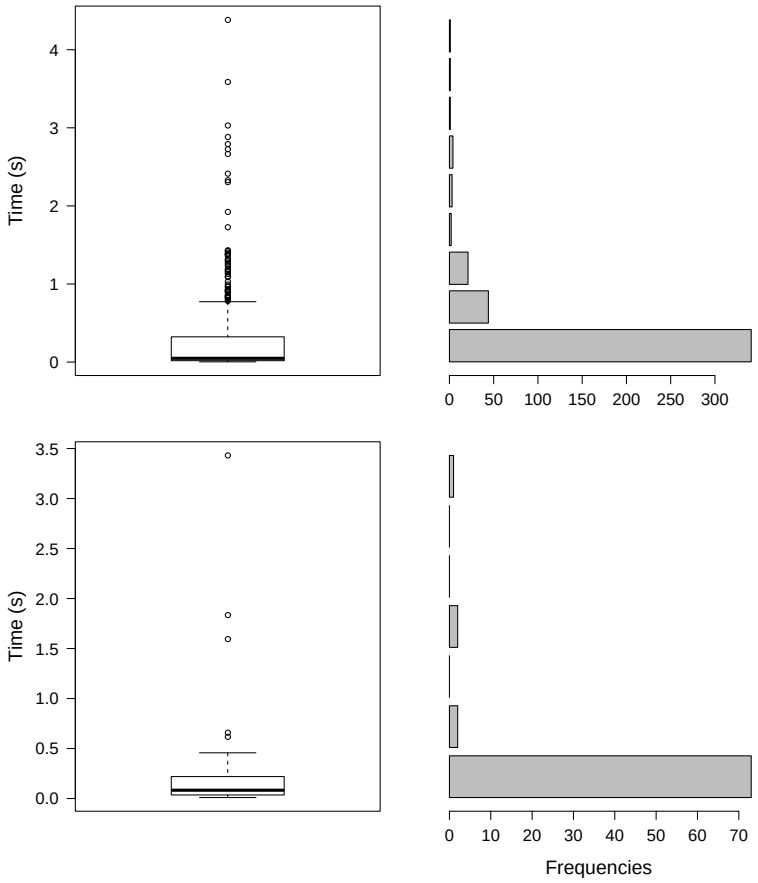

Figure 3.4: Boxplot and frequencies of the ISIs of one trial of the *pre* (top) and *post* (bottom) stages of N1.

shows this analysis for the first recorded trial of neuron N1. We have used 36 s time windows and have slided it every 18 s, which means that every window overlaps in half of its width with the previous one. It can be observed that these densities are very similar. They all have the same two modes and are essentially the same height. These densities do not change across trials. There are other neurons in which the stationarity is not that clear.

Figure 3.11 shows the same type of analysis but for the period of time that starts right after the stimulation. In this cases we have not separated the *post* period from the *final* because in 20 s there are not enough spikes to perform the analysis. Nevertheless, we have used 24 s windows and, therefore, the first panel in the upper left position shows the density for the 24 s right after stimulation, and hence contains the *post* period. The windows move every

12 s and it can be seen how the two modes and height are recovered. Some densities show differences with the ones in the *pre* part, but this fact can also be due to the reduction in the amount of data used to estimate these densities.

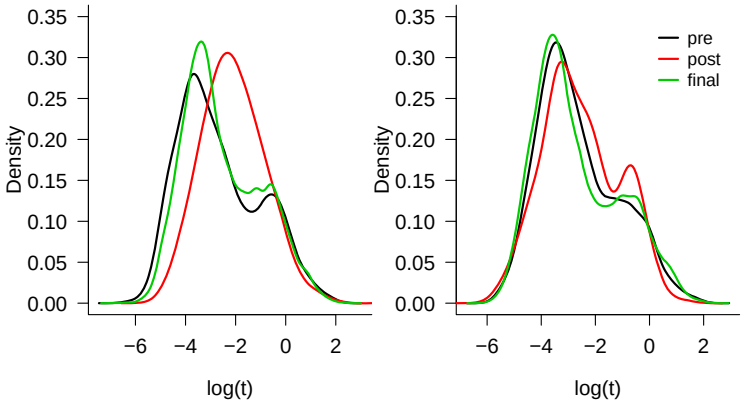

Figure 3.5: Density estimation of the log ISIs of the three stages of N1. Left panel: *bs* stimulation. Right panel: *bf* stimulation.

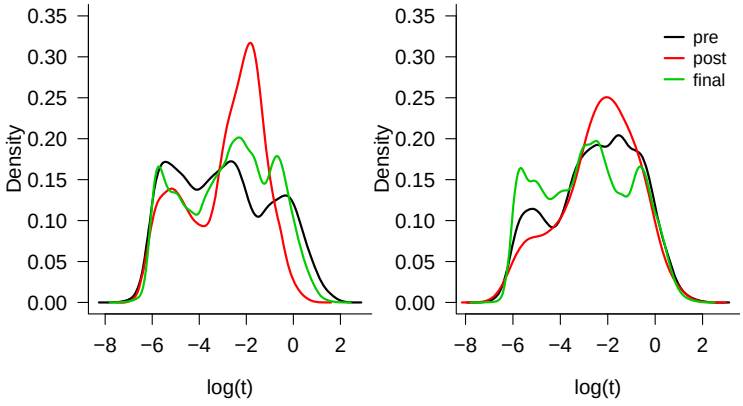

Figure 3.6: Density estimation of the log ISIs of the three stages of N3a. Left panel: *bs* stimulation. Right panel: *bf* stimulation.

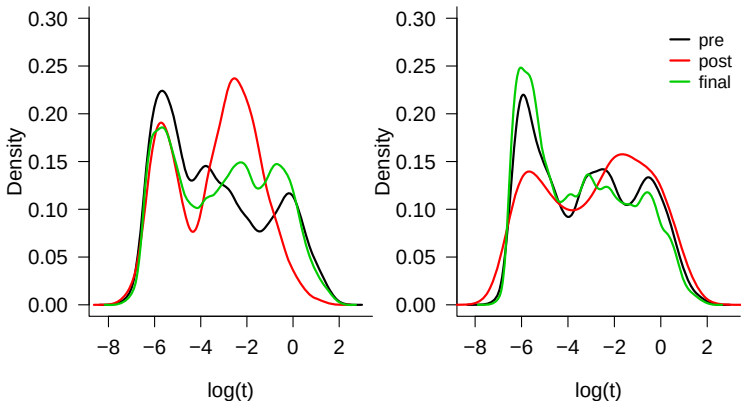

Figure 3.7: Density estimation of the log ISIs of the three stages of N3b. Left panel: *bs* stimulation. Right panel: *bf* stimulation.

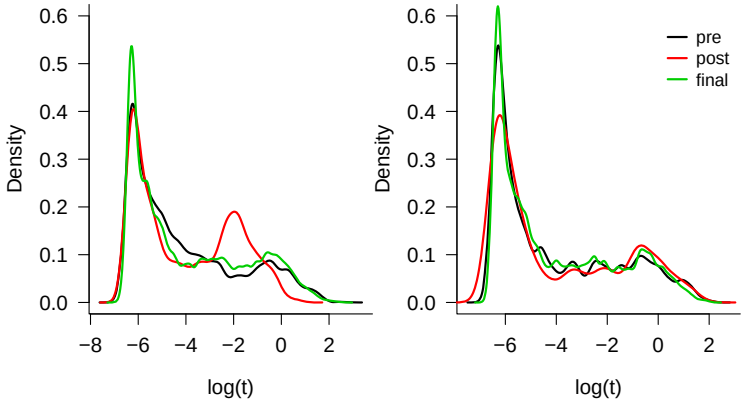

Figure 3.8: Density estimation of the log ISIs of the three stages of N4b. Left panel: *bs* stimulation. Right panel: *bf* stimulation.

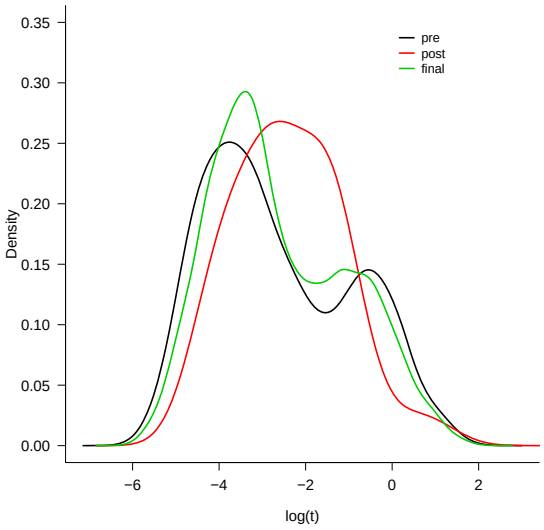

Figure 3.9: Density estimation of the log ISIs of one trial (*bs* stimulation) of neuron N1.

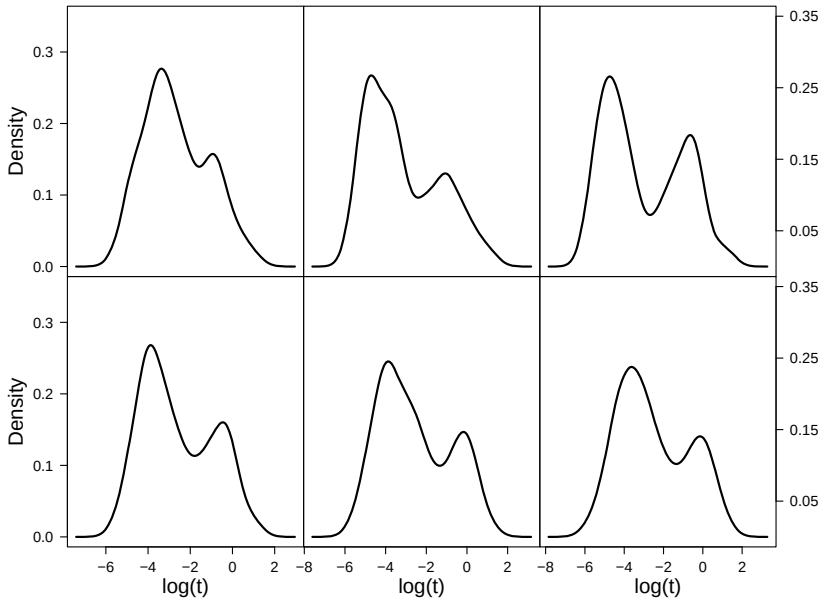

Figure 3.10: Estimated density functions of the log ISIs of the *pre* period for one trial of *bs* of N1 estimated in 36 s windows slided every 18 s. Time course from left to right and up to bottom.

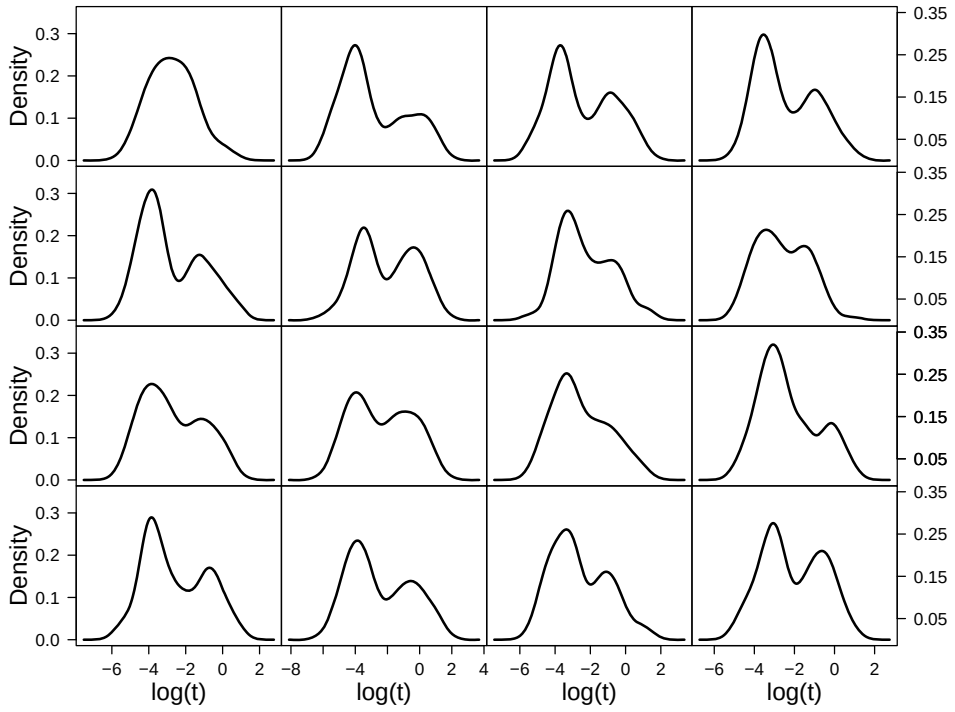

Figure 3.11: Density functions of the log ISIs for 200 s after stimulation for the first trial (*bs* stimulation) of neuron N1 estimated in 24 s windows slided every 12 s. Time course from left to right and up to bottom. The first panel includes the *post* period.

3.3 Autocorrelation measures

In the previous section we defined the inter-spike intervals (ISI) as the time elapsed between consecutive spikes. In this section we introduce autocorrelation measures for spike trains and estimate them for the real data.

Given the ISIs of an observed spike train, $S_1, \ldots, S_n$, we can estimate the serial autocovariance function as

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{i=1}^{n-h}(S_{i+h} - \bar{S})(S_i - \bar{S}), \qquad 0 \leq h < n,$$

where $\bar{S} = \frac{1}{n}\sum_{i=1}^{n} S_i$ is the sample mean. Then, the serial autocorrelation

function is estimated by

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, \qquad 0 \leq h < n \; .$$

Although this is an interesting measure and the most used in the time series literature, it is more common in neuroscience to study a higher order autocorrelation measure between spikes which will be denoted as *higher order inter-spike autocorrelation*,HOISA, (Perkel et al. (1967a)). The term *higher order* results from the fact that distances from any two spikes are taken into account, not only from consecutive spikes, as it will be made clear below. Before introducing the HOISA itself, as it is presented in the neuroscience literature, we will discuss on a very similar measure which can be derived as the serial autocorrelation function of a different time series.

Let us split the total recording time, T, in $Q = \left[\frac{T}{q}\right] + 1$ intervals of length q. Let A_i be the i-th interval, $A_i = [(i-1)q, iq)$. Here it is convenient to note that if only one spike is intended to fall in each interval, q must be sufficiently small. A typical value is $q = 1\,\text{ms}$ since it is known that, due to the refractory period, a neuron needs approximately that time to recover before firing again. Let us define the new series $\{V_i\}_{i=1}^{Q}$:

$$V_i = \sum_{j=1}^{n+1} \mathbb{I}\{X_j \in A_i\}. \tag{3.3}$$

We can estimate its autocovariance function by

$$\hat{\gamma}_V(h) = \frac{1}{Q}\sum_{i=1}^{Q-h}(V_{i+h} - \bar{V})(V_i - \bar{V}), \quad \bar{V} = \frac{1}{Q}\sum_{i=1}^{Q} V_i \; .$$

Now, if $h << Q$ the following approximations, $Q-h \approx Q$ and $\sum_{i=1}^{Q-h} V_{i+h} \approx \sum_{i=1}^{Q} V_i \approx \sum_{i=1}^{Q-h} V_i$ can be used to obtain, after some algebra,

$$\hat{\gamma}_V(h) \approx \frac{1}{Q}\sum_{i=1}^{Q-h} V_{i+h}V_i - \bar{V}^2 \; . \tag{3.4}$$

Since $\bar{V}$ does not depend on h, we can concentrate in $\hat{\gamma}^*_V(h) = \hat{\gamma}_V(h) + \bar{V}^2 = \frac{1}{Q}\sum_{i=1}^{Q-h} V_{i+h}V_i$ without modifying the shape of the function $\hat{\gamma}_V(h)$. Now,

$$V_{i+h}V_i = \sum_{j=1}^{n+1}\sum_{l=1}^{n+1} \mathbb{I}\{X_j \in A_i, X_l \in A_{i+h}\}$$

and then, $\hat{\gamma}_V^*(h)$, as a function of h, is the histogram of these frequencies, though divided by Q.

Observe that in the case of a small enough q (q=1 ms for example), $V_{i+h}V_i = 0$, except when $V_{i+h} = V_i = 1$, in which case the product is 1. This is why, for each h:

$$\hat{\gamma}_V^*(h) \simeq \#\{(V_{i+h}, V_i) : (V_{i+h}, V_i) = (1,1)\}/Q \ ,$$

which is easy to think in terms of time: $V_{i+h} = V_i = 1$ means that there are two spikes that are separated by a distance of, at least $h-1$ and as much as $h+1$. From this follows that $Q\hat{\gamma}_V^*(h)$ counts the number of spike pairs (although some might be missing due to the discretization) that are separated by a distance between $h-1$ and $h+1$. This is the main idea to define the autocorrelation as it follows.

Autocorrelation, as it is used in neuroscience, is very similar to $\hat{\gamma}_V^*(h)$ but it is built in an alternative way. Actually, it is defined as the histogram of relative frequencies (or, sometimes, absolute ones) of the elapsed time between any two spikes of a train that do not surpass a certain $w_{\max}$ chosen by the researcher. This $w_{\max}$ is usually much smaller than T, which allows us to compare with the serial covariance function of $\{V_i\}_{i=1}^Q$, since the approximations in (3.4) are valid.

Given a spike train $\{X_i\}_{i=1}^{n+1}$, let the set of distances between any two spikes be $\{D_m\}_{m=1}^M = \{X_i - X_j/i, j \in \{1,\dots,n+1\}, i \neq j\}$, such that $-w_{\max} \leq D_m \leq w_{\max}$. Moreover, we need to choose b, where $2b$ is the width of the histogram's intervals. In this context, we define the higher order interspike autocorrelation (HOISA) of a spike train at the distance d, by:

$$\hat{g}(d) = \frac{1}{M}\sum_{m=1}^{M} \mathbb{I}\{d-b \leq D_m \leq d+b\}. \tag{3.5}$$

Here b plays a similar role to q in (3.4) and, in fact, this histogram is very similar to that obtained from the serial autocovariance of the series $\{V_i\}_{i=1}^Q$. Some differences might arise from discretization and normalization. Interestingly, for $\hat{\gamma}_V^*(h)$ the discretization is done before calculating the distances, while in the definition of $\hat{g}(d)$ the discretization is carried out when constructing the histogram. On the other hand, to obtain $\hat{\gamma}_V^*(h)$, the absolute frequencies are divided by Q while for $\hat{g}(d)$ the denominator is M. Then the results should be almost proportional. Figure 3.12 shows these two types

of histograms for three lengths of q and b, $q = b = 0.01, 0.1$ and 1 s for the activity in the *pre* period of one trial of neuron N1. The functions $\hat{\gamma}_V^*(h)$ and $\hat{g}(d)$ have been multiplied by Q and M respectively, so that the similarities become more visible. The histograms are practically the same though there are some differences for the three sizes. The differences grow with the size of q and b and are more evident when $q = b = 1$.

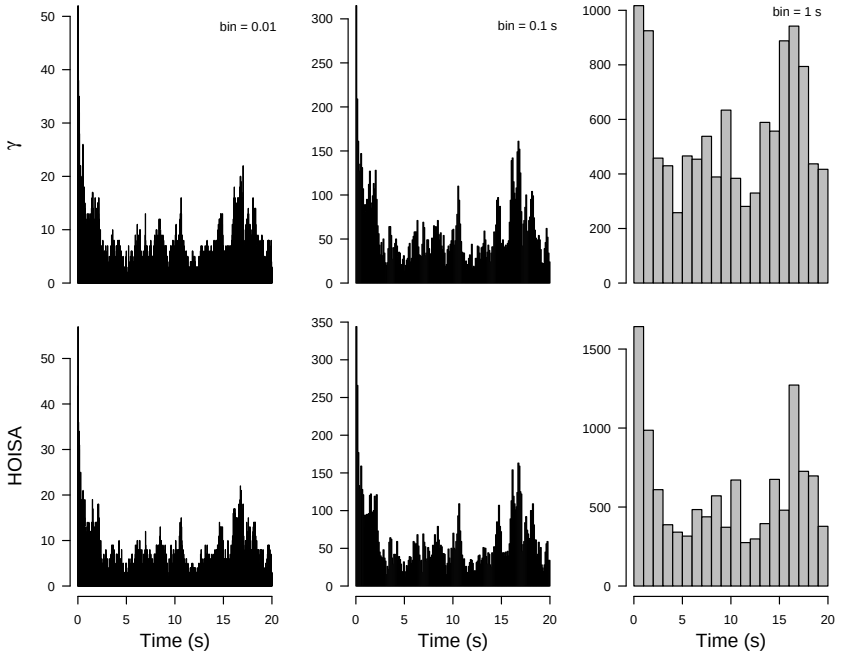

Figure 3.12: Comparison between $Q\gamma_V^*(h)$ (top panel) and $M\hat{g}(d)$ (bottom panel) for three bin sizes, $q = b = 0.01, 0.1, 1$. Data corresponds to one trial of neurons N1.

Note that, in fact, the HOISA is just an estimate of the probability density of time between any two spikes. To get a smoother estimation, a nonparametric kernel estimator is considered:

$$\tilde{g}(d) = \frac{1}{Mh} \sum_{m=1}^{M} K\left(\frac{d - D_m}{h}\right).$$

We have used the Gaussian kernel function, K, and the Sheather and Jones method for bandwidth selection (Seather and Jones (1991)).

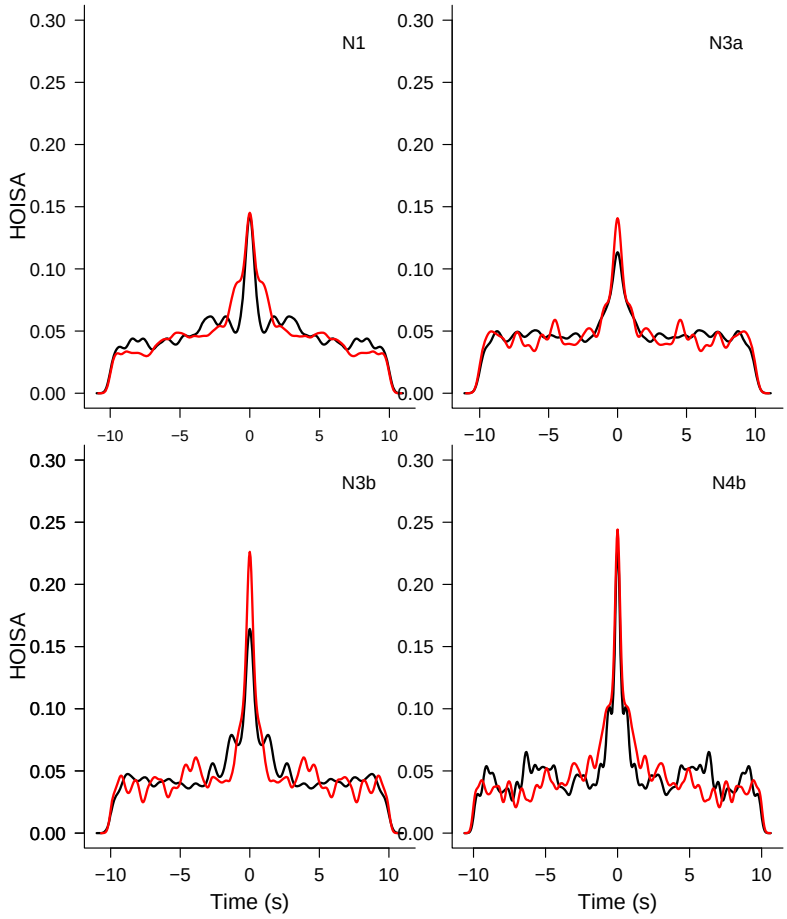

Figure 3.13: Higher order interspike autocorrelation estimated via the kernel method for one trial of neurons N1, N3a, N3b and N4b in the *pre* period. Black lines correspond to trials in which the stimulus was applied in *bs* and red lines to the ones in which *bf* was stimulated.

Figure 3.13 shows the HOISA in the *pre* stage of four neurons, namely, N1, N3a, N3b and N4b. It can be observed that there is a high probability of two spikes occurring very close in time as the high central peaks indicate. Different density estimations are shown for trials that correspond to different stimulation areas to show that, as expected, there are not significant differences between them.

Figure 3.14 shows the HOISA functions in the *post* stage for the same

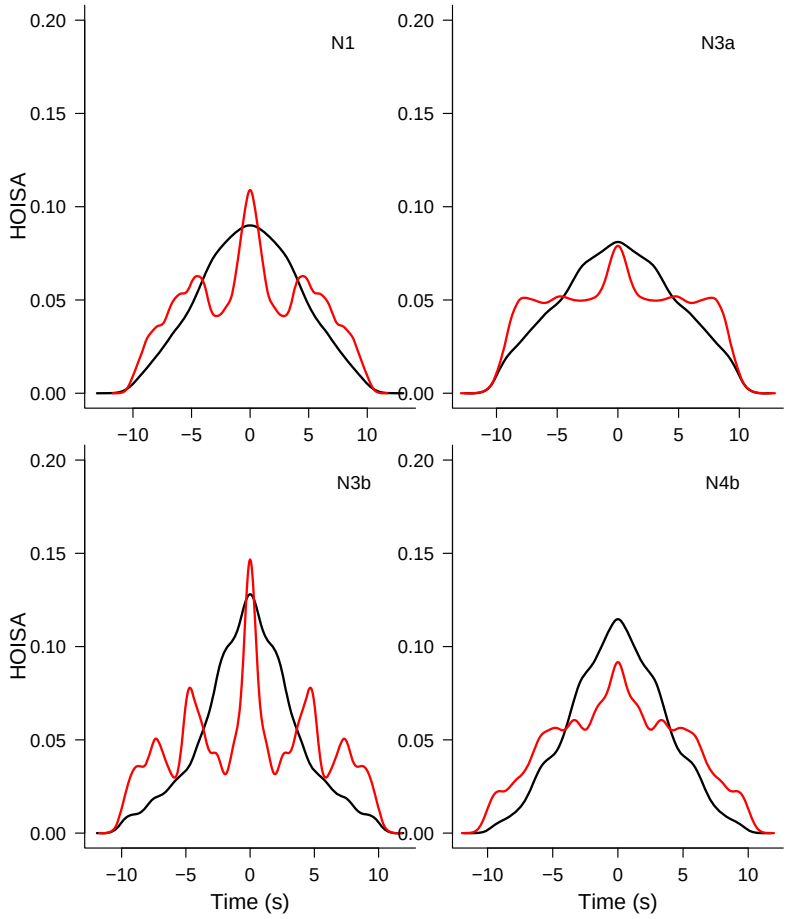

Figure 3.14: Higher order interspike autocorrelation estimated via the kernel method for one trial of neurons N1, N3a, N3b and N4b in the *post* period. Black lines correspond to trials in which the stimulus was applied in *bs* and red lines to the ones in which *bf* was stimulated.

four neurons as above. The differences between the autocorrelation functions are mainly found in their dispersion. For this period of the recorded trials, most of the histograms are unimodal, but there are some trials in which conspicuous secondary peaks can be observed; these are supposed to reflect stimulation-induced oscillations. In the *post* period it makes sense to compare the estimates obtained for each of the two stimuli. In several neurons, autocorrelations for the stimulus *bs* present more dispersion than those for the *bf* stimulus.

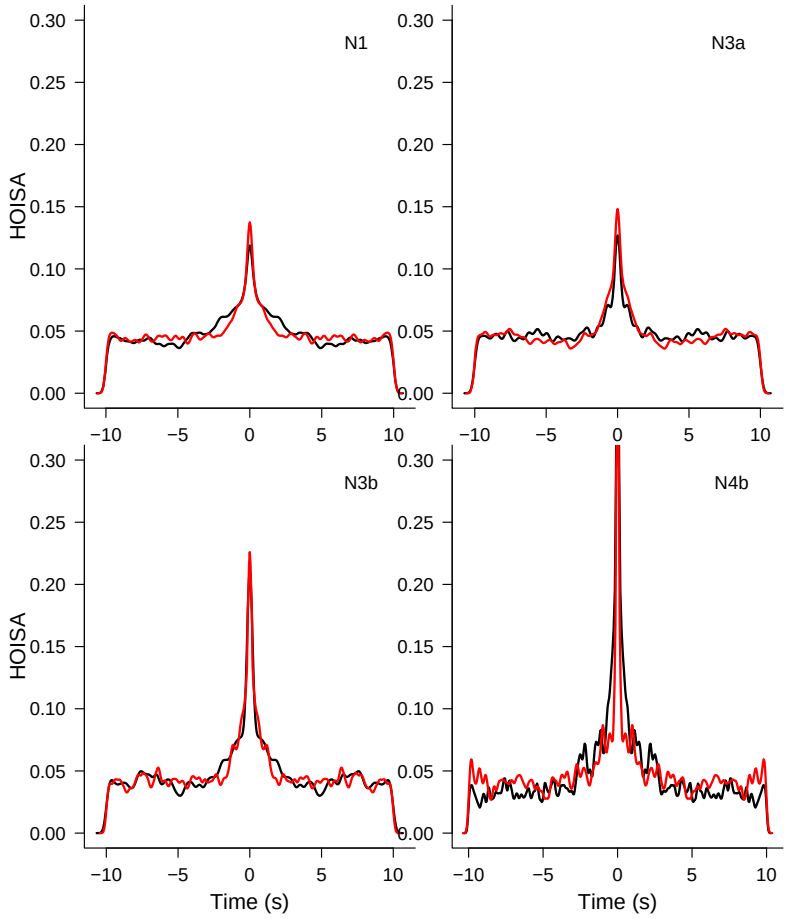

Figure 3.15: Higher order interspike autocorrelation estimated via the kernel method for one trial of neurons N1, N3a, N3b and N4b in the *final* period. Black lines correspond to trials in which the stimulus was applied in *bs* and red lines to the ones in which *bf* was stimulated.

The estimates of the autocorrelation function for the *final* stage of the study can be observed in Figure 3.15 and they are very similar to the corresponding ones of the *pre* condition. The main peak of the probability density remains at zero. There are also other peaks as in *pre*. In the *post* period, distances between spikes were mainly small but when the awake-like pattern is over, the distances return to the behavior they had before the stimulus was applied.

As already mentioned, many of the plots in Figures 3.13, 3.14 and 3.15 exhibit secondary peaks. These secondary peaks are commonly studied in neuroscience as an easy way to detect oscillatory activity. As described in the first chapter, under sleep states or, as in this case, anesthesia, most cortical neurons display an oscillatory activity. Some rhythms have been characterized neurophysiologically in cats, as the slow rhythm (< 1 Hz), the delta rhythm (1–4 Hz) and the spindle oscillation (7–14 Hz). These rhythms are designated as *slow sleep oscillations*. On the other hand, neuronal spike responses are grouped into what are called *bursts*. These features are sequences of action potentials fulfilling certain characteristics, including: a) consecutive spikes within a burst are not separated one from another in more than certain time, and b) between one burst and another there is, at least, a certain time. Time values in this definition may be changed for different areas of the brain depending on the experimental protocols. If there is an oscillation, for example a delta oscillation of 2 Hz, what happens is that after a neuron generates a burst, it is quite likely that the next burst will occur after about 500 ms. In anesthetized cats it is common to record oscillations of about 0.1 Hz (belonging to the so termed *slow rhythm*). Thus, a slow oscillatory activity of about 0.1 Hz could be the cause of the peaks at 10 s. If larger values of $w_{\max}$ were chosen, peaks at around 20 and 30 s could be observed in the in the HOISA. Also, the peaks between 250 ms and 1 s (not clearly observed in the previous figures because of the time used) are indicative of the delta oscillation.

3.4 Testing independence for inter-spike intervals

In this section we will study the existence of dependence among the elapsed times between consecutive spikes. In statistical terms, the question may be stated as a hypothesis test:

- H_0 : the ISIs are independent,
- H_1 : there exists dependence among the ISIs.

Two different tests are proposed. If the ISIs are dependent, this situation will influence the shape of HOISA defined in the previous section. These functions will be used to build the first test. The estimated autocorrelation function for the original train will be compared with another one obtained

from independent spike trains. On the other hand, the Kolmogorov-Smirnov goodness-of-fit test will be used to compare the distribution of the elapsed times between spikes in the original train with the distribution of the times of independent trains.

To obtain the sample of independent ISIs, a random shuffle is performed in the original ISIs. A new iid resample $\{S_i^*\}_{i=1}^n$ is obtained from $\{S_i\}_{i=1}^n$, destroying all the possible order dependence but preserving any other possible features. With this new sample a new spike train is built: $X_1^* = 0$ y $X_i^* = \sum_{j=1}^{i-1} S_j^*$, $i = 2, \ldots, n+1$ whose times between consecutive spikes are independent. The differences between the HOISA function of this independent train and the one of the original train will show how far from independence the train under study is.

The first test statistic is defined as follows. The HOISA function of a registered spike train, $\tilde{g}(x)$, will be compared with the one obtained from a shuffled train. More specifically, N shuffled trains are used, their HOISA functions are computed and averaged to avoid falling in a case that is not representative. This *average HOISA function* is denoted $\bar{g}(t)$. The test statistic is defined as the L_1 distance:

$$T_{HOISA} = \int |\tilde{g}(x) - \bar{g}(x)| \mathrm{d}x \ .$$

H_0 will be rejected for large values of T_{HOISA}.

For the second test, instead of using the HOISA itself, we make use of the observed values of D_m, $m = 1, \ldots, M$. The D_m were introduced for the definition of the HOISA in (3.5). Let $\{d_m\}_{m=1}^M = \{x_i - x_j : i, j \in \{1, \ldots, M\}, i \neq j; -w_{max} < |x_i - x_j| < w_{max}\}$ be the observed distances between any spikes. The K-S statistic involves the empirical distribution of the d_m, $\tilde{F}$, as well as the empirical distribution of their analogues for the shuffled, thus independent, train, $\bar{F}$:

$$T_{KS} = \sup_x |\tilde{F}(x) - \bar{F}(x)| \ .$$

To compute $\bar{F}(x)$, several shuffled trains are used, say N. From each shuffled train, the set of distances between spikes is constructed, and $\bar{F}$ is defined as the empirical distribution of the pooled sample of all these N sets.

To calibrate the distributions of the test statistics a bootstrap method is proposed. The steps for the first test are the following:

1. Sample from $\{s_i = x_{i+1} - x_i\}_{i=1}^n$ to obtain a resample $\{s_i^*\}_{i=1}^n$ of distances between consecutive spikes and build a bootstrap train: $x_1^* = 0$, and $x_i^* = \sum_{j=1}^{i-1} s_j^*$, for $i = 2, \ldots, n+1$.

2. Calculate $\tilde{g}^*(t)$ for this bootstrap train.

3. Resample N times from s^* to obtain: $s^{**(i)}$, $i = 1, \ldots, N$ as before. Build $x^{**(i)}$ as in Step 1 and calculate $\tilde{g}^{**(i)}$ for each train $x^{**(i)}$. Then define $\bar{g}^* = \frac{1}{N} \sum_{i=1}^N \tilde{g}^{**(i)}$.

4. Obtain $T^*_{HOISA} = \int |\tilde{g}^* - \bar{g}^*|$.

5. Repeat Steps 1–4 B times to get $T^*_{HOISA,1}, \ldots, T^*_{HOISA,B}$ and use them to estimate the desired quantiles of the T_{HOISA} distribution or the p-value for T^{obs}_{HOISA}.

For the Kolmogorov-Smirnov test the procedure is very similar:

1. Build the independent spike train $\{x_i^*\}_{i=1}^{n+1}$ as before.

2. Calculate the distances between any two spikes for the bootstrap train: $\{d_m^*\}_{m=1}^M$.

3. Resample N times from x^*, to build N trains $\{x^{**(j)}, j = 1, \ldots, N\}$ and for each train build the set of distances: $\{d_m^{**(j)}\}_{m=1}^{M_j}$.

4. Calculate T^*_{KS} as the Kolmogorov-Smirnov statistic for the samples $\{d_m^*\}_{m=1}^M$ and $(d_1^{**(1)}, \ldots, d_{M_1}^{**(1)}, d_1^{**(2)}, \ldots, d_{M_2}^{**(2)} \ldots, d_1^{**(N)}, \ldots, d_{M_N}^{**(N)})$.

5. Repeat the Steps 1-4 B times to obtain $T^*_{KS,1}, \ldots, T^*_{KS,B}$ and use them to estimate the desired quantiles of the T_{KS} distribution or the p-value for T^{obs}_{KS}.

In general, differences can be observed between the HOISA function of the original train and the one obtained with the resamples. Roughly speaking, the density of the resampled data is more uniformly distributed and then the main peak is lower than in the case of the real data. It is also very common the absence of secondary peaks in the HOISA functions of the resampled trains.

In Table 3.2 the results of the tests for four neurons, N1, N3a, N3b and N4b and three different recordings (one in the *pre* part and two in the *post* part, one for each stimulus) can be observed. The p-values obtained with each test were calculated using the bootstrap method described above. A

Table 3.2: p-values for the independence tests T_{HOISA}, T_{KS} and T_{LB}, constructed using the distances between two spikes.

			T_{HOISA}	T_{KS}	T_{LB}
N1	*pre*		0	0.002	0
	post	*bs*	0.080	0.036	0.205
		bf	0.204	0.668	0.012
N3a	*pre*		0.006	0.266	0
	post	*bs*	0.004	0.002	0
		bf	0.126	0.456	0.018
N3b	*pre*		0	0.002	0
	post	*bs*	0.006	0.004	0
		bf	0.308	0.634	0.980
N4b	*pre*		0	0.001	0.002
	post	*bs*	0.254	0.148	0.320
		bf	0.547	0.514	0.911

total number of 500 bootstrap resamples and $N = 80$ shuffles were used for each bootstrap train in the *pre* part and $N = 100$ in the *post* part. Also, a Ljung-Box (T_{LB}) test was implemented to compare the results.

In general, these results show that, in the *pre* period, the distances between consecutive spikes are not independent. In the *post* period the results depend on the stimulus. We cannot reject the null hypothesis of independence when the *bf* has been applied. On the other hand, for the *bs*, the results are not that clear and depend on the neuron. Figure 3.16 shows the HOISA functions of the *pre* period of four original trains and the average HOISA function for the independent case, averaged over 100 shuffles of the original train. Figure 3.17 shows the same as Figure 3.16 but for the *post* periods of the same trials (*bs* stimulation). In both figures it is easy to recognize the cases where independence is rejected at a level 0.1(neurons N1, N3a, N3b and the *pre* period of neuron N4b) and the one in which it is not (*post* of neuron N4b).

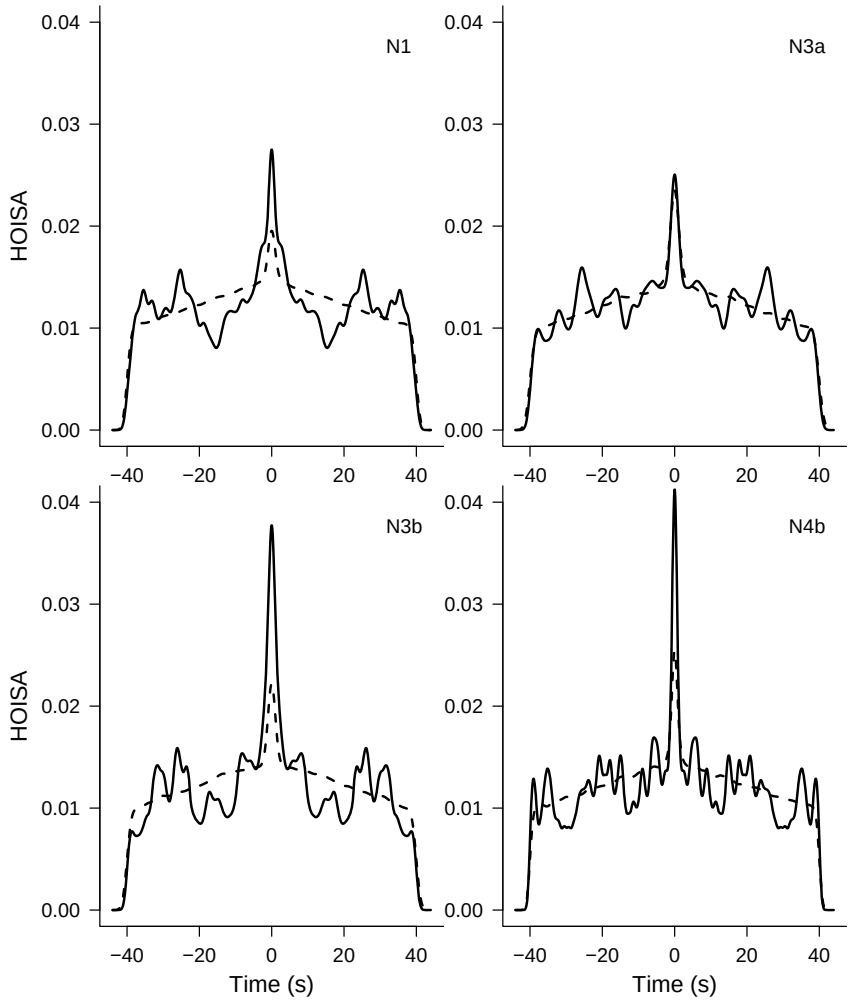

Figure 3.16: Comparison of the HOISA function for the original trains (solid line) and the average HOISA function for independent trains (dashed line). First trial of neurons N1, N3a, N3b and N4b in the *pre* period.

3.5 Chapter conclusions

The densities of the ISIs have been estimated and the stationarity of these variables has been discussed concluding that. Under under sleep-like spontaneous activity, the ISIs are rasonably stationary. Also, the higher order interspike autocorrelation (HOISA) has been presented. This autocorrelation measure is commonly used in neuroscience. The spontaneous activity

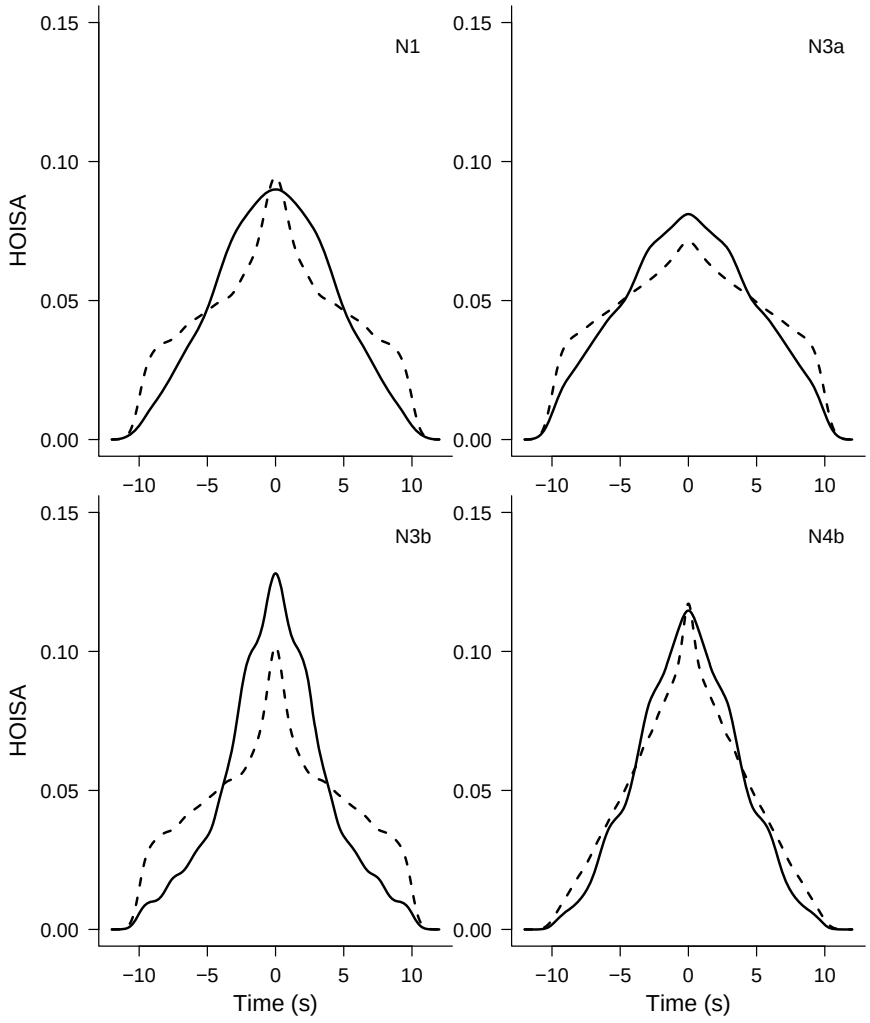

Figure 3.17: Comparison of the HOISA function for the original trains (solid line) and the average HOISA function for independent trains (dashed line). First trial of neurons N1, N3a, N3b and N4b in the *post* period after *bs* stimulation.

of neurons is characterized by the existence of dependence among spikes. Therefore, a test for independence based on the HOISA function has been proposed. As this function is constructed on the basis of a histogram, another test based on the Kolmogorov-Smirnov statistic, has been discussed. The distribution of these statistics under the null hypothesis has been calibrated with a bootstrap procedure. Finally, a Ljung-Box statistic has also

been used for comparison. This last statistic has the inconvenience of being based on the serial autocorrelation which varies very much from one trial to another. In general, it can be observed that dependence exists during the sleep-like *pre* part, reflecting the highly synchronized neuronal oscillatory activity. In the analyzed examples, this dependence is present for some neurons during the period of awake-like activity after the *bs* stimulus, while it does not appear after the *bf* stimulus for most of the neurons. In some cases, the T_{KS} and T_{LB} statistics present values that are not consistent with the ones obtained with the other tests. This does not happen with the T_{HOISA} statistic, what makes it more robust. These results indicate that the HOISA-based test for independence is a useful method for the characterization and analysis of the dynamics of the neuronal oscillatory activity.

Chapter 4

Cross inter spike intervals

Chapter 3 was devoted to make a description of some features of single spike trains. Although each of these isolated trains can carry important information, brain processing highly depends on associations among neurons. Synchrony, oscillations and dynamic associations, among others, play important roles in brain function. These interactions may depend on anatomical connections and on different functional processes. Reliable tools are needed for the quantification of these connectivity properties. In this chapter we introduce joint spiking activity to our analysis.

A method to measure pairwise neural association is introduced in the first section of this chapter. This method is used to construct a test statistic to asses whether stimulation has a significant effect in neural association. This test is used for the spike train data analyzed along this thesis.

4.1 Pairwise neural association measure

The aim of this chapter is to present a method to describe pairwise associations between neurons based on the *cross-inter-spike intervals* (CISIs) that is a generalization of the ISIs. We focus on the problem of quantifying the change neuronal synchronization between the sleep-like and the awake-like periods (a change provoked by the bs/bf stimulation). The first measure we present is a comparative one. By *comparative* we mean the following: this tool measures the association of neurons (CISI-wise) during the awake-like period in comparison to the existing association before the stimulation (i.e., neuronal synchronization in the *post* vs the *pre* periods).

Let $\tilde{S}$ be the random variable denoting the waiting time from one spike of neuron 1 to the following spike of neuron 2. We will call this variable *cross-inter-spike interval*, which we have already mentioned. We decided to work with the logarithms of the CISIs for the same reason as in the previous chapter: there exist a lot of small CISIs but the distribution has a very heavy tail to the right, and logarithms make the results much more interpretable. From now on, we will refer to the variable $\tilde{S}$ conditioned to the occurrence of a spike in neuron 1 at time t as '$\tilde{S}$ at time t'. Now, let $g(s,t)$ be the density function of the natural logarithm of $\tilde{S}$ at time t. We suppose that g is stationary in the period of time before stimulation, $g(s,t) = g_{pre}(s)$. To see how much this density is influenced by the stimulus or how the association structure of the CISIs change, we propose to use a measure of the distance between these densities. So, we define the CISI measure (CM) as the L_1 distance between the density function of $\tilde{S}$ at a time t and the density of the $\tilde{S}$ before the stimulation:

$$CM(t) = \int |g(s,t) - g_{pre}(s)| \, \mathrm{ds}.$$

It is important to notice that this measure is highly dependent on the choice of the first neuron (i.e., it is not symmetric in both neurons) and therefore, it captures causality.

In practice, let $\mathcal{X} = \{0 < X_1 < X_2 < \ldots < X_{J_1} < T\}$ and $\mathcal{Y} = \{0 < Y_1 < Y_2 < \ldots < Y_{J_2} < T\}$ be the spike times of any two simultaneously recorded spike trains. For each spike, X_i, observed in train $\mathcal{X}$, we will have an CISI observation:

$$\tilde{S}_i = \min_{j=1,\ldots,J_2} \{Y_j - X_i : Y_j - X_i > 0\}.$$

Notice that the set from which the minimum is chosen for the definition of the $\tilde{S}_i$ could be the empty set if there are spikes of neuron 1 that are not followed by any spike of neuron 2. The spikes of neuron 1 that correspond to this last case, will not have a $\tilde{S}_i$ associated with them. We can compute the $\tilde{S}_i$ for each spike of neuron 1 and estimate the density function of $\log \tilde{S}$ at time t, say $\hat{g}(s,t)$. On the other hand, in order to capture the dynamics of such process and have a reasonable time resolution, we propose the use of sliding windows. This is, at time t, we will use the spikes that fall in a certain time window around t, $X_i \in (t-v, t+v], i = 1, \ldots, J_1$ and $Y_j \in (t-v, t+v], j = 1, \ldots, J_2$, to compute the CISIs and estimate the density function of their logarithm and therefore estimate $CM(t)$:

$$\widehat{CM}(t) = \int |\hat{g}(s,t) - \hat{g}_{pre}(s)| \, \mathrm{ds}. \tag{4.1}$$

Here, $\hat{g}_{pre}(t)$ is the density function estimator of the logarithms of the CISIs in the period of time before the stimulus onset. For the purpose of the density estimations, the kernel method is used as follows. Let

$$\begin{aligned}\mathcal{A}_{t,v} =&\{i \in \{1,\ldots,J_1\}/X_i \in (t-v,t+v], \exists\, j \in \{1,\ldots,J_2\}/Y_j \geq X_i,\\ &Y_j \in (t-v,t+v]\}\end{aligned}$$

and let $m_{t,v} = \#\mathcal{A}_{t,v}$. Also, let t_{st} be the time when the stimulus is applied,

$$\mathcal{B}_{t_{st}} = \{i \in \{1,\ldots,J_1\}/X_i < t_{st}, \exists\, j \in \{1,\ldots,J_2\}/Y_j \geq X_i, Y_j < t_{st}\}$$

and $m_{t_{st}} = \#\mathcal{B}_{t_{st}}$. So, the density estimators in (9.1) result in

$$\hat{g}(s,t) = \frac{1}{m_{t,v}h} \sum_{i \in \mathcal{A}_{t,v}} K\left(\frac{s - \log(\tilde{S}_i)}{h}\right)$$

and

$$\hat{g}_{pre}(s) = \frac{1}{m_{t_{st}}h} \sum_{i \in \mathcal{B}_{t_{st}}} K\left(\frac{s - \log(\tilde{S}_i)}{h}\right),$$

where K is the kernel function and h is a smoothing parameter. In the following examples the sliding window has length 20 s ($v = 10$) and it is moved every 1 s. Also, a Gaussian kernel has been used and the smoothing parameter was chosen using the Sheather and Jones plug-in method. How the selection of v and h can affect the measure will be discussed in Section 4.3. Figure 4.1 shows the density estimations of $\log(\tilde{S})$ for neurons N1 and N3a, using N1 as reference. For these figure, the density function of $\log(\tilde{S})$ has been estimated in 20 time windows, the centers of which are separated by 10 s. These estimations are shown with solid black lines. The red lines represent the density estimation of $\log(\tilde{S})$ on the *pre* period for comparison. Stimulation has occurred at time zero. The value CM measures the distance between these two functions.

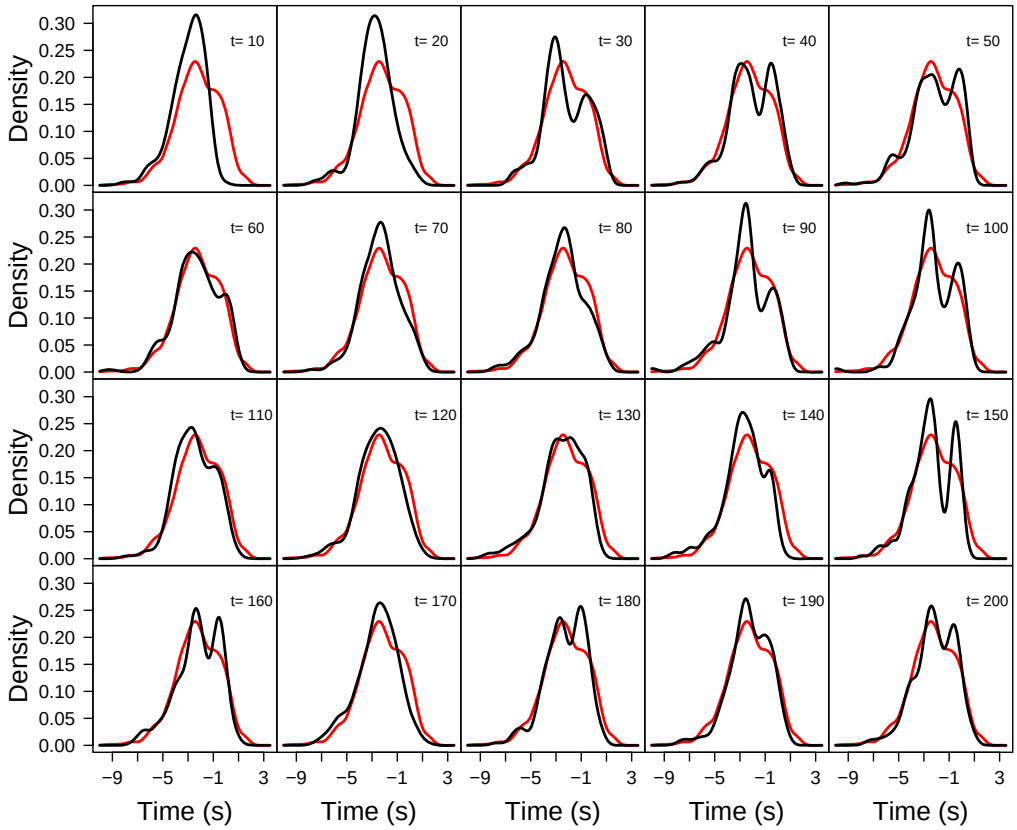

Figure 4.1: Sequence of estimated density functions (black lines) in 20 time windows of length 20 s, centered in t as indicated in the top right corner of each subfigure. The estimated density of the $\log(\tilde{S})$ in the *pre* period is also shown in each window (red lines). Stimulation at time zero.

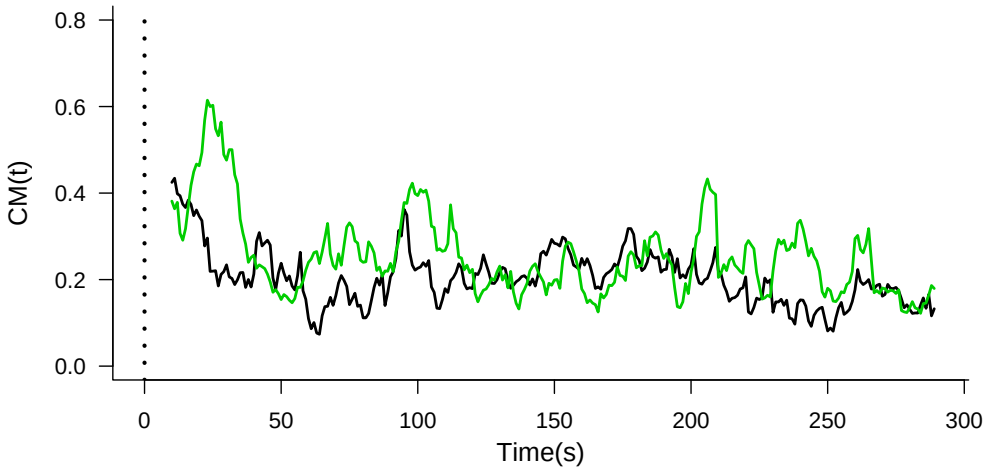

Figure 4.2: $\widehat{CM}(t)$ for neurons N1 and N3a using either N1 as the reference neuron (black line) or N3a as the reference (green line), a time window of length $v = 10$ is used. Dotted line: stimulation time.

Figure 4.2 shows $\widehat{CM}(t)$ for neurons N1 and N3a. The black line is the $\widehat{CM}(t)$ obtained when N1 is used as the reference neuron and the green line when N3a is considered as the reference. It can be observed that there are differences in the curves, showing that some causal effects could be unveiled with this method. Also, the figure shows that the differences between the densities before and after stimulation are greater close to the stimulation time which is shown in the figure by the dotted vertical line and corresponds to time zero.

Figures 4.3, 4.4 and 4.5 show the estimated $CM(t)$ for several pairs of neurons. We only show one possibility of each pair for the sake of brevity. This is, if N1-N3a is shown, then, N3a-N1 is not. When the trials of the *bs* stimulation have been used for the estimation, solid black lines are used and solid magenta lines are used for the estimation when *bf* stimulation has taken place. Although the nature of this measure is very noisy, there are several figures in which we can distinguish that the $\widehat{CM}(t)$ is higher for values of t close to zero.

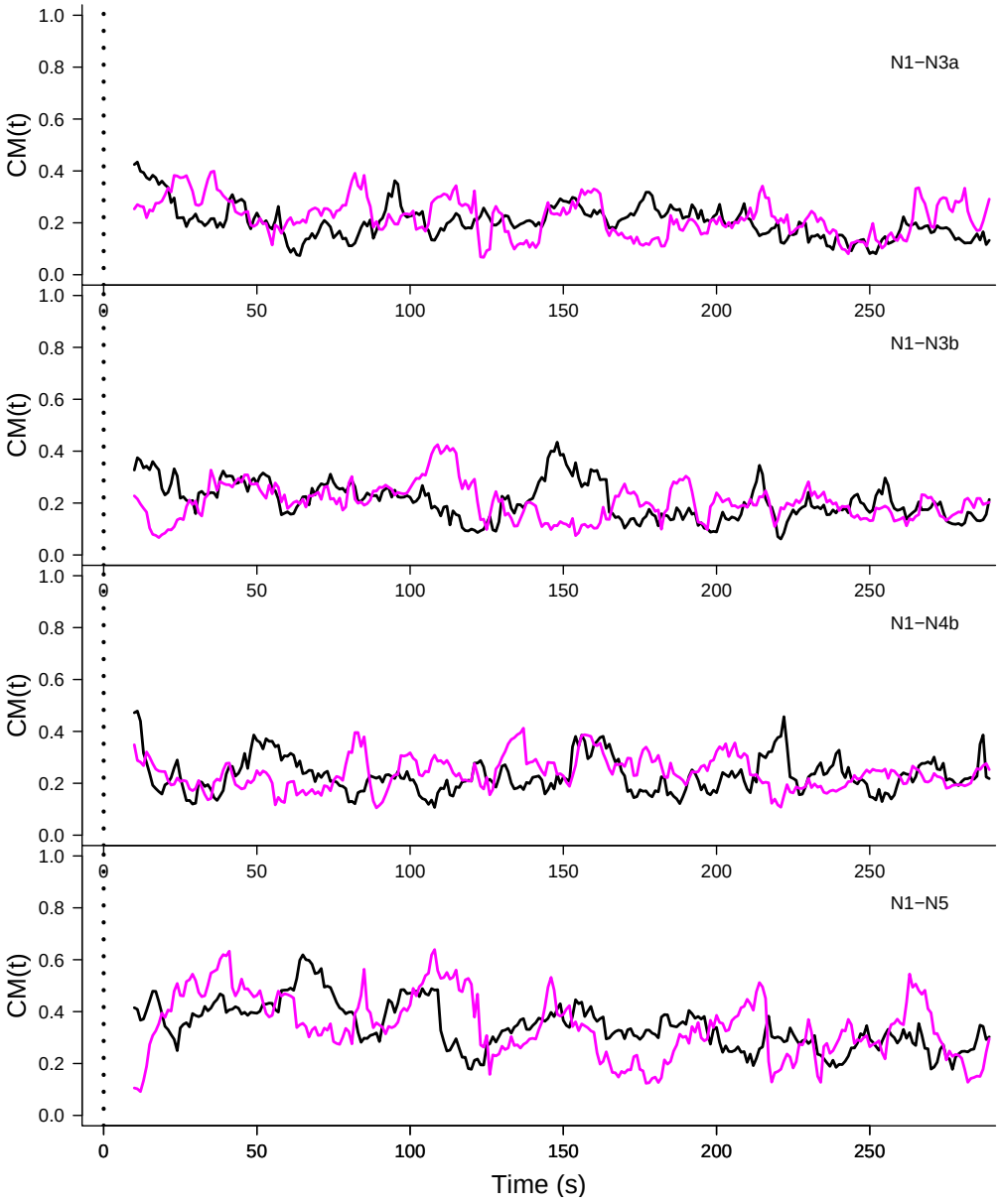

Figure 4.3: $\widehat{CM}(t)$ for all the pairs including neuron N1 for *bs* stimulation (black solid lines) and *bf* stimulation (magenta solid lines). N1 is the reference neuron for every pair. A time window of length $v = 10$ is used. Dotted vertical line: stimulation time.

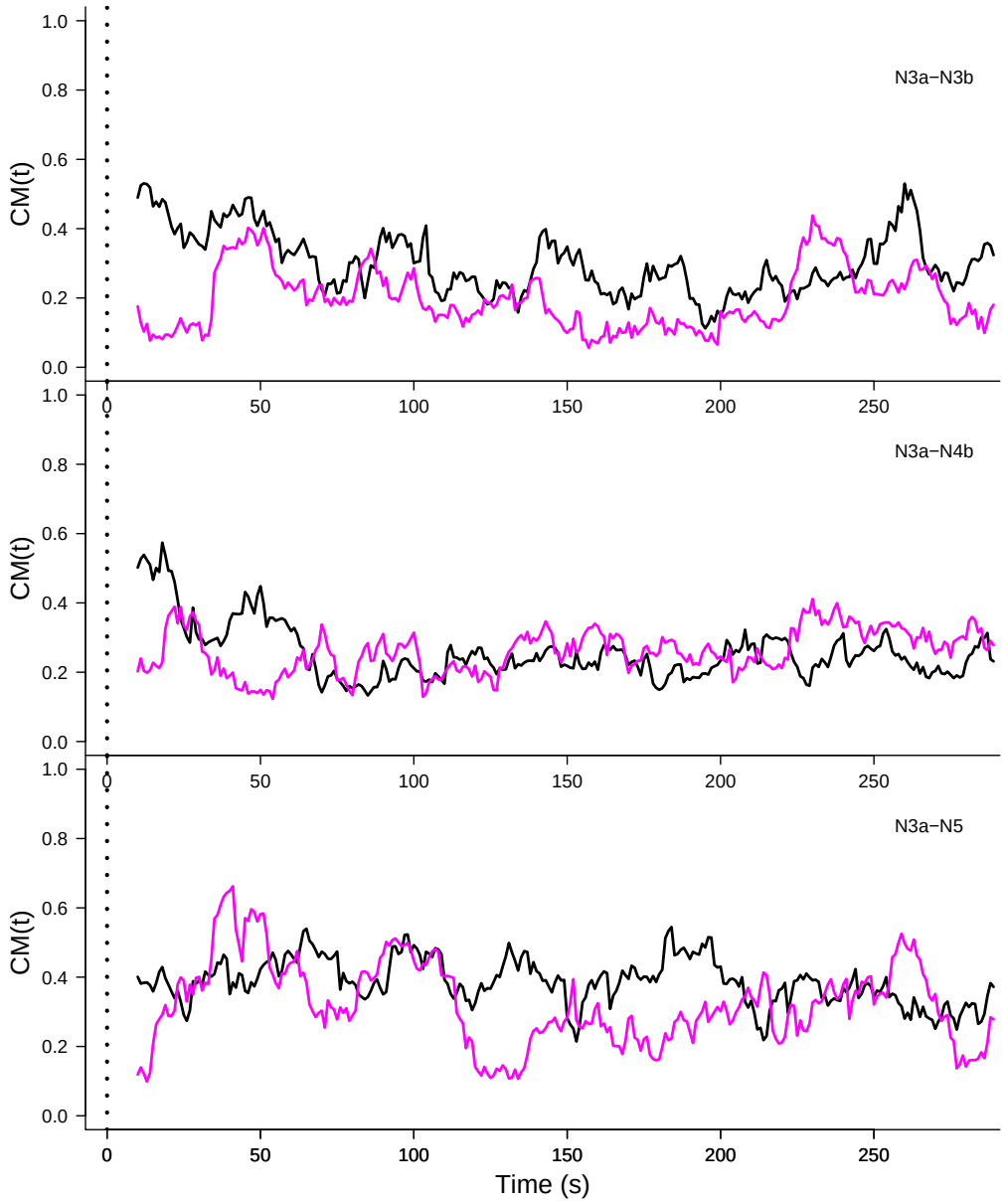

Figure 4.4: $\widehat{CM}(t)$ for all the pairs including neuron N3a (except from N3a-N1) for *bs* stimulation (black solid lines) and *bf* stimulation (magenta solid lines). N3a is the reference neuron for every pair. A time window of length $v = 10$ is used. Dotted vertical line: stimulation time.

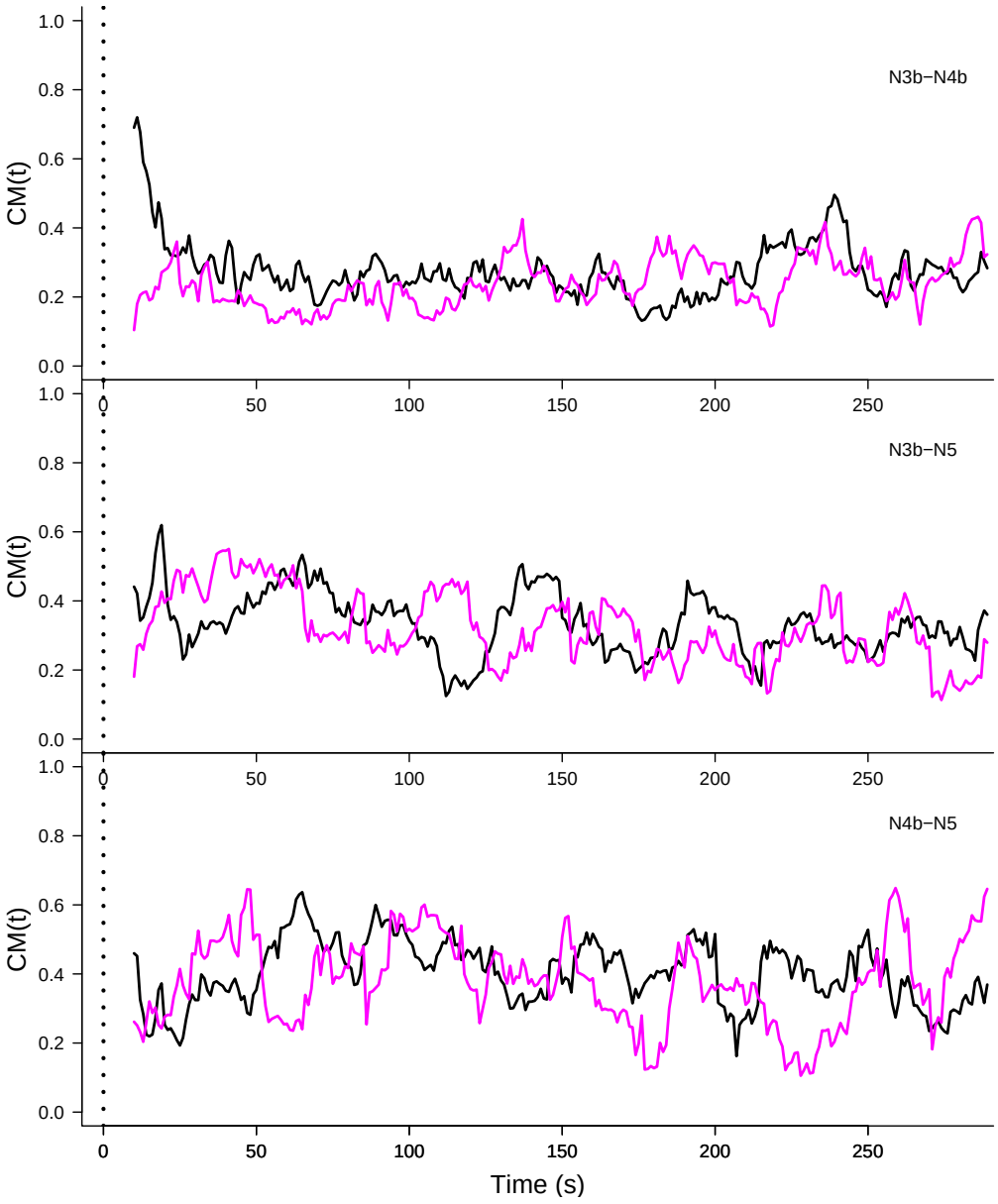

Figure 4.5: $\widehat{CM}(t)$ for all the pairs including neuron N3b and N4b for *bs* stimulation (black solid lines) and *bf* stimulation (magenta solid lines). N3b is the reference neuron in the top two panels, N4b is the reference in the other one. A time window of length 20 s. is used. Dotted vertical line: stimulation time.

4.2 Hypothesis testing

In order to statistically asses significant differences in neuronal synchronization between the pre (sleep-like) period and the synchronization dynamics along the post (awake-like) period, we propose a hypothesis test. The null hypothesis for this particular case is that the CISIs density after bs/bf stimulation and the one in the *pre* stage are the same:

- $H_0 : g(s,t) = g_{pre}(s)\ \forall s$
- $H_1 : g(s,t) \neq g_{pre}(s)\ \forall s$

The test statistic we will use is the estimator of CM defined in (9.1). Since the distribution of the test statistic under the null hypothesis is unknown, we will use a bootstrap procedure to approximate it. Under the null hypothesis, the observed $\tilde{S}$ come from the same density as the $\tilde{S}$ of the *pre* stage, which is $g_{pre}(s)$. As the distribution of these variables are continuous we decided to use a smooth bootstrap procedure described in Chapter 1.

Bootstrap Procedure

Compute the CISIs in the *pre* stage, $\tilde{\mathcal{S}}_n = \{\tilde{S}_1, \ldots, \tilde{S}_n\}$. Compute the CISIs from the observed spikes in the time window $(t-v, t+v]$ and let n_t be the amount of them; $t > t_{st}$.

1. Draw a random sample $U_1^+, \ldots, U_{n_t}^+$ from $\log(\tilde{\mathcal{S}}_n)$.
2. Select a smoothing parameter, h_{boot}, and let the final bootstrap sample be $\{U_i^* = U_i^+ + h_{boot} Z_i\}_{i=1}^{n_t}$ with $Z_1, \ldots, Z_{n_t}$ iid from a distribution with density function $K(z)$.
3. Compute $\widehat{CM}^*(t) = \int |\hat{g}^*(s,t) - \hat{g}_{pre}(s)|\, \mathrm{ds}$, where $\hat{g}^*(s,t)$ is the bootstrap version of the kernel density estimate of $\{U_1^*, \ldots, U_n^*\}$.
4. Repeat Steps 1–3, B times to obtain $\widehat{CM}_b^*(t)$ for $b = 1, \ldots, B$.
5. Compute the desired quantile of the $\widehat{CM}_b^*(t),\ b = 1, \ldots, B$ to use as a critical value at time t.

We eepeat the bootstrap procedure for each time window. Note that, at each window, the resample size is the same as the size of the original sample. In this procedure, K is a suitable kernel function and h_{boot} is a smoothing parameter that is chosen with the method proposed by Bowman et al. (1998)

for the smoothing of distribution functions. Bowman et al. (1998) propose to choose the smoothing parameter to minimize the cross-validation function

$$CV(h) = \frac{1}{n}\sum_{i=1}^{n}\int \left(\mathbb{I}\{(x - x_i) \geq 0\} - \tilde{F}_{-i}(x)\right)^2 \mathrm{dx}$$

where the indicator function given by $\mathbb{I}\{(x - x_i) \geq 0\}$ is a natural characterization of each observation in the distribution functions context and $\frac{1}{n}\sum_{j=1}^{n}\mathbb{I}\{(x - x_j) \geq 0\}$ results in the usual empirical distribution function. The term $\tilde{F}_{-i}(x)$ denotes a kernel estimator of the distribution function constructed neglecting the i-th observation:

$$\tilde{F}_{-i}(x) = \frac{1}{n-1}\sum_{\substack{j=1 \\ j\neq i}}^{n} W\left(\frac{x - x_j}{h}\right)$$

where W is a distribution function and h is smoothing parameter.

4.3 Results

Figures 4.6–4.8 show the results for the hypotheses tests described in the previous section for three pairs of neurons. These pairs are representative of all the studied pairs. Each figure shows the results for one pair of neurons, the trials with *bs* stimulation in the top panel and the trials with *bf* stimulation in the bottom panel. The red lines are the results of the bootstrap test. They represent the 99-percentile of the distribution of the bootstrapped test statistic. For the test, $B = 500$ bootstrap repetitions were performed at each time window.

When the observed $\widehat{CM}(t)$ is below the red line, the density functions of the $\tilde{S}$, before and after stimulation, cannot be considered different. On the other hand, if the black line is above the red one, there is statistical evidence of the differences between these density functions.

In almost all cases it can be observed that differences exist between the distribution of the CISIs before and right after stimulation. Anyhow, these differences are also found along all the time axis. In these particular examples, the results are barely acceptable in some cases (N1-N3a and N1-N3b) but in general are not conclusive.

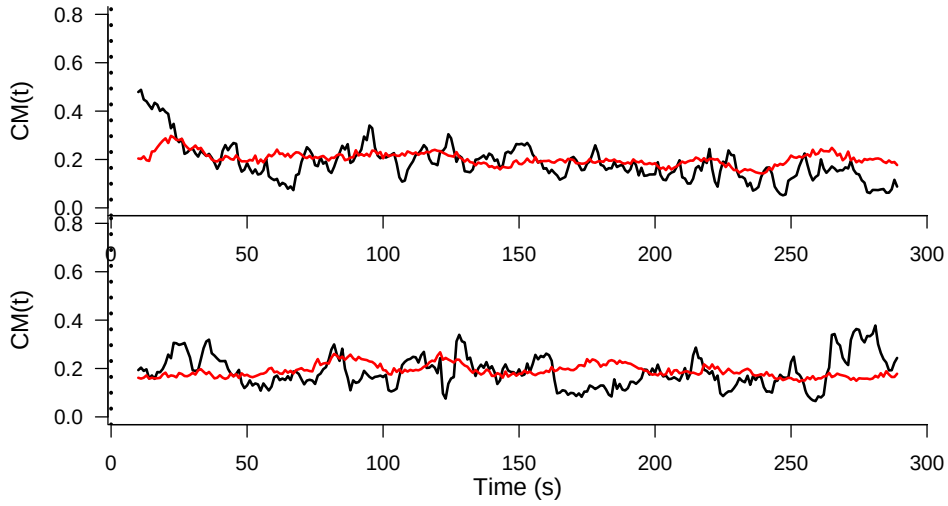

Figure 4.6: $\widehat{CM}(t)$ for the pair N1-N3a after stimulation (black line) for the *bs* stimulation trials (top panel) and *bf* stimulation trials (bottom panel). Bootstrap critical value with significance level $\alpha = 0.01$ (red line) for the null hypothesis $\widehat{CM}(t) = 0$. Stimulation at time zero.

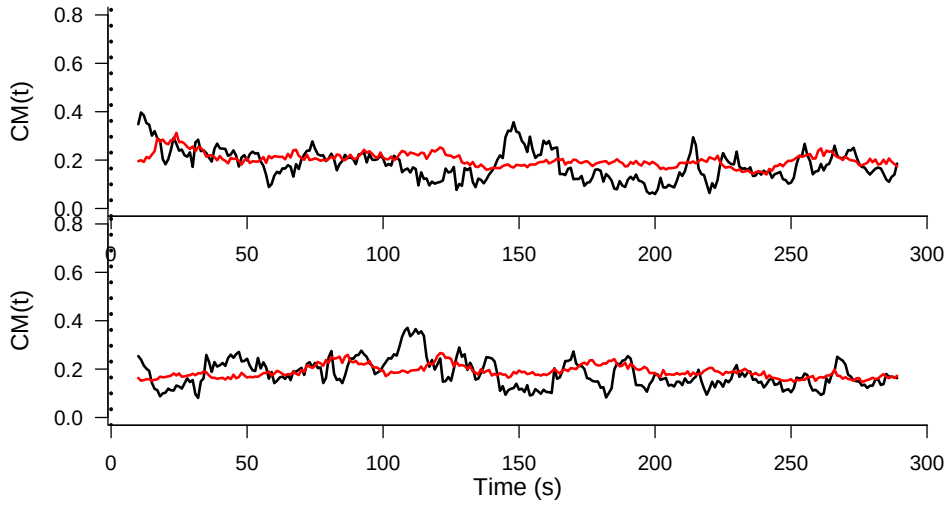

Figure 4.7: $\widehat{CM}(t)$ for the pair N1-N3b after stimulation (black line) for the *bs* stimulation trials (top panel) and *bf* stimulation trials (bottom panel). Bootstrap critical value with significance level $\alpha = 0.01$ (red line) for the null hypothesis $\widehat{CM}(t) = 0$. Stimulation at time zero.

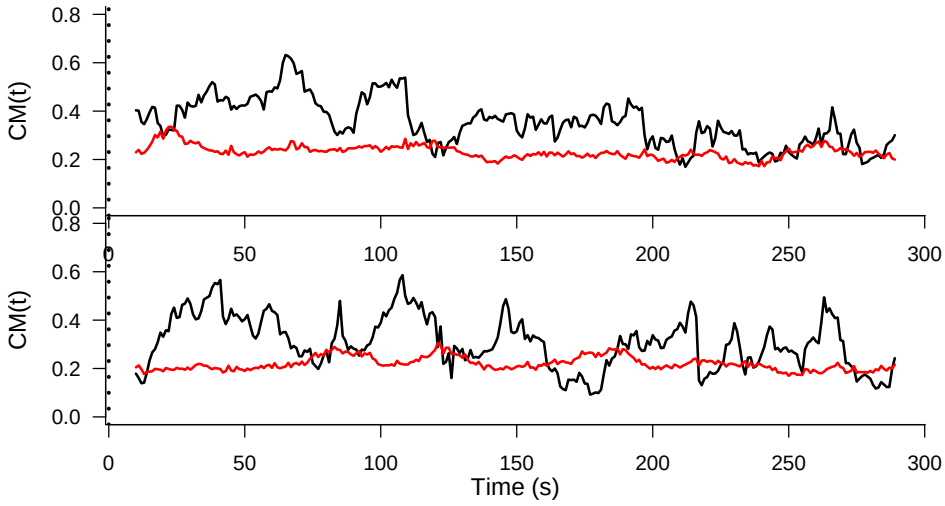

Figure 4.8: $\widehat{CM}(t)$ for the pair N1-N5 after stimulation (black line) for the *bs* stimulation trials (top panel) and *bf* stimulation trials (bottom panel). Bootstrap critical value with significance level $\alpha = 0.01$ (red line) for the null hypothesis $\widehat{CM}(t) = 0$. Stimulation at time zero.

4.3.1 Parameter selection

Figure 4.9 shows the CM curve for three different widths for the sliding window, v: $v = 5$ in black, $v = 10$ in red and $v = 15$ in green. In this case the smoothing parameter has been selected using the Sheather and Jones plug-in method. It is clear that $v = 5$ is too small as the curve results very noisy. Between $v = 10$ and $v = 15$ there are not great differences and therefore we chose the smaller one to minimize the loss of temporal resolution.

Figure 4.10 shows the CM curve for three different choices of the smoothing parameter at each time window $(t - v, t + v]$, say h_t. The choices are h_t equal to the one given by the automatic selector of Sheather and Jones using the Gaussian kernel, say h_t^0, $\frac{h_t^0}{2}$ and $2h_t^0$. For the estimation of the density function in the *pre* period, the same proportion of the corresponding Sheather and Jones parameter is used in each case. The sliding window used is $v = 10$.

As already mentioned, the parameter h_{boot} was chosen using the method presented by Bowman et al. (1998). The parameters resulted in $h_{boot} = 0.177$ for the N1-N3a pair, $h_{boot} = 0.203$ for N1-N3b and $h_{boot} = 0.151$ for N1-N5. Figure 4.11 shows the cross-validation function obtained for the pair N1-N3a.

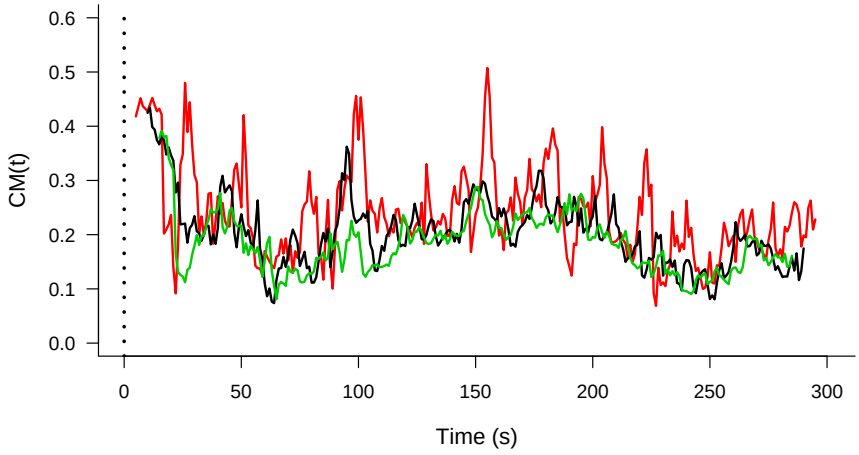

Figure 4.9: $\widehat{CM}(t)$ for neurons N1 and N3a using N1 as the reference neuron and different choices of the sliding window width: $v = 5$ (red), $v = 10$ (black) and $v = 15$ (green). Vertical dotted line: stimulation time.

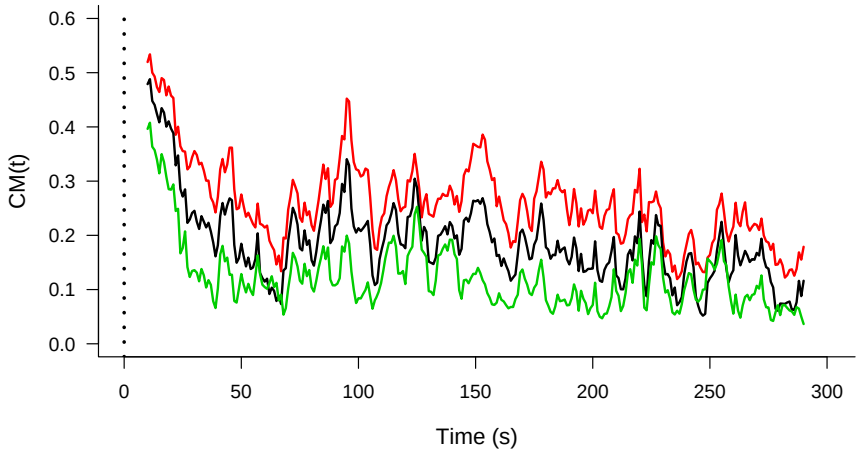

Figure 4.10: $\widehat{CM}(t)$ for neurons N1 and N3a using N1 as the reference neuron and different choices of the smoothing parameter h_t in each time window. $h_t = h_t^0$ given by the Seather and Jones method (black), $h_t = \frac{h_t^0}{2}$ (red) and $h_t = 2h_t^0$ (green). Vertical dotted line: stimulation time.

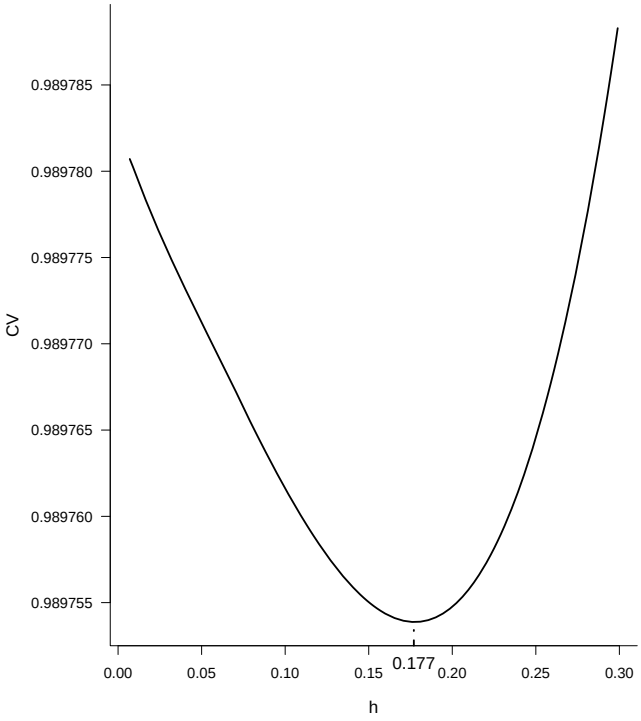

Figure 4.11: Cross-validation function obtained for the pair of neurons N1-N3a in order to choose the optimal smoothing parameter, h_{boot}, for the bootstrap hypothesis test procedure.

4.4 Chapter conclusions

We have presented a method that compares the density functions of the cross-inter-spike intervals between the sleep-like and the awake-like periods. The aim of this approach was to search for differences in these functions and therefore in the associations between the neurons. Statistically significant differences reflect a clear separation in the structure of the dynamical associations between neurons. But this measure resulted extremely noisy. In the studied examples, the differences could be found along the whole time axis with no clear temporal tendencies, as would be expected due to the evident and progressive transformation of the awake-like pattern back to the sleep-like mode that occurs in our experimental model. Although the method is of interest and the visual inspection of the density functions along time is informative, the MC does not seem trustworthy for the intended inference.

Chapter 5

Cross-correlation based synchrony measure

In this chapter we introduce the cross-correlation function for simultaneously recorded spike trains and present a synchrony measure that is based on the estimation of that cross-correlation function. Contrary to the CISI-based method described in the previous chapter, the measure presented here is absolute. It is not relative to the *pre* stage activity. We aim to develop a method to detect significant changes in synchronization dynamics under the two different experimental conditions simulating the sleep-wake cycle. The measure proposed here is flexible in the sense that it can be adapted for different firing rates. We also propose hypothesis tests to asses for the changes in synchrony induced by the differential electrical stimulation on *bs* and *bf*. Two bootstrap procedures are presented. These resampling procedures take into account the dependence between simultaneously spike trains by resampling from the intervals of time that elapse between spikes of a joint spike train built by merging the spike trains. These methods are inspired in the stationary bootstrap proposed by Politis and Romano (1994).

5.1 Integrated cross-correlation synchrony index

Let us consider again, $\mathcal{X} = \{X_i\}_{i=1}^{J_1}$ and $\mathcal{Y} = \{Y_j\}_{j=1}^{J_2}$ two simultaneously recorded spike trains and let us denote $[0, T]$ their common time interval. Let U_i and U_{-i} be the *waiting times* from a spike in train $\mathcal{X}$ to the i-th subsequent and the i-th preceding spike in train $\mathcal{Y}$ respectively. Observe that $U_1 = \tilde{S}$ is the CISI variable. The probability density functions of these

random variables are called the forward and backward cross-interval densities of order i, respectively, and we denote them by $\eta_i(\tau)$ and $\eta_{-i}(\tau)$. The cross-correlation function $\zeta_{XY}(\tau)$ is defined as the weighted sum of cross-interval densities of all orders:

$$\zeta_{XY}(\tau) = \sum_{i\neq 0,\, i=-\infty}^{i=\infty} \psi_i \eta_i(\tau), \tag{5.1}$$

where ψ_i can be thought of as follows. Given a spike on the first train and choosing at random a spike of train 2, ψ_i is the probability that the chosen spike is the i-th subsequent spike to the spike of the first train. Cross-correlation represents the probability density function of the time from an event in train $\mathcal{X}$ to an event randomly chosen in train $\mathcal{Y}$ (Perkel et al. (1967b)).

The cross-interval densities can be estimated from the observed spike trains and, in practice, we can use the empirical normalized cross-correlogram to estimate the cross-correlation function $\zeta_{XY}(\tau)$. The cross-correlogram is built as the histogram of the observed *waiting times* between the spikes of the first neuron and the spikes of the second neuron. Usually, joint firing or close in time firing is the event of interest so only small values of τ in (5.1) really matter. This is why the cross-correlograms are usually built for waiting times smaller than ν. In order to consider a proper density we use the normalized cross-correlogram:

$$\mathcal{T}_{XY}(\tau) = \frac{\zeta_{XY}(\tau)}{\int_{-\nu}^{\nu} \zeta_{XY}(t)\mathrm{d}t} \text{ for } -\nu < \tau < \nu$$

An estimator for this function will be discussed in the next subsection.

Synchrony is usually defined as the event of neurons firing together. In some cases, mostly in scenarios of low firing rate, as in the case of spontaneous activity, spikes of different neurons will not appear exactly at the same time but still be highly synchronized, following a similar firing pattern. If this is the case, to capture synchrony we need a flexible tool that will not only take into account unitary events as regular synchrony measures do.

The measure proposed here is based on the normalized cross-correlation function, calculated in a *cross-correlation window* of length 2ν. More precisely, we use the integral of the cross-correlation function around zero. In this way, we allow synchrony to be based on delayed firing and not only in simultaneous firing. We denote this measure as integrated cross-correlation

synchrony index (ICCSI):

$$\text{ICCSI} = \int_{-\delta\nu}^{\delta\nu} \mathcal{T}_{XY}(\tau)\mathrm{d}\tau, \tag{5.2}$$

where $\delta < 1$ is an arbitrary small number chosen by the researcher. Integrating of $\mathcal{T}_{XY}(\tau)$ in a neighborhood of zero we account for spikes that occur close in time, although not exactly at the same time. In this way, if the ICCSI is large, it means that neurons fire close in time. On the other hand, small values of ICCSI mean that the *waiting times* are not concentrated around zero, and therefore less synchrony exists. Figure 5.1 shows the estimation of the density function of the *waiting times* using a histogram and a kernel density estimator. In grey, the frequency of *waiting times* with absolute value smaller than $\delta\nu = 0.05\,\mathrm{s}$ are highlighted.

5.2 Estimation of ICCSI

In practice, cross-correlation can be estimated using the observed elapsed times from one spike in the first neuron to all the spikes in the second one. Let $\mathcal{D} = \{D_k\}_{k=1}^{N} = \{X_i - Y_j : |X_i - Y_j| < \nu, i = 1, \ldots, J_1, j = 1, \ldots, J_2\}$ be the *waiting times* that lay in a certain cross-correlation window of length 2ν. This is, the set of all possible differences between the spike times of one train and the spike times of the second one, which are smaller than ν. We estimate the normalized cross-correlogram using a kernel estimator of the density function:

$$\hat{\mathcal{T}}_{XY}(\tau) = \frac{1}{Nh}\sum_{k=1}^{N} K\left(\frac{\tau - D_k}{h}\right),$$

with K a kernel function and h the smoothing parameter. The cross-correlogram can also be estimated using a histogram (Perkel et al. (1967b)). Let $\delta \in [0, 1]$, then, the ICCSI in (5.2) can be estimated by

$$\widehat{\text{ICCSI}} = \int_{-\delta\nu}^{\delta\nu} \hat{\mathcal{T}}_{XY}(\tau)\mathrm{d}\tau.$$

5.2.1 ICCSI as a function of time

In the previous discussion, synchrony has been thought as a stationary measure. Stationarity in spike trains is often difficult to assess and it is an

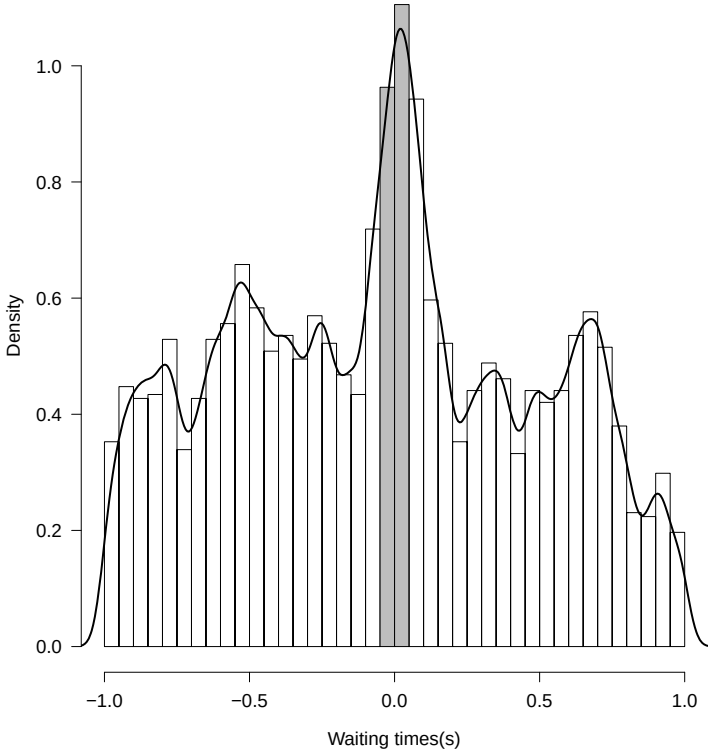

Figure 5.1: Histogram and kernel density estimation for the *waiting times* between neurons N1 and N3a.

object of study in itself. In our context, time-varying properties of ICCSI are of interest, since we want to study how synchrony evolves in two different spontaneous states (sleep vs awake, artificially evoked) and after differential induction of the awake-like state. We take into account time, by making use of moving windows to estimate ICCSI and therefore obtaining $\widehat{ICCSI}(t)$.

At each time point t, let $\eta_i(\tau; t)$ be the cross-interval densities of order i at time t. Then the cross-correlation at time t, $\zeta_{XY}(\tau; t)$, is defined as the weighted sum of the cross-interval densities of all orders at time t and $\mathcal{T}_{XY}(\tau; t)$ as its normalized version. Therefore, we can define $ICCSI(t)$ as

$$\text{ICCSI}(t) = \int_{-\delta\nu}^{\delta\nu} \mathcal{T}_{XY}(\tau; t)\mathrm{d}\tau. \tag{5.3}$$

In practice, to estimate the cross-correlations, and therefore the ICCSI,

we will use the information of spike trains in a neighborhood around t. Let $W_t = [t-w, t+w]$ be a time window of length $2w$ around t, for each $t \in [w, T-w]$. Then, for each t we can define the subtrains $\mathcal{X}_t = \{X_i \in \mathcal{X} : X_i \in W_t, i = 1, \ldots, J_1\}$ and $\mathcal{Y}_t = \{Y_i \in \mathcal{Y} : Y_i \in W_t, i = 1, \ldots, J_2\}$. Moreover, we can define $\mathcal{D}^t = \{D_k^t\}_{k=1}^{N_t}$ as the INISIs of the trains $\mathcal{X}_t$ and $\mathcal{Y}_t$, build the normalized cross-correlogram $\hat{\mathcal{T}}_{XY}(\tau, t)$ and therefore, estimate the ICCSI at time t:

$$\widehat{\text{ICCSI}}(t) = \int_{-\delta\nu}^{\delta\nu} \hat{\mathcal{T}}_{XY}(\tau; t)\mathrm{d}\tau.$$

The ICCSI can be estimated in as many points, t, as desired. It is a continuous measure that will be calculated in a sequence of points, say $t_1, \ldots, t_M$. The amount of waiting times in each time window, W_t, is very variable, specially, when the firing is sparse. Thus, there are windows with very small amount of data. To remedy this we propose, in the next subsection, a kernel smoothing of $\widehat{ICCSI}(t)$.

5.2.2 Nonparametric smoothing of $\widehat{ICCSI}$

The number of spikes at each time window is very variable and when this number is small $\widehat{ICCSI}(t)$ becomes less reliable. To make $\widehat{ICCSI}(t)$ more robust, in order to be able to highlight characteristics of these curves and find patterns due to experimental conditions, we use a regression kernel smoother of the form:

$$\widehat{ICCSI}^{smooth}(t) = \sum_{j=1}^{M} \Psi_j(t)\widehat{ICCSI}(t_j),$$

for some weight functions Ψ_j. We will use the most common kernel estimator: the Nadaraya-Watson estimator, presented in Chapter 1. For this estimator, the weights are:

$$\Psi_j(t) = \frac{K\left(\frac{t_j - t}{h}\right)}{\sum_{r=1}^{M} K\left(\frac{t_r - t}{h}\right)},$$

We use the uniform kernel function $K(u) = 0.5$ if $|u| < 1$ and 0 otherwise. For simplicity of notation $\widehat{ICCSI}^{smooth}(t)$ is actually denoted by $\widehat{ICSI}(t)$ throughout the chapter.

5.3 Testing for synchrony differences

To check whether there are differences between the ICCSI during the awake-like and the sleep-like activity, a hypothesis test is implemented as follows. Under sleep-like activity it is assumed that the mean synchrony do not change with time, i.e., $\text{ICCSI}(t) = \text{ICCSI}_0$ for every $t \in [0, t_{st})$, where t_{st} is the time when the stimulus is applied. Given two values $t_0 \in [0, t_{st})$ and $t_1 \in (t_{st}, T]$ we would like to test if the synchrony index is equal at these two values. This is equivalent to say that, at time t_1 (in the awake-like period, after the stimulus) the synchrony recovered the value it had before the stimulus onset.

- $H_0 : \text{ICCSI}(t) = \text{ICCSI}_0$
- $H_1 : \text{ICCSI}(t) < \text{ICCSI}_0$.

for some fixed $t \in [t_{st}, T]$.

To calibrate the distribution of $\text{ICCSI}(t)$ under the null hypothesis we use a bootstrap procedure. Spontaneous activity in the sleep-like time interval is imitated using a stationary bootstrap (see Politis and Romano (1994)). We propose a resampling procedure that takes into account the dependence that may exist between the spike trains, resampling from a joint spike train. Then, the bootstrapped ICCSIs are calculated and used to calibrate the distribution of the ICCSI under the null hypothesis and test for differences between $\widehat{ICCSI}(t)$ in an awake-like t and the baseline $ICCSI_0$. The bootstrap resamplig procedure is the following:

1. Merge the two observed trains, $\mathcal{X}_1$ and $\mathcal{X}_2$, in one, ordering all the spiking times together in a pooled train. Let this joint train be $\mathcal{X}^p = \{(X_1^p, \gamma_1^p), \ldots, (X_N^p, \gamma_N^p)\}$ where γ_i^p is an indicator variable of the spike train to which the action potential that occurs at time X_i^p belongs.

2. Next, compute the ISIs of this new train: $S_1^p = X_1^p$ and $S_{i+1}^p = X_{i+1}^p - X_i^p, i = 1, \ldots, N-1$ and let $\mathbf{S}^p = \{(S_i^p, \gamma_i^p)\}_{i=1}^N$.

3. Build the sets $\mathbf{S}^1 = \{(S_i^p, \gamma_i^p) : \gamma_{i-1}^p = 1; i = 1, \ldots, N\}$ and $\mathbf{S}^2 = \{(S_i^p, \gamma_i^p) : \gamma_{i-1}^p = 2; i = 1, \ldots, N\}$. This is, $\mathbf{S}^1$ (and $\mathbf{S}^2$) contains the elapsed times from a spike of neuron 1 (respectively 2) to the following spike in the joint train, and their corresponding neuron indicators.

4. Randomly choose $(S_1^{p*}, \gamma_1^{p*})$ from $\mathbf{S}^p$, i.e. $P((S_1^{p*}, \gamma_1^{p*}) = (S_i^p, \gamma_i^p) = \frac{1}{N})$ $i = 1, \ldots, N$.

5. If $S_i^{p*} = S_j^p$ choose $(S_{i+1}^{p*}, \gamma_{i+1}^{p*}) = (S_{j+1}^p, \gamma_{j+1}^p)$, [in the case $j = N$, $(S_{i+1}^{p*}, \gamma_{i+1}^{p*}) = (S_1^p, \gamma_1^p)$], with probability p_{boot} and choose it at random from $\mathbf{S}^{\gamma_j^p}$ with probability $1 - p_{boot}$.

6. Repeat Step 5 until obtaining the first $(S_M^{p*}, \gamma_M^{p*})$ for which $\sum_{i=1}^M S_i^{p*} \geq t_{st}$.

7. Build the ISIs for the first bootstrap train, $\mathcal{X}^{1*}$. Let $L_1 = \min_l\{\gamma_l^{p*} = 1\}$, then $S_1^{1*} = \sum_{k=1}^{L_1} S_k^{c*}$. For $i = 2, \ldots, I_1 = \#\{\gamma_l^{p*} : \gamma_l^{p*} = 1\}$ let $L_i = \min_{l>L_{i-1}}\{\gamma_l^{p*} = 1\}$, and then $S_i^{1*} = \sum_{k=L_{i-1}+1}^{L_i} S_k^{p*}$.

8. Build the first bootstrap train $\mathcal{X}^{1*}$ as $X_i^{1*} = \sum_{k=1}^i S_k^{1*}$ for $i = 1, \ldots, I_1$.

9. Build the second bootstrap train $\mathcal{X}^{2*}$ in a similar way. This consists in repeating Steps 7–8 but with the condition $\gamma_l^{p*} = 2$.

10. Calculate $\widehat{ICCSI}^{*1}(t)$ for the bootstrap trains, $\mathcal{X}^{1*}$ and $\mathcal{X}^{2*}$.

11. Repeat Steps 4–10, B times to obtain $\widehat{ICCSI}^{*b}(t)$, $b = 1, \ldots, B$.

Steps 1–3 in the algorithm are used to build the pooled train. Bootstrap resamples for the ISIs of this joint train are obtained in Steps 4–6. Finally Steps 7–9 separate the pooled bootstrap train to obtain two 'simultaneously recorded' bootstrap trains. This algorithm simulates the distribution of $\widehat{ICCSI}$ under the null hypothesis. In our case, we seek for significant reductions in the awake-like period with respect to the sleep-like period, so we propose to build a critical value as follows.

We are assuming the synchrony is constant in the time period preceding the stimulus onset, $ICCSI(t) = ICCSI_0 \; \forall\, t \in (0, t_{st})$. For each b and at each time t, $ICCSI^{*b}(t)$ is an approximation to that real value $ICCSI_0$ and therefore, the α-percentile of all the bootstraped values is a plausible choice for a critical value. We will denote this value by $ICCSI_\alpha^*$ and the null hypothesis will be rejected at each time point t if $ICCSI(t) < ICCSI_\alpha^*$.

5.4 Testing the differences between two conditions

Apart from detecting differences between synchrony among neurons, before (sleep-like) and after (awake-like) the stimulus onset, the aim of this chapter

is to develop a method to detect differences in synchronization dynamics during the awake-like period induced by the activation of two different pathways. In this context the relevant hypothesis to test is:

$$H_0 : ICCSI^{bs}(t) = ICCSI^{bf}(t)$$

$$H_1 : ICCSI^{bs}(t) \neq ICCSI^{bf}(t),$$

for each $t \in [t_{st}, T]$. In this hypothesis, $ICCSI^{bs}(t)$ and $ICCSI^{bf}(t)$ are the ICCSI under *bs* and *bf* stimulation respectively. Moreover, we develop a test that enables detecting in which time periods there are differences, if any, between the awake-like activity induced by each pathway. The test statistic to use will be $T_{ICCSI}(t) = \widehat{ICCSI}^{bs}(t) - \widehat{ICCSI}^{bf}(t)$. In this case, we need to focus in changes in time so, the bootstrap procedure used in the previous section is not valid. As the test needs to be applied in the time interval after the stimuli onset, the different trials for each stimulus are comparable and we can use the information collected through trials. In this case, the resampling is made at each time point from the different trials taking care for the dependence in a similar way as in the previous procedure. The bootstrap resampling plan is the following:

1. Build a joint train for each recorded trial k, $k = 1, \ldots, K$, $\mathcal{X}_k^p = \{(X_{k1}^p, \gamma_{k1}^c), \ldots, (X_{kN_k}^p, \gamma_{kN_k}^p)\}$ where, as above, γ_{ki}^p is an indicator variable of the spike train to which the action potential that occurs at time X_{ki}^p in trial k belongs.

2. Choose a trial, k_1, at random with equal probability from $\{1, \ldots, K\}$ and define $X_1^{p*} = X_{k_1 1}^c$ and $\gamma_1^{p*} = \gamma_{k_1 1}^c$.

3. If $(X_l^{p*}, \gamma_l^{p*}) = (X_{k_l j}^p, \gamma_{k_l j}^p)$ then, with probability p_{boot} set $k_{l+1} = k_l$, $X_{l+1}^{c*} = X_{k_{l+1}(j+1)}^p$ and $\gamma_{l+1}^{p*} = \gamma_{k_{l+1}(j+1)}^p$. With probability $1 - p_{boot}$, draw k_{l+1} at random with equal probabilities from $\{1, \ldots, K\}$, set $X_{l+1}^{p*} = X_{k_{l+1}m}^p$ so that $X_{k_{l+1}m}^p = \min_\nu \{X_{k_{l+1}\nu}^c > X_l^{p*}\}$ and $\gamma_{l+1}^{p*} = \gamma_{k_{l+1}m}^p$.

4. Increase the index l by one unit and repeat Steps 2–3 while possible, i.e., while there exists some index ν, such that $X_{k_{l+1}\nu}^p > X_l^{p*}$.

5. For each trial, k, and each stimulus, $j = 1, 2$, repeat Steps 2–4 above to obtain the bootstrap train $\mathcal{X}_{jk}^*$. This is, resample a joint spike train an therefore, two trains, one for each neuron, X_{1k}^* and X_{2k}^* from all the registered trials.

6. Compute the bootstrap $\widehat{ICCSI}^{s*}(t)$ for each stimulus, $s = 1, 2$ and $T_{ICCSI}^*(t) = \widehat{ICCSI}^{1*}(t) - \widehat{ICCSI}^{2*}(t)$.

7. Repeat Steps 5–6, B times to obtain $T^{*1}_{ICCSI}(t), \ldots, T^{*B}_{ICCSI}(t)$.

8. Calculate the α and $(1-\alpha)$ quantiles, $T^{*}_{ICCSI,\alpha}(t)$ and $T^{*}_{ICCSI,(1-\alpha)}(t)$, at each t. Reject H_0 at time t if $T_{ICCSI}(t) < T^{*}_{ICCSI,\alpha}(t)$ or $T_{ICCSI}(t) > T^{*}_{ICCSI,(1-\alpha)}(t)$.

If H_0 holds, the trials for both stimuli are generated by the same process. The proposed bootstrap mimics that process using the pooled information in Steps 1–4 above.

5.5 Application to spike trains

In this section we apply the methods described so far to the real data. First of all, we will discuss briefly how the different parameters have been chosen.

5.5.1 Choosing the tunning parameters

Here, we show how the synchrony measure is affected by the choice of the parameters δ and v. As Figure 5.2 shows, the values of the ICCSI are highly influenced by these choices, therefore the importance of being aware of it and having insight of the physiological problem studied. In our problem we have chosen $\delta = 0.025\,\text{s}$. As the neurons are under spontaneous activity, it is plausible to still consider associated two spikes that are separated by 25 ms.

As expected, the larger v the smoother the ICCSI curve, although large values of v also have the drawback of loosing temporal resolution. On the other hand, the changes observed when changing the choice of δ are of another nature. As δ gets larger, the mean value of the curve gets larger and so does the dispersion. This is also expected as the total amount of waiting times are the same, so as the interval of integration gets bigger the area increases. For our analysis, after discussing with the experimentalists and physiologists, we decided to use $v = 10\,\text{s}$ and $\delta = 0.025$. Also we use $\nu = 1\,\text{s}$ and therefore the area of integration will be $\delta\nu = 0.025\,\text{s}$. As we already know, the data come from spontaneous activity recordings. Therefore, the firing rates are very low for each neuron and it is very hard to find spikes occurring at exactly the same moment. Then, synchrony can be considered as the event of neurons firing together in a more general way than unitary events. This is why $\delta = 0.025$ is considered as a good choice. On the other hand, for the nonparametric smoothing of the $ICCSI$ we chose a small value of h, $h = 0.5$. Consequently, since the $ICCSI$ is computed every $0.05\,\text{s}$, we

are averaging one second around each time point where the $ICCSI$ was chosen to gain smoothness.

Figure 5.3 shows the results of the bootstrap test described in Section 5.3 for three different choices of δ (0.01, 0.025 and 0.1) for the pair N1-N3a. It can be observed that, even though the scale of the $ICCSI(t)$ curve changes considerably, the periods of time where the null hypothesis can be rejected are very similar. This means that, when searching for differences in synchrony, the results do not vary significantly depending on the choice of δ.

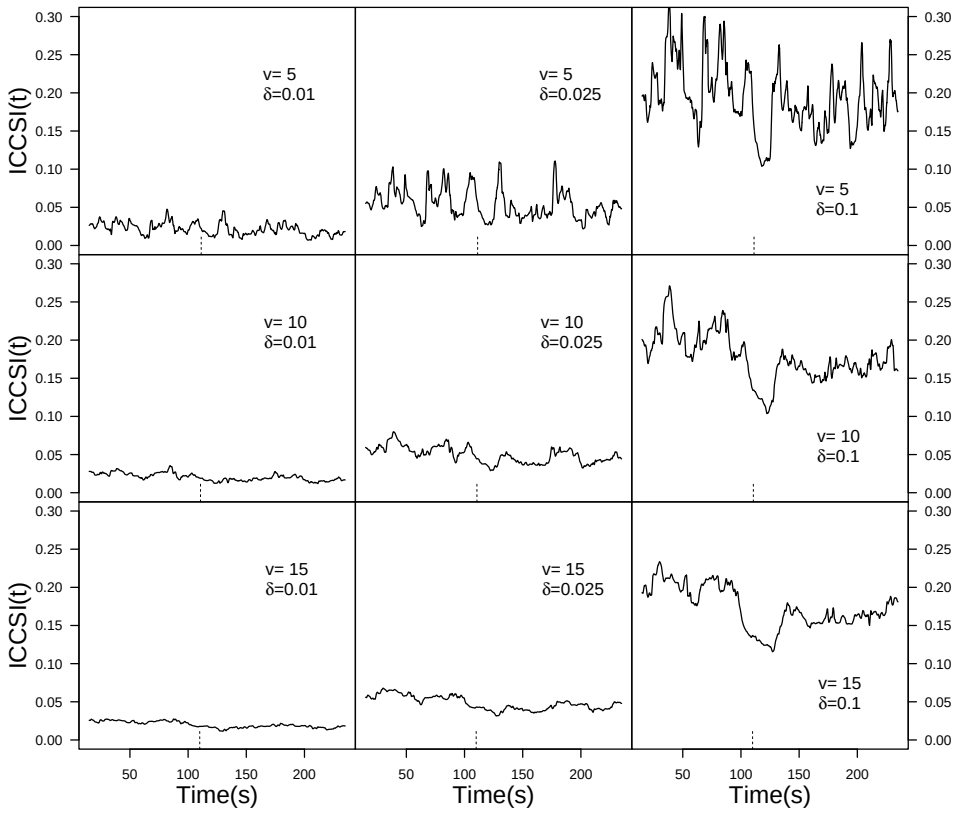

Figure 5.2: ICCSI of pair N1-N3a using different values of v and δ: $v = 5, 10$ and 15, $\delta = 0.01, 0.025$ and 0.1.

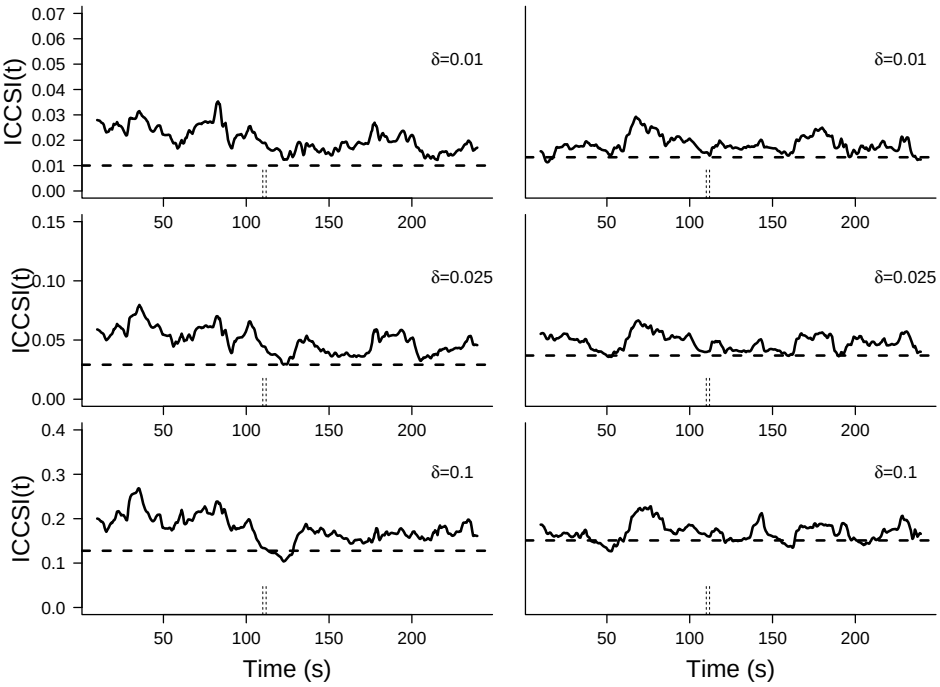

Figure 5.3: Bootstrap significance test for synchrony between neurons N1 and N3a, using different values of δ. Panels in the left column correspond to *bs* and the ones in the right to *bf*. Top panels correspond to $\delta = 0.01$, middle panels to $\delta = 0.025$ and bottom panels to $\delta = 0.1$. The value $v = 10\,\text{s}$. has been considered in all the panels. The horizontal dashed lines correspond to the bootstrap critical value.

5.5.2 Testing for synchrony differences

In this subsection, hypothesis testing for synchrony between two neurons is considered. We show the results for pairs N1-N3a, N1-N3b, N1-N4b, N1-N5, N3a-N3b, N3a-N4b and N3b-N4b.

For the bootstrap tests we have chosen $p_{boot} = 0.97$. We have based our decision on an attempt to reach a balance between imitating the dependence in the data and the variability that the resampling pursues. The stationary bootstrap is based on the sampling of blocks of observations with random length. The length of the k-th block, L_k, follows a geometric distribution, so that $P(L_k = r) = p_{boot}^{r-1}(1 - p_{boot})$. On the other hand, the expected value of the block length is $\frac{1}{1-p_{boot}}$. With this in mind, and knowing that our time series are 460 points long, we chose $p_{boot} = 0.97$ so as to have, in the mean, 14

blocks of length 33. Figure 5.4 shows the test performed with different choices of p_{boot} and $B = 500$. Some differences can be observed, but these are small and, in general, the significant periods of time remain similar. The resulting critical levels are $ICCSI^*_\alpha = 0.0418$ for $p_{boot} = 0.95$, $ICCSI^*_\alpha = 0.0504$ for $p_{boot} = 0.97$, $ICCSI^*_\alpha = 0.0417$ for $p_{boot} = 0.98$ and $ICCSI^*_\alpha = 0.0484$ for $p_{boot} = 0.99$.

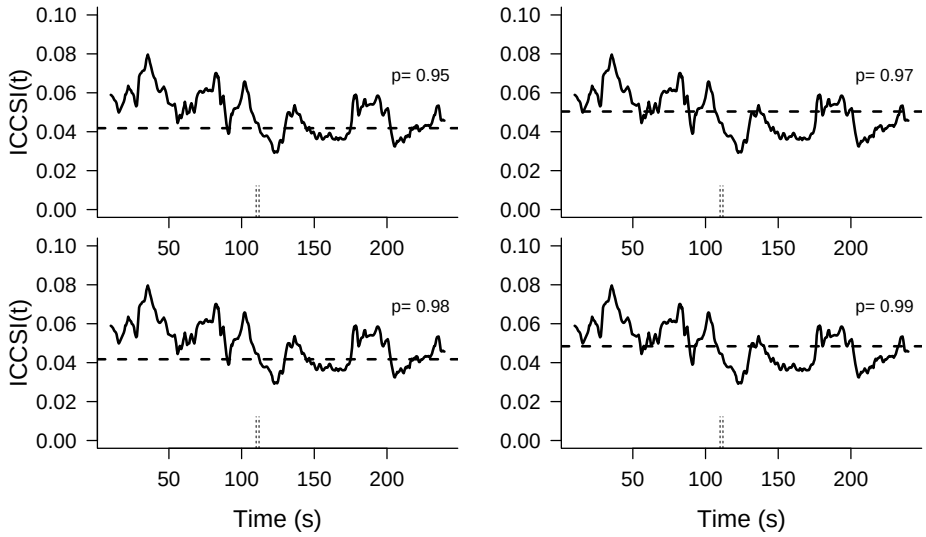

Figure 5.4: Bootstrap significance test for synchrony between neurons N1 and N3a. The horizontal dashed lines correspond to the bootstrap critical value, using different values of p_{boot}: $p_{boot} = 0.95$ (top-left panel), $p_{boot} = 0.97$ (top-right panel), $p_{boot} = 0.98$ (bottom-left panel) and $p_{boot} = 0.99$ (top-right panel).

To test for synchrony variations over time, the bootstrap procedure described in Section 5.3 was used. Figures 5.5 to 5.10 show the results for the existing synchrony between neurons for several pairs of neurons. The significance level $\alpha = 0.05$ has been used in the tests throughout the analysis. The results for the N1-N3a pair were already shown in Figures 5.3 and 5.4. In these cases our method is able to detect subtle changes in the synchronization dynamics. In the *bf* case, we can see a mild effect in some of the pairs of neurons. On the other hand, for *bs*, we can observe an immediate decrease of synchrony after stimulation that can be successfully detected by our test. The decrease in synchrony can be observed even before the stimulation is applied due to the use of moving windows in the method.

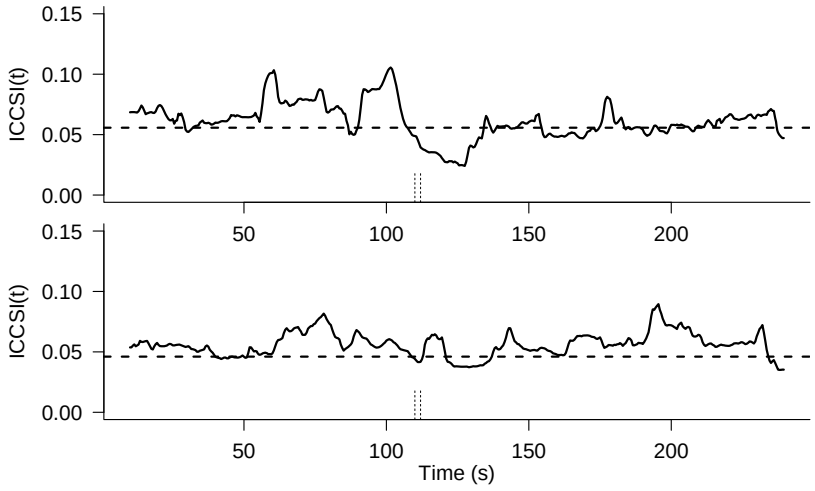

Figure 5.5: Estimated ICCSIs (solid lines) averaging over the three trials of *bs* stimulation (top panel) and *bf* stimulation (bottom panel) of neurons N1 and N3b. The bootstrap critical value is shown in horizontal dashed lines. The period of stimulation is indicated by the dashed vertical lines.

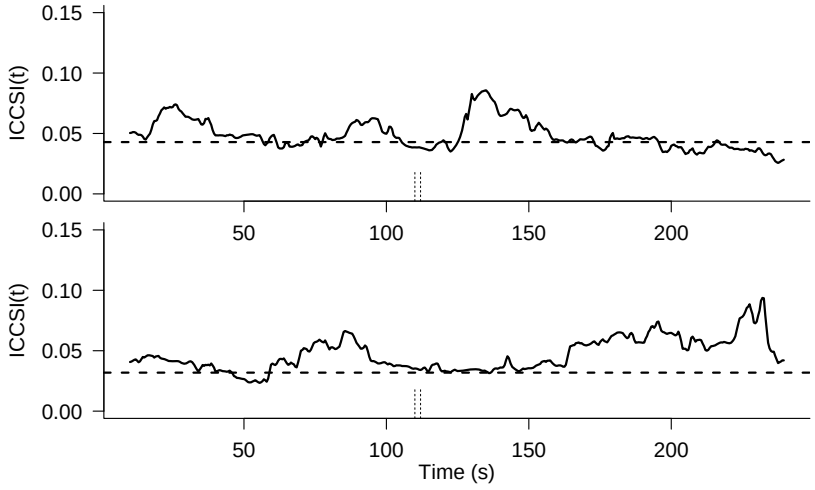

Figure 5.6: Estimated ICCSIs (solid lines) averaging over the three trials of *bs* stimulation (top panel) and *bf* stimulation (bottom panel) of neurons N1 and N4b. The bootstrap critical value is shown in horizontal dashed lines. The period of stimulation is indicated by the dashed vertical lines.

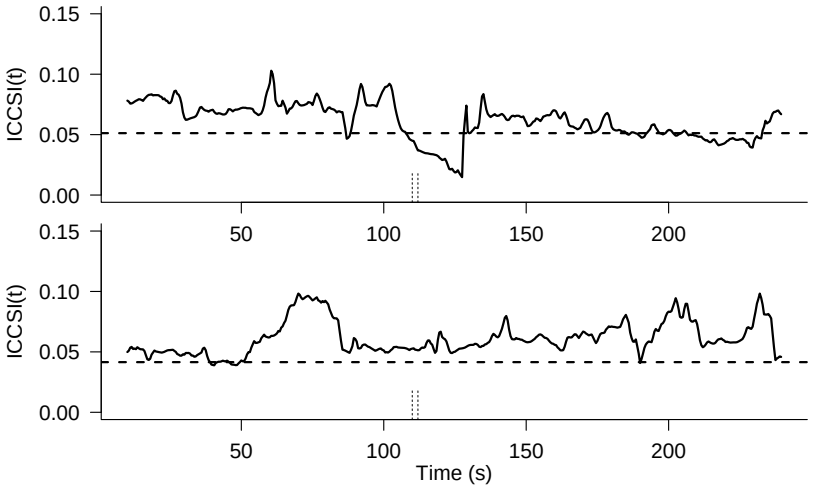

Figure 5.7: Estimated ICCSIs (solid lines) averaging over the three trials of *bs* stimulation (top panel) and *bf* stimulation (bottom panel) of neurons N1 and N5. The bootstrap critical value is shown in horizontal dashed lines. The period of stimulation is indicated by the dashed vertical lines.

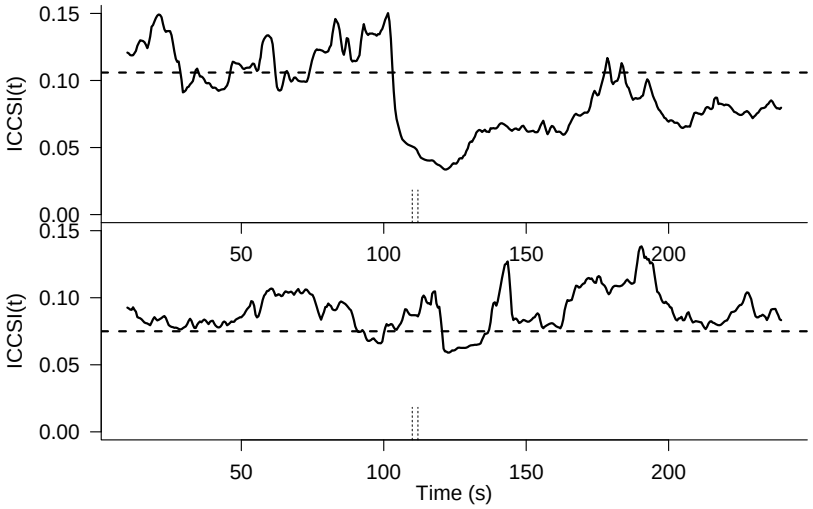

Figure 5.8: Estimated ICCSIs (solid lines) averaging over the three trials of *bs* stimulation (top panel) and *bf* stimulation (bottom panel) of neurons N3a and N3b. The bootstrap critical value is shown in horizontal dashed lines. The period of stimulation is indicated by the dashed vertical lines.

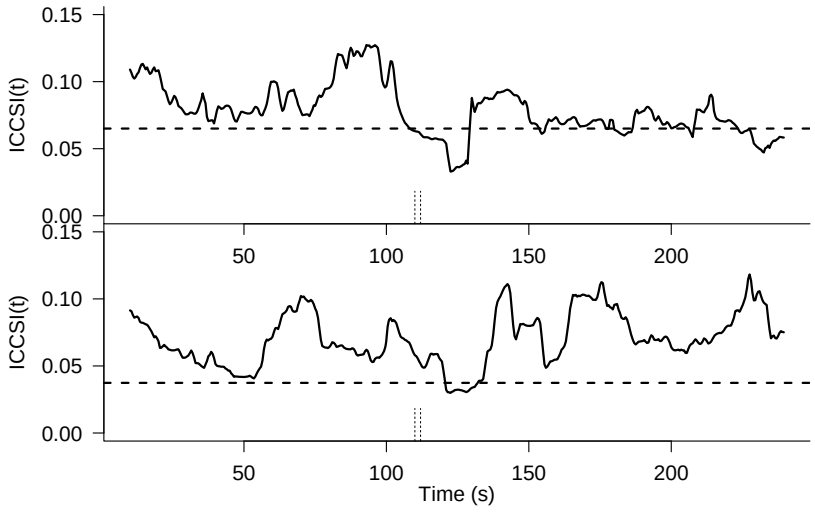

Figure 5.9: Estimated ICCSIs (solid lines) averaging over the three trials of *bs* stimulation (top panel) and *bf* stimulation (bottom panel) of neurons N3a and N4b. The bootstrap critical value is shown in horizontal dashed lines. The period of stimulation is indicated by the dashed vertical lines.

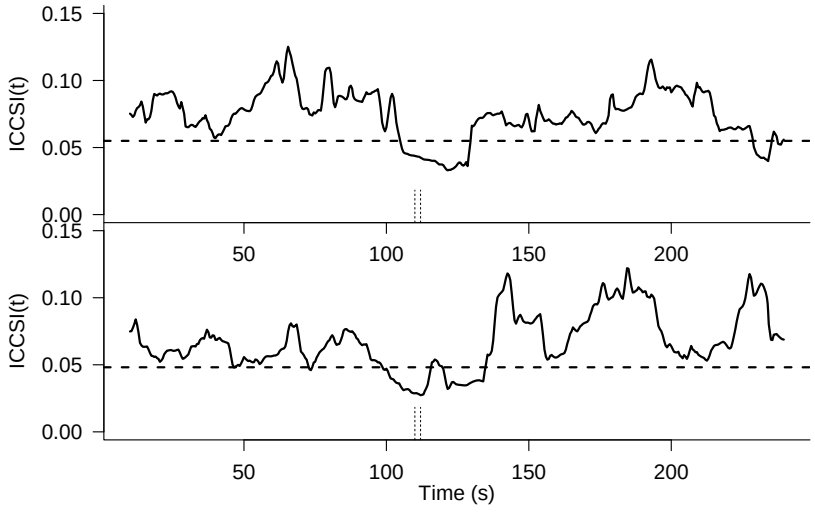

Figure 5.10: Estimated ICCSIs (solid lines) averaging over the three trials of *bs* stimulation (top panel) and *bf* stimulation (bottom panel) of neurons N3b and N4b. The bootstrap critical value is shown in horizontal dashed lines. The period of stimulation is indicated by the dashed vertical lines.

In Figures 5.11 to 5.13 we show the results obtained when testing for the effect of the applied stimulation (bs/bf) in the difference between the $\widehat{ICCSI}$. Difference $\widehat{ICCSI}$ curves are shown together with their 95% critical bands obtained by the bootstrap procedure described in Section 5.4. The figures show that our test can effectively detect differences between synchrony under stimulus bs and under stimulus bf for almost every pair of neurons. As expected, these differences are found mostly right after the stimulation.

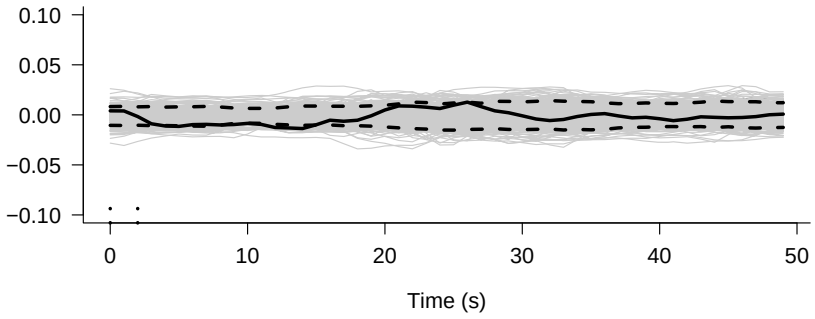

Figure 5.11: Differences (bs-bf) of $\widehat{ICCSI}(t)$ (solid black lines) with critical bands (dashed black lines) built with 500 bootstrap replications (grey lines) for neurons N1 and N3a. Stimulation time at $t = 0$.

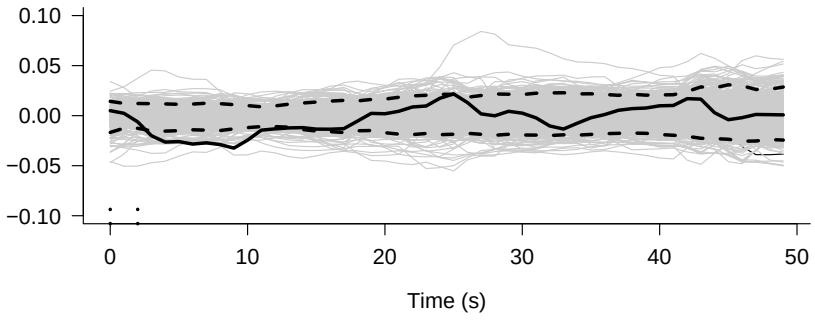

Figure 5.12: Differences (bs-bf) of $\widehat{ICCSI}(t)$ (solid black lines) with critical bands (dashed black lines) built with 500 bootstrap replications (grey lines) for neurons N1 and N3b. Stimulation time at $t = 0$.

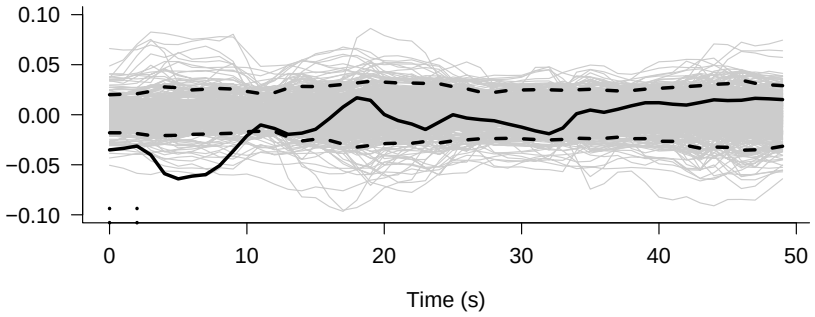

Figure 5.13: Differences (*bs-bf*) of $\widehat{ICCSI}(t)$ (solid black lines) with critical bands (dashed black lines) built with 500 bootstrap replications (grey lines) for neurons N3a and N5. Stimulation time at $t = 0$.

5.6 Simulation study

A simulation study was carried out to study the sensibility of our measure to sudden changes in synchrony. For this aim, we simulated pairs of spike trains controlling their association. We used an underlying Poisson process with rate $\lambda(t)$, say $M^0(t)$. To generate two spike trains from this underlying process, we assumed to have a realization of $M^0(t)$ with events at $X_1^0, \ldots, X_N^0$ and two vectors of random errors $\mu^1 = (\mu_1^1, \ldots, \mu_N^1)$ and $\mu^2 = (\mu_1^2, \ldots, \mu_N^2)$ with μ_i^j sampled from a uniform distribution chosen accordingly to the firing rate as explained later. Let $M^1(t)$ and $M^2(t)$ be the pair of spike trains induced by $M^0(t)$ as follows:

$$P(M^j(t) - M^j(t^-) = 1) = \begin{cases} p^j(t) & \text{if} \quad t = X_i^0 + \mu_i^j \quad \text{for some} \quad i = 1, \ldots, N \\ 0 & \text{otherwise} \end{cases}$$

for $j = 1, 2$, $i = 1, \ldots, N$ and $p^j(t)$ a certain probability function defined for each train.

Since we wanted to study dynamic changes in synchrony, we considered a time point, t_0, as the time where the association between the trains change. So, the probabilities were set constant before t_0 and also constant, although with a different value, after t_0. Also, for simplicity, we used the same probabilities for both trains. This is:

$$p^1(t) = p^2(t) = \begin{cases} p_1 & \text{if} \quad t < t_0 \\ p_2 & \text{if} \quad t \geq t_0 \end{cases} .$$

On the other hand, we defined the firing rate of the trains as constant throughout the trial, say λ_0. Therefore, the firing rate of the process $M^0(t)$ was defined as

$$\lambda(t) = \begin{cases} \lambda_0 p_1 & \text{if} \quad t < t_0 \\ \lambda_0 p_2 & \text{if} \quad t \geq t_0 \end{cases} .$$

In practice, we drew random numbers $\rho_i^j \in [0, 1]$ and then selected $X_i^0 + \mu_i^j$ as a spike for train j if $\rho_i^j \leq p_i^j$ (which occurs with probability p_i^j).

Finally, in our simulation study, 80 s spike trains with constant rate of $\lambda_0 = 4$ Hz were generated: 40 s were simulated with probability p_1 of acquiring the spikes from the underlying process and another 40 s with probability p_2. We used $\mu_i^j \sim U(-1/(10\lambda_0), 1/10\lambda_0))$ for all $i = 1, \ldots, N$ and $j = 1, 2$, in order to have a controlled error which shifts the spikes in a small amount but so that it was not likely that one spike would be shifted so much that it goes over another spike. The choices for the parameters p_1 and p_2 are, $p_1 = 0.9$ with $p_2 = 0.1, 0.3, 0.5$ and 0.7 on one hand and $p_1 = 0.7$ with $p_2 = 0.1, 0.3, 0.5$ and 0.65 on the other. We simulated 500 pairs of trains and estimated the $ICCSI$ function for them using the same parameters as for the real data ($\delta\nu = 0.025$ ms, $v = 10$ s.). Then we performed the bootstrap test described in Section 5.3 with $B = 500$ and $p_{boot} = 0.97$ the same as for the real data.

Figure 5.14 shows eight ICCSI curves from eight pairs of simulated spike trains with $p_1 = 0.7$ and $p_2 = 0.65$ in the top panel, whereas in the bottom panel the average of 500 of these curves are shown for different choices of p_2. Figures 5.15 and 5.16 show the rejection percentage of the null hypothesis using the bootstrap test in Section 5.3. Figure 5.15 shows the simulations for $p_1 = 0.9$ and Figure 5.16 for $p_1 = 0.7$. As it can be observed the test can easily detect the changes in synchrony. Of course the use of sliding windows provokes the existence of a period of time where the rejection percentage grows slowly but this is one of the prices we have to pay because of the low firing rates.

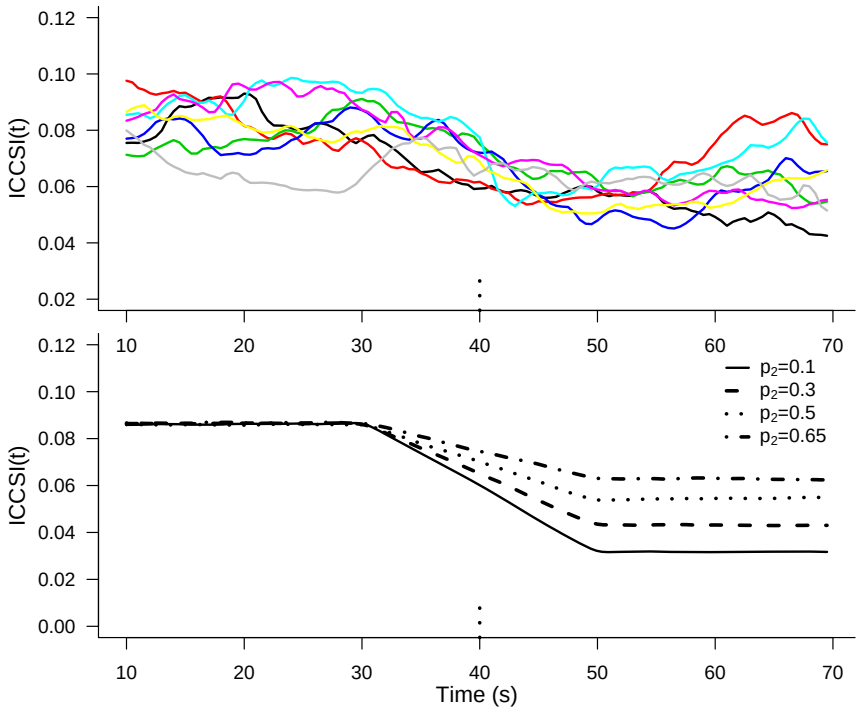

Figure 5.14: $ICCSI$ curves for eight simulated pairs of neurons using $p_1 = 0.7$ and $p_2 = 0.65$ (top panel) and average of 500 $ICCSI$ curves for $p_1 = 0.7$ and $p_2 = 0.1$ (solid line), 0.3 (dashed line), 0.5 (dotted line) and 0.65 (dashed-dotted line) (bottom panel). The stimulus is simulated at $t = 40$.

5.7 Chapter conclusions

The method presented in this chapter, the ICCSI, is based on the cross-correlation function. It is a flexible method because it permits the tunning of its parameters to better fit the problem. In order to asses for differences in synchronization between the sleep-like and the awake-like periods and between the two stimulation conditions, we proposed bootstrap based hypothesis tests. Two resampling and bootstrap procedures were presented. These resampling procedures take into account the dependence between simultaneous spike trains by resampling from the intervals of time that elapses between spikes of a joint spike train built by merging the spike trains. A simulation study illustrates the good behavior of these tests. When applied to the real data, the methods prove to be useful to address the problems posited by the neuroscientists. In the sample used for our study, significant decreases in neural synchrony are found after the stimulation of *bs*. Never-

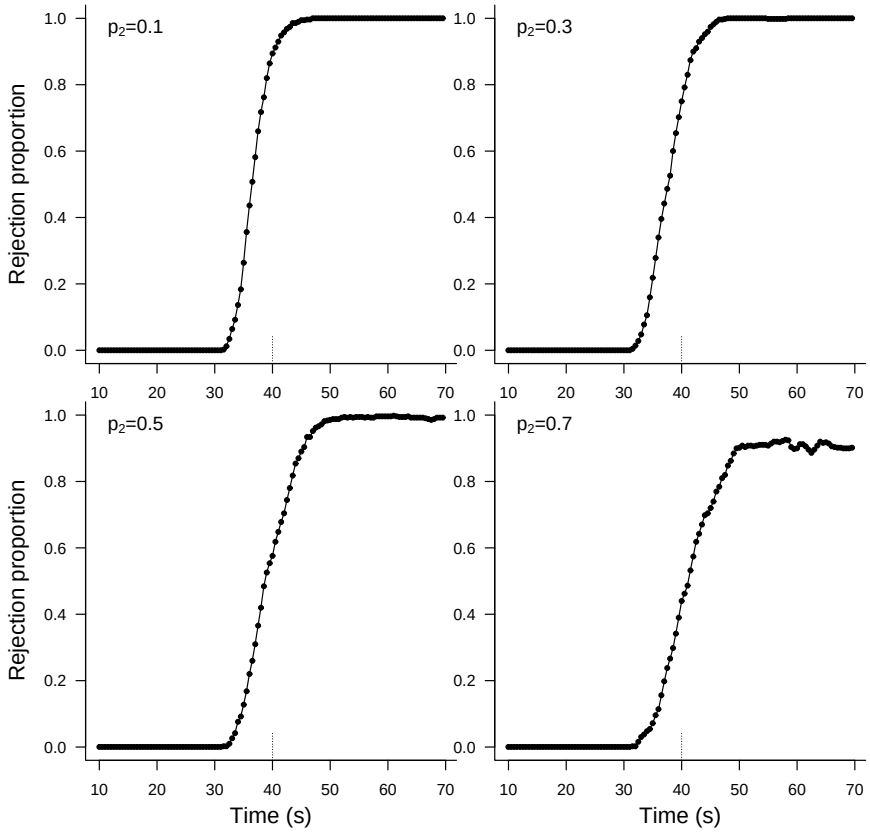

Figure 5.15: Rejection proportion of the bootstrap test for changes in synchrony using $p_1 = 0.9$ and different $p_2 = 0.1, 0.3, 0.5$ and 0.7. The stimulus is simulated at $t = 40$.

theless, there is not enough evidence to detect these kind of differences in every pair of neurons in the *bf* case. On the other hand, a second test permits the observation of differences between the two used stimuli. This is, the bootstrap test results allow to state that the observed synchrony during the awake-like period is different depending on the stimulus.

Although we did not observe the same behavior in all the studied pairs of neurons, we can asses that the stimuli disrupts, in a differential way, the neural synchrony for a short period of time, and then the normal synchrony is slowly recovered.

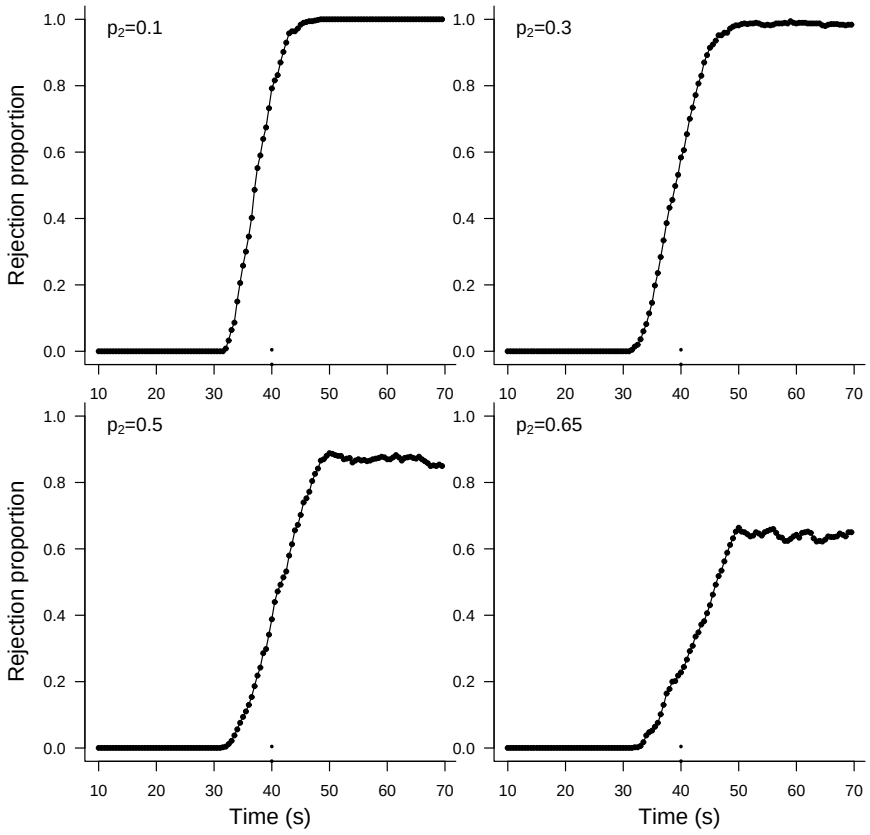

Figure 5.16: Rejection proportion of the bootstrap test for changes in synchrony using $p_1 = 0.7$ and different $p_2 = 0.1, 0.3, 0.5$ and 0.65. The stimulus is simulated at $t = 40$.

Chapter 6

Cross nearest spike interval based method to measure synchrony

In this chapter we present a different method to measure synchrony between pairs of neurons. This method is also based on spikes times rather that binned trains. The aim is to solve the same problem as with the ICCSI but using a different approach.

Along this chapter, cross nearest spike intervals are defined, depending on some tuning parameters whose selection is also discussed. A synchrony measure based on these intervals is proposed and fitted to the data using generalized additive models. Relevant hypothesis tests concerned with synchrony are proposed and calibrates using a theoretical approximation and also by simulation. The bootstrap method is also used to construct confidence bands. All these methods are applied to the data studied in this thesis.

6.1 Synchrony measure based on cross nearest spike intervals

Let $\mathcal{X} = \{X_j\}_{j=1}^{J_1}$ and $\mathcal{Y} = \{Y_j\}_{j=1}^{J_2}$ be two simultaneously recorded spike trains as in the previous chapters. Let $N_X(t)$ and $N_Y(t)$ for $t \in [0, T)\}$ be the counting processes associated to $\mathcal{X}$ and $\mathcal{Y}$ respectively. Recall from Chapter 3 that $N_X(t) = \#\{X_j \leq t \; j = 1, \ldots, J_1\}$ and $N_Y(t) = \#\{Y_j \leq t \; j = 1, \ldots, J_2\}$. Also, let $n = J_1 + J_2$ be the total amount of spikes.

Define the cross nearest spike interval (CNSI) as the time that elapses between the spikes of one neuron and the closest spike of the other neuron, not necessarily forward. If we denote the CNSI variable as $\tilde{U}$ then, $\tilde{U} = \min\{U_{-1}, U_1\}$, where U_{-1} and U_1 are the *waiting times* defined in Chapter 5 and U_1 is CISI defined in Chapter 4.

As already discussed, when the firing rate of spike trains is low, we need a broader definition of synchrony because exact firing matches hardly ever occur in 1 ms windows, not even in 10 ms windows. Here, synchrony is defined as the event of AP occurring within a time window of width δ that can be selected by the researcher according to the problem, just as with the ICCSI. Issues about the selection of δ are discussed in the next section. Let n_δ be the number of CNSIs smaller than or equal to δ, counting from both neurons. So, we define the following measure, which we will call CNSI-Synchrony Measure (CSM) and denote with p^δ:

$$p_\delta = \frac{n_\delta}{n},$$

where n is total amount of spikes, over both neurons, and therefore, the total amount of CNSIs too. The measure p_δ is a global measure of synchrony. It is symmetric and it does not pick up causality. In terms of the counting processes, n_δ can be expressed as:

$$n_\delta = \sum_{j=1}^{J_1} \mathbb{I}\{N_Y(X_j+\delta) - N_Y(X_j-\delta) \geq 1\} + \sum_{j=1}^{J_2} \mathbb{I}\{N_X(Y_j+\delta) + N_X(Y_j-\delta) \geq 1\}$$

where, recall that, $\mathbb{I}\{A\}$ is the indicator function of event A.

In order to define CSM as a function of time, $p_\delta(t)$, to take into account the non stationarity of the processes we consider $n_\delta(t)$ as follows:

$$\begin{aligned} n_\delta(t) = &\sum_{j=1}^{J_1} \mathbb{I}\{N_Y(X_j+\delta) - N_Y(X_j-\delta) \geq 1)\}\mathbb{I}\{X_j = t\} + \\ &+ \sum_{j=1}^{J_2} \mathbb{I}\{N_X(Y_j+\delta) + N_X(Y_j-\delta) \geq 1\}\mathbb{I}\{Y_j = t.\} \end{aligned}$$

Of course, $X_j = t$ and $Y_j = t$ are events of probability zero but, even in practice, when working with binned trains, observing a spike in a given bin is very difficult when firing rates are low. So, in order to solve this problem,

we will work with the information provided by a neighborhood of t. Let $V_t = (t-v, t+v]$ be a symmetric time window of length $2v$ around t, with v chosen by the researcher depending on the context of the study. For higher firing rates, smaller windows could be used. Therefore, we define

$$\begin{aligned} n_\delta(t,v) = &\sum_{j=1}^{J_1} \mathbb{I}\{N_Y(X_j+\delta) - N_Y(X_j-\delta) \geq 1\}\mathbb{I}\{X_j \in (t-v, t+v]\} + \\ &+ \sum_{j=1}^{J_2} \mathbb{I}\{N_X(Y_j+\delta) + N_X(Y_j-\delta) \geq 1\}\mathbb{I}\{Y_j \in (t-v, t+v]\} \end{aligned} \tag{6.1}$$

and let $n(t,v)$ be the number of spikes (and CNSIs) that fall in V_t:

$$n(t,v) = \sum_{j=1}^{J_1} \mathbb{I}\{X_j \in (t-v, t+v]\} + \sum_{j=1}^{J_2} \mathbb{I}\{Y_j \in (t-v, t+v]\}$$

For each t, we compute the CSM in the time window V_t obtaining $p_\delta(t)$. Consequently, to measure synchrony at a time point t_0, the probabilities $P(N_Y(X_j+\delta) - N_Y(X_j-\delta) \geq 1)$ and $P(N_X(Y_j+\delta) - N_X(Y_j-\delta) \geq 1)$ are considered constant along V_{t_0} and estimated using the information of the whole interval. In this way, the local low firing rate is outweighed by neighborhood activity.

Population-wise talking, $p_\delta(t)$ is an estimator of the probability, $\pi_\delta(t)$, of (given two trains) finding two spikes (one of each neuron) closer than the quantity δ, given that occurred one spike at time t. In general, π_δ can be considered as the probability of success of a certain Bernoulli trial, where the trial corresponds with one observation of the $\tilde{U}$ with success probability that $\tilde{U}$ is smaller than δ. Therefore, n_δ becomes the observation of a binomial variable, η_δ, that counts the number of successes of the Bernoulli trials, $\eta_\delta \sim B(N, \pi_\delta)$ and consequently, p_δ is an estimator of the probability π_δ, $p_\delta = \hat{\pi}_\delta$. Finally, the same argument can be used at each time window V_t to obtain $p_\delta(t)$ as the estimator of $\pi_\delta(t)$ where $\eta_\delta(t) \sim B(N(t), \pi_\delta(t))$. In the rest of the chapter we will drop the subscript δ except for the quantity n_δ, because there is no possibility of confusion as everything depends on this quantity in the same way.

6.2 Selection of V_t and δ

When the firing rate is high, or a large amount of trials are available, both windows V_t and δ can be chosen to be small. On the other hand, if there are few trials or the firing rates are low, wider windows need to be chosen to ensure that enough spikes are present in the intervals. How to choose the quantities V_t and δ is a matter of discussion. We propose a compromise between V_t and δ and the use of the smallest values that will allow computations of the CSM. Figure 6.1 shows an example carried out with real data. In this figure, we observe the amount of zeros (no presence of CNSIs smaller than δ) encountered while computing the CSM on a 150 s trial of a real neuron. We have repeated computations for 45 pairs (v,δ): $v = 3, 4, 4.5, 5, 5.5$ seconds and $\delta = 0.005, 0.01, 0.025, 0.03, 0.035, 0.04, 0.045, 0.05, 0.1$. Each line corresponds to one value of v. We see that the amount of zeros for small windows together with small δ is large. Figure 6.2 shows CSM for four different selections of the pair (δ, v) for the real data example. As the firing rates of these neurons are very low, we can see that very small δ and v would result in a lot of zeros and even sometimes not allow for the calculation of CSM because of lack of spikes in a given window. When a large v, $v = 5$, and a large δ, $\delta = 0.1$ are chosen, the CSM grows considerably because the proportion of CNSIs smaller than δ is large. Finally, when a large v ($v = 10$) and $\delta = 0.05$ are chose, the CSM flattens probably due to the fact that in a wide range of time points, very large pieces of the spike trains are used repeatedly. In this case we have chosen $v = 5$ and $\delta = 0.05$ with the results as shown in the top-right panel of the figure.

The values of CSM are clearly influenced by the choice of the parameters v and δ. Therefore it is important to asses how much of the synchrony observed is real and how much of it is a product of the choice of these parameters or due to independent neuron firing. We will discuss this issue in Section 6.4.

6.3 Model formulation

In this section we present a model for our measure. As already discussed, at each time window, the proposed measure is an estimator of the probability of a binomial process where success is the event of observing $\tilde{U}$ smaller than δ and the total number of CNSIs is the number of Bernoulli trials, $\eta(t) \sim B(n(t), \pi(t))$. We propose to use binomial generalized additive models (GAM) with a logit link to explain the proportion of 'small' CNSIs in time. The convenience of using GAMs lies in the flexibility of non-parametric

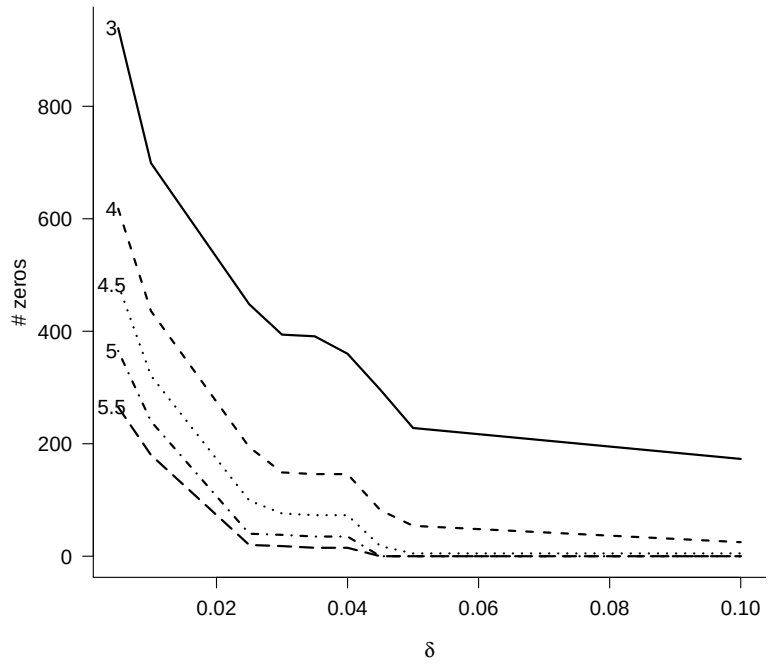

Figure 6.1: Number of zeros obtained in the calculation of CSM when using different δ (x-axis) and different v for the time windows ($(t-v, t+v]$): $v = 3$ (solid line), $v = 4$ (short-dashed line), $v = 4.5$ (dotted line), $v = 5$ (dot-dashed line) and $v = 5.5$ (long dashed line).

functions. It would be not reasonable to choose a parametric model for the response function. The additive terms in the model allows for an easy interpretation.

In general, GAMs can be represented as follows:

$$g(\mu) = \mathbf{S}^*\alpha + f_1(x_1) + f_2(x_2) + \ldots + f_m(x_m)$$

where $\mu = E(Y)$, Y is a random variable belonging to the exponential family of distributions, g is called the link function, $\mathbf{S}^*$ is the model matrix for the parametric component of the model, α is the corresponding parameter vector and the f_j are smooth functions of the covariates x_j (Hastie and Tibshirani (1990); Wood (2006)).

For our problem, we propose the use of logistic GAMs. In the particular case of a binomial GAM, the canonical link function used is the logit:

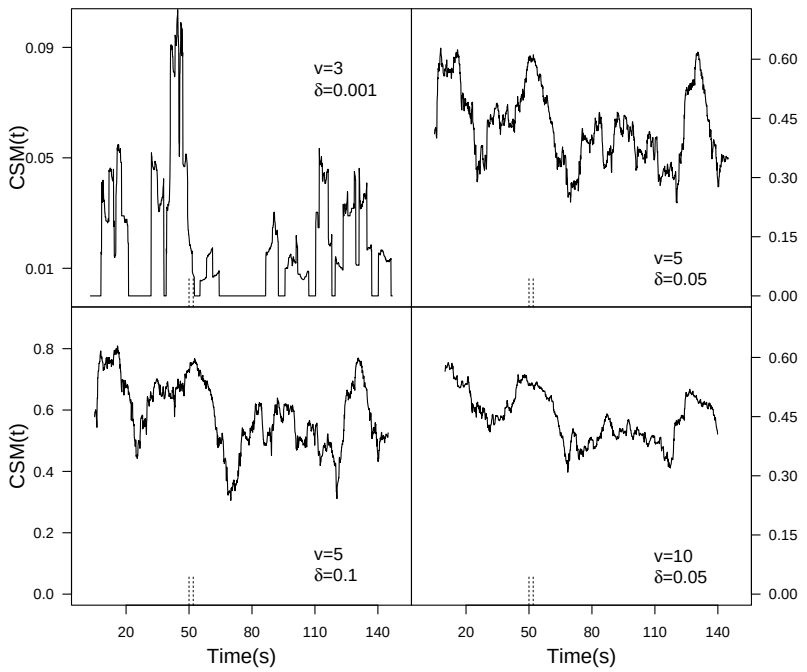

Figure 6.2: $CSM(t)$ for different δ and v. $\delta = 0.001$, $v = 3$ (top-left panel); $\delta = 0.05$, $v = 5$ (top-right panel); $\delta = 0.1$, $v = 5$ (bottom-left panel); $\delta = 0.05$, $v = 10$ (bottom-right panel).

$\text{logit}(\mu) = \frac{\mu}{1-\mu}$. Also, it is important to recall that we have several experimental conditions (induced by the two stimuli) which are suppesed to drive to different probabilities of joint firing. The model we propose is a *mean curve plus difference* one, which we will call Model 1:

$$\text{logit}(\pi_j(t)) = \beta_0 + \sum_{l=1}^{J-1} \beta_l \mathbb{I}\{j = l\} + f_0(t) + \sum_{l=1}^{J-1} f_l(t)\mathbb{I}\{j = l\} + \epsilon_t,$$

where $\pi_j(t)$ stands for the success probability in $\eta_j(t) \sim B(N_j(t), \pi_j(t))$ under the j-th experimental condition. The parametric part is reduced only to a intercept and $f_0(t)$ is a common curve for all the conditions, while $f_j(t)$ represents the difference from the curve $f_0(t)$ to the one corresponding to condition $j = 1, \ldots, J-1$. Finally, the ϵ_t are the errors of the model which will be discussed later on.

In practice, we only have two conditions, $J = 2$, and we have three repetitions for each condition. Let $n_{\delta ji}(t)$ and $n_{ji}(t)$ be the observed n_δ and

total number of CNSIs observed in the time window V_t under the i-th trial of the j-th condition respectively, for $j = 1, 2$ and $i = 1, 2, 3$. Therefore, $p_{ji}(t) = \frac{n_{\delta ji}(t)}{n_{ji}(t)}$ is the CSM at time t for the i-th trial of the j-th condition. So, the model to fit is:

$$\text{logit}(p_{ji}(t)) = \beta_0 + \beta_1 \mathbb{I}\{j = 1\} + f_0(t) + f_1(t)\mathbb{I}\{j = 1\}, \tag{6.2}$$

where, after obtaining $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{f}_0$ and $\hat{f}_1$ we will have an estimator of $\pi_j(t)$, $\hat{\pi}_j(t)$, for each condition.

We are interested in the time evolution of synchrony and therefore in the differences that could occur in certain periods of time. Nonetheless, a first overall study of the synchrony profile is cautious. To do this, we will fit a simple model, a *single curve* model (Model 2), and compare it with Model 1:

$$\text{logit}(p_{ji}(t)) = \beta_0 + f_0(t). \tag{6.3}$$

Model 2 in (6.3) corresponds to the hypothesis of no differences between experimental conditions over the whole time period. It comprises a single function for both conditions, while Model 1 in (6.2) allows the smooth functions to differ. With the present approach, in a two experimental conditions model, a possible difference between the two conditions is easier to observe because, if there are no differences between conditions then $\hat{f}_1(t)$ will be essentially flat.

To represent the smooth functions we have used penalized regression splines, with cubic splines basis. Smoothing parameters are chosen by the generalized cross validation (GCV) criterion and the number of knots using the akaike information criterion (AIC) value. The `mgcv` R package was used for this aim (Wood (2006) pp.128–133).

An important issue to note is that no dependency between time points is taken into account with these proposed models. Nevertheless, we are aware of the high dependency that exists between consecutive time points. The dependency will be taken into consideration when building confidence bands for the estimators and when developing the hypothesis tests. For this aim, we consider that the dependency is gathered in the errors of the model and we propose to model this dependency with an AR(1) model:

$$\epsilon_t = \alpha\epsilon_{t-1} + a_t \quad \text{with} \quad a_t \sim N(0, \sigma^2).$$

We can estimate α by ordinary least squares. The estimated errors are

$\hat{\epsilon}_{ji}(t) = p_{ji}(t) - \hat{\pi}_j(t)$ and we need to find α so as to minimize the sum of squares:

$$\sum_{j=1}^{2}\sum_{i=1}^{3}\sum_{t=2}^{T}(\hat{\epsilon}_{ji}(t) - \alpha\hat{\epsilon}_{ji}(t-1))^2.$$

Differentiating this expression in α and making the derivative equal to zero it is straight forward to prove that the $\hat{\alpha}$ we are looking for is

$$\hat{\alpha} = \frac{\sum_{j=1}^{2}\sum_{i=1}^{3}\sum_{t=2}^{T}\hat{\epsilon}_{ji}(t)\hat{\epsilon}_{ji}(t-1)}{\sum_{j=1}^{2}\sum_{i=1}^{3}\sum_{t=2}^{T}\hat{\epsilon}_{ji}^2(t-1)}.$$

Now we can estimate σ^2 by $\hat{\sigma}^2 = \widehat{Var}(\hat{a}_{ji}(t))$, with $\hat{a}_{ji}(t) = \hat{\epsilon}_{ji}(t) - \hat{\alpha}\hat{\epsilon}_{ji}(t-1)$.

6.3.1 Hypothesis testing

Let us state the hypothesis regarding the electrical stimulation in *bs* and *bf* to be tested in the context of the present chapter. Considering the two experimental conditions, *bs*, denoted by 1, and *bf*, denoted by 2, the two main null hypotheses to test are

- $H_0^1 : \pi_1(t) = \pi_0$ for every t after the condition onset. Here, π_0 represents the baseline synchrony before the stimulus. A similar hypothesis can be formulated for condition 2.
- $H_0^2 : \pi_1(t) = \pi_2(t)$. With this test we aim to detect differences in the synchrony profile induced by the two different experimental conditions.

To test these hypotheses we will use CSM as a test statistic and bootstrap critical bands to find significant differences. The bootstrap procedure will be discussed in Section 6.7.

A third important aspect to investigate is whether the observed synchrony is different from the one that would be expected just by chance. To test this we will compare the observed synchrony with the expected one. This problem is presented in the next section.

6.4 Synchrony due to firing rate

Even in the case of two independent spike trains, CSM will not be zero, and its value could be large due to random close firing. It seems reasonable that

this synchrony obtained by chance will increase with increasing firing rates. In the way the CSM is built, it could be affected by the firing rate and by its changes in time. To study this aspect we will compare the observed CSM with the one expected only by chance, due to the firing rate. For this purpose we need to calculate what CSM values are expected for the observed activity if the two trains were independent. These values will be approximated in this chapter in two different ways: 1) approximating the theoretical expected synchrony by chance and 2) by simulating independent trains and modeling their synchrony.

6.5 Theoretical approximation

Let us start with the theoretical expression for the CSM:

$$A_i(t,\nu) = \{s \in (t-\nu, t+\nu]/N_i(s) > \overline{N}_i(s)]\}$$

where $\overline{N}_i(s) = \lim_{\tau \to s^-} N_i(\tau)$ and $i \in \{X, Y\}$. Observe that $N_i(t+\nu) - N_i(t-\nu) \geq 1$ if and only if $A_i(t,\nu) \neq \emptyset$. Therefore, we can think of the quantity

$$\sum_{j=1}^{J_1} \mathbb{I}\{N_Y(X_j+\delta) - N_Y(X_j-\delta) \geq 1\}\mathbb{I}\{X_j \in (t-v, t+v]\}$$

in (9.3) as an estimator of the expected value of:

$$\#\{s \in A_X(t,v)/A_Y(s,\delta) \neq \emptyset\} = \sum_{s \in A_X(t,v)} \mathbb{I}\{A_Y(s,\delta) \neq \emptyset\}.$$

Taking expectations and using the property $E(A) = E(E(A|B))$ we have:

$$E\left[\sum_{s\in A_X(t,v)} \mathbb{I}\{A_Y(s,\delta) \neq \emptyset\}\right] = E\left[E\left(\sum_{s\in A_X(t,v)} \mathbb{I}\{A_Y(s,\delta) \neq \emptyset\}\middle| N_X\right)\right] =$$
$$= E\left[\sum_{s\in A_X(t,v)} E\left(\mathbb{I}\{A_Y(s,\delta) \neq \emptyset\}\middle| N_X\right)\right],$$

which, under the assumption of independence, becomes:

$$E\left[\sum_{s\in A_X(t,v)} E\left(\mathbb{I}\{A_Y(s,\delta) \neq \emptyset\}\right)\right] = E\left(\sum_{s\in A_X(t,v)} \rho_{\delta Y}(s)\right)$$

where $\rho_{\delta Y}(s) = E(\mathbb{I}\{A_Y(s,\delta) \neq \emptyset\})$ which is the expected value for the indicator that there is a 'jump' in process 2 in $(s-\delta, s+\delta]$. As we are using the information of the whole interval $(t-v, t+v]$ to compute n^δ, we can assume the process is stationary in this interval, so, $\rho_{\delta Y}(s) \equiv \rho_{\delta Y}$ constant in $(t-v, t+v]$. If this is not the case, but v is sufficiently small, then, $\rho_{\delta Y}(s)$ can be approximated by a constant value, for example, the value of $\rho_{\delta Y}(s)$ at the middle point of the interval: $\rho_{\delta Y}(s) \approx \rho_{\delta Y}(t)$.

From the previous paragraph we have,

$$E\left(\sum_{s\in A_X(t,v)} \mathbb{I}\{A_Y(s,\delta) \neq \emptyset\}\right) \approx \rho_{\delta Y}(t) E\left(\sum_{s\in A_X(t,v)\neq\emptyset} 1\right) = \rho_{\delta Y}(t) r_X(t,v)$$

where $r_X(t,v) = E(N_X(t+v) - N_X(t-v))$.

The previous discussion also holds for the quantity

$$\sum_{j=1}^{J_1} \mathbb{I}\{N_Y(X_j+\delta) - N_Y(X_j-\delta) \geq 1\}\mathbb{I}\{X_j \in (t-v, t+v]\}$$

in (9.3). So finally we get that, under independence,

$$E(n_\delta) \approx \rho_{\delta Y} r_X(t,v) + \rho_{\delta X} r_Y(t,v),$$

where, $\rho_{\delta X} = E(\mathbb{I}\{A_X(s,\delta) \neq \emptyset\})$ and $r_Y(t,v) = E(N_Y(t+v) - N_Y(t-v))$.

Finally, we estimate $\rho_{\delta i}$, $i \in \{X, Y\}$ by,

$$\hat{\rho}_{\delta i} = \frac{1}{2v} \int_{t-v}^{t+v} \mathbb{I}\{N_i(s+\delta) - N_i(s-\delta) \geq 1\}\mathrm{ds}$$

and $r_i(t,v)$, $i \in \{X, Y\}$ by,

$$\hat{r}_i(t,v) = N_i(t+v) - N_i(t-v).$$

On the other hand, we have that the total number of CNSIs, $n(t,v)$, in the time window $(t-v, t+v]$ is $n(t,v) = r_X(t,v) + r_Y(t,v)$ and therefore,

$$\frac{n_\delta(t,v)}{n(t,v)} \approx \frac{\rho_{\delta Y} r_X(t,v) + \rho_{\delta X} r_Y(t,v)}{r_X(t,v) + r_Y(t,v)} \tag{6.4}$$

is an approximation to the CSM under the independence hypothesis.

6.6 Simulation approximation

The expected random synchrony under independence given certain neuronal activity can be also approximated using simulation. For this, we assume there exists an underling function, g, of synchrony dependent on the firing rates of the two spike trains,r_1 and r_2, this is:

$$g(r_1, r_2) = E\left(\text{synchrony between trains } X \text{and } Y \text{ under independence} | r_1, r_2\right)$$

We smooth this function with two dimensional splines after approximating it by simulation as follows. Given two values of instant probability of firing, p_1 and p_2, we simulate independent spike trains with the corresponding firing rates and compute the CSM for the simulated trains. This procedure is repeated for every pair (p_1, p_2) in a two dimensional grid, P, spanning from 1 Hz to 100 Hz. Going from 2 to 10 Hz every 1 Hz, from 12 to 60 every 2 Hz and from 65 to 100 every 5 Hz.

To simulate a spike train with a global firing rate of p_1, 500 s time intervals are divided in 1 ms bins. In each bin, a Bernoulli trial with success probability p_1 (spiking) is drawn. Let $\{b_i\}_{i=1}^{M}$ be the indices of the bins in which spikes are randomly assigned. Then, the spike times (in ms) for the simulated train are $X_j^* = 0.001 b_j \; j = 1, \ldots, M$. The same procedure is used to simulate trains with global firing rate equal to p_2. Using the two simulated spike trains with firing probabilities p_1 and p_2, the CSM is calculated and an average over the 500 s is considered as the approximation to the value of g at (p_1, p_2). Notice that this simulation scheme is not independent of δ and v and therefore it needs to be carried out for each choice of these two parameters.

Finally, the smoothed g is used to calculate the expected synchrony, under the assumptions of independence, for the real pair of neurons. At each time point, the instant firing rate of each real train are estimated and those values evaluated in $g(r_1, r_2)$ to obtain the expected value.

6.7 Bootstrap confidence bands and testing for differences

A bootstrap procedure is carried out to build confidence bands for the predictions of the selected model and also for the predictors under the null hypothesis described in Section 6.3.1. The bootstrap procedure is now described to build confidence bands for the estimators and later the procedure

is slightly changed to test the hypotheses. The procedure for I trials is the following:

1. Fit the penalized regression model to the response data $p_{ji}(t)$ and obtain the fitted probabilities $\hat{\pi}_j(t)$ for each condition j.

2. Compute the errors for each trial i: $\hat{\epsilon}_{ji}(t) = p_{ji}(t) - \hat{\pi}_j(t)$, $i = 1, \ldots, I$.

3. Estimate α and σ^2 in the AR(1) model for the residuals: $\hat{\epsilon}_j(t) = \alpha\hat{\epsilon}_j(t-1) + a_t$ (as described in Section 6.3).

4. Build bootstrap errors $\epsilon^*_{ji}(t) = \hat{\alpha}\epsilon^*_{ji}(t-1) + z_t$, with $\epsilon^*_{ji}(1) = \hat{\pi}_j(1) - \frac{Y_j}{N_{ji}(1)} + z_1$, $Y_j \sim B(N_1, \hat{\pi}_j(1))$ and $z_t \sim N(0, \hat{\sigma}^2)$.

5. Compute the bootstrap data by adding the bootstrapped errors to the fitted model: $p^*_{ji}(t) = \hat{\pi}_j(t) + \epsilon^*_{ji}(t)$.

6. Fit the regression model from Step 1 to the bootstrap data to obtain the bootstrap synchrony curve $\pi^*_j(t)$.

7. Repeat Steps 4–6 B times to obtain B bootstrap curves for each condition: $\pi^{*1}_j(t), \ldots, \pi^{*B}_j(t)$.

To build the confidence bands for the estimators the only step left is to compute the quantiles $\alpha/2$ and $1-\alpha/2$ at each time point t to obtain $(1-\alpha)\%$ confidence bands at each condition $j = 1, 2$.

For the hypothesis tests, the procedure changes slightly mainly because the model to be fitted has to be the one under the null hypothesis. It is worth to note that the errors used to fit the AR model that will be used to build the bootstrap samples are the ones obtained by fitting the general model (6.2). Therefore, the procedure for hypothesis testing with I trials is the following:

To fit the AR(1) for the errors:

1. Fit the general penalized regression model (as before) to the response data $p_{ji}(t)$ and obtain the fitted probabilities $\hat{\pi}^g_j(t)$ for each condition j.

2. Compute the errors for each trial i: $\hat{\epsilon}_{ji}(t) = p_{ji}(t) - \hat{\pi}^g_j(t)$, $i = 1, \ldots, I$.

3. Estimate α and σ^2 of the AR(1) model for the errors: $\epsilon_j(t) = \alpha\epsilon_j(t-1) + a_t$ (as described in Section 6.3).

Build the bootstrap samples under the null:

4. Fit the penalized general regression model under the null hypothesis to the response data $p_{ji}(t)$ and obtain the fitted probabilities $\hat{\pi}_j^0(t)$ for each condition j.

5. Build bootstrap errors $\epsilon_{ji}^*(t) = \hat{\alpha}\epsilon_{ji}^*(t-1) + z_t$, with $\epsilon_{ji}^*(1) = \hat{\pi}_j(1) - \frac{Y_j}{N_{ji}(1)} + z_1$, $Y_j \sim B(N_1, \hat{\pi^g}_j(1))$ and $z_t \sim N(0, \hat{\sigma}^2)$.

6. Compute the bootstrap data by adding the bootstrapped errors to the null model: $p_{ji}^*(t) = \hat{\pi}_j^0(t) + \epsilon_{ji}^*(t)$

7. Fit the regression model from Step 1 to the bootstrap data to obtain the bootstrap synchrony curve $\pi_j^*(t)$.

8. Repeat Steps 5–7, B times to obtain B averaged bootstrap curves for each condition: $\pi_j^{*1}(t), \ldots, \pi_j^{*B}(t)$.

For the first hypothesis test, the model fitted in Step 4 is the one that corresponds to the null hypothesis of constant synchrony before stimulation, the baseline model:

$$\text{logit}(\pi_0(t)) = \beta_0$$

and it is fitted using only the data from the period previous to stimulation. Once the bootstrap curves are obtained, compute the desired quantiles to build confidence bands around the baseline model for each condition.

For the second hypothesis test we need a further step. The null hypothesis H_0^2 states that $\pi_1(t) = \pi_2(t)$. This means that $\pi_1(t) - \pi_2(t) = 0$. So, in this case, we want to build confidence bands for the difference of the synchrony curves. First of all, in Step 4, we need to fit the model which represents this hypothesis. This is Model (6.3) described in Section 6.3. Then, we will subtract the bootstrap curves to obtain:

$$\bar{\pi}_{dif}^{*b}(t) = \bar{\pi}_1^{*b}(t) - \bar{\pi}_2^{*b}(t), \qquad m = 1, \ldots, B.$$

After this additional step we will compute the desired quantiles to build confidence bands for zero (difference under the null hypothesis) and see whether the observed difference curve $\hat{\pi}_{dif}(t) = \hat{\pi}_1(t) - \hat{\pi}_2(t)$ falls inside the band.

Table 6.1: *AIC values for the two models.*

	Model 1	Model 2
AIC	145134	163071.8

6.8 Results

The main interest of studying synchrony between neurons lies in its time evolution. In this section we present the results when the methods are applied to the pair of neurons N1 and N3a. The aim of the study, just as in Chapter 5, is to determine in what periods of time the CSM profiles differ.

The models described in Section 6.3 were fitted and the results are presented here. In this context, we have two experimental conditions which stand for stimulation in *bs* and *bf*. In Table 6.1 the AIC values for each model can be observed and, based on them, we can say that Model 1 is a better one.

In Figure 6.3 we can observe that the curve that accounts for the difference between experimental conditions in the GAM, $f_1(t)$, is different from zero, this means that there are differences between the conditions and different smooth terms are needed for each stimulus. This hypothesis is tested statistically below using the bootstrap procedure discussed in Section 6.7. Figure 6.4 shows the bootstrap confidence bands built for each of the estimators using $B = 500$ bootstrap replicates.

Figure 6.5 shows the results for the first hypothesis test using $B = 500$ bootstrap replicates. The CSM at baseline is represented as the horizontal dotted line. It is very clear how the CSM for both stimuli depart from this baseline value showing a difference in synchrony before (sleep-like mode) and after stimulation (awake-like mode).

Finally, Figure 6.6 shows the results for the second hypothesis test using $B = 500$ bootstrap replicates. Although the critical band for zero (expected synchrony difference under the null hypothesis) is quite wide, we can still find a period of time, right after stimulation, where the observed difference lies outside the band. This means that the differences in synchrony observed between stimuli are significant.

Next, we show the results for the comparison of the observed synchrony

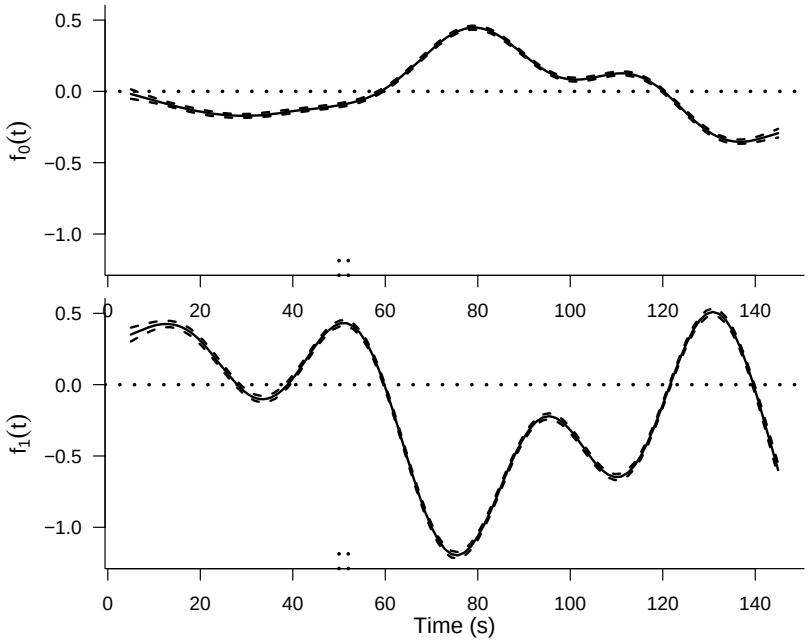

Figure 6.3: Smooth terms for GAM for CSM. Mean curve for both experimental conditions $f_0(t)$ (top panel) and smooth curve that accounts for the difference between experimental conditions, $f_1(t)$, (bottom panel). Two standard error lines above and below the estimate of the smooth curves are shown (dashed lines).

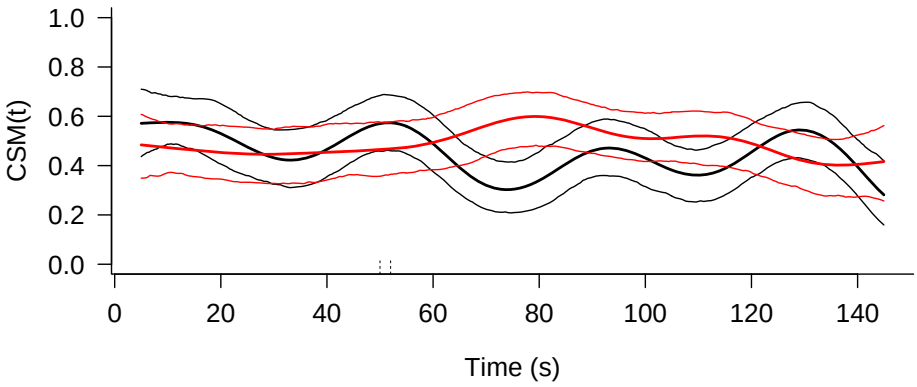

Figure 6.4: CSM profiles for the two stimuli with bootstrap confidence bands, averaged over three trials. Results for *bs* trials are represented in black and for *bf* trials in red.

to the expected one by chance. Figures 6.7 and 6.8 show these results for the real data. The first uses the calculations from the analytical expression obtained in Section 6.5 and the second one the results obtained with the simulations described in Section 6.6. In both cases we can observe that the expected synchrony under independence is smaller than the observed one. There are short periods of time where the synchrony observed by chance gets close to the observed one. This happens, for example, in both figures, for the *bs* stimulus, right after the stimulation has taken place. This uprise in the expected synchrony is due to the increase of the firing rates so the real synchrony might not have increased but actually have decreased. According to our working hypothesis we would expect the synchrony to decrease when the stimuli are applied. In fact, for *bs* stimulation there is a short time interval (right after stimulation) where the observed synchrony could be explained just from the increase in the firing rate. This is not the case for *bf* stimulation.

6.9 Chapter conclusions

An alternative method to measure neural synchrony, specially under low firing rate scenarios, has been proposed in this chapter. The method is based

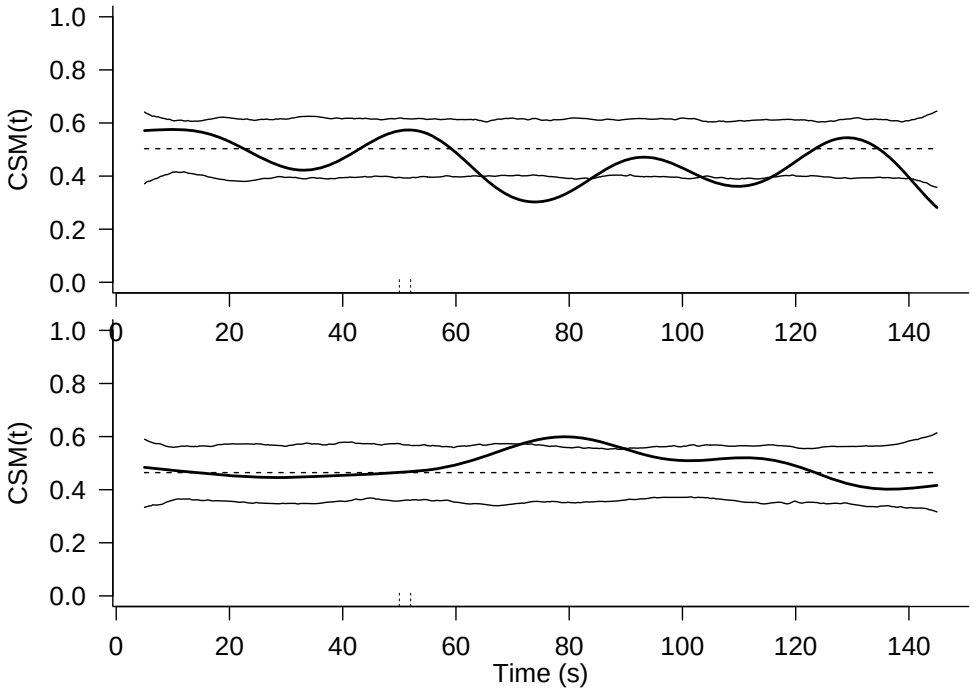

Figure 6.5: CSM profiles for the two stimuli averaged over three trials: *bs* (top panel) and *bf* (bottom panel). Baseline CSM estimated from the pre-stimuli time period (dotted lines) with bootstrap critical region.

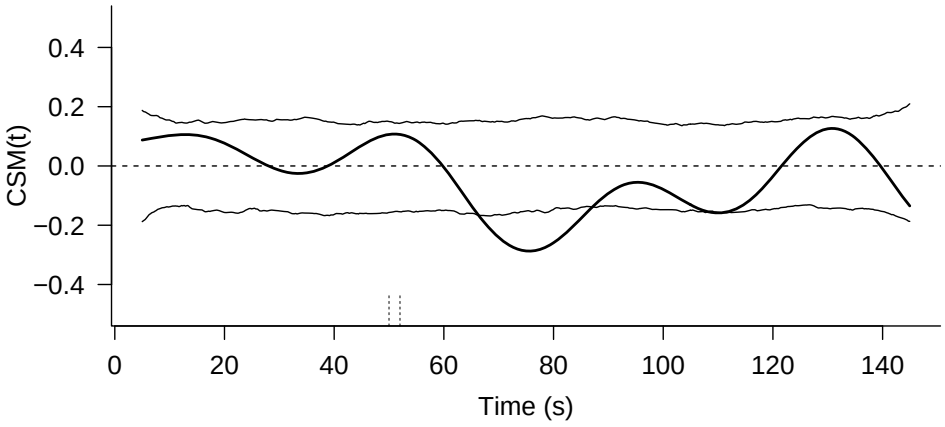

Figure 6.6: Difference of CSM profile for the two stimuli averaged over three trials and bootstrap confidence bands for the null hypothesis of difference equal to zero.

on the computation of times between nearest spikes of two spike trains. This is a flexible method that allows the researcher to decide how close two spikes have to be, to be considered synchronous. The method uses the information of time windows to estimate synchrony at a given time point. Although temporal resolution is lost with the use of this kernel-like method, it allows to study synchrony even with very few -or even only one- trials. Generalized additive models are proposed to fit these curves, giving flexibility to their shape by the use of nonparametric functions. Nevertheless, the dependency between consecutive time points is taken into account by modeling the error term with an autoregressive process.

The synchrony expected by chance has been computed showing that the CSM is able to distinguish true synchrony from the one that arises just due to the firing rate. Two hypothesis tests have been proposed and bootstrap procedures have been developed to calibrate the distribution of the test statistics. The results show that the differential stimulations have different effects on the synchrony profile. They also show that the observed synchrony in a short time interval right after stimulation can be partially explained by the increasing in the firing rate in that period.

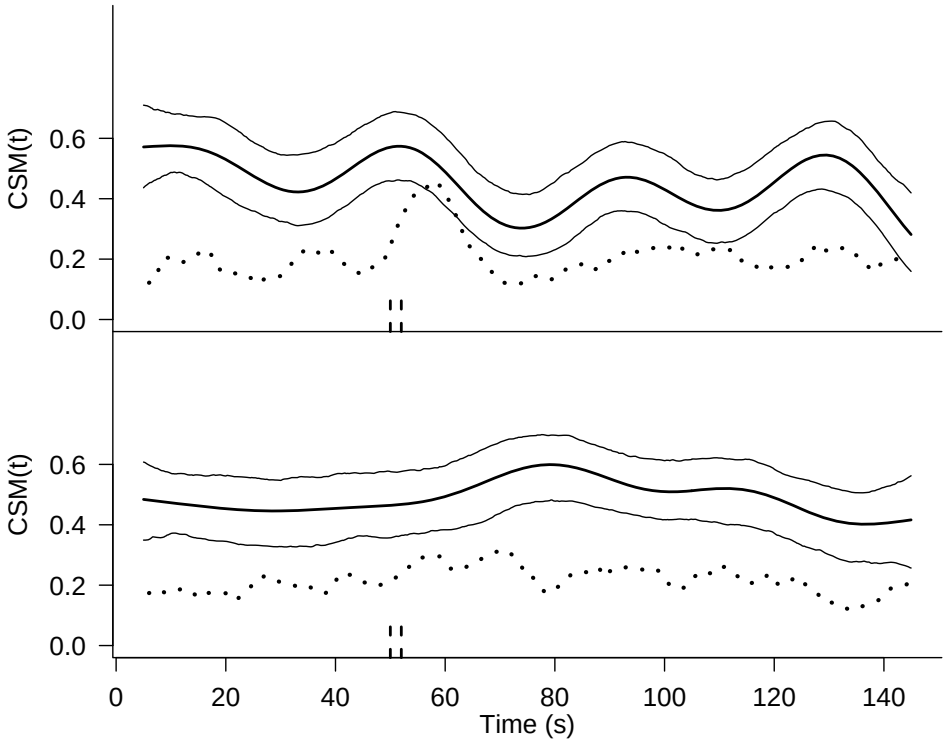

Figure 6.7: CSM profiles (solid lines) for *bs* stimulation (top panel) and *bf* stimulation (bottom panel) compared with the theoretical expected synchrony under independence (dotted lines). The moment of stimulation is indicated with dashed vertical lines. Bootstrap confidence bands are shown using thin solid lines.

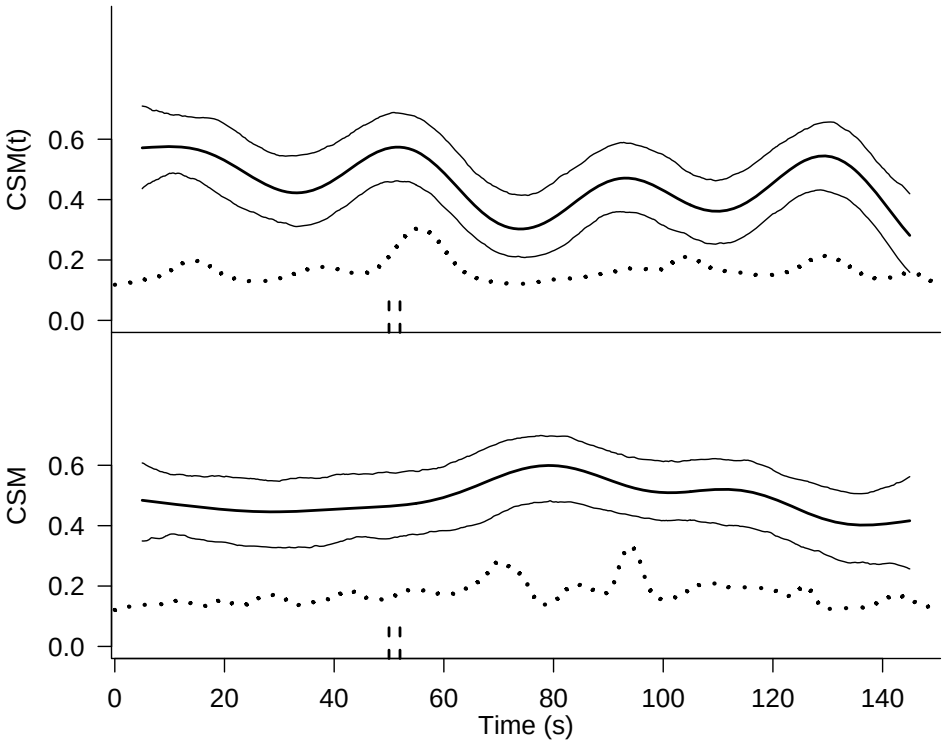

Figure 6.8: CSM profile (solid lines) for *bs* stimulation (top panel) and *bf* stimulation (bottom panel) compared with the simulated expected synchrony under independence (dotted lines). The moment of stimulation is indicated with dashed vertical lines. Bootstrap confidence bands are shown using thin solid lines.

Chapter 7

Stimuli and neural orientation selectivity effects using functional data analysis

Until now we have devoted our analysis to pairs of neurons and to compare the possible effects of differential stimulation on synchrony for a given cell pair. It is also important to check if these differences are significant at a population level. This is, if the differences are consistent throughout a group or a population of neurons. In this chapter we will zoom out from a single pair of neurons and study a group of cells.

For the analysis performed in this part of our study we have organized our pool of cells according to an important functional property of V1 neurons, the orientation selectivity, i.e., the existence of a preferred orientation of the visual stimulus (Hubel and Wiesel (1962)). This property is important in neurophysiological studies because it can provide important clues regarding the functional architecture of the striate cortex. The neuroscientists of the Neurocom group are interested in the synchronization dynamics of V1 cells during the two periods of spontaneous activity (sleep-like and awake-like) generated with their experimental model, taking also into account the orientation selectivity of the recorded neurons. This can render important information regarding the underlying functional connectivity. So, we will compare the synchrony profiles among different pairs of neurons, including the orientation selectivity in the analysis. We will study whether a second factor, regarding the affinity in orientation selectivity of a given pair, has a determinant influence in how neuronal synchronization evolves.

The experiment in which our data was recorded included in a first step the

characterization of each neuron regarding their preferred orientation. Several lines, with different angles of orientation, were shown to the cat and the firing was recorded, so each recorded neuron was associated with one specific orientation, corresponding to the highest firing rate. In this way, when paired up, we can define a new variable that is the difference between the favorite orientations of two neurons.

It is worth mentioning that, although the *orientation selectivity* should be a continuous variable, the way in which the experiments are carried out, make it a discrete variable, since only 16 orientations are shown, 8 angles with two different directions each. The use of two directions of movement with the same angle is also of interest in order to define another property of V1 neurons, the selectivity to direction. For the purposes of the present work, we will only take into account the orientation selectivity, and hence will just consider 8 orientations: $0°, 22.5°, 45°, 67.5°, 90°, 112.5°, 135°$ and $157.5°$. For this aim, we will identify the orientation $180°$ with $0°$, $202.5°$ with $22.5°$, and so on. Finally, let us define G as the random variable that measures the difference between preferred directions of two given neurons: $G = \min\{|O_1 - O_2|, 180 - |O_1 - O_2|\}$. Given the previous considerations, G can take one of these five possible outcomes: $0°, 22.5°$, $45°$, $67.5°$ and $90°$.

We will use the synchrony measure defined in Chapter 5 for the analyses. Our data are functions: the ICCSI profiles. So, in this chapter, the analysis are carried out using a functional data analysis approach. Details on functional data analysis are not given here, although we will give the basic notions that are necessary to understand our analysis. For details and theory on this subject, refer to, for example, the books by Ramsay and Silverman (2002, 2005) or Ferraty and View (2006).

We will use a group of eight simultaneously recorded neurons ($n = 8$). Therefore there are 28 pairs to work with. In this particular experiment four trials for each stimulation were carried out, providing us with $N = 224$ curves. In this case, the number of pairs in each of the categories given by G is:

$0°$	$22.5°$	$45°$	$67.5°$	$90°$
5	12	6	3	2

and there are 8 trials for each pair, 4 for each stimulus (bs/bf). Therefore there are 20, 48, 24, 12 and 8 curves in each category defined by stimulus

and orientation. On the other hand, we can obtain as many points in the curves as we like. Points in an equispaced grid $0 < t_1 < \ldots < t_M = T$ are considered: from 10 s to 230 s every 0.1 s. Therefore, each synchrony curve is evaluated in 2300 points. The curves, $ICCSI(t)$ are bounded, since $0 \leq ICCSI(t) \leq 1 \quad \forall t \in [0, T]$. Figure 7.1 shows the data averaged over trials. The top panel shows the functions that correspond to *bs* stimulation and in the bottom panel shows the ones for *bf* stimulation. Different colors are used for the different levels of G.

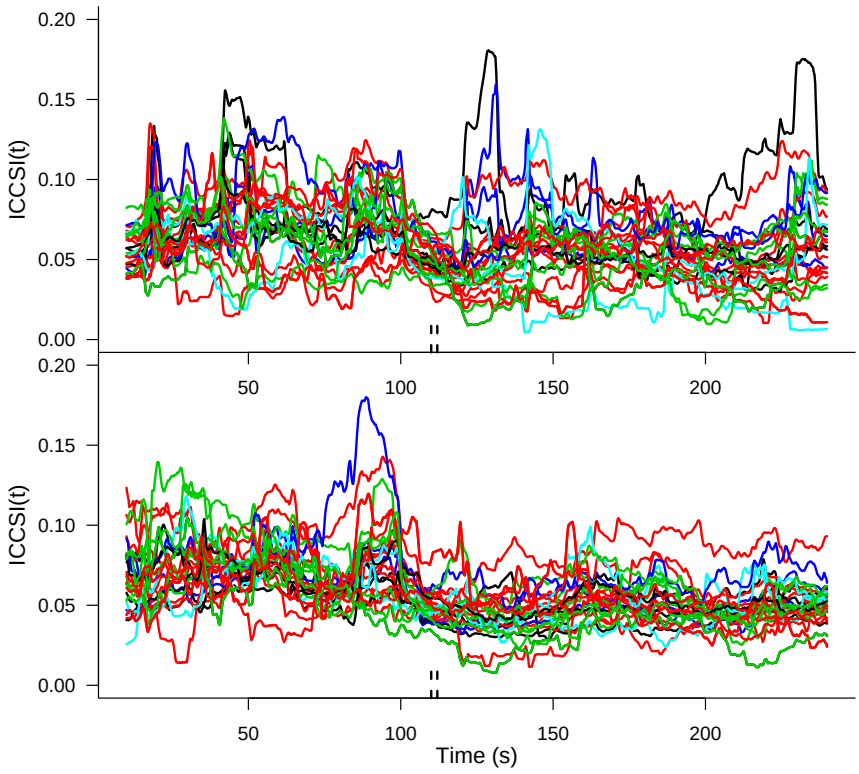

Figure 7.1: Functional data. Top panel: $ICCSI(t)$ averaged over trials for the first stimulus, *bs*. Bottom panel: $ICCSI(t)$ averaged over trials for the second stimulus, *bf*. Difference in orientation selectivity groups are defined in colors: 0° (black), 22.5° (red), 45° (green), 67.5° (blue) and 90° (cyan).

We will search for population differences in the dynamics of the awake-like period induced by each stimuli taking into account the possible effect of the other factor: difference between orientations selectivity. This problem can be dealt with as a two-way ANOVA with a two level factor: *stimulus* and a

five-level factor: G. Since the response is $ICCSI(t)$, a functional variable, we will use the method proposed by Cuesta-Albertos and Febrero-Bande (2010).

7.1 Functional two-way ANOVA

As already mentioned, the aim of this chapter is to search for differences on synchrony dynamics relative to two factors: *stimulus* and *difference in orientation selectivity*. As the difference in orientation selectivity can take values in a given finite set of values, the problem is one of a two way analysis of variance with fixed effects in which the response variable is functional. In the following subsection, we outline the methods but, for a more detailed explanation on the random projections methodology please refer to Cuesta-Albertos et. al (2007) and, on the specific case of functional ANOVA, refer to Cuesta-Albertos and Febrero-Bande (2010).

7.1.1 The random projections method for the ANOVA model

The random projections approach to solve problems in the functional data context is based on the ideas of Cuesta-Albertos et al. (2007). In that paper, the authors give an extension, on Hilbert spaces, to the Cramer-Wold theorem, which characterizes a probability distribution in terms of one-dimensional projections. Their Theorem 4.1 states that if two distributions are different, and we choose a marginal of them, those marginals will be almost surely different. Based on this fact, Cuesta-Albertos and Febrero-Bande (2010) propose a method for hypothesis testing in infinite dimensional spaces. We will state their result more formally, particularizing it to our problem. Roughly speaking, what they say is the following. Let us assume the data belong to a Hilbert space, $\mathcal{H}$, with a scalar product $\langle , \rangle$, and let μ be a Gaussian distribution in $\mathcal{H}$. Suppose the hypothesis to be tested is whether certain parameters, say γ_1 and γ_2, are equal ($H_0 : \gamma_1 = \gamma_2$). If $\gamma_1 \neq \gamma_2$, then the set of random directions, ν, from μ in $\mathcal{H}$, for which $\langle \nu, \gamma_1 \rangle = \langle \nu, \gamma_2 \rangle$, has probability zero. This is, if H_0 fails, then it also fails in its projected version for μ-almost every $\nu \in \mathcal{H}$. Therefore, a test at level α in the one dimensional space is a test at the same level to test H_0.

We now present the functional two-way ANOVA model for our problem and state the methodology more formally. Let $A = \{(i, j) : i, j \in$

$1, \ldots, n$ and $i < j\}$ and denote the pairs of neurons with indices $(i, j) \in A$. Therefore, $r = \#A = \frac{n(n-1)}{2}$ is the total amount of neuron pairs. We consider the following linear model for the synchrony curves:

$$y_{kl}^{(i,j)}(t) = m(t) + \alpha_k(t) + \beta_{g(i,j)}(t) + \gamma_{kg(i,j)}(t) + \epsilon_{kl}^{(i,j)}(t), \tag{7.1}$$

for $k = 1, 2$ and $l = 1, 2, 3, 4$. The function $\hat{y}_{kl}^{(i,j)}(t)$ is the integrated cross-correlation index ($ICCSI(t)$) for trial l of the pair given by neurons i and j under stimulus k; $m(t)$ is the global effect, $\alpha_k(t)$ is the effect of stimulus k. The function $g : A \mapsto \{1, 2, \ldots, 5\}$ indicates the level of the factor G that corresponds to the pair given by neurons i and j, identifying level 1 to the level 0°, level 2 to 22.5° and so on. Therefore, $\beta_{g(i,j)}$ is the effect of level $g(i, j)$ to the synchrony. The effect of a possible interaction between the factors is gathered by $\gamma_{kg(i,j)}$ and, finally, $\epsilon_{kl}^{(i,j)}(t)$ is the random error term. For identifiability of the parameters, we assume:

$$\alpha_1 + \alpha_2 = 0, \ \sum_{g=1}^{5} \beta_g = 0, \ \text{ and } \ \sum_{k=1}^{2} \sum_{g=1}^{5} \gamma_{kg} = 0.$$

The relevant null hypotheses to be tested are:

$$H_0^{\alpha} : \alpha_1 = \alpha_2 = 0,$$

which means that there is no effect of the stimulus,

$$H_0^{\beta} : \beta_1 = \cdots = \beta_5 = 0,$$

which states that there is no effect of the orientation selectivity. Also, a hypothesis for the interactions is reasonable:

$$H_0^{\gamma} : \gamma_{kg} = 0 \ \ \forall\, k = 1, 2, \forall\, g \in \{1, 2, \ldots, 5\}.$$

Theorem 2.1 in Cuesta-Albertos and Febrero-Bande (2010) states that, if the data belong to a Hilbert space, $\mathcal{H}$, endowed with a scalar product $\langle , \rangle$, μ is a Gaussian distribution on $\mathcal{H}$ such that its one-dimensional projections are nondegenerate, then,

1. If $\exists k \in \{1, 2\}$, such that $\alpha_k \neq 0$, then

$$\mu(\{\nu \in \mathcal{H} : \langle \nu, \alpha_k \rangle = 0 \ \forall k \in \{1, 2\}\}) = 0$$

2. If $\exists g \in \{1, 2, 3, 4, 5\}$, such that $\beta_g \neq 0$, then

$$\mu(\{\nu \in \mathcal{H} : \langle \nu, \beta_g \rangle = 0 \ \forall g \in \{1, 2, \ldots, 5\}\}) = 0$$

3. If $\exists k \in \{1,2\}$ and $g \in \{1,2,3,4,5\}$ such that $\gamma_{kg} \neq 0$, then

$$\mu(\{\nu \in \mathcal{H} : \langle \nu, \gamma_{gk} \rangle = 0 \ \forall k \in \{1,2\}, \forall g \in \{1,2,\ldots,5\}\}) = 0$$

Therefore, the proposed procedure is to randomly project the data on the one-dimensional space and to test the hypotheses in that context. Given a random function, $\nu(t)$, we consider the projected model:

$$y_{kl}^{(i,j)(\nu)} = m^{(\nu)} + \alpha_k^{(\nu)} + \beta_{g(i,j)}^{(\nu)} + \gamma_{kg(i,j)}^{(\nu)} + \epsilon_{kl}^{(i,j)(\nu)}, \tag{7.2}$$

and the hypotheses in the one dimensional problem:

$$\begin{aligned} &H_0^{\nu,\alpha} : \alpha_1^{(\nu)} = \alpha_2^{(\nu)} = 0 \\ &H_0^{\nu,\beta} : \beta_1^{(\nu)} = \cdots = \beta_5^{(\nu)} = 0 \\ &\quad \text{and} \\ &H_0^{\nu,\gamma} : \gamma_{kg}^{(\nu)} = 0 \ \ \forall\, k \in \{1,2\} \text{ and } g \in \{1,2,\ldots,5\}. \end{aligned} \tag{7.3}$$

The tests defined in the one-dimensional response case are clearly conditional to the random projection used, ν. To reduce the error introduced by the choice of the random projection, we will use the correction that implies controlling the false discovery rate (FDR) introduced by Benjamini and Hochberg (1995). This is also recommended by Cuesta-Albertos and Febrero-Bande (2010). We will use a procedure that arises from the paper of Benjamini and Yekuteli (2001). Given the ordered p-values, $p_{(1)} < \ldots < p_{(d)}$, obtained using d random projections, we will choose the corrected p-value as the quantity $\inf\{\frac{d}{i} p_{(i)}, i = 1, \cdots, d\}$.

So far, this test can help to search for global differences between the two groups of curves although we would rather study how these differences change in time. To do this, we propose to use moving windows along time. For each time point, t, we will consider an interval of time, centered at t, and project the resultant curve pieces so as to perform the ANOVA test with a better time resolution.

The hypothesis in (7.3) can be tested with any regular ANOVA approach. We will consider the classical ANOVA statistic to test these hypotheses, which we will denote by Q. This statistic has the form:

$$Q = \frac{(SSR_0 - SSR_q)/(q - q_0)}{SSR_T/(N - (q+1))}$$

where q is the total number of parameters of the model, q_0 is the number of parameters of the reduced model, i.e. the model under the null hypothesis. SSR_q is the total sum of squares of the residuals of the complete model:

$$\begin{aligned} SSR_q &= \sum_{\substack{(i,j)\in A \\ k=1,2 \\ l=1,2,3,4}} (y_{kl}^{(i,j)} - \hat{y}_{kl}^{(i,j)})^2 \\ &= \sum_{\substack{(i,j)\in A \\ k\in\{1,2\} \\ l\in\{1,2,3,4\}}} \left(y_{kl}^{(i,j)} - (\hat{m} + \hat{\alpha}_k + \hat{\beta}_{g(i,j)} + \hat{\gamma}_{kg(i,j)})\right)^2 \end{aligned}$$

and SSR_0 is the sum of squares of the residual for the reduced model, under the null hypothesis. For example, for hypothesis $H_0^{\nu,\alpha}$:

$$SSR_0 = \sum_{\substack{(i,j)\in A \\ k\in\{1,2\} \\ l\in\{1,2,3,4\}}} \left(y_{kl}^{(i,j)} - (\hat{\hat{m}} + \hat{\hat{\beta}}_{g(i,j)} + \hat{\hat{\gamma}}_{kg(i,j)})\right)^2$$

where $\hat{\hat{m}}, \hat{\hat{\beta}}_{g(i,j)}$ and $\hat{\hat{\gamma}}_{kg(i,j)}$ are the estimates of the parameters of the reduced model.

The test statistic Q is referreded to as the F-statistic because, in classical ANOVA analysis, under conditions of independence, normality and homocedasticity, it follows an F distribution with $q - q_0$ and $N - (q + 1)$ degrees of freedom. In the next section we will discuss why we cannot trust the F-distribution.

7.2 ANOVA with dependent data

In this section we introduce the problem of dependence that is present in our data. This dependency comes from the fact that the data are observed at neuron pairs level. So, it is only fair to think that the curves obtained from two pairs of neurons with one cell in common could be correlated.

7.2.1 ANOVA model with dependent errors

We need to be very careful in the way we write the model that accounts for dependence. Since we are going to work with the projections of the ICCSI

functions, we can forget about the infinite dimensional problem and focus in the one-dimensional one. Let us consider the model in (7.2) dropping the superscript (ν):

$$y_{kl}^{(i,j)} = m + \alpha_k + \beta_{g(i,j)} + \gamma_{kg(i,j)} + \epsilon_{kl}^{(i,j)}, \tag{7.4}$$

with $(i,j) \in A$, $k \in \{1,2\}$ and $g \in \{1,2,\ldots,5\}$. Model (7.4) is a two-way ANOVA model with a two-level first factor and a five-level second factor, with unbalanced cells, as we do not observe the same amount of pairs of neurons with difference in orientation selectivity. The following interpretation of the problem is more convenient. Consider each $Y_k^{(i,j)}$, i.e., the synchrony of a given pair of neurons under each stimulus, as a random variable. In this way, we have four realizations of each variable, so, our linear model becomes another linear model with $2r = m$ variables with $L = 4$ observations each. The model can be written in a matrix form, as follows:

$$\mathbf{y} = \mathbf{X}\zeta + \epsilon$$

where $\mathbf{y} \in \mathbb{R}^{4m\text{x}1}$ are the data ordered in a convenient form, $\mathbf{X} \in \mathbb{R}^{4m\text{x}q}$ is the design matrix, $\zeta \in \mathbb{R}^{q\text{x}1}$ is the vector of parameters and $\epsilon \in \mathbb{R}^{4m\text{x}1}$ is the vector of errors. Here, we consider the data ordered as follows:

$$\mathbf{y} = (y_{1,1}^{(1,2)}, y_{1,1}^{(1,3)}, \ldots, y_{1,1}^{(7,8)}, y_{2,1}^{(1,2)}, \ldots, y_{2,1}^{(7,8)}, y_{1,2}^{(1,2)}, \ldots, y_{1,2}^{(7,8)}, \ldots, y_{2,4}^{(1,2)}, \ldots, y_{2,4}^{(7,8)})$$

and therefore the vector of errors, ϵ, has the form:

$$\epsilon = (\epsilon_{1,1}^{(1,2)}, \epsilon_{1,1}^{(1,3)}, \ldots, \epsilon_{1,1}^{(7,8)}, \epsilon_{2,1}^{(1,2)}, \ldots, \epsilon_{2,1}^{(7,8)}, \epsilon_{1,2}^{(1,2)}, \ldots, \epsilon_{1,2}^{(7,8)}, \ldots, \epsilon_{2,4}^{(1,2)}, \ldots, \epsilon_{2,4}^{(7,8)})$$

The assumptions of normality and homocedasticity for the errors are reasonable in this context, but the fundamental problem in this study comes from the presence of dependence among the data that comes from pairs of neurons sharing a cell. The data obtained from the first trials of pair N1-N2 and N1-N3 are clearly dependent as they share neuron N1. This is, $Y_k^{(i,j)}$ and $Y_k^{(i',j')}$ are dependent if $\{i,j\} \cap \{i',j'\} \neq \emptyset$. So, the errors of the model are considered normally distributed with zero mean and covariance matrix Σ, and we will make some additional assumptions on Σ. We assume that the variance of $\epsilon_k^{(i,j)}$ is the same for all (i,j) and all k and equal to σ^2. On the other hand, we will also assume that, if $\#(\{i,j\} \cap \{i',j'\}) = 1$ then $\text{cov}(\epsilon_k^{(i,j)}, \epsilon_k^{(i',j')}) = \rho\sigma^2$. Let $B = \{(i,j,k,l,i',j',k',l') : \#(\{i,j\} \cap \{i',j'\}) = 1, k = k', l = l'\}$ then, in

summary:

$$\mathrm{Cov}(\epsilon_{kl}^{(i,j)}, \epsilon_{k'l'}^{(i,j)}) = \begin{cases} \sigma^2 & \text{if } (i,j,k,l) = (i',j',k',l') \\ \sigma^2\rho & \text{if } (i,j,k,l,i',j',k',l') \in B \\ 0 & \text{otherwise} \end{cases} \tag{7.5}$$

Therefore, Σ results in a very special matrix, with, σ^2 in the diagonal, $\sigma^2\rho$ where the variable in the column and the variable in the row share a neuron and they also share the trial (and, thus, the stimulus). In our particular example Σ is a 224×224 ($224 = 2 \cdot 4 \cdot 28$) matrix, composed by a diagonal of eight 28×28 blocks and the rest of the elements are zeros:

$$\Sigma = \begin{pmatrix} \sigma^2\mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma^2\mathbf{C} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \sigma^2\mathbf{C} \end{pmatrix} \tag{7.6}$$

The blocks in the diagonal are all equal and equal to the covariance matrix, $\sigma^2\mathbf{C}$, of the data that corresponds to one level of the first factor (stimulus). Let us explicitly show $\mathbf{C}$:

$$\mathbf{C} = \begin{pmatrix}
1 & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\rho & 1 & \rho & \rho & \rho & \rho & \rho & \rho & 0 & 0 & 0 & 0 & 0 & \rho & \rho & \rho & \rho & \rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\rho & \rho & 1 & \rho & \rho & \rho & \rho & 0 & \rho & 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 & 0 & \rho & \rho & \rho & \rho & 0 & 0 & 0 & 0 \\
\rho & \rho & \rho & 1 & \rho & \rho & \rho & 0 & 0 & \rho & 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 & \rho & 0 & 0 & 0 & \rho & \rho & \rho & 0 \\
\rho & \rho & \rho & \rho & 1 & \rho & \rho & 0 & 0 & 0 & \rho & 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 & \rho & 0 & 0 & \rho & 0 & \rho & \rho \\
\rho & \rho & \rho & \rho & \rho & 1 & \rho & 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 & \rho & 0 & 0 & \rho & 0 & \rho \\
\rho & \rho & \rho & \rho & \rho & \rho & 1 & 0 & 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 & \rho & 0 & 0 & \rho & \rho \\
\rho & \rho & 0 & 0 & 0 & 0 & 0 & 1 & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\rho & 0 & \rho & 0 & 0 & 0 & 0 & \rho & 1 & \rho & \rho & \rho & \rho & \rho & 0 & 0 & 0 & 0 & \rho & \rho & \rho & \rho & 0 & 0 & 0 & 0 \\
\rho & 0 & 0 & \rho & 0 & 0 & 0 & \rho & \rho & 1 & \rho & \rho & \rho & 0 & \rho & 0 & 0 & 0 & \rho & 0 & 0 & 0 & \rho & \rho & \rho & 0 \\
\rho & 0 & 0 & 0 & \rho & 0 & 0 & \rho & \rho & \rho & 1 & \rho & \rho & 0 & 0 & \rho & 0 & 0 & 0 & \rho & 0 & 0 & \rho & 0 & \rho & \rho \\
\rho & 0 & 0 & 0 & 0 & \rho & 0 & \rho & \rho & \rho & \rho & 1 & \rho & 0 & 0 & 0 & \rho & 0 & 0 & 0 & \rho & 0 & 0 & \rho & 0 & \rho \\
\rho & 0 & 0 & 0 & 0 & 0 & \rho & \rho & \rho & \rho & \rho & \rho & 1 & 0 & 0 & 0 & 0 & \rho & 0 & 0 & 0 & \rho & 0 & 0 & \rho & \rho \\
0 & \rho & \rho & 0 & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & 0 & 1 & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & 0 & 0 & 0 & 0 \\
0 & \rho & 0 & \rho & 0 & 0 & 0 & \rho & 0 & \rho & 0 & 0 & 0 & \rho & 1 & \rho & \rho & \rho & 0 & 0 & 0 & \rho & \rho & \rho & 0 & 0 \\
0 & \rho & 0 & 0 & \rho & 0 & 0 & \rho & 0 & 0 & \rho & 0 & 0 & \rho & \rho & 1 & \rho & \rho & 0 & \rho & 0 & 0 & \rho & 0 & 0 & \rho & \rho & 0 \\
0 & \rho & 0 & 0 & 0 & \rho & 0 & \rho & 0 & 0 & 0 & \rho & 0 & \rho & \rho & \rho & 1 & \rho & 0 & 0 & \rho & 0 & 0 & \rho & 0 & \rho \\
0 & \rho & 0 & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & 0 & \rho & \rho & \rho & \rho & \rho & 1 & 0 & 0 & 0 & \rho & 0 & 0 & \rho & \rho \\
0 & 0 & \rho & \rho & 0 & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & 1 & \rho & \rho & \rho & \rho & \rho & 0 & 0 \\
0 & 0 & \rho & 0 & \rho & 0 & 0 & 0 & \rho & 0 & \rho & 0 & 0 & \rho & 0 & \rho & 0 & 0 & \rho & 1 & \rho & \rho & \rho & 0 & 0 & \rho \\
0 & 0 & \rho & 0 & 0 & \rho & 0 & 0 & \rho & 0 & 0 & \rho & 0 & \rho & 0 & 0 & \rho & 0 & \rho & \rho & 1 & \rho & 0 & \rho & 0 & \rho \\
0 & 0 & \rho & 0 & 0 & 0 & \rho & 0 & \rho & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & \rho & \rho & \rho & \rho & 1 & 0 & 0 & \rho & \rho \\
0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & \rho & \rho & 0 & 0 & \rho & \rho & 0 & 0 & 1 & \rho & \rho & 0 \\
0 & 0 & 0 & \rho & 0 & \rho & 0 & 0 & 0 & \rho & 0 & \rho & 0 & 0 & \rho & 0 & \rho & 0 & \rho & 0 & \rho & 0 & \rho & 1 & \rho & \rho \\
0 & 0 & 0 & \rho & 0 & 0 & \rho & 0 & 0 & \rho & 0 & 0 & \rho & 0 & \rho & 0 & 0 & \rho & \rho & 0 & 0 & \rho & \rho & \rho & 1 & 0 & \rho & \rho \\
0 & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & \rho & \rho & 0 & 0 & \rho & \rho & 0 & \rho & \rho & 0 & 1 & \rho & \rho \\
0 & 0 & 0 & 0 & \rho & 0 & \rho & 0 & 0 & 0 & \rho & 0 & \rho & 0 & 0 & \rho & 0 & \rho & 0 & \rho & 0 & \rho & \rho & 0 & \rho & \rho & 1 & \rho \\
0 & 0 & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & 0 & \rho & \rho & 0 & 0 & 0 & \rho & \rho & 0 & 0 & \rho & \rho & 0 & \rho & \rho & \rho & \rho & 1
\end{pmatrix}$$

With this definition, Σ results positive definite for small values of ρ. The eigen values of the matrix in (7.6) result in only three different values: $1-2\rho$, $1+4\rho$ and $1+12\rho$, which, for $\rho \in [0,1]$, are all greater than zero at the same time, whenever $\rho < 0.5$. This means that this correlation structure for the data is plausible only for $\rho \in [0, 0.5)$.

In this context, we cannot assume that the ANOVA statistic follows a F distribution as in classical analysis. If the correlation coefficient and variance of the errors were known we could solve this problem in a very simple way. Aitken (1935) and Rao (1973), Chapter 4, explain it in detail, but, the procedure would be to transform the data to an independent setup. As Σ is positive definite, there exist $\mathbf{Q}$ and $\mathbf{D}$ matrices such that $\left(\frac{1}{\sigma^2}\Sigma\right)^{-1} = \mathbf{Q}\mathbf{D}\mathbf{Q}^t$ with $\mathbf{Q}$ an orthogonal matrix and $\mathbf{D}$ a diagonal one, with positive numbers in the diagonal and therefore $\mathbf{R} = \mathbf{Q}\mathbf{D}^{\frac{1}{2}}\mathbf{Q}^t = \sigma\Sigma^{-\frac{1}{2}}$. So, if ρ were known, we could multiply our matrix model by $\mathbf{R}$ to have

$$\mathbf{R}\mathbf{y} = \mathbf{R}\mathbf{X}\zeta + \tilde{\epsilon} \tag{7.7}$$

where now, $\mathbf{R}\mathbf{y}$ is a vector of independent data since $\tilde{\epsilon} \sim N(0, \sigma^2 I)$.

As we do not know neither the variance nor the real correlation coefficient we will use a different approach. We will calibrate the distribution of the ANOVA test statistic, under the null hypothesis, using a parametric bootstrap as described in the following subsection.

7.2.2 Estimation of the correlation coefficient

Since we assume that the covariance between the different pairs of error terms are equal, provided the pairs belong to B, we will estimate the correlation coefficient as the average of the Pearson correlation coefficient of the corresponding pairs, which is equivalent to,

$$\hat{\rho} = \frac{1}{\hat{\sigma}^2}\frac{1}{\#B}\sum_{(i,j,k,l,i',j',k',l')\in B} \hat{\epsilon}_{kl}^{(i,j)} \cdot \hat{\epsilon}_{k'l'}^{(i',j')}, \tag{7.8}$$

where $\hat{\epsilon}_{kl}^{(i,j)}$ are the elements of the residual vector: $\hat{\epsilon} = \mathbf{y} - \mathbf{X}\hat{\zeta}$, with $\hat{\zeta}$ the estimated model parameters, and $\hat{\sigma}^2$ is the estimated variance of these residuals.

7.2.3 Bootstrap calibration of the distribution of the ANOVA test statistic

As the data are considered normally distributed, we propose the use of a parametric bootstrap to calibrate the distribution of the ANOVA test statistic under the null hypothesis. We will now describe the procedure for a general null hypothesis $H_0^i : \zeta_i = 0$.

Once the linear model has been fitted to obtain $\hat{\zeta}$ and the classical ANOVA statistic has been computed (denoting the result with Q^{obs}) we estimate σ^2 and ρ from the residuals to build the estimated covariance matrix, $\hat{\Sigma} = \Sigma(\hat{\sigma}^2, \hat{\rho})$. We can proceed with the following bootstrap algorithm:

1. Replace the i-th parameter in the estimated $\hat{\zeta}$ by zero (null hypothesis). This set of parameters will be denoted by $\hat{\zeta}_0$. Build a bootstrap sample: $\mathbf{y}^* = \mathbf{X}\hat{\zeta}_0 + \mathbf{z}$ with $\mathbf{z} \sim N(0, \hat{\Sigma})$.

2. Fit the linear model to the sampled data, obtain the bootstrap version of the estimated parameters $\hat{\zeta}^*$ and compute Q^*, the bootstrap version of Q.

3. Repeat Steps 1–2, B times to obtain $Q^{*1}, \ldots, Q^{*B}$.

4. Compute the desired $(1-\alpha)$-percentile of the bootstrap statistics, $Q^*_{1-\alpha}$. We reject the null hypothesis if $Q^{obs} > Q^*_{1-\alpha}$.

7.3 Results

The random projections method can be used to solve several problems for functional data. For example, Cuevas et. al. (2007) discuss the application of these ideas in the classification and estimation frameworks. This methodology can be also used to define a functional depth. The depth notion makes reference to how central is a point in a given set of points. In $\mathbb{R}$, for example, the depth of a point is often defined as the number of intervals with extremes in data values that contain that point. With this notion of depth, the median is the deepest point in a given data set. We will use a method based on random projections proposed by Cuevas et al. (2007), based in the results we have already described for the ANOVA problem. Another method to measure functional depth can be found in Fraiman and Muñiz (2001), for example. Consider an independent direction ν and a process $\mathcal{F}$. Let $f_1(t), \ldots, f_n(t)$ be realizations of the functional process $\mathcal{F}$ and let $D(r)$

be a measure of depth in $\mathbb{R}$. The *deepest* element of the set $f_1(t), \ldots, f_n(t)$ is defined as $f_j(t)$ for which

$$j = \text{argmax}_{i=1\ldots,n}\{D\left(\langle \nu, f_i(t)\rangle\right)\}.$$

Figure 7.2 shows the deepest curve of each group in terms of G.

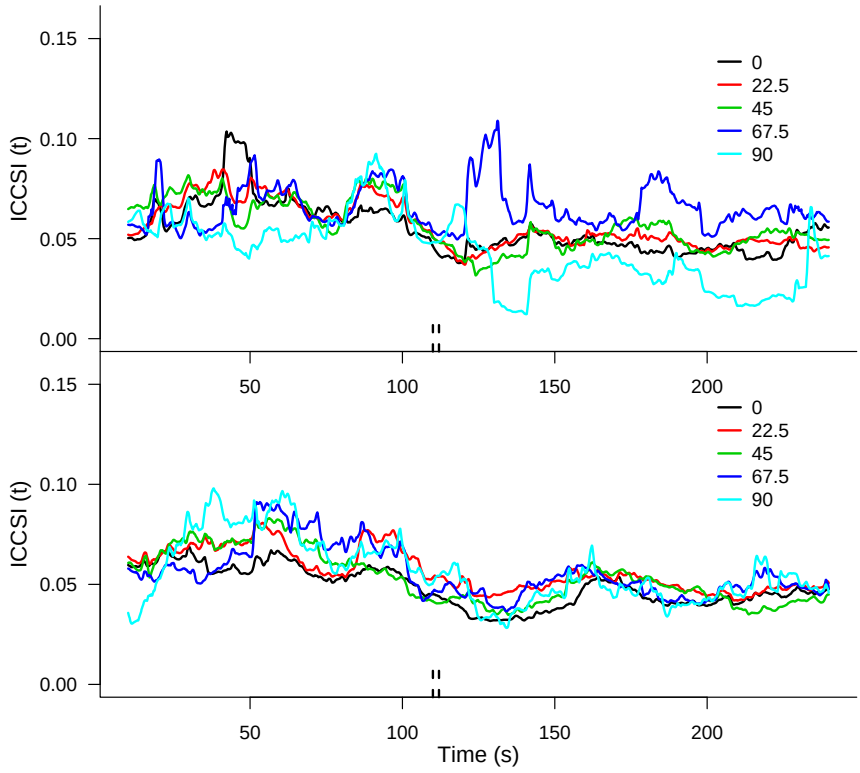

Figure 7.2: Top panel: deepest $ICCSI(t)$ curves for the first stimulus, *bs*. Bottom panel: deepest $ICCSI(t)$ curves for the second stimulus, *bf*. Difference in orientation selectivity groups are defined in colors: 0° (black), 22.5° (red), 45° (green), 67.5° (blue) and 90° (cyan). The functional depth is computed using the random projections method.

To draw the random vectors we use Brownian motions or, more precisely, approximations to standard Brownian motions by sequences of partial sums of normally distributed random variables. We only need to compute the values of the random vectors in the equidistant time points, $t_1, \ldots, t_M$, where the functions, $\widehat{ICCSI}(t)$, are defined. For this, we consider M independent

and identically distributed standard normal random variables, $Z_1, \ldots, Z_M$, and define a trajectory ν_1 as

$$\nu_1(t_1) = 0 \text{ and } \nu_1(t_k) = \nu_1(t_{k-1}) + (t_k - t_{k-1})Z_k \text{ for } k = 2, \ldots, M. \quad (7.9)$$

On the other hand, we would like to have directions without tendency and such that their variability throughout the trajectory is not large so as not to give more importance to a certain period of the functions. For this aim we will define the random trajectories as the sum of two Brownian motion, as just defined, where one of them has been "flipped" so as to be equal to zero in the last time point t_M. This is, let ν_1 and ν_2 be defined as in (7.9) and let $\nu_3(t_k) = \nu_2(t_{M-k+1})$. In this way, the final directions we will use are defined as $\nu(t) = \nu_1(t) + \nu_3(t)$.

A preliminary analysis, fitting model (7.4), showed that the interaction between factors are not significant. Therefore, the final model considered is:

$$y_{kl}^{(i,j)} = m + \alpha_k + \beta_{g(i,j)} + \epsilon_{kl}^{(i,j)}, \quad (7.10)$$

with, $k \in \{1, 2\}$, $g(i, j) \in \{1, 2, \ldots, 5\}$, $l \in \{1, 2, 3, 4\}$. Figure 7.3 shows the p-values obtained by using sliding windows along time to study the development of the effects of both factors. Forty seconds time windows were considered, moving along the time axis from the second 20 of recording to the second 215. In the time period between 110 and 150 this was done every second, in the rest, it was done every 5 s. At each window, 30 random directions have been used to project the data (the same ones in every window) and the FDR correction was applied to come up with just one p-value. It is clear that there are differences between the two approaches used. When dependence is not taken into account (dashed lines) the test is less conservative than when the dependence is included. Although for the effect of the stimuli there is a period of time at the beginning of the awake-like period for which both tests reject the null hypothesis, next there is another period in which it would be rejected if the dependence was not accounted for. The results show that there exists a period of time during the awake-like mode when the differences between the effects of the kind of stimulus are significant. On the other hand, the differences in synchrony among the levels of the factor G are also found significant after the stimulus.

The estimation of the correlation parameter changes for each window, nevertheless, the estimation is not very variable, not even from one projection to the other. Figure 7.4 shows the $\hat{\rho}$ as a function of time. This figure also

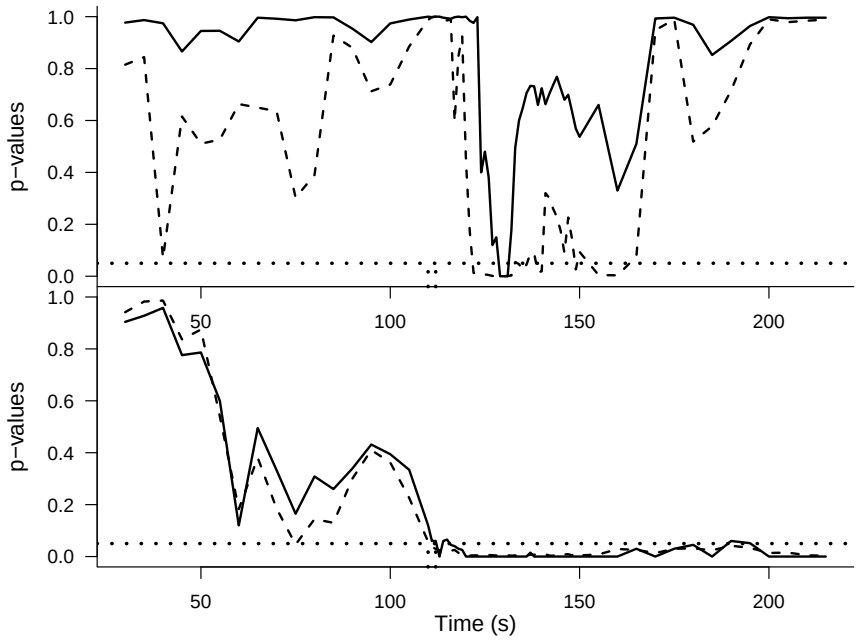

Figure 7.3: p-values for the two-way ANOVA as a function of time. p-values for the null hypotheses of no effect of the stimuli (top panel) and for the null hypothesis of no effect of the difference in orientation selectivity (bottom panel). p-values obtained using the F distribution are shown with dashed lines while the p-values obtained using the bootstrap are shown in solid lines. The horizontal dotted line is the constant value 0.05 for reference and the vertical dotted lines depicts the stimulation time.

shows the mean $\bar{\hat{\rho}}$ across projections. We can observe that, at the beginning of the recording, correlation coefficients were estimated to be greater than 0.5 and they were truncated so as the covariance matrix, $\widehat{\Sigma}$, resulted positive definite.

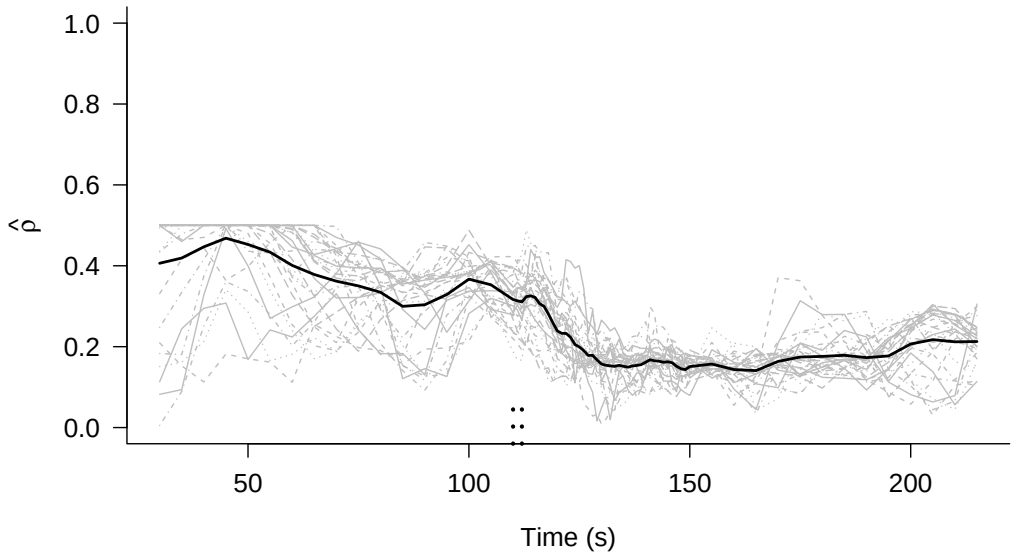

Figure 7.4: Evolution of the estimate of the correlation coefficient ρ. Estimations for different random projections (grey lines), mean $\bar{\hat{\rho}}$ (black line). The vertical dotted lines depicts the stimulation time.

If the case of large correlation was the most common, if pairs sharing a neuron were very correlated ($\rho > 0.5$) an alternative, nonparametric, bootstrap can be carried out. For example, in the following procedure, the trials are shuffled, assigning to each trial the residuals of other trial chosen with equal probability from the eight possible ones. Having fitted model (7.10):

1. For each $k \in \{1, 2\}$ and $l \in \{1, 2, 3, 4\}$ draw the bootstrap pair (k^*, l^*) with equal probability from $\{1, 2\} \times \{1, 2, 3, 4\}$, this is $P(k^* = k', l^* = l') = \frac{1}{8}$ for all $k' \in \{1, 2, \}$ and all $k' \in \{1, 2, 3, 4\}$.

2. Define $\hat{\epsilon}_{k,l}^{(i,j)*} = \hat{\epsilon}_{k^*,l^*}^{(i,j)} \ \ \forall (i, j)$.

This bootstrap procedure has the drawback that, in our case, the vector of bootstrap residuals can only take eight possible values. A possible improvement of the method is to use a smoothed version. For this, a smoothing

parameter h, small with respect to the standard deviation of the residuals, would be chosen and Step 2 would be replaced by:

2. Define $\hat{\epsilon}_{k,l}^{(i,j)*} = \hat{\epsilon}_{k^*,l^*}^{(i,j)} + hZ_{k,l}^{(i,j)*}$ with $Z_{k,l}^{(i,j)*} \sim N(0,1)$ iid $\forall(i,j)$.

7.3.1 Distribution of the test statistic Q

To visualize the differences between the distribution of the test statistic taking into account the dependence and the one when independence between observations is assumed (F-distribution), Figure 7.5 shows the estimated density of Q obtained by simulation. We simulated data from (7.10) using the parameters estimated for one particular random projection: 10000 data vectors were generated from $\mathbf{y} = \mathbf{X}\zeta + \epsilon$, where $\zeta = (m, \alpha^t, \beta^t)$, with $m = 1.6$, $\alpha_1 = 0.1$, $\beta_1 = -0.06$, $\beta_2 = 0.2$, $\beta_3 = -0.08$ and $\beta_4 = 0.27$. We simulated the case of 8 neurons, $i, j \in \{1, \cdots, 8\}$, and 4 trials for each stimulus. The second factor was chosen exactly as in the real case and, therefore, the design matrix, $\mathbf{X}$, was the same as for the real data. The errors were sampled from a multivariate normal distribution, $\epsilon \sim N(0, \Sigma)$ with Σ defined as in (7.6) and $\sigma^2 = 1.42$. The values for the correlation coefficient were chosen as $\rho = 0, 0.1, 0.15$ and 0.4. Figure 7.5 shows that the null distribution of the test statistic is well approximated by the F distribution (as it should) when $\rho = 0$ and, on the other hand, it departs from it when ρ increases. This is an evidence for the need of using the bootstrap to calibrate de distribution of Q instead of using the F distribution. When ρ increases, the peak of the density decreases and the tale of the distribution becomes more heavy. This explains why, sometimes, the test using the classical approach rejects while the bootstrap approach does not.

7.4 Simulation study

To evaluate the performance of our test we developed a small simulation study. We simulated data, similar to the real ones, using different (known) model parameters and correlation coefficients. With the simulated data, we computed Q and followed the procedure used to calibrate the distribution of the test statistic for a level 0.05 test. Finally, we compared the results (reject\non-reject of the null hypothesis) with the true case. The data were generated as if three trials under two stimulation conditions of a group of seven neurons had been recorded. Each neuron was assigned with a given fixed characteristic (orientation selectivity) to be able to compute the second factor (difference in orientation selectivity). In this case the preferred

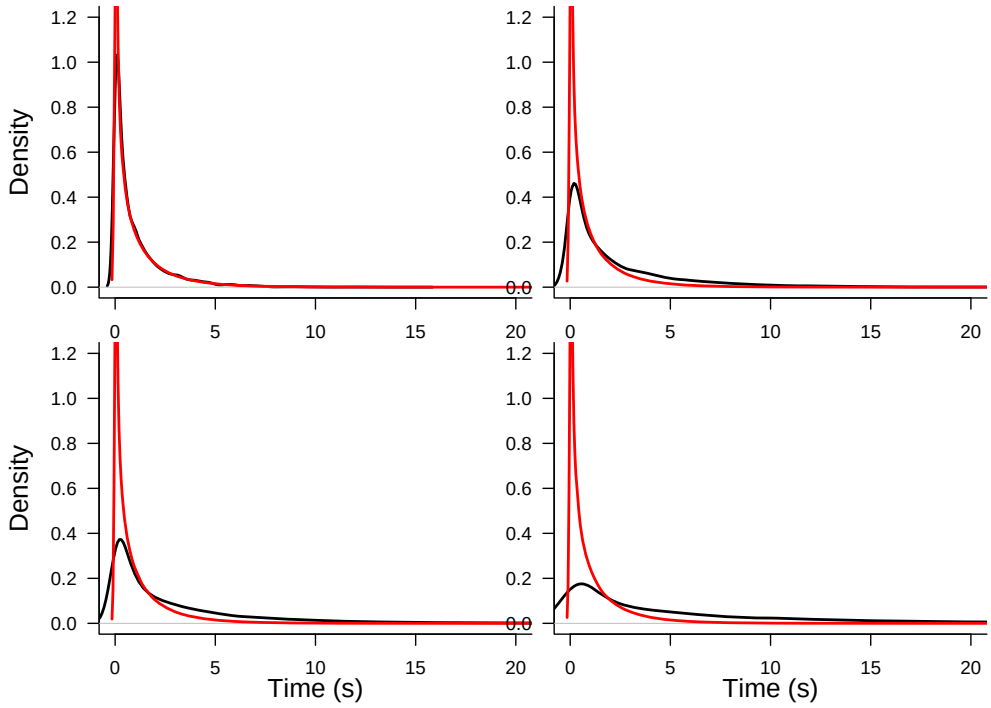

Figure 7.5: Density of the the test statistic Q under H_0 (black lines) compared with the F distribution (red lines) under different dependence scenarios: $\rho = 0$ (top-left panel), $\rho = 0.1$ (top-right panel), $\rho = 0.15$ (bottom-left panel) and $\rho = 0.4$ (bottom-right panel).

orientations were defined with only two possible values that were assigned arbitrarily to each neuron. Then,

$$\mathbf{y} = \mathbf{X}\zeta + \epsilon, \;\; \epsilon \sim N(0, \Sigma),$$

where, $\zeta = (\mu, \alpha^t, \beta^t)^t$, where $\alpha^t = (\alpha_1, -\alpha_1)$ and $\beta^t = (\beta_1, -\beta_1)$. The covariance matrix, Σ, was defined as in (7.5) for four values of ρ. The values used for the simulations are shown in Table 7.1.

Table 7.1: Parameters used in the simulation study.

m	0
α_1	$0, 0.5, 1$
β_1	$0, 0.5, 1$
ρ	$0, 0.05, 0.15, 0.35$

On the other hand, a large variance ($\sigma^2 = 1$) with respect to the parameters was chosen to reflect a typical situation for the real data. For most of the projections, the signal to noise ratio is rather large. This large variance could affect the ability of the test to detect differences between the parameters. $M = 1000$ Montecarlo simulations were performed and for each one, $B = 500$ bootstrap trials were used. The results are shown in Tables 7.2, 7.3 and 7.4. Table 7.2 shows the proportion of rejections for the situation of no effect of the first factor. It can be observed that the test meets the level satisfactorily almost all the times, although it rejects more than expected in some cases, when the correlation coefficient is positive. Table 7.3 shows the proportion of rejection of the null hypothesis that there is no effect of the second factor. In this case the behavior under the null hypothesis is good, although it seems to be a little conservative for large correlations. Regarding the power of the test, we can observe 100% rejection under the alternative in almost all the cases. For the first factor, when $\alpha_1 = 0.05$, proportions of rejection decrease with the increase of ρ. Surprisingly, the value of the correlation parameter does not seem to influence much the results for β_1.

Finally, Table 7.4 shows similar simulation results for the interaction between the two factors. In this case we show the proportion of rejections under the null hypothesis for different values of ρ and 1000 Montecarlo replications. We can see that the level of 0.05 is successfully met, although for the case of large correlation values, the test seems to be conservative. It is important to notice that, regarding the power of the test, in these simulations the test

Table 7.2: Proportion of rejections for H_0^α at level 0.05.

		$\rho = 0$	$\rho = 0.05$	$\rho = 0.15$	$\rho = 0.35$
$\alpha_1 = 0$	$\beta_1 = 0$	0.046	0.079	0.077	0.078
	$\beta_1 = 0.5$	0.047	0.068	0.084	0.069
	$\beta_1 = 1$	0.055	0.077	0.095	0.076
$\alpha_1 = 0.5$	$\beta_1 = 0$	0.999	0.995	0.944	0.791
	$\beta_1 = 0.5$	0.999	0.996	0.951	0.779
	$\beta_1 = 1$	1	0.990	0.946	0.770
$\alpha_1 = 1$	$\beta_1 = 0$	1	1	1	1
	$\beta_1 = 0.5$	1	1	1	1
	$\beta_1 = 1$	1	1	1	1

Table 7.3: Proportion of rejections for H_0^β at level 0.05.

		$\rho = 0$	$\rho = 0.05$	$\rho = 0.15$	$\rho = 0.35$
$\alpha_1 = 0$	$\beta_1 = 0$	0.058	0.056	0.057	0.045
	$\beta_1 = 0.5$	0.058	0.049	0.046	0.036
	$\beta_1 = 1$	0.053	0.046	0.041	0.045
$\alpha_1 = 0.05$	$\beta_1 = 0$	1	1	1	1
	$\beta_1 = 0.5$	0.999	1	1	1
	$\beta_1 = 1$	0.999	1	1	1
$\alpha_1 = 1$	$\beta_1 = 0$	1	1	1	1
	$\beta_1 = 0.5$	1	1	1	1
	$\beta_1 = 1$	1	1	1	1

Table 7.4: Proportion of rejections for H_0^γ at level 0.05.

		$\rho = 0$	$\rho = 0.05$	$\rho = 0.15$	$\rho = 0.35$
$\alpha_1 = 0$	$\beta_1 = 0$	0.054	0.033	0.032	0.034
	$\beta_1 = 0.5$	0.039	0.041	0.035	0.034
	$\beta_1 = 1$	0.040	0.038	0.035	0.031
$\alpha_1 = 0.05$	$\beta_1 = 0$	0.031	0.037	0.034	0.035
	$\beta_1 = 0.5$	0.043	0.052	0.041	0.035
	$\beta_1 = 1$	0.042	0.043	0.037	0.030
$\alpha_1 = 1$	$\beta_1 = 0$	0.029	0.038	0.043	0.038
	$\beta_1 = 0.5$	0.035	0.033	0.036	0.032
	$\beta_1 = 1$	0.047	0.052	0.035	0.041

rejected the null hypothesis 100% of the times (results not shown), using values of $\gamma = 0.5, 1$ to simulate data in the alternative.

7.5 Chapter conclusions

In this chapter we have applied up-to-date methodology for functional data analysis to the synchrony curves. Random projections techniques have been used. These methods are very easy to implement and interpret, what makes them appealing for the application in many problems. In particular, we have used them for a design of experiments model, a two-way functional ANOVA. In this context, the method showed to be useful as it allowed us to find significance in the effects of the factors under study.

The model under study involved the synchrony curves obtained by the method described in Chapter 5. The curves are separated in groups given by the stimuli and the difference in preferred orientation between the two neurons involved in each curve. The aim of the chapter was to study if the differences during the awake-like period regarding the applied stimuli (bs/bf), found in Chapter 5, were still significant at a group level. Moreover, a factor regarding difference in preferred orientation of neurons was brought into the analysis. Differences between the levels of this second factor were also of interest. Although there were groups with very few elements, several conclusions can be established.

It was also shown the importance of the incorporation of the dependence

between curves into the analysis. The distribution of the test statistic was approximated using a parametric bootstrap on the residuals of the model allowing for dependence. It was shown that the classical F-test statistic can lead to false positives, as the distribution of the test statistic has a heavier tail than the F distribution. Also, a nonparametric bootstrap was proposed for the cases of large correlation coefficients, although the results of its implementation were not shown.

The interactions between the factors did not result significant. An effect of the factor *stimulus* can be found at the beginning of the awake-like period (few seconds after the stimulus onset), allowing us to extend the results of Chapter 5 to a more general setting. Although, the analysis of more neuron groups is necessary, we can affirm that we were able to find evidence of the differential effect that electrical stimulation in *bs* and *bf* have in the synchrony between neurons during the subsequent awake-like mode of operation at a population level. On the other hand, the effect of the factor *difference in orientation selectivity* results significant throughout all the generated awake-like activity.

Chapter 8

Population study

In this chapter we gather data from many recordings of several experiments. All the experiments were developed under the same conditions and following the same protocol so the data can be merged in a large data base and analyzed all together. The aim of the chapter is to study differences between the levels of the factor *difference in orientation selectivity* as described in Chapter 7. So, we will study the data from each stimulus separately. In other words, for each stimulus, we would like to compare the synchrony functions observed for each group given by the second factor. We will use a regression approach where the neural synchrony is the response variable while *time* is the independent variable. Within each group defined by the stimulus, we observe the sample $\{(t_{li}, y_{li}) : l = 1, \ldots, L; i = 1, \ldots, n_l\}$, where l gives the level of the factor *difference in orientation selectivity*, n_l is the number of time instants observed for the l-th level of difference in orientation selectivity and t_{li} are the time points where the synchrony is estimated. In our particular problem, the synchrony estimate was computed at the same time points for every level. Thus, we can drop the subscript l for the design points as well as for the sample sizes. We will denote by n the sample size of every level. The data sample is then denoted by $\{(t_l, y_{li}) : l = 1, \ldots, L;\ i = 1, \ldots, n\}$.

We assume the following regression models for the data:

$$y_{li} = m_l(t_i) + \epsilon_{li}, \ \ l = 1, \ldots, L \text{ and } i = 1, \ldots, n, \tag{8.1}$$

where m_l are smooth curves and $\{\epsilon_{li}\}_{i=1}^{n}$ are random errors with finite variance. The aim of the study is to test whether the regression functions are equal:

$$H_0 : m_1 = \cdots = m_L \ \text{ versus } \ H_1 : m_l \neq m_j \ \text{ for some } l \text{ and } j. \tag{8.2}$$

8.1 Estimation of the regression functions

To estimate the regression functions in (8.1) we will use nonparametric regression methods, in particular, local polynomial kernel estimators. As briefly described in Chapter 1, recall that nonparametric regression functions estimators are of the form:

$$\hat{m}_l(t) = \sum_{i=1}^{n} w_i y_{li}, \text{ with } \sum_{i=1}^{n} w_i = 1$$

where w_i is the weight given to the point t by the i-th observation. These weights depend on a kernel function K and a bandwidth h but not on the index l.

Suppose that the $(p+1)$-th derivative of the function $m_l(t)$ exists at the point t_0. The idea of local polynomial estimation lies in approximating the function $m_l(t)$ by a polynomial of degree p in a neighborhood of t_0. The polynomial is locally fitted by solving a weighted least squares problem. In this way, the ν-th derivative of the function m_l at t_0 is estimated by

$$\hat{m}_l^{(\nu)}(t_0) = \nu!\hat{\gamma}_{l\nu} \ \nu = 1, \ldots, p,$$

where the parameter vector $\hat{\gamma}_l^t = (\hat{\gamma}_{l0}, \hat{\gamma}_{l1}, \ldots, \hat{\gamma}_{lp})$ is the one that minimizes:

$$\Psi_l(\gamma) = \sum_{i=1}^{n} \left(y_{li} - \sum_{j=0}^{p} \gamma_j (t_i - t_0)^j \right)^2 K\left(\frac{t_i - t_0}{h} \right), \tag{8.3}$$

where K is a kernel function and h is a suitably chosen bandwidth . The term $\sum_{j=0}^{p} \gamma_j (t_i - t_0)^j$ in (8.3) comes from a Taylor expansion of the function $m(t)$ around $t = t_0$. Note that, in particular, $\hat{\gamma}_{l0}$ estimates $m_l(t_0)$.

To solve the minimization problem in (8.3), it is better to use matrix notation. Let $\mathbf{X}$ be the $n \times (p+1)$ design matrix:

$$\mathbf{X} = \begin{pmatrix} 1 & (t_1 - t_0) & \cdots & (t_1 - t_0)^p \\ 1 & (t_2 - t_0) & \cdots & (t_2 - t_0)^p \\ \vdots & \vdots & & \vdots \\ 1 & (t_n - t_0) & \cdots & (t_n - t_0)^p \end{pmatrix}$$

and let $\mathbf{y}_l \in \mathbb{R}^{n \times 1}$ be the vector of observed data for level l:

$$\mathbf{y}_l = \begin{pmatrix} y_{l1} \\ y_{l2} \\ \vdots \\ y_{ln} \end{pmatrix}$$

and, let $\mathbf{W}$ be the $n \times n$ matrix of weights:

$$\mathbf{W} = \text{diag}\left\{ K\left(\frac{t_i - t_0}{h}\right)\right\}.$$

Observe that neither $\mathbf{X}$ nor $\mathbf{W}$ depend on l.

Finally, the weighted least squares problem in (8.3) can be rewritten as:

$$\hat{\gamma}_l = \text{argmin}_\gamma (\mathbf{y}_l - \mathbf{X}\gamma)^t \mathbf{W} (\mathbf{y}_l - \mathbf{X}\gamma)$$

which, by least squares theory, can be explicitely written as:

$$\hat{\gamma}_l = (\mathbf{X}^t \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{W} \mathbf{y}_l.$$

8.1.1 Bandwidth selection

As in any other kernel procedure, some elements have to be chosen, such as the kernel function, the degree of the polynomial and the smoothing parameter. Although it is very computationally intensive, we have used local variable smoothing parameters because of the sudden changes that occur close to the stimuli. After trying many choices, we decided to choose these parameters by a leave-15-out cross-validation method. We chose the leave-15-out because we preferred to take the risk of oversmoothing than severe undersmoothing the original data. Given a time window of length $2w$, $V_{t_0} = [t_0 - w, t_0 + w]$ around a time point, t_0, we compute the cross-validation function:

$$\text{CV}_{t_0}(h) = \frac{1}{n} \sum_{i=1}^{n} (y_{li} - \hat{m}_{l,-i}(t_i, h))^2 \mathbb{I}\{t_i \in V_{t_0}, \} \tag{8.4}$$

where $\hat{m}_{l,-i}(t_i, h)$ is the local polynomial kernel regression estimator (with smoothing parameter h), computed without the values $t_{i-7}, \ldots, t_i, \ldots, t_{i+7}$ and evaluated at t_i. Observe that only the design points inside the time window V_{t_0} are used to select the optimal h for t_0. The cross-validation smoothing parameter, $h_{\text{CV}}(t_0)$, is the one that minimizes $\text{CV}_{t_0}(h)$ in V_{t_0}.

8.2 Comparison of the regression functions

To test the hypothesis of equality of regression functions given in (8.2) we will consider three different test statistics. The first two tests compare the estimated curves (or a characteristic of the curves) for every level with the pooled one, computed using all the data in a single regression model. This is, if the null hypothesis is true, all the data comes from the same model, and, therefore, it can be used all together to estimate the common regression function: $m = m_1 = \cdots = m_L$.

8.2.1 Estimation of the pooled regression function

We will denote by $\hat{m}_P(t)$ the local polynomial kernel estimator of the function $m(t)$ when all the pooled data sample was used to compute it. To fit $\hat{m}_P$, another weighted least square problem must be solved.

Let $\mathbf{y}^P = (\mathbf{y}_1^t, \ldots, \mathbf{y}_n^t)^t \in \mathbb{R}^{Ln\times 1}$ be the pooled data sample and let $\mathbf{X}^P$ be the $Ln \times (p+1)$ design matrix for this new problem. Note that $\mathbf{X}^P$ is formed by L blocks of matrices equal to $\mathbf{X}$:

$$\mathbf{X}^P = \left.\begin{pmatrix} \mathbf{X} \\ \mathbf{X} \\ \vdots \\ \mathbf{X} \end{pmatrix}\right\} L \text{ times}$$

Now, let us denote by $\mathbf{W}^P$ the $nL \times nL$ block-diagonal matrix of weights for the combined problem. Observe that $\mathbf{W}^P$ can be constructed by repeating the elements of the diagonal of $\mathbf{W}$ L times:

$$\mathbf{W}^P = \left.\overbrace{\begin{pmatrix} \mathbf{W} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{W} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{W} \end{pmatrix}}^{L \text{ times}}\right\} L \text{ times}$$

Finally, let $\hat{\gamma}^P$ be the solution vector for the weighted least squares problem,

which can be written as

$$\hat{\gamma}_l = \text{argmin}_\gamma (\mathbf{y}^P - \mathbf{X}^P\gamma)^t \mathbf{W}^P (\mathbf{y}^P - \mathbf{X}^P\gamma).$$

Therefore,

$$\hat{\gamma}^P = \left[(\mathbf{X}^P)^t \mathbf{W}^P \mathbf{X}^P\right]^{-1} (\mathbf{X}^P)^t \mathbf{W}^P \mathbf{y}^P. \tag{8.5}$$

Observe that,

$$\begin{aligned}(\mathbf{X}^P)^t \mathbf{W}^P \mathbf{X}^P = \left(\mathbf{X}^t \cdots \mathbf{X}^t\right) \begin{pmatrix} \mathbf{W} & \cdots & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \cdots & \mathbf{W} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \vdots \\ \mathbf{X} \end{pmatrix} = \\ = \mathbf{X}^t\mathbf{W}\mathbf{X} + \cdots + \mathbf{X}^t\mathbf{W}\mathbf{X} = L\mathbf{X}^t\mathbf{W}\mathbf{X}\end{aligned}$$

then,

$$\left[(\mathbf{X}^P)^t \mathbf{W}^P \mathbf{X}^P\right]^{-1} = \frac{1}{L}\left[\mathbf{X}^t\mathbf{W}\mathbf{X}\right]^{-1}.$$

On the other hand,

$$\begin{aligned}(\mathbf{X}^P)^t \mathbf{W}^P \mathbf{y}^P = \left(\mathbf{X}^t \cdots \mathbf{X}^t\right) \begin{pmatrix} \mathbf{W} & \cdots & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \cdots & \mathbf{W} \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_L \end{pmatrix} = \\ = \mathbf{X}^t\mathbf{W}\mathbf{y}_1 + \cdots + \mathbf{X}^t\mathbf{W}\mathbf{y}_L = \mathbf{X}^t\mathbf{W}\sum_{l=1}^{L}\mathbf{y}_l.\end{aligned}$$

And, therefore, $\hat{\gamma}^P$ in (8.5) is

$$\begin{aligned}\hat{\gamma}^P = \left[(\mathbf{X}^P)^t \mathbf{W}^P \mathbf{X}^P\right]^{-1} (\mathbf{X}^P)^t \mathbf{W}^P \mathbf{y}^P = \frac{1}{L}\left[\mathbf{X}^t\mathbf{W}\mathbf{X}\right]^{-1}\mathbf{X}^t\mathbf{W}\sum_{l=1}^{L}\mathbf{y}_l = \\ = \left[\mathbf{X}^t\mathbf{W}\mathbf{X}\right]^{-1}\mathbf{X}^t\mathbf{W}\bar{\mathbf{y}}\end{aligned}$$

with $\bar{\mathbf{y}} = \frac{1}{L}\sum_{l=1}^{L}\mathbf{y}_l$. Then, $\hat{\gamma}^P$ is the solution for the weighted least squares problem for the local polynomial estimation of $\bar{\mathbf{y}}$. In summary, the local polynomial estimator for the combined sample can be obtained by fitting the estimator to the sample of the averages of the L curves at each time point: $\{(t_i, \bar{y}_i) : i = 1, \ldots, n\}$, with $\bar{y}_i = \frac{1}{L}\sum_{l=1}^{L} y_{li}$.

8.2.2 Hypothesis tests

We now present three tests which are studied in detail, for the dependent errors case, by Vilar-Fernández et al. (2007).

Test A. The first test we consider compares the nonparametric variance estimator of the pooled sample, $\hat{\sigma}_P^2$, and a convex combination of the nonparametric variance estimators of each sample, $\hat{\sigma}_C^2$. The test statistic is

$$\hat{Q}^{(A)} = \hat{\sigma}_P^2 - \hat{\sigma}_C^2,$$

where the estimators are defined by

$$\hat{\sigma}_P^2 = \frac{1}{n}\sum_{l=1}^{L}\sum_{i=1}^{n}(y_{li} - \hat{m}_P(t_i))^2 \text{ and}$$

$$\hat{\sigma}_C^2 = \frac{1}{n}\sum_{l=1}^{L}\sum_{i=1}^{n}(y_{li} - \hat{m}_l(t_i))^2.$$

This last estimator was introduced by Hall and Marron (1990) and the test was discussed in Dette and Neumeyer (2001) for the independence case.

Test B. The second test was proposed by Young and Bowman (1995) and comprises the difference between the function estimators themselves:

$$Q^{(B)} = \frac{1}{n}\sum_{l=1}^{L}\sum_{i=1}^{n}(\hat{m}_P(t_i) - \hat{m}_l(t_i))^2.$$

Test C. The third, and last, test we consider is a Crámer-von-Mises type test. The test statistic is

$$Q^{(C)} = \sum_{l=2}^{L}\sum_{s=1}^{l-1}\int(\hat{m}_l(t) - \hat{m}_s(t))^2\omega_{ls}(t)\mathrm{dt},$$

where $\omega_{ls}(t)$ are weight functions defined in the support of the design variables. This test has been discussed in the independence case by King et al. (1991) and Kulasekera (1995), among others.

To calibrate the distribution of these test statistics under the null hypothesis, we will use a bootstrap procedure. As it seems natural to assume that the errors of the model in (8.1) are correlated, we propose a stationary bootstrap procedure. Vilar-Fernández et al. (2007) discuss these and other

bootstrap tests. The bootstrap procedure we use to calibrate the test statistics is slightly different from one of their proposals and it is described in the next section.

8.3 Bootstrap procedure

To calibrate the distributions of the test statistics, and test the null hypothesis at a level α, we use a stationary bootstrap procedure (Politis and Romano (1994)) for the residuals of model (8.1). We decided to use a nonparametric bootstrap, since no parametric structure seems plausible for the residuals. Note that, under the null hypothesis, all the regression curves are equal. This is the reason why the bootstrap samples are built adding resamples of the residuals to the regression function estimate obtained from the combined sample. This is, the observations from the L groups are averaged and used to estimate the regression function $m(x)$, using the local polynomial kernel estimator, $\hat{m}_P(t)$, already described. Nevertheless, the residuals are computed from the individually estimated curves, $\hat{m}_l$. With this in mind, let us describe the bootstrap algorithm to be used for the test $\hat{Q}^{(\bullet)}$, with $\bullet = A, B, C$:

1. Compute the test statistic, $\hat{Q}^{(\bullet)}_{obs}$ for the observed sample

$$\{(t_i, y_{li}) : i = 1, \ldots, n;\ l = 1, \ldots, L\}.$$

2. Obtain the residuals for each individual estimation

$$\hat{\epsilon}_{li} = y_{li} - \hat{m}_l(t_i),\ l = 1, \ldots, L;\ i = 1, \ldots, n.$$

3. Obtain a bootstrap sample of the residuals, for each group, as follows:

   3.1 Fix a real number $p_{boot} \in [0, 1]$.

   3.2 Draw $\hat{\epsilon}^*_{l1}$ randomly from $\{\hat{\epsilon}_{l1}, \ldots, \hat{\epsilon}_{ln}\}$.

   3.3 If $\hat{\epsilon}^*_{li} = \hat{\epsilon}_{lj}$ for $j = 1, \ldots, n$ and $i < n$ has been drawn, then, $\hat{\epsilon}^*_{l(i+1)}$ is chosen as $\hat{\epsilon}_{l(j+1)}$ with probability p_{boot} and drawn randomly from $\{\hat{\epsilon}_{l1}, \ldots, \hat{\epsilon}_{ln}\}$ with probabilty $1 - p_{boot}$. In the particular case of $j = n$, $\hat{\epsilon}_{l(j+1)}$ is replaced by $\hat{\epsilon}_{l1}$.

4. Obtain the bootstrap resamples by

$$y^*_{li} = \hat{m}_P(t_i) + \hat{\epsilon}^*_{li}\ l = 1, \ldots, L;\ i = 1, \ldots, n.$$

   where $\hat{m}_P$ is the regression function estimated from the pooled sample.

Table 8.1: Number of synchrony curves obtained for every group of neurons.

	0°	22.5°	45°	67.5°	90°
bs	550	687	568	367	154
bf	588	740	562	390	152

5. Compute $\hat{Q}^{(\bullet)*}$ using the bootstrap resample $\{(t_i, y^*_{li}) : i = 1, \ldots, n;\ l = 1, \ldots, L\}$.

6. Repeat Steps 3–5 a large number, B, of times to obtain $\hat{Q}^{(\bullet)*1}, \ldots, \hat{Q}^{(\bullet)*B}$.

7. Finally, compute the desired percentile, $\hat{Q}^{(\bullet)*}_{(1-\alpha)}$, of the bootstrap test statistics and reject the null hypothesis if

$$\hat{Q}^{(\bullet)}_{obs} > \hat{Q}^{(\bullet)*}_{(1-\alpha)}.$$

8.4 Results

We performed the analysis using data from 9 experiments. The synchrony was estimated for each possible pair of neurons and then an average was computed for each level of the factors *stimulus* and *difference in orientation selectivity*. The total amount of synchrony curves obtained for each group are shown in Table 8.1

The synchrony functions were estimated, every second, in a 210 s interval including the time point when the stimulus was applied. The curves were averaged within each group. These averages are the *raw* data for our subsequent analyses and can be observed in Figure 8.1.

Regarding the regression estimator, a Gaussian kernel was used to locally fit degree $p = 3$ polynomials. As the first derivative of our functions are also of interest and odd degrees are often recomended (Fan and Gijbels (1996)), $p = 3$ was a natural choice.

The local bandwidths were selected using the cross-validation procedure described in Section 8.1.1, using 40 s overlapping sliding windows. This is, at each time point t_0, the data that fell in the time window $V_{t_0} = [t_0 - 20, t_0 + 20]$ were used to find the cross-validation smoothing parameter, $h_{\mathrm{CV}}(t_0)$. More precisely, $h_{\mathrm{CV}}(t_0)$ was chosen as the value that minimized $\mathrm{CV}(h)$ from a grid of values spanning from 1.5 to 5. The grid comprised values from 1.5 to

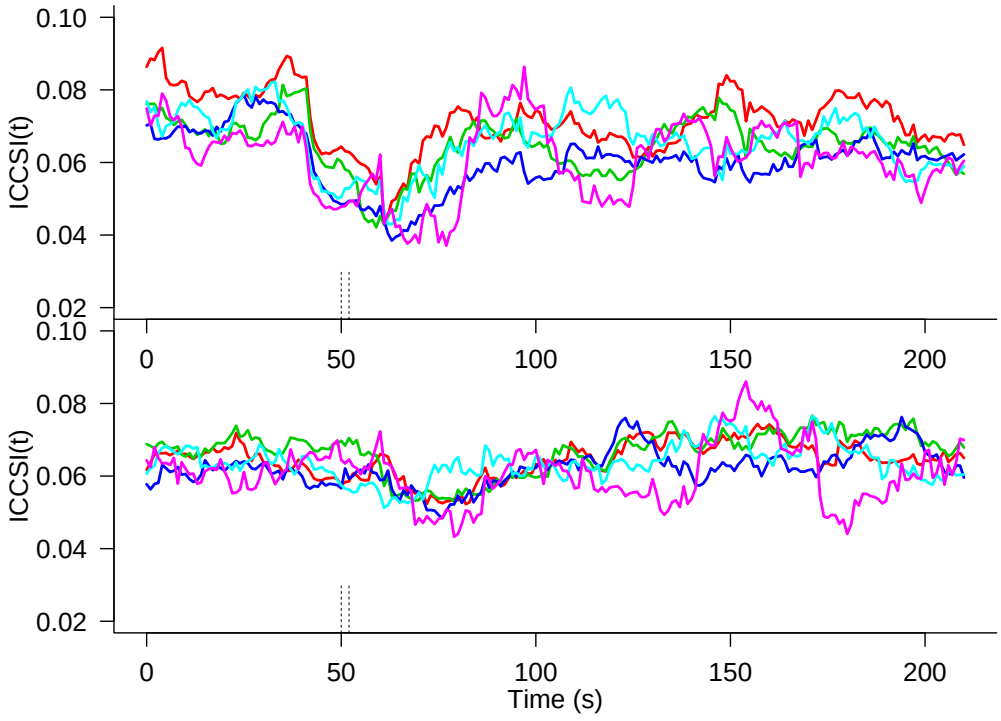

Figure 8.1: ICCSI functions averaged in each group defined by the stimuli *bs* (top panel) and *bf* (bottom panel) and the difference in orientation selectivity. Difference in orientation selectivity levels are shown in different colors: 0° (red), 22.5° (green), 45° (blue), 67.5° (cyan) and 90° (magenta).

1.9 every 0.1 s, from 2 to 20 every 0.5 s and from 25 to 50 every 5 s. After the optimal bandwidth was found for each time point the curve $h_{\mathrm{CV}}(t)$ was smoothed using a moving average procedure as it is natural to think that the optimal bandwidths should evolve smoothly on time. Figure 8.2 shows the local polynomial estimates for the regression functions in (8.1).

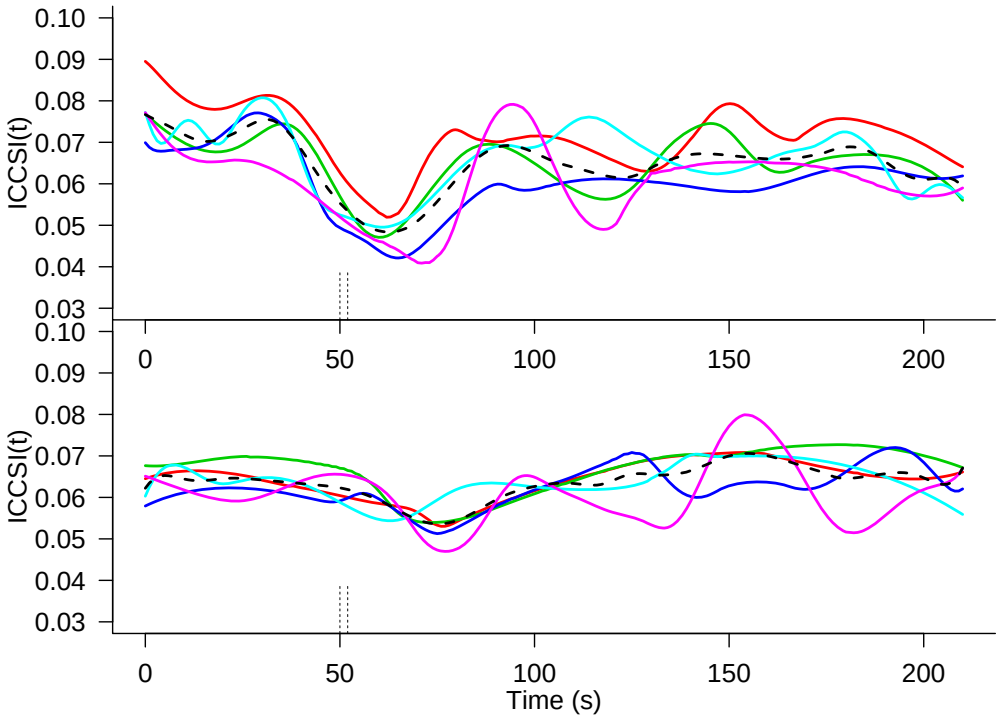

Figure 8.2: Local polynomial estimates of the ICCSI curves using local variable bandwidths chosen by leave-15-out cross validation. Top panel: curves corresponding to the *bs* stimulus. Bottom panel: curves corresponding to the *bf* stimulus. Difference in orientation selectivity levels are shown in different colors: 0° (red), 22.5° (green), 45° (blue), 67.5° (cyan) and 90° (magenta). The estimate of the pooled regression curve is also shown (black dashed line).

In order to perform the hypothesis test, the pooled regression function is also estimated. For this aim, $\hat{m}_P(t)$ was computed with the average of the original 5 curves. Again, to fit the estimator, the local bandwidths were selected via a leave-15-out cross-validation procedure. This estimate is shown in Figure 8.2 with a dashed black line.

For the resampling procedure, $B = 500$ bootstrap samples were used. Once the residuals, $\hat{\epsilon}^*_{li}$ were sampled and the bootstrap curves built, variable local bandwidths were searched just as for the original data. On the other hand, the pooled bootstrap curves were also obtained and their bandwidths selected. Although the resampling procedure itself is not very time consuming, the optimization procedure for the bandwidth for each curve is somehow slow. Regarding the parameter p_{boot} of the bootstrap algorithm, $p_{boot} = 0.94$ was chosen. In this way, as the sample lengths are 210, the blocks in the resampling procedures are of length 16 on the mean and therefore, around 13 blocks are used.

The three tests lead to the same conclusion: there exist enough evidence to reject the null hypothesis of equality of regression functions. Taking into account the whole time interval, the p-values obtained were smaller than 0.002. These results hold for both stimuli. It seemes reasonable to check whether these differences are still found if we just look at a period of time right after the stimulus onset. For this aim we repeated the tests A, B and C for a period of time from 10 s before the stimulus onset to 30 s after the stimulus ($[40\,s, 80\,s]$). Again, the p-values found are smaller than 0.002.

To check whether the results were merely a consequence of a vertical shift of the curves, this is, if the general shape of the functions is the same but some presented regularly more synchrony than the others, we performed the test after aligning the curves. For this aim, the curves were shifted vertically so that they all coincided at the time were the stimulus occured and the tests were carried out again. Figure 8.3 shows these estimated aligned regression functions. The results were again the same. We can reject the hypothesis of equality of regression functions, for both stimuli, when we look at the whole time interval as well as when we only take into account a smaller interval including the stimulus.

One of the advantages of local polynomial fitting is that it allows to easily estimate the derivatives of the regression functions at the same time as the function itself. We have used degree 3 polynomials in order to obtain the estimates of the first derivative of the regression functions for inspection. Anyway, to estimate the derivatives we have increased the bandwidths by a factor of five, to be able to see the global tendencies more than the local features. Figure 8.4 shows the derivatives in a time period of 130 s, including the stimulus, of the regression curves shown in Figure 8.2, with the same color code.

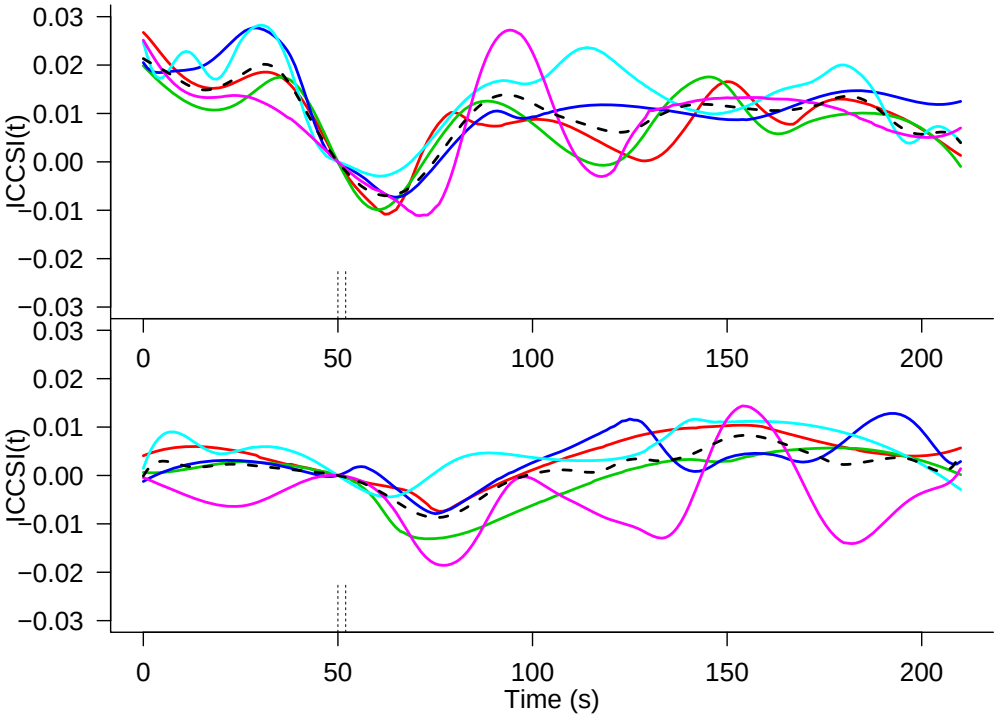

Figure 8.3: Local polynomial estimates of the aligned ICCSI curves using local variable bandwidths chosen by leave-15-out cross validation. Top panel: curves corresponding to the *bs* stimulus. Bottom panel: curves corresponding to the *bf* stimulus. Difference in orientation selectivity levels are shown in different colors: 0° (red), 22.5° (green), 45° (blue), 67.5° (cyan) and 90° (magenta). The estimate of the pooled regression curve is also shown (black dashed line).

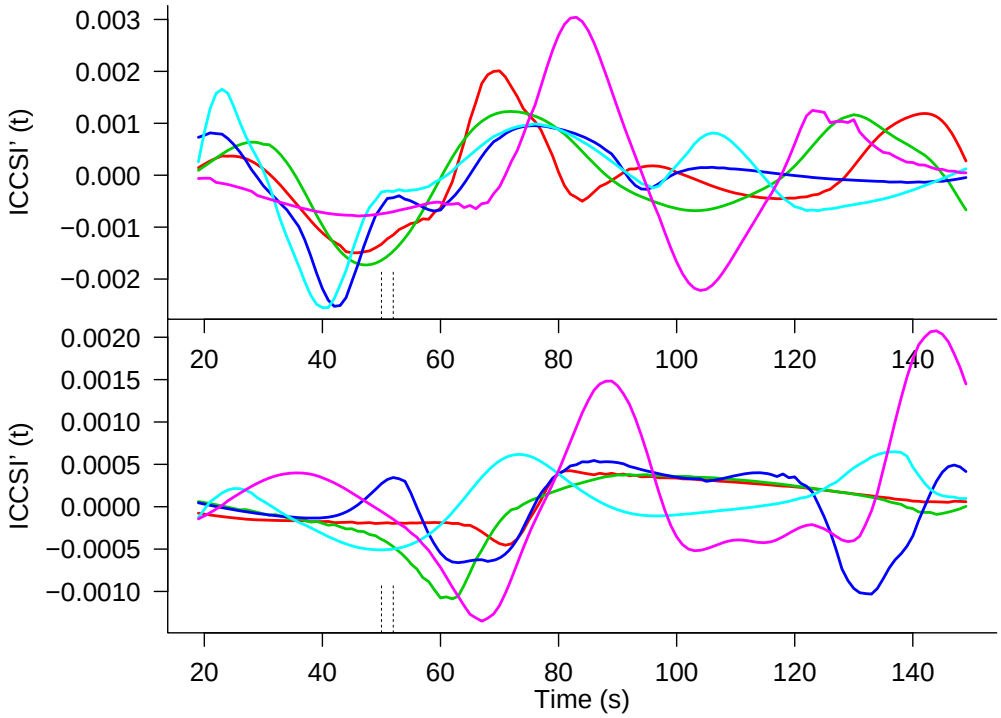

Figure 8.4: Local polynomial estimates of the derivatives of the ICCSI curves. Top panel: curves corresponding to the *bs* stimulus. Bottom panel: curves corresponding to the *bf* stimulus. Difference in orientation selectivity levels are shown in different colors: 0° (red), 22.5° (green), 45° (blue), 67.5° (cyan) and 90° (magenta).

Let us focus, first, in the case of the *bs* stimulus. From Figures 8.2 and 8.4 we can make a few statements. The decrease in synchrony observed around the stimulus onset is more rapid for the groups of differences 0°, 22.5° and 45° than for the groups 67.5° and 90°. Although from the regression curves it seems that group 22.5° desynchronizes faster than level 0°, Figure 8.4 shows that, when computed with a wider bandwidth, the derivatives are very similar and, therefore, they desynchronize at the same rate. On the other hand, level 67.5° (blue curve) seems to drop more rapidly than the other ones, as its derivative is the smallest one at some point around 10 s before the stimulus. But, it is also true that, at the stimulus, the velocity of the drop diminishes and, also, the interval of time until resynchronization is longer and the drop

is the most profound, which suggests a larger desynchronization. Let us look now at around 25 s after the stimulus. Disregarding level 90°, it can be observed that, the smaller the difference in orientation selectivity, the larger the speed of the increase in synchrony. This fact could be explained by a stronger functional or anatomical relation between neurons with the same preferred orientation. From all this, we can state that the smaller the difference in favorite orientation, the neurons desynchronize faster as well as resynchronize faster.

The case of the 90° level is strange. It looks like the drop in synchrony is much slower than for the other levels, but, on the contrary, after 20–25 s have passed from the stimulus, the resynchronization occurs very rapidly. This is surprising as it contradicts the previous statement that the more related the cells are regarding orientation selectivity, the faster they would recover from the desynchronization provoked by the stimulus. This fact, and the fact that the curve has a lot of ups and downs indicate that there might be something special about these data. It is important to notice that this level has a much smaller sample size than the other ones. This fact could be affecting the estimation, making the curve a lot less smoother than it should, maybe because of a large variability in the original curves.

The case of the stimulus *bf* is more complicated than the one of *bs*. To begin with, the effect of this stimulus is not as clear as for the other one. If some, the effect can be observed around 10 s after the stimulus onset. We can observe that for the levels 0°, 22.5° and 45° the development of the synchrony is very similar. The rates of resynchronization are practically the same. Also, the decreasing of synchrony occurs practically at the same time, although for the level 0° it seems to be less abrupt than for the other two. Again, the behavior of the level 90° seems strange as the curve is very wiggly, and, the same occurs in this case for 67.5°.

8.5 Chapter conclusions

In this chapter we carried out a population analysis in order to study differences in synchrony strength between the levels of the factor *difference in orientation selectivity*. A nonparametric regression approach was used and three different test statistics computed to test the coincidense of the regression functions. The results show that the differences are statistically significant proving that the associations and the dynamics of those associations are related to the functional affinity among neurons. Moreover, we are

in position of stating a description of those relations although a deeper statistical analysis would be needed to prove this fact: neurons desynchronize and resynchronize faster when their favorite orientation differ less.

Chapter 9

Discussion and conclusions

The aim of this thesis has been to present statistical methods useful to measure synchrony dynamics of pairs of neurons under low firing rate activity. Nonparametric methods have been used to study characteristics of spike trains, to present synchrony measures and to develop hypothesis tests to better understand the synchrony mechanisms.

9.1 Discussion

The biological problem consists in investigating the synchrony dynamics between neurons in V1 of anesthetized cats and its relationship with certain factors. The first factor under study is defined by a controlled electrical stimulation in any of two precise sites of the cortex: the brainstem and the basal forebrain. These two regions regulate the transitions of the sleep-wake cycle and, when stimulated, provoke the change from the sleep-like activity to the typical awake-like activity. The stimulations are performed one at a time and, therefore, the effect of those stimulations can be differentially analyzed. On the other hand, the second factor emerges from an intrinsic characteristic of each neuron: orientation selectivity. This factor is defined as the difference between preferred orientations and can be only included in the analysis of groups of neurons.

Before studying relations between pairs of neurons, a single spike train analysis seems reasonable. Chapter 3 is devoted to study the stationarity of neural activity under spontaneous activity and controlled conditions as the ones guaranteed by the experimental setting. The findings are that stationarity is acceptable for our data during the sleep-like period (before the

electrical stimulation is applied) and that this stationarity is recovered after the stimulus effect has vanished. This is an interesting fact that gives some solid ground where to start studying pairwise associations from.

Chapters 4, 5 and 6 are devoted to the study of methods to capture the nature of neural synchrony under unfavorable situations such as low firing rates and small amount of trials. Usually, if many trials are available, neural associations are estimated using binned trains methods across trials but, this approach is not possible in our scenarios due to the lack of many trials. To overcome this drawback, our methods are based on a *sliding windows* procedure. That is, the information of small neighborhoods of the time point under study is used to estimate the activity at that point. This approach has the obvious disadvantage that the temporal resolution is affected and therefore, conclusions can be in a general basis although not with a very precise timing. Anyway, informative results have been obtained which encourage us to continue in this line of study.

Single spike trains analysis is based mainly on the study of inter spike intervals. Therefore, a natural way of moving on to pairwise associations is the study of times between the spikes of the two trains. The cross inter spike intervals are presented in Chapter 3 as a first approach to neural synchrony. The results are interesting regarding the comparison of the density functions of the CISI before and after the stimulus onset. These comparisons show a clear change in the density functions when the stimuli appears. Despite this fact looked promising, the final performance of the CISI based measure is not satisfactory. The resulting curves are too rough and the hypothesis tests are not able to find any relevant information.

The results presented in Chapters 5 and 6 are considerably more revealing. The ICCSI method presented in Chapter 5 is based on the cross-correlation function and accounts for the area under this function in a neighborhood of zero. On the other hand, the CSM, presented in Chapter 6, counts the amount of small cross nearest spike intervals over the total count, again in small time windows. Both methods are flexible in the sense that the time that occurs between two spikes to be considered synchchronous can be chosen by the researcher depending on the context.

Different approaches are used with the previously mentioned methods. For the ICCSI, the estimated *raw* values are used, developing a complete nonparametric bootstrap hypothesis tests. In the case of the CSM, a semiparametric approach is used: the values obtained by the synchrony estimator

are modeled with generalized additive models. Thus, the resulting curves obtained by the CSM method are much smoother than the ones obtained with the ICCSI. This fact can result in a drawback or in an advantage depending on the context.

The results obtained with ICCSI and CSM are comparable. Although the applications in Chapter 6 are only shown for one pair of neurons, namely, N1 and N3a, we can still make comparisons of the outcomes based on that pair. In both cases, significant differences are found between the synchrony during the sleep-like and the awake-like periods with the application of the *bs* stimulus. In the ICCSI case, the period when these differences are significant starts few seconds after the stimulation time (Figures 5.3 and 5.4). With CSM, the results are somehow different. The period in which the differences are found to be significant starts around 15 s after the stimulus. This delay could be an artifact caused by the extra smoothing that the splines give (Figure 6.5). On the other hand, in the *bf* case, ICCSI cannot find any differences between the sleep-like and awake-like synchrony while CSM finds them in a period of 18 s starting around 20 s after the stimulation occurs. Regarding the differential effect of the stimuli, ICCSI barely finds differences between the synchrony induced the the two stimuli. The period in which these differences are roughly accepted lasts for around 15 s right after the stimuli onset (Figure 5.11). In the CSM case, the differences are found with much more evidence for a period of 20 s starting from around 17 s after the stimuli appear.

An important fact, discussed in Chapter 6, is that the estimated synchrony is not a mere effect of the spiking activity. The estimate of the expected synchrony in the independence case results different from the observed one. Anyhow, there exists a short time interval. right after stimulation where the observed synchrony can be explained by some increasing in the firing rate.

The differences just described are also studied at group and population levels. In Chapter 7 we use an existing functional data analysis technique to study the effect of the stimuli together with a second factor: *difference in orientation selectivity*. For this analysis we use the synchrony curves estimated using the ICCSI method presented in Chapter 5 from one group of simultaneously recorded neurons. As the variable *difference in orientation selectivity* is presented as a fixed effect, the problem can be considered as a two way ANOVA model with functional dependent response. The results indicate that the functional relationship between neurons, given by the orientation selectivity affinity, has a great effect in the reaction to the stimuli.

However, this effect does not depend on the applied stimulus, as the interaction between the two factors does not result significant.

To go a little bit further and get a global overview, in Chapter 8 we gather the recordings of nine experiments. As we found out in Chapter 7 that the interaction between the stimulus and the level of the factor *difference in orientation selectivity* is not relevant, the population analysis is carried out for each stimulus separately. The ICCSI curves for each group defined by the difference in orientation selectivity are averaged together and a nonparametric regression framework is used for the comparison of the average curves. The results are conclusive in the sense that the differences are found significant with extremely low p-values. Moreover, a simple overview of the resulting regression curves and their derivatives, close to the stimuli, suggests a relation between the affinity between neurons regarding their preferred orientation and the strength of their association. The functional affinity between two neurons seems to be represented in the way their synchrony behaves. The hypothesis that remains to be proven is: the stronger the functional affinity of two neurons regarding their orientation selectivity, the faster they will transit from a synchronized state to a desynchronized state and viceversa.

Overall, this work is a contribution to the development of statistical tools for neuroscience. Although the methods were thought for and applied to a very particular problem, we believe the methods here discussed can be used in many other contexts where low firing rates are an issue. On the other hand, this work also emphasizes the usefulness of nonparametric methods and the bootstrap in this context. Nonparametric statistic is a natural choice when no parametric model seems plausible for the data, as commonly occurs with spike data. On the other hand, bootstrap techniques are powerful tools which are easy to implement and, although they can be time consuming, they provide a great alternative to parametric inference using minimal mathematical assumptions.

9.2 Conclusions

Considering the objectives of the study given in Section 2.1 and the discussion above, we can conclude that:

- A bootstrap test to detect dependence among the inter-spike interval of a single neuron has been proposed. From its the application to the experimental data we conclude that it is able to successfully detect the presence of dependence among the ISI.

- The proposed method to detect fluctuations in neuronal synchronization based on the density function of the cross inter-spike intervals is not stable enough to analyze low firing neuronal activity. Nevertheless, the visual inspection of these densities is informative.

- We have developed two methods, CSM and $ICCSI$, that are able to successfully estimate neuronal synchrony under low firing rate scenarios. Stimulation-induced variations in this measured synchrony strength were differentiated by means of bootstrap hypothesis tests.

- Differences in neural pairwise synchronization depending on the activation of the brainstem vs the basal forebrain were assessed through the development of bootstrap hypothesis tests. Analyses were carried out both at a single pair level and at a group level arriving to the same conclusions.

- We have proposed bootstrap tests that allow to affirm that there exist significant differences in the synchronization dynamics between the sleep-like state and the awake-like state regarding the functional affinity between neurons given by their preferred orientation.

9.3 Future work

Although the main objectives of the study have been reached, there still remain many aspects to improve and different analyses to implement. We will point out some of these aspects in this section:

- Apply our synchrony measures to controlled *in-vitro* data in which the correlation is known in order to be able to test the methods. These type of data has been already generated by researchers of the Theoretical Neuroscience/Neuroinformatics research group of the Freie Universität at Berlin. A collaboration with this group is planned for the coming months.

- Study the possible implementation of regression methods where the response variable is the synchrony curve and the dependent variable is the difference in orientation selectivity. The orientation selectivity could be computed with a better resolution using nonparametric density estimation techniques for circular data. Therefore, it could be included in the models as a random effect. The prediction power of a model like this would be very useful in the sense that our predictions could be possibly extended to other areas of the cortex other than V1.

- Implement the nonparametric bootstrap stated in Chapter 7 for comparison with the parametric bootstrap.
- Study new models for the derivatives presented in Chapter 8 in order to statistically asses the relationship between the functional affinity of neurons and their synchrony strength.
- Study the velocity of desynchronization and resynchronization at a pairwise level developing methods to test for differences in such profiles.
- Apply our methods in other contexts and compare the results with the obtained by other standard methods.

References

Aertsen, A.M. (1989). Dynamics of neuronal firing correlation: modulation of "effective connectivity". *Journal of Neurophysiology* 61, 900–917.

Aitken, A. (1935). On least squares and linear combination of observations. *Proc. Royal Society of Edinburgh* 55, 42–48.

Barbieri, R., Quirk, M.C., Frank, L.M., Wilson, M.A. and Brown, E.N. (2001). Construction and analysis of non-Poisson stimulus-response models of neural spiking activity. *Journal of Neuroscience Methods* 105, 25–37.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B* 57, 289–300.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* 29, 1165–1188.

Bowman, A., Hall, P. and Prvan, T. (1998). Bandwidth selection for the smoothing of distribution functions. *Biometrika* 85, 799–808.

Brillinger, D.R. (1992). Nerve cell spike train data analysis: A progression of technique. *Journal of the American Statistical Association* 87, 260–271.

Brown, N.E., Kass, R.E. and Mitra, P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neuroscience* 7, 456–461.

Burlet, S., Tyler, C.J. and Leonard, C.S. (2002). Direct and indirect excitation of laterodorsal tegmental neurons by hypocretin/orexin peptides: Implications for wakefulness and narcolepsy. *The Journal of Neuroscience* 22, 2862–2872.

Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74, 829–836.

Cuevas, A., Febrero, M. and Fraiman, R. (2007). Robust estimation and classification for functional data via projection-based depth notions. *Computational Statistics* 22, 481–496.

Cuesta-Albertos, J. and Febrero-Bande, M. (2010). A simple multiway ANOVA for functional data. *Test* 19, 537–557.

Cuesta-Albertos, J.A., Fraiman, R. and Ransford, T. (2007). A sharp form of the Cramer–Wold theorem. *Journal of Theoretical Probability* 20, 201–209.

Dayan, P. and Abbot, LF. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, The MIT Press.

Dette, H. and Neumeyer, N. (2001). Nonparametric analysis of covariance. *The Annals of Statistics* 29, 1361–1400.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* 7, 1–26.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall.

Faes, C., Geys, H., Molenberghs, G., Aerts, M., Cadarso-Suárez, C., Acuña, C. and Cano, M. (2008). A flexible method to measure synchrony in neuronal firing. *Journal of the American Statistical Association* 103, 149–161.

Fan, J. and Gijbels, I. (1995). Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society: Series B* 57, 371–394.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman and Hall.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. Springer.

Fraiman, R. and Mũniz, G. (2001). Trimmed means for functional data. *Test* 10, 419–440.

Gerstein, G.L. and Mandelbrot, B. (1964). Random walk models for the spike activity of a single neuron. *Biophysical Journal* 4, 41–68.

Gerstein, G.L. and Perkel, D.H. (1969). Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science* 164, 828–830.

Gerstner, W. and Kistler, W. (2002). *Spiking Neuron Models, Single Neurons, Populations, Plasticity*. Cambridge University Press.

González-Montoro, A.M., Cao, R., Espinosa, N., Mariño, J. and Cudeiro, J. (2011). Autocorrelation Measures and Independence Tests in Spike Trains. In: *Modern Mathematical Tools and Techniques in Capturing Complexity* (Pardo, L., Narayanaswamy, B. and Gil, M.A.) 471–484. Springer: Complexity.

Grün, S. (1996). *Unitary Joint-Events in Multiple-Neuron Spiking Activity: Detection, Significance, and Interpretation*. Reihe Physik, Verlag Harri Deutsch.

Grün, S., (2009). Data-driven significance estimation for precise spike correlation. *Journal of Neurophysiology* 101, 1126–1140.

Grün, S., Diesmann, M. and Aertsen, A. (2002). Unitary events in multiple single-neuron spiking activity: I. Detection and significance. *Neural Computation* 14, 43–80.

Hall, P. and Marron, J.S. (1990). On variance estimation in nonparametric regression. *Biometrika* 77, 415–419.

Hastie, T.J. and Tibshirani, R.J. (1990). *Generalized Additive Models*. Chapman and Hall.

Henry, G.H., Dreher, B. and Bishop P.O. (1974). Orientation specificity of cells in cat striate cortex. *Journal of Neurophysiology* 37, 1394–1409.

Hu, B., Steriade, M. and Deschênes, M. (1989). The effects of brainstem peribrachial stimulation on perigeniculate neurons: the blockage of spindle waves. *Neuroscience* 31, 1–12.

Hubel, D.H. and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cats visual cortex. *Journal of Physiology* 160, 106–154.

Kass, R.E. and Ventura, V. (2001). A spike-train probability model. *Neural Computation* 13, 1713–1720.

Kass, R.E., Ventura, V., Brown, E.N. (2005). Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology* 94, 8–25.

King, E., Hart, J.D. adn Wehrly, T.E. (1991). Testing the equality of two regression curves using linear smoothers. *Statistics and Probability Letters* 12, 239–247.

Kulasekera, K.B. (1995). Comparison of regression curves using quasi-residuals. *Journal of the American Statistical Association* 90, 1085–1093.

Kruskal, P.B., Stanis, J.J., McNaughton, B.L. and Thomas, P.J. (2007). A binless correlation measure reduces the variability of memory reactivation estimates. *Statistics in Medicine* 26, 3997–4008.

Künsch, H.R. (1989). The Jackknife and the bootstrap for general stationary observations. *The Annals of Statistics* 17, 1217–1241.

Liu, R.Y. and Singh (1992). Moving blocks jacknife and bootstrap capture weak dependence. In: *Exploring the Limits of Bootstrap* (LePage, R., Billard, L.) 225–248. John Wiley and Sons.

Mariño, J. and Cudeiro, J. (2003). Nitric oxide-mediated cortical activation: A diffuse wake-up system. *The Journal of Neuroscience* 23, 4299–4307.

Nadaraya, E.A. (1964). Remarks on nonparametric for density functions and regression functions. *Theory of Probability and its Applications* 15, 134–137.

Nawrot, M., Aertsen, A. and Rotter, S. (1999). Single-trial estimation of neuronal firing rates: From single-neuron spike trains to population activity. *Journal of Neuroscience Methods* 94, 81–92.

Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics* 33, 1065–1076.

Perkel, D.H., Gerstein G.L. and Moore, G.P. (1967a). Neuronal spike trains and stochastic processes. I. The single spike train. *Biophysical Journal* 7, 391–418.

Perkel, D.H., Gerstein G.L. and Moore, G.P. (1967b). Neuronal spike trains and stochastic processes. II. Simultaneous spike trains. *Biophysical Journal* 7, 419–440.

Politis, D.N. and Romano, J.P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association* 89, 1303–1313.

Priestley, M.B. and Chao, M.T. (1972). Nonparametric function fitting. *Journal of the Royal Statistical Society: Series B* 34, 385–392.

Quian Quiroga, R., Kraskov, A., Kreuz, T. and Grassberger, P. (2002). Performance of different synchronization measures in real data: A case study on electroencephalographic signals. *Physical Review E* 65, 041903.

Ramsay, J.O. and Silverman, B.W. (2002). *Applied Functional Data Analysis: Methods and Case Studies*. Springer.

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, Second Edition. Springer Series in Statistics.

Rao, C.R. (1973). *Linear Statistical Inference and its Applications*, Second Edition. John Wiley and Sons, New York.

Ruppert, D. and Wand, M.P. (1994). Multivariate locally weighted least squares regression. *The Annals of Statistics* 22, 1346–1370.

Reich, D.S., Victor, J.D. and Knight, B.W. (1998). The power ratio and the interval map: spiking models and extracellular recordings. *Journal of Neuroscience* 18, 10090–10104.

Rodieck, R.W., Kiang, N.Y.S. and Gerstein, G.L. (1962). Some quantitative methods for the study of spontaneous activity of single neurons. *Biophysical Journal* 2, 351–368.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics* 27, 832–837.

Shadlen, M.N. and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience* 18, 3870–3896.

Shadlen, M.N., and Movshon, J.A. (1999). Synchrony unbound: a critical evaluation of the temporal binding hypothesis. *Neuron* 24, 67–77.

Sheather, S.J. and Jones, M.C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B* 53, 683–690.

Silverman, B.W. and Young, G.A. (1987). The Bootstrap: To smooth or not to smooth? *Biometrika* 74, 469–479.

Simonoff, J.S. (1996). *Smoothing Methods in Statistics*. Springer.

Singer, W., Engel, A.K., Kreiter, A.K., Munk, M.H.J., Neuenschwander, S. and Roelfsema, P.R. (1997). Neuronal assemblies: necessity, signature and detectability. *Trends in Cognitive Sciences* 1, 252–261.

Singer, W. (1999). Neuronal synchrony: A versatile code review for the definition of relations? *Neuron* 24, 49–65.

Smith, D.R. and Smith, G.K. (1965). A statistical analysis of the continual activity of single cortical neurones in the cat unanaesthetized isolated forebrain. *Biophysical Journal* 5, 47–74.

Steriade, M. (1994). Sleep oscillations and their blockage by activating systems. *Journal of Psychiatry and Neuroscience* 19, 354–358.

Steriade, M., Gloor, P., Llins, R.R., Lopes da Silva, F.H. and Mesulam, M.M. (1990). Basic mechanisms of cerebral rhythmic activities. *Electroencephalography and Clinical Neurophysiology* 76, 481–508.

Steriade, M., Jones, E.G. and McCormick, D.A. (1997). *Thalamus, Vol I: Organization and function*. Oxford: Elsevier.

Steriade, M., McCormick, D.A. and Sejnowski, T.J. (1993). Thalamocortical oscillations in the sleeping and arousal brain. *Science* 262, 679–685.

Stone, C.J. (1977). Consistent nonparametric regression. *The Annals of Statistics* 5, 595–620.

Tuckwell, H.C. (1988). *Introduction to theoretical neurobiology: Volumen 2, nonlinear and stochastic theories*. Cambridge University Press.

Vilar-Fernández, J.M., Vilar-Fernández, J.A. and González-Manteiga, W. (2007). Bootstrap tests for nonparametric comparison of regression curves with dependent errors. *Test* 16, 123–144.

Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer.

Watson, G.S. (1964). Smooth regression analysis. *Sankhya: Series A* 26, 359–372.

Wand, M.P. and Jones, M.C. (1995). *Kernel Smoothing*. Chapman and Hall.

Wood, S. (2006). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC Texts in Statistical Science.

Wu, C.F.J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *The Annals of Statistics* 14, 1261–1295.

Young, S.G. and Bowman, A.W. (1995). Non-parametric analysis of covariance. *Biometrics* 51, 920–931.

Resumen en Español

El objetivo de esta tesis es presentar herramientas estadísticas para tratar algunos problemas metodológicos en el análisis de datos electrofisiológicos; específicamente, la estimación de la dinámica de la sincronía entre pares de neuronas con bajas tasas de disparo. Los métodos son aplicados a datos reales y se hace inferencia acerca de la sincronía estimada con ellos. El problema biológico fue propuesto al grupo de estadística, MODES, de la Universidade da Coruña, por investigadores del grupo Neurocom, de la misma Universidad. Tanto las preguntas como los datos resultaron un reto desde el punto de vista estadístico. Esta tesis describe los distintos enfoques y propuestas para resolver el problema y el proceso para alcanzar los objetivos.

La neurociencia es el campo de conocimiento que estudia la estructura y función del sistema nervioso, en particular, del cerebro humano. Tiene numerosas áreas de estudio y aúna a muchas disciplinas como, medicina, psicología, biología e ingeniería, entre otras. La electrofisiología es la rama de la neurociencia que estudia las propiedades eléctricas y la actividad eléctrica de las neuronas. Los avances tecnológicos han hecho posible el registro electrofisiológico simultáneo de grupos de neuronas, generando grandes cantidades de datos que requieren de herramientas y metodologías potentes para un posterior tratamiento y análisis. Ciencias como matemáticas, física, estadística y computación se involucran cada día más con la neurociencia, para hacer frente a la gran demanda de métodos para el análisis de la electrofisiología.

Las neuronas son células especializadas, que, junto con las células gliales, son las unidades básicas estructurales y funcionales del sistema nervioso. Estas células están organizadas en grandes y complejas redes y dan forma y conectan las componentes del sistema nervioso. Esto es, transportan información desde las zonas sensoriales, la analizan y llevan las respuestas hacia otras células o zonas musculares.

Las neuronas se caracterizan por su capacidad de propagar información

muy rápidamente a través de largas distancias. La información se transporta en forma de impulsos eléctricos, llamados *potenciales de acción* o *espigas*. Estos impulsos nerviosos son bastante fáciles de registrar dado que son cambios abruptos en el potencial de la membrana de las neuronas, tienen una amplitud relativamente alta ($\sim$100 mV) y duran aproximadamente 1 ms. Para registrarlos, los electrofisiólogos colocan electrodos cerca o dentro de las neuronas. Como los potenciales de acción son todos muy similares, se cree que la información está codificada en las secuencias de ellos. Estas secuencias se denominan *trenes de potenciales de acción* o *trenes de espigas* y son el objeto de estudio principal de esta tesis. Como los principios del procesamiento neuronal de la información no se entienden en su totalidad, la forma en que estos trenes llevan la información es un tema de debate y existen varias propiedades a tener en cuenta. Las tasas de disparos y los tiempos exactos de disparo son las dos líneas más importantes de investigación. Por otro lado, las asociaciones entre neuronas y la sincronía son características clave para entender el código neuronal.

Los datos con los que se trabaja a lo largo de esta tesis provienen de un posible estado del cerebro, denominado actividad espontánea. La actividad espontánea es la actividad que se observa en ausencia de estímulos y puede pensarse como la actividad cerebral en un estado de reposo. En este trabajo, estudiamos la sincronización de las neuronas bajo dos tipos de actividad espontánea: el *estado de sueño inducido por la anestesia* y el *estado de vigilia inducido eléctricamente*. Si bien utilizamos el término *estímulo* para referirnos a la micro-estimulación eléctrica que se utiliza para inducir el estado de vigilia, esta micro-estimulación no altera el estado de actividad espontánea.

Durante el sueño la actividad global del cerebro se caracteriza por ser muy sincronizada, con ondas de mucha amplitud y baja frecuencia. Esta actividad oscilatoria global puede inducirse mediante anestésicos permitiendo así el estudio de propiedades neuronales características de este estado. Durante la fase de vigilia la actividad global del encéfalo no presenta dichas oscilaciones. La actividad global típica durante la vigilia también puede inducirse mediante la aplicación de micro-estimulaciones en ciertas áreas del cerebro, denominadas vías ascendentes activadoras que regulan el paso del estado de sueño a vigilia y viceversa. Estás vías están ubicadas en el tronco encefálico (*te*) y el área peribraquial (*pb*) (Steriade et al. (1997)).

Aunque el sueño es una parte fundamental de nuestro día a día, todavía existen muchas preguntas sin respuesta. Una de esas preguntas es cómo es

regulado el ciclo de sueño-vigilia por las redes de neuronas. ¿Cómo es la dinámica espontánea de la sincronización entre neuronas corticales? ¿Cómo es esta sincronización interrumpida por los sistemas ascendentes? ¿Cómo son los patrones temporales de la sincronización durante la vigilia? ¿Cómo evolucionan hacia el estado de sueño? Éstas son algunas preguntas que guían el trabajo de investigación del grupo Neurocom.

En pocas palabras, el proyecto experimental busca analizar los efectos de la micro-estimulación eléctrica en *te* y *pb* en la sincronía entre neuronas de la corteza visual primaria de gatos anestesiados. Se presentan métodos para medir la sincronía y contrastes de hipótesis con respecto a los efectos de dicha estimulación. También se introduce otro factor que involucra una característica específica de las neuronas al análisis: su selectividad a la orientación.

La hipótesis más importante de ese proyecto es que es posible extraer información sobre la arquitectura funcional cortical del comportamiento espontáneo de las neuronas, un comportamiento obtenido de la actividad de disparo y reflejada en la fuerza de la sincronización entre pares de células. En su trabajo, los investigadores de Neurocom se interesan en medir la dinámica de la sincronización entre pares de neuronas en el sueño y, también, durante la vigilia, usando un modelo experimental concreto. Dicho modelo consiste en cambiar de los patrones del sueño a los de la vigilia a través de estimulando eléctricamente tanto en *te* como en *pb*. Pero, existen problemas metodológicos que hacen que sea complicado encontrar diferencias significativas: el escaso número de potenciales de acción, que es típico de la actividad espontánea.

El presente trabajo es el resultado de la búsqueda de técnicas estadísticas para definir (bajo las condiciones experimentales ya mencionadas):

- la dinámica de la sincronización de pares de neuronas bajo actividad espontánea.
- las diferencias en la fuerza de la sincronización entre los estados de sueño y vigilia.
- la eficacia de *te* y *pb* provocando la transición del estado de sueño al de vigilia, y la diferencia relativa en ese efecto.
- la dinámica de la sincronización de pares de neuronas con respecto a su selectividad a la orientación.

A lo largo del trabajo se definen técnicas, fundamentalmente no paramétricas para la estimación de la sincronización y se proponen contrastes de hipótesis para las hipótesis que se desprenden de los objetivos recién mencionados. Los métodos utilizados se centran, mayormente, en técnicas de suavizado tipo núcleo y bootstrap. De todas maneras, también se utilizan métodos semiparamétricos, específicamente modelos aditivos generalizados y técnicas de análisis de datos funcionales.

Los métodos se aplican a grupos de neuronas simultáneamente registradas en la corteza visual de gatos anestesiados que han sido sometidos a la micro-estimulación tanto en *te* como en *pb* de forma separada y aleatoria. Por otro lado, antes de las micro-estimulaciones, los gatos son sometidos a estímulos visuales consistentes en barras de luz orientadas y así se detecta la orientación preferida de cada neurona registrada.

Uno de los métodos más utilizados para medir asociación neuronal es el análisis de cros-correlaciones. Por ejemplo, el *joint peristimulus time histogram (JPSTH)* (Gerstein and Perkel (1969); Aertsen (1989)) muestra la dinámica de la correlación entre neuronas a partir de un estímulo dado. Este método es una generalización del *peristimulus time histogram (PSTH)*, que acumula los disparos, a través de los ensayos, de una sola célula. El JPSTH en un histograma bidimensional de la frecuencia de disparos conjunto de una neurona en el tiempo t y de la otra en el tiempo u. Su versión normalizada es el coeficiente de correlación de Pearson, calculado a través de los ensayos, de la frecuqncia de disparos de ambas neuronas con cierto retardo. Esta medida asume que los ensayos son indistinguibles y, por lo tanto, no tiene en cuenta la variabilidad entre ensayos. El cros-correlograma es la suma de las diagonales del JPSTH y, por lo tanto, es un histograma de la actividad conjunta de las neuronas como función de retardos en el tiempo.

Otros métodos ampliamente utilizados para capturar sincronización entre neuronas se basan en los *eventos unitarios* (Grün (1996); Grün et al. (2002); Grün (2009)). Estos métodos se basan en la discretización de los trenes, partiendo el tiempo en pequeños bins y asignando un uno a los bins en los que se haya producido una espiga y un cero a los que no. El análisis de eventos unitarios estima la probabilidad de que dos neuronas disparen juntas bajo la hipótesis de independencia de los trenes. Pueden definirse contrastes para la sincronía en términos de la diferencia entre las frecuencias esperadas y las observadas. Faes et al. (2008) propusieron otro índice de sincronía, la *medida de sincronía condicional*, que se basa en la estimación de la probabilidad de que ocurran disparos conjuntos dado que ha habido actividad en una de las

neuronas. Otros métodos han sido propuestos por Quiroga et al. (2002), Kruskal et al. (2007) y otros.

Generalmente, los experimentos se registran muchas veces bajo las mismas condiciones, contando as con muchos ensayos. De esta manera, y si la actividad de disparo es alta, se pueden utilizar métodos con trenes discretizados, y usando la información de cada bin a lo largo de todos los ensayos reduciendo la variabilidad dentro de los ensayos y trabajando con buena resolución temporal. En el caso de esta tesis, se han llevado a cabo pocos ensayos en cada condición debido, entre otras causas, a la larga duración de los ensayos y a la necesidad de limitar el número de estimulaciones. Por otro lado, los trenes de espigas contienen baja actividad, situación característica de la actividad espontánea en la corteza visual primaria, bajo la cual se registran las neuronas. De ahí la necesidad de desarrollar herramientas específicas para la estimación de la sincronía bajo estas circunstancias. A lo largo de la memoria se presentan tres métodos para estimar la sincronía. Nuestros métodos se basan en ventanas móviles. De esta forma, la sincronía puede estimarse en cada instante de tiempo, t, utilizando un solo ensayo, haciendo uso de la información de un pequeño entorno de t. Obviamente, estos métodos tienen peor resolución temporal pero, de todas maneras, se pueden obtener conclusiones acerca de la dinámica general de la sincronización.

El primer paso en nuestro estudio es el de describir la actividad de neuronas aisladas. Se estudia la estacionariedad de la actividad espontánea para un grupo de neuronas, encontrando que ésta es aceptable cuando las células no se encuentran bajo el efecto del estímulo. Esta estacionariedad se recupera cuando ha pasado el efecto del estímulo.

El análisis de neuronas aisladas se basa, principalmente, en el estudio de los tiempos entre disparos. El primer método presentado para medir asociaciones entre neuronas, se basa en una generalización de estos tiempos.

Sea $\tilde{S}$ la variable aleatoria que denota el tiempo de espera entre un disparo de la neurona 1 hasta el siguiente disparo de la neurona 2. Esta variable se denota por *intervalos entre espigas cruzados* (cros-inter-spike interval (CISI)). Se trabaja con los logaritmos de los CISIs para interpretar más fácilmente los resultados. Se denota por $g(s,t)$ la función de densidad del logaritmo de $\tilde{S}$ condicionado a que hubo un disparo de la neurona 1 en el instante t. Suponemos que g es estacionaria en el intervalo de tiempo antes de la estimulación, $g(s,t) = g_{pre}(s)$. La medida propuesta, mide la influencia del estímulo en la estructura de la densidad de CISI mediante la distancia L_1

entre $g_{pre}(s,t)$ y la densidad de CISI después de la estimulación:

$$CM(t) = \int |g(s,t) - g_{pre}(s)| \, \mathrm{ds}.$$

En la práctica, utilizamos la actividad de los trenes en ventanas de tiempo para estimar estas densidades, y por lo tanto CM, con métodos de tipo núcleo:

$$\widehat{CM}(t) = \int |\hat{g}(s,t) - \hat{g}_{pre}(s)| \, \mathrm{ds}, \tag{9.1}$$

donde, $\hat{g}_{pre}(t)$ es el estimador tipo núcleo de la función de densidad de CISI antes del estímulo y $\hat{g}(s,t)$ es el estimador tipo núcleo de la densidad de CISI en una ventana de tiempo que contiene al instante t.

La comparación de las densidades estimadas para los períodos de sueño y vigilia resulta muy interesante. Se observa cómo, tras la transición provocada por la estimulación, las densidades son distintas a aquellas observadas en el período de sueño. Sin embargo, cuando el tiempo pasa, las densidades se vuelven a asemejar a las originales. De todas maneras, el método que se sugiere para medir la diferencia en la sincronización en el estado de vigilia con respecto a la original, resulta muy ruidoso y los contrastes de hipótesis no son capaces de llegar a conclusiones satisfactorias con respecto a las hipótesis planteadas.

Tambin se presentan en la tesis otros dos métodos para medir sincronía. Un primer método es el llamado $ICCSI$. El valor de $ICCSI(t)$ se define a través de la integral en un vecindario alrededor de cero de la función de correlación cruzada entre los trenes, X e Y, en el instante t, $\mathcal{T}_{XY}(\tau;t)$:

$$\mathrm{ICCSI}(t) = \int_{-\Delta}^{\Delta} \mathcal{T}_{XY}(\tau;t)\mathrm{d}\tau. \tag{9.2}$$

Para estimar $ICCSI$, se utilizan ventanas alrededor de cada punto t y se estima la función de correlación cruzada con el correlograma normalizado.

El segundo método, al que denominamos CSM, se define a partir de los intervalos de tiempo entre un disparo de una neurona y el más cercano de la otra. A estos intervalos los llamamos *intervalos entre espigas más cercanas cruzadas* (cross-nearest spike interval (CNSI)). El CSM mide la proporción de estos tiempos que son menores que cierto valor, δ, elegido convenientemente por el investigador. Formalmente, sean X e Y dos trenes de espigas con un número de J_1 y J_2 espigas respectivamente y $N_X(t) = \#\{X_j \leq$

$t,\ j = 1, \ldots, J_1\}$ y $N_Y(t) = \#\{Y_j \leq t,\ j = 1, \ldots, J_2\}$ con $t \in [0, T)\}$ los procesos de contar asociados a X e Y respectivamente. Definimos,

$$n_\delta(t, v) = \sum_{j=1}^{J_1} \mathbb{I}\{N_Y(X_j + \delta) - N_Y(X_j - \delta) \geq 1\}\mathbb{I}\{X_j \in (t - v, t + v]\}+ \\ + \sum_{j=1}^{J_2} \mathbb{I}\{N_X(Y_j + \delta) + N_X(Y_j - \delta) \geq 1\}\mathbb{I}\{Y_j \in (t - v, t + v]\} \tag{9.3}$$

donde $2v$ es el ancho de la ventana móvil a utilizar. Si, además, $n(t, v)$ es el número total de espigas, de X y de Y en conjunto, que caen en la ventana $(t - v, t + v]$, CSM se define, en el instante t, como:

$$p_\delta(t, v) = \frac{n_\delta(t, v)}{n(t, v)}.$$

Estos dos métodos, $ICCSI(t)$ y $CSM(t)$, son flexibles dado que permiten el ajuste de sus parámetros según el contexto. Por otro lado, en ambos casos se hace uso de ventanas móviles para medir la sincronización local a lo largo del tiempo. De ésta forma, las bajas tasas de disparo y la baja cantidad de ensayos se compensa con el uso de la información de pequeños vecindarios.

Para cada uno de los métodos recién mencionados, se utilizan metodologías diferentes. Para $ICCSI$, las curvas estimadas se utilizan para llevar a cabo contrastes de hipótesis completamente noparamétricos. Además, se presentan dos contrastes bootstrap que tienen en cuenta la dependencia de los datos. Estos métodos están basados en el bootstrap estacionario de Politis y Romano (1994). Por otro lado, para el CSM se utiliza un enfoque semiparamétrico. La medida se ajusta a los datos utilizando modelos aditivos generalizados, resultando así, funciones de sincronización mucho más suaves que las obtenidas con $ICCSI$. Esta suavidad *extra* puede resultar una ventaja o una desventaja según el objeto de estudio. También se proponen contrastes bootstrap para el estudio de las diferencias en la sincronización. En este caso, los contrastes bootstrap utilizados son paramétricos.

Los resultados obtenidos con $ICCSI$ y CSM son comparables. Con ambos métodos encontramos períodos, después de la disrupción del estado de sueño, en que la sincronización difiere de la estimada antes de la estimulación. Con CSM estas diferencias se encuentran después de que varios segundos hayan transcurrido desde la aplicación de la micro-estimulación.

Con respecto a la diferencia entre el efecto de la estimulación *te* y *pb*, CSM tiene mayor éxito que $ICCSI$, sobre todo en el caso de *te* y, otra vez, las diferencias encontradas por un método se encuentran desfasadas con respecto al otro y esto puede ocurrir debido a la suavización extra que conlleva CSM. También, y en particular para CSM, se estimó la sincronía que se observaría si dos trenes fueran independientes, encontrando que ésta sería menos a la realmente observada, concluyendo así, que el método es capaz de detectar sincronía más allá de la que surgiría simplemente por estar las neuronas en actividad.

Las diferencias descritas en el párrafo anterior también se estudian tanto a un nivel de grupo de neuronas como a nivel poblacional. A nivel de un grupo de neuronas, se utiliza una metodología de análisis de datos funcionales, basada en proyecciones aleatorias, para realizar un análisis de la varianza de dos factores con variable de respuesta funcional (Cuesta-Albertos y Febrero-Bande (2010)). Las funciones utilizadas son las estimadas con el método $ICCSI$ para un grupo de 8 neuronas. Los dos factores en cuestión, son, por un lado, las vías ascendentes activadoras estimuladas: *te* y *pb*, y, por otro, la diferencia entre la orientación preferida de cada neurona del par. Esta variable se presenta como un efecto fijo y se calcula a partir de la orientación preferida de cada neurona, definida en el procedimiento experimental. Los resultados muestran que la relación funcional entre neuronas, dada por la afinidad en la selectividad a la orientación, repercute en el cambio que ocurre en la sincronización al aplicar el estímulo. Sin embargo, este hecho no depende del área estimulada, ya que la interacción entre factores no resulta significativa.

Para obtener una visión global de la relación entre la selectividad de la orientación y el efecto de la estimulación en la sincronización entre pares de neuronas, se realiza un análisis poblacional. Teniendo en cuenta que no la interacción entre factores no resulta significativa, este análisis se lleva a cabo para cada una de las áreas que provocan la transición del estado de sueño al de vigilia por separado. Se estima la sincronía entre pares de neuronas utilizando los registros de nueve experimentos. Las curvas obtenidas con ICCSI se promedian dentro de cada nivel del factor diferencia en orientación preferida y las curvas promediadas se modelan mediante el método de regresión polinómica local, utilizando ventanas de suavizado locales. Se utilizan varios contrastes para comparar las curvas (Vilar-Fernández et al. (2007)) y los resultados son concluyentes. Los p-valores obtenidos para la hipótesis de igualdad de funciones de regresión son muy pequeños para ambos estímulos, corroborando así los resultados encontrados con un solo grupo

de neuronas. Más aún, observando las derivadas de dichas curvas, también estimadas por el método de regresión polinómica local, podemos esbozar una hipótesis con respecto a la relación entre la diferencia en orientación preferida y la fuerza de la sincronización: mientras más sea la afinidad de las neuronas con respecto a su selectividad a la orientación, más rápidamente se desincronizarán estas neuronas y, a su vez, más rápido se resincronizarán. De todas maneras, serían necesarios análisis más profundos para comprobar estadísticamente esta hipótesis.

En conjunto, este trabajo es una contribución al desarrollo de herramientas estadísticas para la neurociencia. Aunque los métodos han sido propuestos para un problema en particular, son aplicables en otros muchos contextos donde las bajas tasas de disparos resulten un problema. Por otro lado, este trabajo enfatiza la utilidad de los métodos no paramétricos y el bootstrap. La estadística no paramétrica es una elección natural cuando ningún modelo paramétrico resulta adecuado, como suele suceder con los datos de trenes de potenciales de acción. Además, las técnicas bootstrap son muy potentes, fáciles de implementar y, aunque pueden ser costosas computacionalmente, son una gran alternativa a la inferencia paramétrica usando supuestos matemáticos mínimos.

Las principales conclusiones de esta tesis pueden resumirse en las siguientes líneas:

- Se propuso un contraste bootstrap para detectar dependencia entre los intervalos entre espigas. De la aplicación a los datos experimentales se desprende que el contraste detecta satisfactoriamente la dependecia existente entre estos intervalos.

- El método basado en la densidad de los intervalos entre espigas cruzados para detectar fluctuaciones en la sincronía no resulta lo suficientemente estable para analizar bajas tasas de disparo. Sin embargo, la inspección visual de la evolución de dichas densidades resulta informativa.

- Se definieron dos medidas de sincronía, CSM e $ICCSI$. Estas medidas estiman satisfactoriamente la sincronización neuronal en escenarios de baja actividad de disparo. Las variaciones en la sincronía estimada inducida por la estimulación se diferenciaron utilizando contrastes bootstrap.

- Se propusieron contrastes bootstrap que permitieron encontrar diferencias significativas en la sincronización inducida por la activación de

las vías ascendentes localizadas en el tronco encefálico versus las localizadas en el área peribraqueal. Estos contrastes se llevaron a cabo tanto a nivel de pares de neruonas como a nivel de grupos de células.

- Hemos propuesto contrastes bootstrap que permiten encontrar diferencias significativas en la dinámica de la sincronización entre los estados de sueño y vigilia relacionadas con la afinidad entre neuronas dada por su orientación preferida.

References

Aertsen, A.M. (1989). Dynamics of neuronal firing correlation: modulation of "effective connectivity". *Journal of Neurophysiology* 61, 900–917.

Cuesta-Albertos, J. and Febrero-Bande, M. (2010). A simple multiway ANOVA for functional data. *Test* 19, 537–557.

Faes, C., Geys, H., Molenberghs, G., Aerts, M., Cadarso-Suárez, C., Acuña, C. and Cano, M. (2008). A flexible method to measure synchrony in neuronal firing. *Journal of the American Statistical Association* 103, 149–161.

Gerstein, G.L. and Perkel, D.H. (1969). Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science* 164, 828–830.

Grün, S. (1996). *Unitary Joint-Events in Multiple-Neuron Spiking Activity: Detection, Significance, and Interpretation.* Reihe Physik, Verlag Harri Deutsch.

Grün, S., (2009). Data-driven significance estimation for precise spike correlation. *Journal of Neurophysiology* 101, 1126–1140.

Grün, S., Diesmann, M. and Aertsen, A. (2002). Unitary events in multiple single-neuron spiking activity: I. Detection and significance. *Neural Computation* 14, 43–80.

Kruskal, P.B., Stanis, J.J., McNaughton, B.L. and Thomas, P.J. (2007). A binless correlation measure reduces the variability of memory reactivation estimates. *Statistics in Medicine* 26, 3997–4008.

Politis, D.N. and Romano, J.P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association* 89, 1303–1313.

Quian Quiroga, R., Kraskov, A., Kreuz, T. and Grassberger, P. (2002). Performance of different synchronization measures in real data: A case study on electroencephalographic signals. *Physical Review E* 65, 041903.

Steriade, M., Jones, E.G. and McCormick, D.A. (1997). *Thalamus, Vol I: Organization and function*. Oxford: Elsevier.

Vilar-Fernández, J.M., Vilar-Fernández, J.A. and González-Manteiga, W. (2007). Bootstrap tests for nonparametric comparison of regression curves with dependent errors. *Test* 16, 123–144.