Neuronal networks, spike trains statistics and Gibbs distributions

Rodrigo Cofré

UNIVERSITY OF NICE - SOPHIA ANTIPOLIS

DOCTORAL SCHOOL STIC

SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

P H D T H E S I S

to obtain the title of

PhD of Science

of the University of Nice - Sophia Antipolis

Specialty : Control, Signal and Image Processing

Defended by

Rodrigo Cofré

Neuronal Networks, Spike Trains Statistics and Gibbs Distributions

Thesis Advisor: Bruno Cessac

prepared at INRIA Sophia Antipolis, Neuromathcomp Team

defended on November 5, 2014

Jury :

Reviewers :	Pierre Collet	-	École Polytechnique (Paris)
	Roberto Fernández	-	Utrecht University (Utrecht)
	Fred Wolf	-	Max Planck Institute (Göttingen)
President :	Olivier Faugeras	-	INRIA (Sophia Antipolis)
Examinators :	Jean-Marc Gambaudo	-	Institut Non Linéaire (Nice)
	Alain Destexhe	-	UNIC (Gif-sur-Yvette)
	Adrian Palacios	-	CINV (Valparaíso)
	Olivier Marre	-	Institut de la Vision (Paris)

UNIVERSITÉ DE NICE SOPHIA-ANTIPOLIS

ÉCOLE DOCTORALE STIC
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE

Pour obtenir le titre de

DOCTEUR DE SCIENCES

de l'Université de Nice Sophia-Antipolis
Discipline : Automatique Traitement du signal et des images

Présentée et soutenue par

Rodrigo Cofré

Statistique de potentiels d'action et distributions de Gibbs dans les réseaux de neurones

Thèse dirigée par: Bruno Cessac

preparée a l'INRIA Sophia Antipolis, Neuromathcomp Team

soutenue le 5 Novembre, 2014

Jury :

Rapporteurs :	Pierre Collet	-	École Polytechnique (Paris)
	Roberto Fernández	-	Utrecht University (Utrecht)
	Fred Wolf	-	Max Planck Institute (Göttingen)
President :	Olivier Faugeras	-	INRIA (Sophia Antipolis)
Examinateurs :	Jean-Marc Gambaudo	-	Institut Non Linéaire (Nice)
	Alain Destexhe	-	UNIC (Gif-sur-Yvette)
	Adrian Palacios	-	CINV (Valparaíso)
	Olivier Marre	-	Institut de la Vision (Paris)

Acknowledgments

I want to thank my advisor Bruno Cessac. It has been an honor to be his Ph.D. student. Bruno has been incredibly generous with his time and energy. He provided me with many interesting and accessible problems to study. This thesis is the result of my selection among them. Bruno made my Ph.D. experience productive and stimulating. His ability to capture the essence of the problems we face will have a lasting influence on my work. I hope that our collaboration will continue in the future.

I would like to thank the jury of this thesis for accepting their role. It is for me a great honor to be assessed by these eminent researchers. I am grateful to Antonio Galves and Errico Presutti for being such wonderful people and for inviting me to spend two weeks at the GSSI in L'Aquila, where I learned a lot from them. Thanks to all the people who discuss my thesis with me or invited me to give talks (Roberto Fernandez, Jean-Rene Chazottes, Jean-Marc Gambaudo, Samuel Petite, Wulfram Gerstner, Jean-Pierre Eckmann, Michèle Thieullen, Mathieu Lerasle, Etienne Tanré). I would like to acknowledge the financial support I have received from the French ministry of Research and University of Nice. I am grateful of the project Keops (ANR-CONICYT) that funded all my travel to conferences and meetings. I enjoyed a very nice and supportive atmosphere at INRIA, I shared inspiring discussions with other INRIA members; mostly from the Neuromathcomp, Athena and Tosca teams.

My time in Nice was made enjoyable and unforgettable in large part due to the many friends, groups and places that became part of my life. I am proud that my children Diego and Vicente were born in this country and in this region.

I would like to thank Rolando Rebolledo, his example motivated me to follow the path of science.

Special thanks to my parents, "la Nonna", the rest of my family and my parents in law. Lastly, I would like to thank my wife Consuelo, she has been part of every important thing in my life, including this thesis: for all your love and encouragement.

Thank you.

Abstract

Sensory neurons respond to external stimulus using sequences of action potentials ("spikes"). They convey collectively to the brain information about the stimulus using spatio-temporal patterns of spikes (spike trains), that constitute a "neural code". Since spikes patterns occur irregularly (yet highly structured) both within and over repeated trials, it is reasonable to characterize them using statistical methods and probabilistic descriptions. However, the statistical characterization of experimental data present several major constraints: apart from those inherent to empirical statistics like finite size sampling, 'the' underlying statistical model is unknown. In this thesis we adopt a complementary approach to experiments. We consider neuromimetic models allowing the study of collective spike trains statistics and how it depends on network architecture and history, as well as on the stimulus. First, we consider a conductance-based Integrate-and-Fire model with chemical and electric synapses. We show that the spike train statistics is characterized by non-stationary, infinite memory, distribution consistent with conditional probabilities (Left interval specifications), which is continuous and non null, thus a Gibbs distribution. Then, we present a novel method that allows us to unify spatio-temporal Maximum Entropy models (whose invariant measure are Gibbs distributions in the Bowen sense) and neuro-mimetic models, providing a solid ground towards biophysical explanation of spatio-temporal correlations observed in experimental data. Finally, using these tools, we discuss the stimulus response of retinal ganglion cells, and the possible generalization of the concept of receptive field.

Resumé

Les neurones sensoriels réagissent à des stimuli externes en émettant des séquences de potentiels d'action ("spikes"). Ces spikes transmettent collectivement de l'information sur le stimulus en formant des motifs spatio-temporels qui constituent le code neural. On observe expérimentalement que ces motifs se produisent de façon irrégulière, mais avec une structure qui peut être mise en évidence par l'utilisation de descriptions probabilistes et de méthodes statistiques. Cependant, la caractérisation statistique des données expérimentales présente plusieurs contraintes majeures: en dehors de celles qui sont inhérentes aux statistiques empiriques comme la taille de l'échantillonnage, 'le' modèle statistique sous-jacent est inconnu. Dans cette thèse, nous abordons le problème d'un point de vue complémentaire à l'approche expérimentale. Nous nous intéressons à des modèles neuro-mimétiques permettant d'étudier la statistique collective des potentiels d'action et la façon dont elle dépend de l'architecture et l'histoire du réseau ainsi que du stimulus. Nous considérons tout d'abord un modèle de type Intègre-et-Tire à conductance incluant synapses électriques et chimiques. Nous montrons que la statistique des potentiels d'action est caractérisée par une distribution non stationnaire et de mémoire infinie, compatible avec les probabilités conditionnelles (*left interval-specification*), qui est non-nulle et continue, donc une distribution de Gibbs. Nous présentons ensuite une méthode qui permet d'unifier les modèles dits d'entropie maximale spatio-temporelle (dont la mesure invariante est une distributions de Gibbs dans le sens de Bowen) et les modèles neuro-mimétiques, en fournissant une base solide vers l'explication biophysique des corrélations spatio-temporelles observée dans les données expérimentales. Enfin, en utilisant ces outils, nous discutons la réponse des cellules ganglionnaires de la rétine à des stimulus, et la possible généralisation du concept de champ récepteur.

Contents

List of Figures

List of Abbreviations

MaxEnt:	Maximum Entropy
GLM:	Generalized Linear Model
PSP:	Post Synaptic Potential
IF:	Integrate and Fire
LIF:	Leaky Integrate and Fire
MEA:	Multi-Electrode-Array
LGN:	Lateral Geniculate Nucleus
HMC:	Homogeneous Markov Chain
MCMC:	Markov Chain Monte Carlo
RF:	Receptive Field
LNL:	Linear Non Linear Model

List of symbols

$\omega_k(n)$:	Spike event
$\omega(n)$:	Spike pattern
$\omega_{n_1}^{n_2}$:	Spike Block
ω :	Spike Train
T :	Time length of the Spike Train
N :	Number of Neuron
R :	Range of the model
D :	Memory of the model ($R = D - 1$)
$m_l(\omega)$:	Monomial number l
L :	Total number of monomials
$\mathcal{H}$:	Potential
Z :	Partition Function
$\mathcal{S}$:	Entropy
$\mathcal{P}$:	Free Energy or Topological Pressure
$\pi_\omega^{(T)}$:	Empirical probability measured in the spike train ω of length T
μ :	Gibbs measure
σ :	Shift operator
$\mathcal{M}_{inv}$:	Space of invariant measures under the shift operator
C_{ik} :	Correlation between monomial i and k
χ	Susceptivility matrix

Chapter 1

Introduction

Overview

In this chapter, we present neurons, biological neural networks, the visual pathway and the retina. We introduce the in-vitro multi-electrode array method used to capture neural activity in living neural networks. We present the models of spike train statistics that we use and analyze along this thesis. We finish this chapter introducing the main questions addressed in this thesis.

Contents

1.1 Neurons

Neurons are electrically excitable cells that form the basic components of information processing in the central nervous system (Gerstner and W.Kistler, 2002). They are essentially made up of three different parts (see figure 1.1) : the dendrites, the soma, and the axon.

Dendrites extend from neuron cell body. They receive messages from other neurons, in the form of spikes that are pulse-like stereotyped waves of voltage of high intensity (see figure 1.1). The input electrical signals that a given neuron receives are added during the integration process (see fig 1.2). Neurons have a membrane that

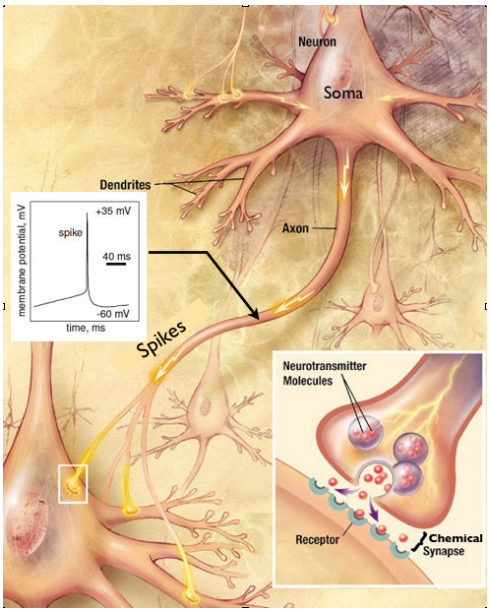


Figure 1.1: Figure showing the different parts of a neuron, the spikes and chemical synapses. (Left middle) spike; characterized as a rapid change in the membrane potential; (Right bottom) a chemical synapse (Image source: Modified from (National Institut on Aging-brochure) and (Izhikevich, 2007)

.

separates the extracellular and the intracellular fluid or cytoplasm. The membrane of the neurons act as a capacitor and therefore has a time constant which prevents rapid changes in the membrane potential (Difference in electric potential between the interior and exterior of the cell). The membrane potential of a neuron shows fluctuations in time. If the resulting voltage is raised above a threshold, an action potential, or "spike", is generated at the beginning of the axon. It travels along the axon until its end, at a synapse with another neuron, where it usually causes a new synaptic transmission. Just after a spike has been emitted, the cell can not emit any other spike during an interval of a few milliseconds; this time is called the refractory period. The very brief duration of spikes (a few milliseconds) allows us to consider them as instantaneous events. Therefore, from now we assume that neurons communicate with a discrete "code", made of a temporal series of spikes called "spike trains". Neurons communicate electrical and chemical signals with another cells using two types of synapses: Chemical synapses that relay the information via the release of neurotransmitters and electrical synapses that pass ions directly through gap junctions (see figure 1.3).

1.1.1 Chemical synapses

In chemical synapses the presynaptic neuron (neuron conducting an action potential toward the synapse) releases neurotransmitters captured by a postsynaptic neuron (the neuron receiving the signal) causing a brief change in the membrane potential called post-synaptic potential (PSP). Chemical synapses may be excitatory or inhibitory. While excitatory synapses cause a temporary increase of postsynaptic membrane potential, inhibitory synapses cause the opposite effect (see figure 1.2). The decay or growth after the spike is modeled using the called *alpha profiles*, which are functions governing the inter-neural transmission delays in the synaptic connections of the network. Neurons are constantly adding up the excitatory and the inhibitory synaptic input in time. If that summation is at or above a threshold the neuron fires a spike (see figure 1.2).

1.1.2 Electrical synapses

An electrical synapse is a conductive link between two adjacent neurons that is formed between them. The anatomical basis of electrical synapses are called gap junctions; they contain connexons that allow flow of ions directly between the cells, electrically coupling them (see figure 1.3). Electrical synapses are fast compared with chemical synapses, usually are bidirectional and can be found in many parts of the nervous system (Bennett and Zukin, 2004).

At the network level, electric synapses have several prominent effects such as neurons synchronization (Beierlein et al., 2000, Galarreta and Hestrin, 1999), and the generation of neural rhythms (Bennett and Zukin, 2004, Hormuzdi et al., 2004). Further on this thesis we will discuss in more detail what is the role of electric synapses in the spike train statistics while studying a neural network model.

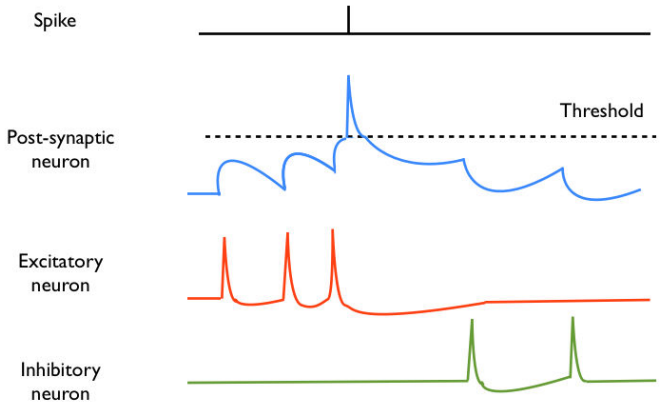


Figure 1.2: Input electrical signals coming from chemical synapses are added during the integration process causing fluctuations in the membrane potential of the post-synaptic neuron. If the resulting voltage is raised above a threshold, an action potential, or spike, is generated.

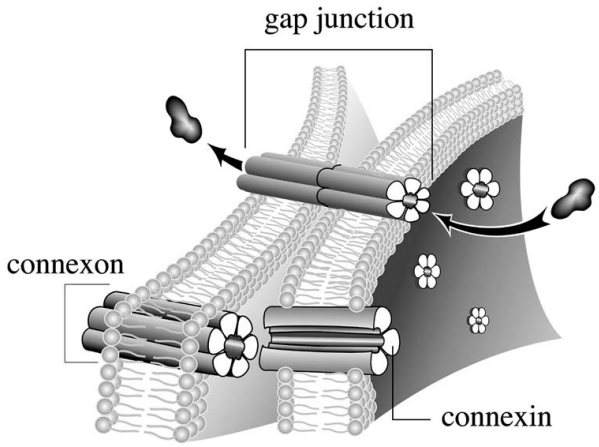


Figure 1.3: Gap junctions: When the connexons in the membranes of two cells in contact are aligned, they form a continuous channel that connects the two cell interiors, allowing the flow of ions from one cell to the other (image source: Wihelma Echevarria)

1.2 Neural Networks

Animal nervous systems are composed of thousands or millions of interconnected cells that form neural networks. The activity of a neuron directly influences other neurons to which it is connected through a synapse (chemical or electrical). Massive and hierarchical multi-layered networking of neurons seems to be the fundamental precondition for the information processing capabilities of nervous system (Rojas, 1996). In this thesis, we will focus on one example of biological neural network: the retina. This is mainly due the following reasons:

1. Sight is a very important sense in mammals. Nevertheless, it is still a mystery how the visual perception in generated. Since the spikes on the retina are the only source that the brain have to create the visual perception, there is considerable interest in the power of retinal processing and how it shapes our visual perception (Meister and Berry II., 1999).

2. It is possible to remove the retina from an animal. The living retina can be attach it into an array of multi electrodes. Light can be projected into the retina and the ganglion cells spiking activity can be recorded in vitro (Kriegeskorte and Kreiman, 2010) (see figure 1.7). This technology provides reliable data to study how the retina respond to different types of light stimulation.

3. One of the main characteristics of the retina is to operate exclusively as a feed-forward network to higher areas of the brain. There is no feedback mechanism by which the brain changes how the retina transduces light into spikes (the eyes saccades may be considered as a form of feed-back, though).

4. The retina is the most extensively studied piece of the central nervous system (Hegger, 2006). It is possible to compare results with previous achievements done by the scientific community.

5. The retina provides a model for learning how the vertebrate nervous system works, thus scientist are using the retina to understand the brain. A large part of the cortical tissue is dedicated almost exclusively to the initial processing of visual information (Roederer, 2005).

To put in context the role of the retina in the visual perception, let us first introduce the visual pathway.

1.2.1 Visual pathway

The process of "seeing" is very complex and not yet well understood. Vision just begins in the eye. Light entering the eye is projected onto the back of the eyeball, where a complex signaling by cells within the layers of the retina is converted into spikes. These spikes are sent through the optic nerves to the lateral geniculate nucleus (LGN), there the message is processed. Then, neurons of the LGN send

their axons to higher centers in the visual cortex of the brain for further processing necessary for visual perception.

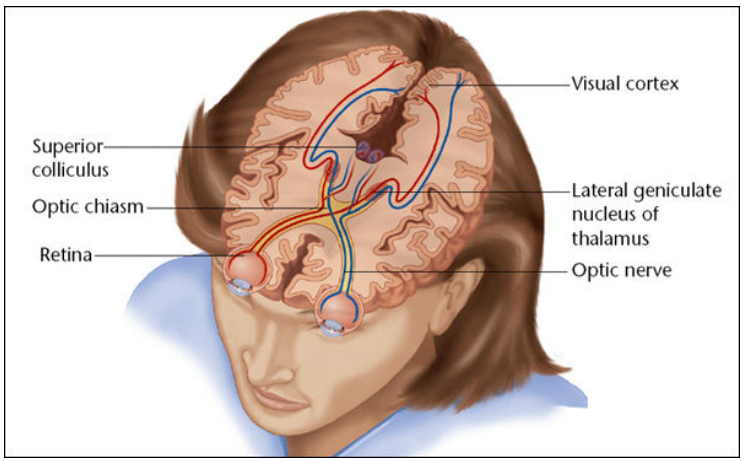


Figure 1.4: Visual Pathway: Illustration of the main steps in visual perception (image source: www.photigy.com)

1.2.2 Retina

The retina is a complete piece of brain located at the back of the eye (see fig 1.4). It transduces spatial and temporal variations in light intensity and transmits them to the brain. The retina present a huge anatomical complexity. It is composed of at least 50 clearly distinct cell types (Gollisch and Meister, 2010). They differ in size, temporal and spatial properties. The neurons are arranged in cellular layers that are interconnected (see fig 1.5).

The retina is composed of five layers of cells: Photoreceptors, Horizontal cells, Bipolar cells, Amacrine Cells and Ganglion cells. The distinct cell types in the retina and its specific connections suggests that each has a specific function. Photoreceptors (cells which convert light into nerve impulses) are connected to bipolar cells which arborize to receive multiple synaptic contacts from them. Horizontal cells are laterally interconnecting photoreceptors and bipolar cells. Amacrine cells extend their dendrites laterally to contact bipolar, other amacrine cells, and ganglion cells. This last type of cells collect information from bipolar cells and amacrine cells. Amacrine and bipolar cells, like many types of neuron in the brain, are widely coupled by gap junctions to their neighbors (see figure 1.5). Since many amacrine cells fire action potentials, it is likely that gap junctions allow them to synchronize their firing. The visual information arrives to the ganglion cells is in the form of chemical messages sensed by its receptors. This message is integrated within the ganglion-cell and "digitized" into spikes. Axons of retinal ganglion-cells form the optic nerve (see figure 1.4). Information about the visual stimulus is encoded by ganglion cells patterns of spikes. Spikes conserve their shape when traveling along

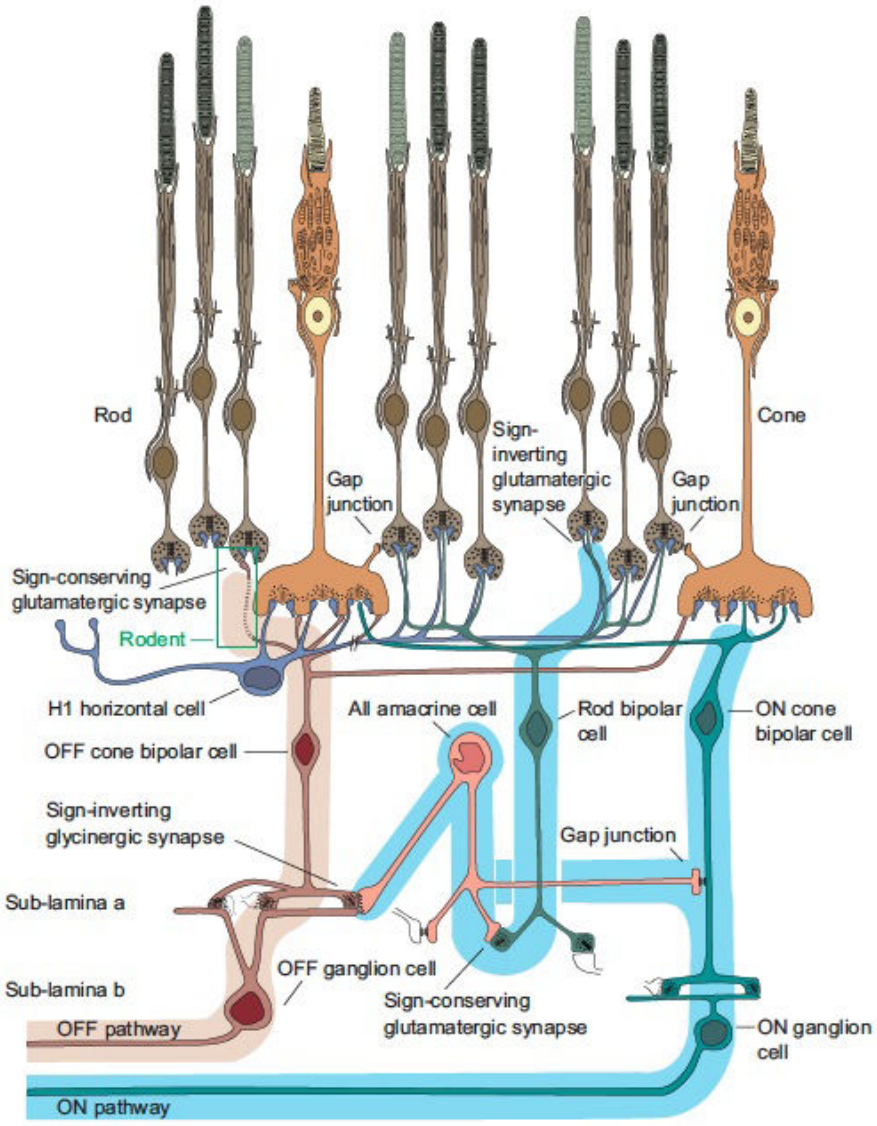


Figure 1.5: Several distinct neural circuits within the retina transmit the signals from the photoreceptors (rods and cones) to the retinal ganglion cells. The different types of cells communicate information between them through chemical and electrical synapses (image source: (Sharpee and Stockman, 1999))

axons and signals travel rapidly over large distances. As a consequence, retinal ganglion cells are capable to convey spike information reliably and rapidly to the brain.

1.2.2.1 Receptive Fields of retinal ganglion cells

All visual information from the external world reaching the brain is transmitted by retinal ganglion cells, each of them is sensitive to light reaching a small region of the retina called *receptive field* (Gauthier et al., 2009). When this region is stimulated the firing of the retinal ganglion cell is influenced. Therefore, retinal ganglion cells fire action potentials in response to certain types of retinal stimulation. The receptive fields of retinal ganglion cells are composed of inputs from many rods and cones (see figure 1.6). The most common receptive fields are arranged into a central disk, the "center", and a concentric disk, the "surround", each region responding oppositely to light. There are on-center, off-surround ganglion cells which fire most strongly when there is light in the center of the field, but no light in the surround. There are also off-center, on-surround ganglion cells that fire most strongly when the opposite occur. There exist also more complex receptive fields that are not center surround.

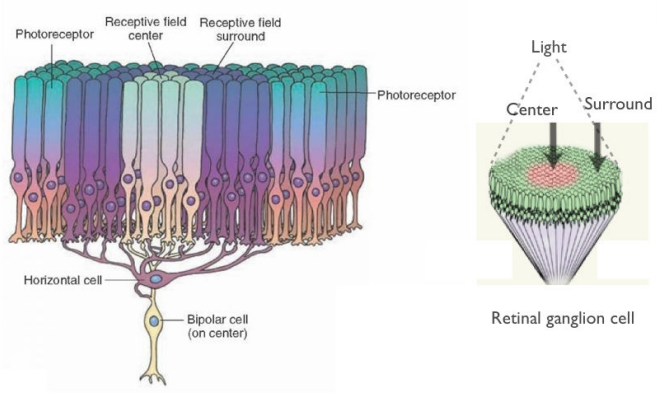


Figure 1.6: Center surround antagonism in receptive fields of retinal ganglion cells (image modified from: www.studyblue.com and www.webexhibits.org)

1.3 Encoding stimuli by spikes in the retina: "Neural code"

Our perception of the world is driven by inputs from the sensory neurons (Rieke et al., 1996). Deciphering how populations of spiking neurons represents sensory information is one of the main problems in computational neuroscience and central for our understanding of perception. Characterizing the relationship between sensory stimuli and the spike responses of neurons is referred in the literature as "neural

coding problem" (Rieke et al., 1996). Clearly, this code is not deterministic. It has been observed in retinal ganglion cells that even when presenting the same stimulus under very controlled experimental conditions the neural activity changes from trial to trial (Shadlen and Newsome, 1998). The reason of this variability can either be non-observed inputs, intrinsic noise of the individual neuron associated with various biophysical processes that affect neuronal firing. It can also be collective dynamics. This last remark is one of the reasons motivating this thesis. Thus, there is no one-to-one mapping between visual stimulus and neural response. However, we can assign a probability to each possible response given the stimulus. In this setting, the problem can be framed as characterizing the probabilistic relationship between stimuli and spike train response.

1.3.1 Multi-Electrode Arrays

Multi-Electrode arrays are devices that contain multiple electrodes through which neural extracellular recordings are obtained attaching a living neural tissue into it (see figure 1.7). For the particular case of the retinal ganglion cells, electrodes record voltages generated in the extracellular array by the current fields outside the cells in the local region when they generate action potentials (see figure 1.7). Therefore, each electrode records the activity of several cells. Nowadays it is possible to record a large number of simultaneously active neurons through multi-electrode arrays (up to 4096 electrodes). Then, spike-sorting algorithms allows to identify which neuron emits a spike and recover discrete time spikes from continuous extracellular recordings (Marre et al., 2009). Spikes trains are usually "binned". This consist in discretizing the time by choosing time windows Δt_b (ranging from 1 to 20 milliseconds) considering that into these windows there is a spike or not; depending on the binning there are cases where there is more than one spike in a bin, it these cases the convention is to consider them as a single spike. After this procedure a spike train or raster plot is obtained as a result of a visual stimulus presentation (see figure 1.7).

1.4 Models of Spike train statistics

We have argued that the link between stimuli and responses cannot be expressed by a one-to-one mapping, that is, we cannot predict the neural response exactly. Statistical techniques and probabilistic descriptions are needed to characterize the spike response. Models of spike train statistics are the central problem of this thesis, because they provide a possible way to characterize the "neural code". In this section we present briefly the mathematical setting we will use and analyze further in this thesis. We will come back to these models to analyze them mathematically, specially characterizing the spike train statistics they produce.

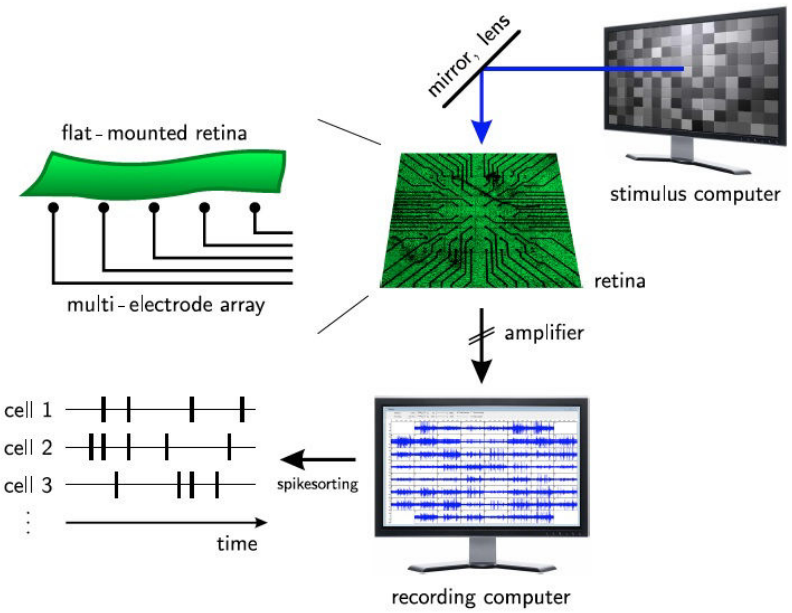


Figure 1.7: A retina is mounted in the multi-electrode array, light stimulus is projected onto it via a computer in which images are produced. The membrane potential of neurons in the multi-electrode array are collected. Then, a complex task of signal processing is performed, called spike sorting, which detects the time at which a spike is produced, finally a binning size is chosen previous to obtain a spike train (Source: C. Mendl Phd thesis 2011)

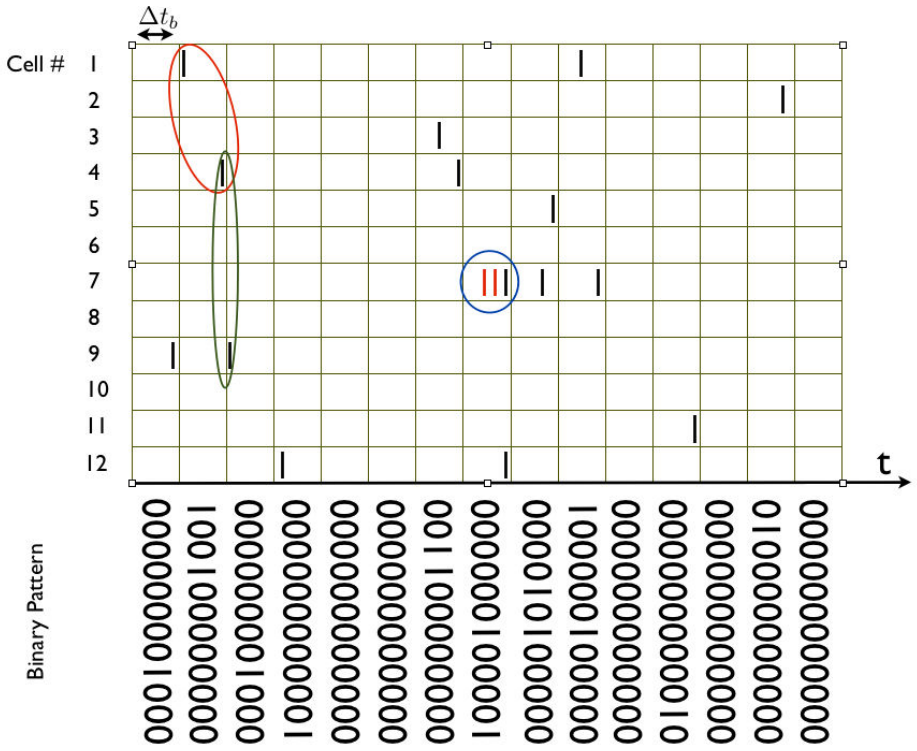


Figure 1.8: After having choose a binning size Δt_b, a given neuron either fires a spike or it doesn't, so it's state is described by a binary variable $\{0, 1\}$. The blue circle represent the case of more than one spike is fired in the same time bin mapped into a single spike in the binary pattern. The red ellipse shows two spikes much more separated in time than the green ellipse. Binned binary data consider spikes in the red ellipse as simultaneous and those in the green ellipse occur in successive times.

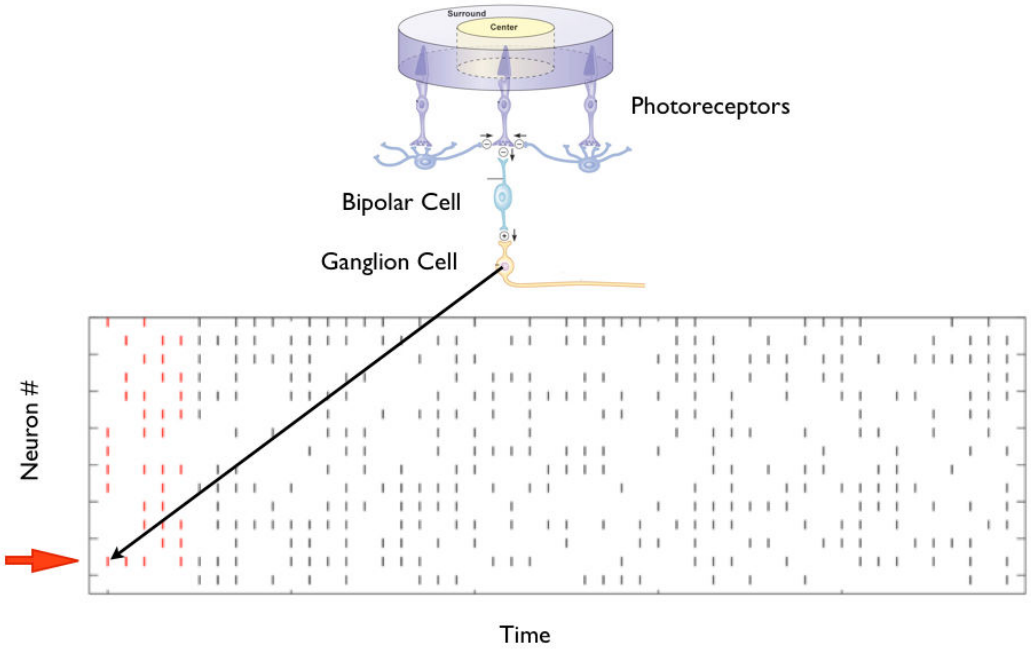


Figure 1.9: Spike train from retinal ganglion cells. Each row represent the activity of one ganglion cell. While some of the spikes can be generated by direct excitation of the receptive field of the respective ganglion cell, others can be generated by intrinsic noise, or by indirect synaptic inputs not necessarily from their own receptive field pathway (Image modified from: www.droualb.faculty.mjc.edu)

1.4.1 Maximum Entropy method in the context of spike train statistics

The maximum entropy principle (MaxEnt) is a statistical method that have been successfully applied to characterize the spike train statistics of retinal ganglion cells data obtained from MEA (Schneidman et al., 2006, Marre et al., 2009, Vasquez et al., 2012). It consists in choosing a set of characteristic events in the data and measure their empirical values. These empirical values are the constraints of the problem. Maximum entropy method consist of finding, in the space of probability distributions that matches these values, the one with maximum entropy (Jaynes, 1957). Finding a distribution with maximal entropy which is still consistent with the data, amounts to finding the simplest or minimally structured model. This method does not make any assumptions about the underlying neural tissue that have generated the data. Maximizing the entropy given those constraints provides a unique probability, called a Gibbs distribution[1]. The MaxEnt relies on the assumption of stationarity (time translation invariance of statistics) as well as an a priori and arbitrary choice of constraints. This severely constrains the statistical model. In particular, memoryless models focuses on synchronous events, hiding temporal causality.

1.4.2 Generalized Linear models

Generalized Linear models (GLM) focus on the point process[2] nature of spike trains. They are commonly used for modeling the relationship between neural population activity and presented stimuli. In this modeling approach, the instantaneous spike rate of the point process is described by a stimulus filter, a post-spike filter and a constant offset which sets the baseline firing rate of the neuron (see figure 2.7). Although these types of models are strictly phenomenological, their components can be broadly compared to biophysical mechanisms. The stimulus filter approximates the receptive field of the ganglion cell. The post-spike filter mimics voltage- currents following a spike, whereas the nonlinearity implements a soft threshold converting membrane potential to instantaneous spike probability (Calabrese et al., 2011). For each neuron, a static nonlinear function is then applied to the summed filter responses. This method have also been used to predict retinal ganglion cell responses to light stimulus.

1.4.3 Conductance Based Integrate and Fire

In 1907 Lapicque developed a neuron model that is still broadly used nowadays (Lapicque, 1907). In this approach, a neuron is modeled using an equivalent electric circuit composed of a parallel capacitor and a resistor, that represents the capacitance and leak resistance of the cell membrane (see figure 1.10). When the membrane capacitor is charged to a certain threshold potential, an action potential is fired and the capacitor discharges, resetting the membrane potential (Abbott, 1999). The

[1] Later on we shall make a distinction between "equilibrium states" and "Gibbs distributions"
[2] Random process giving random configurations of points in time and space.

model is known as Leaky Integrate-and-Fire neuron model (LIF). It is the simplest dynamical model that captures the basic properties of neurons, including the temporal integration of noisy sub-threshold inputs, all-or-nothing spiking. It postulates a differential equation describing the behavior of the network motivated by the microscopic picture of how the system is assumed to work. Noise is important in the cases of interest to us, as a stochastic description is required. Although action potentials can vary in duration, amplitude and shape, integrate-and-fire models treat them as identical stereotyped events. This simple model is capable of being analyzed mathematically while at the same time being sufficiently complex to capture many of the essential features of neural processing.

LIF models implement a reset mechanism on the membrane potential: If neuron k has been reset at time τ, then the voltage at time $t > \tau$ depends only on the voltage from the last reset time and not on previous values. In the conductance based Integrate and Fire model proposed by (Rudolph and Destexhe, 2006), contrarily to LIF, there is also a dependence in the past via the conductance and this dependence is not erased by the reset. These models may have infinite memory.

One important application of these models in this thesis will be the study of how neuronal collective dynamics affects the encoding properties of multiple neurons and how the consideration of the gap junctions term incorporate information about neuronal interdependencies.

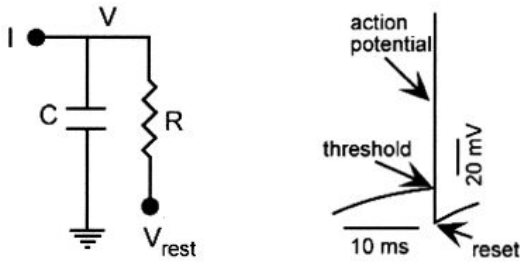


Figure 1.10: Left: Equivalent circuit of the Integrate-and-Fire model. (Right). Voltage trace of the Integrate and fire model. When reaching the threshold instantaneously a spike is emitted and the voltage is reset (Image source (Abbott, 1999))

1.5 What is this thesis about?

The problem of the statistical characterization of experimental spiking data presents several major constraints. Apart from those inherent to empirical statistics such as finite size sampling, there is the fact that the true underlying statistical model is

unknown. As a consequence researchers are seeking canonical forms of probabilities from general principles (either based on statistical inference or biophysics) to fit them to spiking data available thanks to MEA.

This thesis addresses this problem from a different point of view. We build and analyze neuro-mimetic models providing a mathematical description of neural dynamics. In particular, we focus on models suitable to reproduce spike trains produced by retinal ganglion cells.

The main advantage of this approach is that we may control the parameters characterizing the dynamics and spike train statistics of the neural network model. Another important advantage of this approach is that we do not assume a priori the form of the probability distributions, instead, they arise naturally from the dynamics of the neural network.

From these neuro-mimetic models it is possible to obtain a probabilistic mapping between the network architecture, the stimuli, the spiking history of the network and the spiking response in terms of conditional probabilities of spike patterns given the network history, allowing a mechanistic and causal understanding of the origin of correlations.

This approach allows us to address interesting questions about the way in which structured networks of interconnected neurons respond collectively to a stimulus. In particular, this approach not only allows us to predict spike responses to applied stimulus, but also, to explain "why" neurons respond the way they do. More precisely, this thesis attempts to answer the following questions:

Question 1: *Is it possible to characterize the network dynamics and the population spike train statistics in a neural network model?*

This question has been previously addressed in (Cessac, 2011b), where only the effects of chemical synapses over the dynamics and statistics were considered in a conductance based neural network model. While communication between neurons involves chemical and electric synapses, the role of electric synapses in shaping collective dynamics has been quite less studied than the role of chemical synapses. They are supposed to play a key role in producing correlations in the spike train statistics. We address question 1 by building and analyzing a biologically realistic model based on (Rudolph and Destexhe, 2006) and including electric synapses. The main motivation of this work is to better understand how ganglion cells in the retina supporting both chemical and electric synapses coordinate spatio-temporal spike patterns to convey information to the brain. We address this question in chapter 3.

Question 2: *Are Gibbs distributions good candidates to analyze the spike train statistics for experimental data?*

Although is impossible to categorically answer this question, we may address it through presenting three examples of methods used by the computational

neuroscience community (Maximum Entropy Principle, Generalized Linear Models and Conductance based Integrate-and-Fire with and without electric synapses), which correspond to Gibbs distributions. We investigate under which conditions these methods can correspond to the same spike train characterization, and how to compare them. We address this question in chapter 2.

Question 3: *When attempting to characterize spike train statistics, Gibbs distributions arise from data-based approaches (Maximum Entropy Models) and from neuro-mimetic model-based approaches (Generalized Linear Models and Conductance based Integrate-and-Fire). Is it possible to take advantage of properties of Gibbs distributions to link both approaches?*

We investigate whether it is possible to associate a unique canonical Maximum entropy potential (a potential with a minimal number of parameters written as a linear combination of spiking events) to the potential associated to the Gibbs distribution corresponding to a given neuro-mimetic model, like Generalized Linear Model or Integrate-and-Fire model. Also, we investigate if the canonical Maximum Entropy potential can be explicitly constructed using the neural network parameters shaping the neuro-mimetic model. Thus we look for a deterministic and exact link between neuro-mimetic models and Maximum Entropy Models, revealing the underlying mechanistic origin of correlations observed in experiments of spike trains in the retina using the Maximum Entropy approach. This question is addressed in chapter 4.

Question 4: *Can we derive a more general notion of receptive fields looking at the difference between the spontaneous and stimulus evoked spiking response of a network of neurons using properties of Gibbs distributions?*

While neurons respond collectively to a stimulus, the most classical notion of receptive field is written in terms of firing rates of single neurons. We look for a more general definition of the receptive fields in terms of spatio-temporal patterns. We address this question in chapter 5 using neural network models and linear response theory.

1.6 De quoi traite cette thèse

Le problème de la caractérisation statistique des données expérimentales de potentiel d'action présente plusieurs contraintes majeures. En dehors de ceux inhérents aux statistiques empiriques comme la taille fini de l'échantillon, il y a le fait que le vrai modèle statistique sous-jacent est inconnu. En conséquence, les chercheurs cherchent des formes canoniques, des probabilités en suivant des principes généraux (soit sur l'inférence statistique ou sur la biophysique) pour ajuster les données qui s'obtiennent gråce à MEA.

Cette thèse aborde le problème d'un point de vue différent. Nous construisons et analysons des modèles neuro-mimétiques qui fournissent une description mathématique de la dynamique des neurones. En particulier, nous nous concentrons sur des modèles adaptés à reproduire des potentiels d'action produits par des cellules ganglionnaires de la rétine .

Le principal avantage de cette approche est ce qui nous permet de contrôler les paramètres du modèle de réseau neuronal, que caractérisent la statistique de potentiel d'action et la dynamique. Autre avantage important de cette approche est que nous ne supposons pas à priori la forme des distributions de probabilité, à la place, ils 'se produisent naturellement à partir de la dynamique du réseau de neurones.

A partir de modèles neuro-mimétiques, il est possible d'obtenir une correspondence probabiliste entre l'architecture du réseau, les stimuli, l' histoire du réseau et la réponse, en termes de probabilités conditionnelles, étant donné l' histoire du réseau. Cela permet une compréhension mécaniste et causale de l'origine des corrélations.

Cette approche nous permet de répondre à des questions intéressantes sur la façon dont des réseaux structurés de neurones interconnectés collectivement répondent aux stimulus. Cette approche nous permet non seulement de prédire les réponses à un stimulus appliqué , mais aussi, d'expliquer "pourquoi" les neurones répondent de la façon dont ils le font. Plus précisément, cette thèse se propose de répondre aux questions suivantes:

Question 1: *Est-il possible de caractériser la dynamique du réseau et les statistiques de potentiel d'action dans un modèle de réseau de neurones?*

Cette question a déjà été abordée dans (Cessac, 2011b), où seulement les effets des synapses chimiques sur la dynamique et les statistiques ont été considérés dans un modèle de réseau de neurones sur la base de la conductance. Tandis que la communication entre les neurones implique des synapses chimiques et électriques. Le rôle des synapses électriques dans la dynamique collective a été moins étudiée, que le rôle des synapses chimiques. Ils sont censés jouer un rôle clé dans la production des corrélations dans la statistique de potentiel d'action. Nous abordons la question à partir de la constriction d'un modèle biologiquement réaliste basé sur ((Rudolph and Destexhe, 2006)) comprenant les synapses électriques. La principale motivation de ce travail est de mieux comprendre comment les cellules ganglionnaires de la rétine ayant des synapses chimiques et électriques peuvent coordonner motifs spatio-temporel et modèles pour transmettre des informations vers le cerveau. Cette question nous l'abordons dans le chapitre 3.

Question 2: *Sont-ils les distributions de Gibbs de bons candidats pour analyser les statistiques de potentiel d'action pour données expérimentales?*

Bien qu'il soit impossible de répondre à cette question catégoriquement, nous abordons cette question en présentant trois exemples de méthodes utilisées

par la communauté des neurosciences (Principe d'entropie maximale, modèles linéaires généralisés et Intégre-et-tire avec et sans synapses électriques). Tous correspond à des distributions de Gibbs. Nous cherchons les conditions dans lesquelles, ces méthodes peuvent correspondre à la même statistique de potentiel d'action, et comment les comparer. Nous abordons cette question dans le chapitre 2.

Question 3: *En essayant de caractériser les statistiques de potentiel d'action, les distributions de Gibbs apparaissent naturellement tant dans des approches basées sur des données (modèles d'entropie maximale) que dans des approches basées sur des modèles neuro-mimétiques (modèles linéaires généralisés et Intégre-et-tire). Est-il possible de profiter des propriétés des de Gibbs approches pour relier les différentes approches?*

Nous étudions la possibilité d'associer un unique potentiel d'entropie maximal canonique (avec un nombre minimal de paramètres écrits comme combinaison linéaire des événements), associée au potentiel de Gibbs correspondant à un modèle neuro-mimétique (comme le modèle GLM ou Intégre- et-tire). Aussi, nous étudions si le potentiel canonique d'entropie maximale peut être construit explicitement en utilisant les paramètres du réseau de neurones qui déterminent le modèle neuro-mimétique. Ainsi, nous regardons le lien déterministe entre les modeles neuro-mimétique et modèles d'entropie maximale, ce qui permet de révéler l'origine mécaniste sous-jacente de corrélations observées dans les expériences de potentiel d'action dans la rétine en utilisant l'approche d'entropie maximale. Cette question est abordée dans le chapitre 4.

Question 4: *Peut-on tirer une notion plus générale de champs réceptifs à partir de la différence entre le regime spontané et celui évoqué par le stimulus dans une modèle de réseau de neurones en utilisant les propriétés des distributions de Gibbs?*

Tandis que les neurones respondent collectivement à un stimulus, la notion plus classique de réceptifs champ est écrit en termes de taux de décharge de neurones isolés. Nous recherchons une définition plus générale des champs réceptifs en termes de modèles spatio-temporels.

Cette question est abordée dans le chapitre 5 en utilisant des modèles de réseaux de neurones et la théorie de la réponse linéaire .

Chapter 2

Gibbs distributions as canonical models for spike train statistics

Overview

In this chapter, we show how Gibbs distributions arise from the models introduced in the first chapter. We start introducing the notations and definitions used throughout this thesis. We introduce the thread of this thesis: the transition probabilities of spike blocks. We then present Markov chains as a particular example where stationary and fixed past transition probabilities bring useful information about spike train statistics. Finally we move to the mathematical analysis of the methods presented in chapter 1. We conclude that Gibbs distributions can be considered as canonical models for characterizing spike train statistics.

Contents

2.1 Setting

In order to set a common ground for the analysis of spike train statistics, we introduce here the notations used throughout this thesis. We consider a network of N neurons. We assume that there is a minimal time scale, δ, set to 1 without loss of generality such that a neuron can at most emit one spike within a time window of size δ. This provides a time discretization labeled with an integer time n. To each neuron k and discrete time n one associates a *spike state* $\omega_k(n) = 1$ if neuron k as emitted a spike in the time window $[n\delta, (n+1)\delta[$ and $\omega_k(n) = 0$ otherwise. The spike-state of the entire network in time bin n is thus described by a vector $\omega(n) \stackrel{\text{def}}{=} [\,\omega_k(n)\,]_{k=1}^{N}$, called a *spiking pattern*. A *spike block* is a finite ordered list of such vectors, written:

$$\omega_{n_1}^{n_2} = \{\,\omega(n)\,\}_{\{n_1 \leq n \leq n_2\}}\,,$$

where spike times have been prescribed between time n_1 to n_2. The *length* of a block is $n_2 - n_1 + 1$, the number of time steps from n_1 to n_2. Thus, there are 2^{Nn} possible blocks with N neurons and range n.

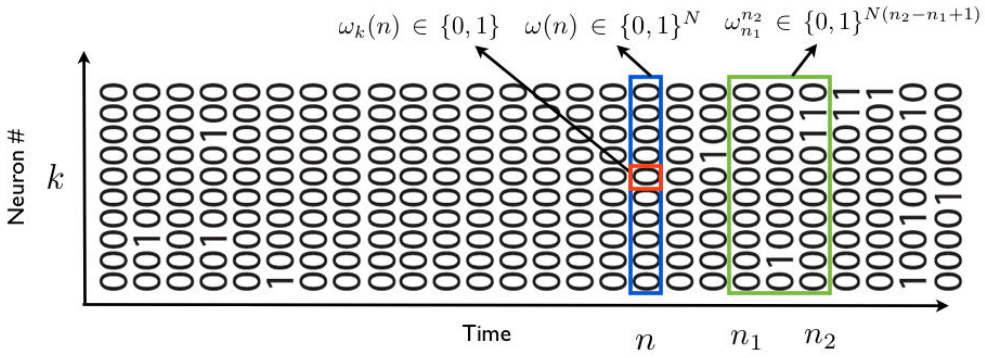


Figure 2.1: Notation used throughout this thesis. In red there is a spike state, in blue a spiking pattern and in green a spike block of range 3.

A *spike train* is a spike block $\omega_{-\infty}^{+\infty}$. Obviously, experimental spike trains start from some initial time $t_0 > -\infty$ and end at some final time $T < +\infty$, but, on mathematical grounds the consideration of bi-infinite sequences simplifies the analysis. To alleviate notations we simply write ω for a spike train. The set of spiking pattern is denoted $\mathcal{A} = \{\,0,1\,\}^N$. The set of spike blocks ω_m^n is denoted $\mathcal{A}_m^n$. The set of spike trains is denoted $\Omega = \mathcal{A}^{\mathbb{Z}}$. To each spike train ω and each neuron index $j = \{1 \ldots N\}$ we associate an ordered (generically infinite) list of "firing times" $\{t_j^{(r)}(\omega)\}_{r=1}^{+\infty}$ such that $t_j^{(r)}(\omega)$ is the r-th time of firing of neuron j in the spike train

ω. $\tau_k(t,\omega)$ denote the last firing time of neuron k, before time t, in the spike train ω.

2.1.1 Transition probabilities

The transitions between spike blocks are probabilistic. They depend on the neural network characteristics such as neurons conductances, synaptic responses or external currents. They give information about the dynamics that takes place in the observed neural networks. Especially, they have a *causal* structure i.e. the probability of an event depends on the past. The transition probabilities reflects underlying biophysical mechanisms in the neural network, which are also causal. It is impossible to know explicitly the history dependence of a neuron since it depends on the past evolution of all variables determining the neural network. A possible simplification is to consider that this probability depends *only* on the spikes emitted in the past by the network. In this way, we are seeking a family of transition probabilities of the form $\mathbb{P}\left[\,\omega(n)\,\middle|\,\omega_{n-D}^{n-1}\,\right]$, the probability that the firing pattern $\omega(n)$ occurs at time n, given a past spike block ω_{n-D}^{n-1}. Here, D is the *memory depth* of the probability, i.e., how far in the past does the transition probability depends on ω. We say that those transition probabilities are *time-translation invariant* or *stationary* if for all n, $\mathbb{P}\left[\,\omega(n)\,\middle|\,\omega_{n-D}^{n-1}\,\right] = \mathbb{P}\left[\,\omega(0)\,\middle|\,\omega_{-D}^{-1}\,\right]$ whenever $\omega_{n-D}^{n} = \omega_{-D}^{0}$.

2.1.2 Different types of transition probabilities used to characterize spike trains statistics

These transition probabilities depending on D can take four different forms:

1. Memory-less: $D = 0$ (Independent of the past). This is the case when successive spike patterns are independent. We use the convention that $\mathbb{P}\left[\,\omega(n)\,\middle|\,\omega_{n-D}^{n-1}\,\right] = \mathbb{P}\left[\,\omega(n)\,\right]$, when $D = 0$.

2. Fixed length memory: Fixing $D > 0$ and assuming that transition probabilities $\mathbb{P}\left[\,\omega(n)\,\middle|\,\omega_{n-D}^{n-1}\,\right]$ do not to depend explicitly on time (stationarity assumption), they define a homogeneous Markov chain.

3. Variable length memory: Here D is not fixed. In the context of neural networks models it appears naturally as memory of the neurons goes back up to its last spike, which is variable for each neuron and for each time. To define them properly we need to define a probabilistic context tree and a family of infinite order transition probabilities in this context tree. In this thesis we do not use this approach (It has been already introduced in neuroscience e.g. in (Galves and Löcherbach, 2013)).

4. Infinite memory: They arise in the context of the neural network models that we analyze in this thesis. The models we examine consider memory up to the last firing time, which can be unbounded, thus we have to consider cases

where the last firing time can go arbitrary far in past, this correspond to transition probabilities of the form: $\mathbb{P}\left[\,\omega(n)\,\middle|\,\omega_{-\infty}^{n-1}\,\right]$. These transition probabilities define non Markovian processes called chains with "complete connections" (Fernandez and Maillard, 2005) described further in this chapter.

2.2 Discrete Time Homogeneous Markov Chains and spike train statistics

We start presenting Markov chains and how they are defined from transition probabilities. We show how an invariant probability distribution can be obtained from the transition probabilities and when it is unique. The invariant probability distribution of the Markov chain is the fundamental object of this section, because it is supposed to characterize the spike train statistics when the spike train is stationary. We discuss the main disadvantage of estimating transition probabilities from data and how this issue can be solved using MaxEnt method. We first show how this approach can be used in the case when successive spikes are independent, then we move to the more general case of fixed memory that includes the memory-less case.

2.2.1 Block coding

The properties of Markov chains considered in this thesis are easily expressed using matrix/vectors representation. For this purpose, we choose a symbolic representation of spike blocks of length L (which are the states of the Markov chain we construct). There are $M = 2^{NL}$ such possible spike blocks. Each spike block of length $L = D + 1$: ω_0^D is associated to a unique integer (index) $l = \sum_{k=1}^{N}\sum_{n=0}^{D} 2^{n\,N+k-1}\,\omega_k(n)$, where neurons $k = 1, .., N$ are considered from top to bottom and time $n = 0, .., D$ from left to right in the spike train. We denote $\omega^{(l)}$ the spike block corresponding to the index l. Here an example with $N = 2$ and $L = 3$, $\omega^{(6)}= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

2.2.2 Markov chains and transition probabilities

We start defining Markov chains. This presentation is based in the classical textbook (Bremaud, 1999).

A sequence $\{X_n\}_{n\geq 0}$ of random variables with values in a set E is called a *discrete-time stochastic process* with state space E if n is a discrete set. If $X_n = l$, the process is said to be in state l at time n, or to visit state l at time n. Here $E = \mathcal{A}_0^{D-1} = \mathcal{A}_n^{D-1+n}$

Homogeneous Markov Chain
Let $\{X_n\}_{n\geq 0}$ be a discrete-time stochastic process with state space E. If for all integers $n > 0$ and all states $i_0, \ldots, i_{n-1}, l$

$$\mathbb{P}[X_n = l \mid X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \ldots, x_0 = i_0] = \mathbb{P}[X_n = l \mid X_{n-1} = i_{n-1}] \tag{2.1}$$

whenever both sides are well-defined, this stochastic process is called a *Markov chain.* i.e., the state of the system at time n depends only on the state of the system at time $n-1$, and does not depend on any other state before. It is called a homogeneous Markov chain (HMC) if, in addition, the right-hand side of (2.1) is independent of n. In this thesis the states are spike blocks of finite length L, $X_n = \omega_n^{n+L}$. States are denoted by either block notation $\omega^{(l)}$ or by integer numbers following the previously introduced block coding. Here an example of the state space $E = \mathcal{A}_0^1$ with $N = 2$ and $L = 2$.

$$E = \{\underbrace{\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]}_{0}, \underbrace{\left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]}_{1}, \ldots, \underbrace{\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]}_{15}\}$$

2.2.3 Forbidden and allowed transitions

We are considering transition probabilities of the form $\mathbb{P}[\omega(n) \mid \omega_{n-D}^{n-1}]$. It is convenient to represent these probabilities as transitions between blocks $\omega_{n-D}^{n-1} \to \omega_{n-D+1}^{n}$ where the block ω_{n-D+1}^{n-1} is common. This allows indeed a matrix representation of the Markov chain. However, the price to pay is to artificially enlarge the set of transitions, because one is lead to consider, transition between any 2 blocks $\omega^{(l)}, \omega^{(l')}$.

Consider two spike blocks $\omega^{(l)}, \omega^{(l')}$ of length $L \geq 1$. The transition $\omega^{(l)} \to \omega^{(l')}$ is *legal* if $\omega^{(l)}, \omega^{(l')}$ have a common sub-block ω_1^D of $\omega^{(l)}$ and ω_0^{D-1} of $\omega^{(l)}$. Here is an example taking blocks with $N = 2$ and $L = 3$ of a legal transition:

$$\omega^{(l)} = \left[\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{smallmatrix}\right] \to \omega^{(l')} = \left[\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{smallmatrix}\right]$$

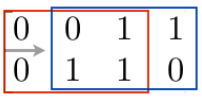


Figure 2.2: Allowed transition: The red block shifts to the blue block.

Thus, the probability to go from $\omega^{(l)}$ to $\omega^{(l')}$ is non-negative.

$$\mathbb{P}[\left[\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{smallmatrix}\right] \mid \left[\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{smallmatrix}\right]] = \mathbb{P}[\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] \mid \left[\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{smallmatrix}\right]] \geq 0$$

For forbidden transition, the transition probability is zero, as the event is impossible.

$$\omega^{(l)} = \left[\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{smallmatrix}\right] \nrightarrow \omega^{(l')} = \left[\begin{smallmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \end{smallmatrix}\right]$$

Then,

$$\mathbb{P}[\begin{smallmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \end{smallmatrix} \mid \begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{smallmatrix}] = 0$$

We construct the *transition matrix* $\mathcal{T}$ as follows:

$$\mathcal{T}_{\omega^{(l)},\omega^{(l')}} = \begin{cases} \mathbb{P}[\omega^{(l')} \mid \omega^{(l)}] \geq 0 & \text{if } \omega^{(l)} \to \omega^{(l')} \text{ is legal} \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

Note that although $\mathcal{T}$ has $(2^{NL})^2$ entries, it is a sparse matrix since each line has, at most, 2^N non-zero entries.

Since its entries are probabilities, and since a transition from any state $\omega^{(l)}$ must be to some state, it follows that

$$\mathbb{P}[\omega^{(l')} \mid \omega^{(l)}] \geq 0; \quad \sum_{\omega^{(l')} \in E} \mathbb{P}[\omega^{(l')} \mid \omega^{(l)}] = 1 \tag{2.3}$$

for all states $\omega^{(l)}, \omega^{(l')}$. A matrix $\mathcal{T}$ indexed by E and satisfying the above properties is called a stochastic matrix (for an illustration see 2.4).

A square stochastic matrix $\mathcal{T}$ is called *primitive* if for some ≥ 1 the matrix $\mathcal{T}^n$ has no entries equal to 0.

Remark: Any block ω_0^D of length $L = D+1$ can be viewed as a legal transition from the block $\omega^{(l)} = \omega_0^{D-1}$ to the block $\omega^{(l')} = \omega_1^D$ and in this case we write $\omega_0^D \sim \omega^{(l)}\omega^{(l')}$.

Transition Diagrams

A Markov chain transition matrix can be represented graphically as a transition-probability diagram where each node represents a state (or spike block), directed arc connects state $\omega^{(l)}$ to state $\omega^{(l')}$ if a one-step transition from $\omega^{(l)}$ to $\omega^{(l')}$ is allowed. The one-step transition probability $\mathbb{P}[\omega^{(l')} \mid \omega^{(l)}]$ is written next to the arc. A transition from a state to itself is represented by a loop (for an illustration see 2.3).

2.2.4 Distribution of an homogeneous Markov Chain

The random variable $X_0 = \omega_0^{D-1}$ is called the initial state, and its probability distribution ν, is the initial distribution

$$\nu(i) = \mathbb{P}[X_0 = i], \tag{2.4}$$

From the homogeneous Markov property and the definition of the transition matrix;

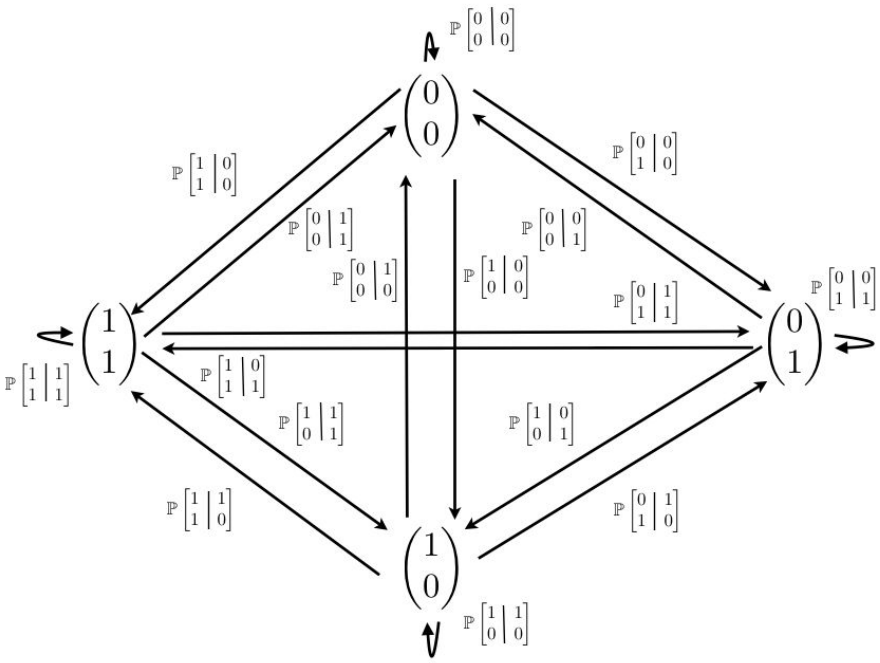


Figure 2.3: Transition diagram: for a simple network of two neurons and 1 step memory. In this case all states are accessible in one time-step. This is not the case any more if $D > 1$.

$$\mathcal{T}_{\omega^{(l)},\omega^{(l')}} = \begin{array}{c} \omega^{(l)}/\omega^{(l')} \\ \end{array} \begin{array}{c} \begin{matrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{matrix} \\ \begin{matrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{matrix} \begin{bmatrix} \mathbb{P}\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right] \\ \mathbb{P}\left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right] \\ \mathbb{P}\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right] \\ \mathbb{P}\left[\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right] & \mathbb{P}\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] \end{bmatrix} \end{array}$$

Figure 2.4: Transition matrix corresponding to the transition diagram 2.3

$$\mathbb{P}[X_0 = i_0, X_1 = i_1, \ldots, X_k = i_k] = \nu(i_0)\mathcal{T}_{i_0,i_1} \cdots \mathcal{T}_{i_{k-1},i_k} \tag{2.5}$$

for all $k > 0$, for all states $i_0, i_1, \ldots, i_k$, equation (2.5) constitute the probability law of the homogeneous Markov chain. The distribution of a discrete-time homogeneous Markov chain is determined by its initial distribution and its transition matrix.

Invariant Probability

We now introduce the central notion of the theory of discrete-time homogeneous Markov chains. A probability vector μ indexed by E satisfying

$$\mu^T = \mu^T \mathcal{T} \tag{2.6}$$

(where T denotes the transpose) is called an *invariant or equilibrium distribution* of the transition matrix $\mathcal{T}$ (or of the corresponding homogeneous Markov chain). The global balance equation (2.6) says that for all states $\omega^{(l)}$

$$\mu[\omega^{(l)}] = \sum_{\omega^{(l')} \in E} \mathbb{P}[\omega^{(l)} \mid \omega^{(l')}]\mu[\omega^{(l')}] \tag{2.7}$$

Detailed Balance

From equation (2.3) we get:

$$\sum_{\omega^{(l')}} \mathbb{P}[\omega^{(l')} \mid \omega^{(l)}]\mu[\omega^{(l)}] = \mu[\omega^{(l)}]$$

and from equation (2.7):

$$\sum_{\omega^{(l')}} \left(\mathbb{P}[\omega^{(l)} \mid \omega^{(l')}]\mu[\omega^{(l')}] - \mathbb{P}[\omega^{(l')} \mid \omega^{(l)}]\mu[\omega^{(l)}]\right) = 0 \tag{2.8}$$

A sufficient, but not necessary condition to have (2.8) equal to 0 is to have the quantity inside the parenthesis equal to zero. We call *detailed balance* this condition.

$$\mathbb{P}\left[\omega^{(l')} \,\middle|\, \omega^{(l)}\right] \mu[\omega^{(l)}] = \mathbb{P}\left[\omega^{(l)} \,\middle|\, \omega^{(l')}\right] \mu[\omega^{(l')}], \quad \forall \omega^{(l)}, \omega^{(l')} \in E. \tag{2.9}$$

If the detailed balance equations are satisfied, μ is therefore a stationary distribution of $\mathcal{T}$.

Remarks:

- Detailed balance is a sufficient condition to have an invariant measure, but it is not necessary.

- Detailed balance is a common assumption in statistical physics, standard MCMC (Markov chain Monte Carlo) methods are able to estimate (reconstruct) the invariant measure of a Markov process assuming that the detailed balance condition is satisfied.

- The framework we use in this thesis does not make this assumption, only a weaker condition of primitivity (see section 2.2.7.1) of the transition matrix is required.

Conditions for uniqueness

It may exist several distributions satisfying (2.6); the invariant probability is not necessarily unique. Take for example the identity as transition matrix (note that is a stochastic matrix). Then any probability distribution on the state space is a stationary distribution. Uniqueness requires additional assumptions.

We call *Return time* to the state $\omega^{(l)} \in E$, the time:

$$T_{\omega^{(l)}} = \inf\{n \geq 1 : X_n = \omega^{(l)}\}$$

A state $\omega^{(l)} \in E$ is called *recurrent* if

$$P(T_{\omega^{(l)}} < \infty) = 1$$

and otherwise it is called transient. A recurrent state $\omega^{(l)} \in E$ is called *positive recurrent* if

$$\mathbb{E}[T_{\omega^{(l)}}] < \infty.$$

A state $\omega^{(l)}$ is said to communicate with state $\omega^{(l')}$ (written $\omega^{(l)} \leftrightarrow \omega^{(l')}$) if both can be accessible from each other, denoted $\omega^{(l)} \to \omega^{(l')}$ and $\omega^{(l')} \to \omega^{(l)}$. A set of states B is a communicating class if every pair of states in B communicates with each other, and no state in B communicates with any state not in B.

A Markov chain is *irreducible* if its state space is a single communicating class, i.e., if it is possible to get to any state from any state (not necessarily in one step).

Theorem 1 (Uniqueness of the invariant probability) (Bremaud, 1999) *An irreducible Markov Chain is positive recurrent if and only if there exist an invariant measure μ. Then, μ is the unique invariant measure.*

In terms of the transition matrix only a condition of primitivity is needed to ensure the uniqueness of the invariant probability.

2.2.5 Ergodic Theorem

A state $\omega^{(l)}$ is said to be *ergodic* if it is aperiodic (returns to state $\omega^{(l)}$ occur at irregular times) and positive recurrent. If all states in an irreducible Markov chain are ergodic, then the chain is said to be ergodic.

Theorem 2 (Ergodic Theorem) (Bremaud, 1999) *Let $\{X_n\}_{n\geq 0}$ be an ergodic Markov chain with the stationary distribution μ, and let $f : E \to \mathbb{R}$ be such that*

$$\sum_{\omega^{(l)}\in E} |f(\omega^{(l)})|\mu(\omega^{(l)}) < \infty$$

Then, for any initial distribution $\nu, P_\nu - a.s:$

$$\lim_{T\to\infty} \frac{1}{T}\sum_{k=1}^{T} f(X_k) = \sum_{\omega^{(l)}\in E} f(\omega^{(l)})\mu(\omega^{(l)})$$

This theorem tell us that the time average of an observable f over a sequence of events $X_k = \omega_k^{k+L}$ is the same as the average value of the observable with respect to the stationary distribution μ, independent of initial probability distribution ν. Further in this chapter, when we discuss about statistical estimation we come back to this point.

2.2.6 Observables and Potentials

In this section we introduce the concept of potentials. We derive how to get form a potential their associated transition probabilities and invariant measure, we show in 2.2.7.2 an explicit example where the computations can be done analytically. Let us start with some definitions.

Observables

An *observable* is a function, $\mathcal{O}(\omega)$, that associates a real number to a spike train. We say that an observable $\mathcal{O}$ has *range* R if it depends on R consecutive spike patterns, e.g. $\mathcal{O}(\omega) = \mathcal{O}(\omega_0^{R-1})$. We consider here that observables do not depend explicitly on time (*time-translation invariance of observables*). As a consequence, for any time n, $\mathcal{O}(\omega_0^{R-1}) = \mathcal{O}(\omega_n^{n+R-1})$ whenever $\omega_0^{R-1} = \omega_n^{n+R-1}$. Prominent examples of observables are products of the form:

$$\mathcal{O}(\omega) = m_{p_1,\ldots,p_r}(\omega) = \prod_{u=1}^{r} \omega_{k_u}(n_u), \tag{2.10}$$

where p_u, $u = 1 \ldots r$ are pairs of neuron-time events $(k_u, n_u), k_u = 1 \ldots N$ being the neuron index and $n_u = 0 \ldots D$ being the time index. These observables are called *monomials* and take values $\{0, 1\}$. Typical choices of monomials are $\omega_{k_1}(n_1)$ which is 1 if neuron k_1 fires at time n_1 and which is 0 otherwise; $\omega_{k_1}(n_1)\,\omega_{k_2}(n_2)$ which is 1

if neuron k_1 fires at time n_1 and neuron k_2 fires at time n_2 and which is 0 otherwise. Another example is $\omega_{k_1}(n_1)\,(1-\omega_{k_2}(n_2))$ which is 1 is neuron k_1 fires at time n_1 *and neuron* k_2 *is silent at time* n_2. For N neurons and a time range R there are thus 2^{NR} possible monomials. Any observable of range R can be represented as a linear combination of products (2.10). Monomials constitute therefore a canonical basis for observable representation.

To alleviate notations, instead of labeling monomials by a list of pairs, as in (2.10), we label them by an integer index, l (the index is defined in the same way as the block index). So a monomial reads m_l.

Potential

An example of observable is called a *potential*. Any potential of range R can be written as a linear combination of monomials, which constitute the "effective interactions". The range of the potential is the maximum of the range of the monomials m_l. A potential of range R is therefore written as follows:

$$\mathcal{H}(\omega_0^D) \;=\; \sum_{l=1}^{2^{NR}} h_l m_l(\omega_0^D). \tag{2.11}$$

where the coefficients h_l are finite[1] real numbers. Some coefficients in the expansion may be zero. We assume throughout this thesis that $h_l > \infty$.

2.2.7 Markov Chains and normalized potentials

We assume here that the memory depth of the chain D is constant and finite, although an extension of the present formalism to variable length Markov chains (D variable) or chains with complete connections (D infinite) is possible (Fernandez and Maillard, 2005). We also assume that the chain is homogeneous (transition probabilities do not depend explicitly on time) and primitive (In the present context this property is ensured by the assumption $\mathcal{H} > -\infty$ (sufficient condition)). Then, the Markov chain admits a unique invariant probability μ which obeys the Chapman-Kolmogorov relation: $\forall n_1, n_2, n_2 + D - 1 > n_1$,

$$\mu\left[\,\omega_{n_1}^{n_2}\,\right] = \prod_{n=n_1}^{n_2-D} \mathbb{P}\left[\,\omega(n+D)\;\middle|\;\omega_n^{n+D-1}\,\right]\mu\left[\,\omega_{n_1}^{n_1+D-1}\,\right]. \tag{2.12}$$

Introducing:

$$\phi(\omega_n^{n+D}) = \log \mathbb{P}\left[\,\omega(n+D)\;\middle|\;\omega_n^{n+D-1}\,\right], \tag{2.13}$$

we have

$$\mu\left[\,\omega_{n_1}^{n_2}\,\right] = e^{\sum_{n=n_1}^{n_2-D}\phi(\omega_n^{n+D})}\mu\left[\,\omega_{n_1}^{n_1+D-1}\,\right]. \tag{2.14}$$

[1] Thus, we do not consider here hard core potentials with forbidden configurations.

ϕ is called a *normalized potential*. As we will see in the next section to each bounded $\mathcal{H}$ of the form (2.11) corresponds a unique normalized potential ϕ and a unique invariant measure μ. Although this correspondence can be found in many textbooks (see for example (Georgii, 1988, Keller, 1998)), we summarize it here since it is the core of the methods we use in this thesis.

2.2.7.1 Transfer Matrix

A generalization of the concept of transition matrix is the *transfer matrix* $\mathcal{L}$:

$$\mathcal{L}_{\omega^{(l)},\omega^{(l')}} = \begin{cases} e^{\mathcal{H}(\omega_0^D)} & \text{if } \omega_0^D \sim \omega^{(l)}\omega^{(l')} \\ 0, & \text{otherwise.} \end{cases} \tag{2.15}$$

Remarks:

- The transfer matrix is not necessarily a stochastic matrix (only when the potential $\mathcal{H}$ is normalized).
- From the assumption $\mathcal{H} > -\infty$, each legal transition corresponds to a positive entry in the matrix $\mathcal{L}$. Therefore $\mathcal{L}$ is primitive and satisfies the Perron-Frobenius theorem (Gantmacher, 1998).

Theorem 3 (Perron Frobenius) (Gantmacher, 1998)
$\mathcal{L}$ has a unique maximal and strictly positive eigenvalue s associated with a right eigenvector $R(\cdot)$ and a left eigenvector $L(\cdot)$, with positive and bounded entries, such that $\mathcal{L}R = sR$, $L\mathcal{L} = sL$. Those vectors can be chosen such that $\langle L, R\rangle = 1$ where $\langle\cdot,\cdot\rangle$ is the scalar product. The remaining part of the spectrum is located in a disk in the complex plane, of radius strictly lower than s. As a consequence, for all v ,

$$\frac{1}{s^n}\mathcal{L}^n v \to \langle L, R\rangle v$$

as $n \to \infty$
The Gibbs-probability of a spike block w is:

$$\mu(w) = \langle L_w, R_w\rangle$$

where L_w and R_w are the $w-th$ component of $R(\cdot)$.

As a consequence of the Perron-Frobenius theorem, $\mathcal{L}$ has a unique real positive eigenvalue s, strictly larger in modulus than the other eigenvalues. This unique eigenvalue and its associated eigenvectors define uniquely the invariant probability measure associated to $\mathcal{L}$.

The following holds:

(a) The potential:

$$\phi(\omega_0^D) = \mathcal{H}(\omega_0^D) - \underbrace{\log R\left(\omega_0^{D-1}\right) + \log R\left(\omega_1^D\right) - \log s}_{\zeta(\omega_0^D)} \tag{2.16}$$

where $\zeta(\omega_0^D)$ is called *normalization function* is normalized i.e. it defines via (2.13) an homogeneous Markov chain with transition probability $\mathbb{P}\left[\omega(D) \,\middle|\, \omega_0^{D-1}\right] = e^{\phi(\omega_0^D)}$.

(b) The unique invariant probability of this Markov chain can be written:

$$\mu(\omega_0^{D-1}) = \frac{R\left(\omega_0^{D-1}\right) L\left(\omega_0^{D-1}\right)}{\langle L, R\rangle}. \tag{2.17}$$

(c) It follows from Chapman-Kolmogorov equation (2.12) and from (2.16,2.17) that, for $D > 0$:

$$\mu\left[\omega_0^n\right] = \frac{e^{\sum_{k=0}^{n-D} \mathcal{H}\left(\omega_k^{k+D}\right)}}{s^{n-D+1}} R\left(\omega_{n-D+1}^n\right) L\left(\omega_0^{D-1}\right). \tag{2.18}$$

This can be easily seen:

Write the Chapman-Kolmogorov equation 2.12 using the normalized potential:

$$\mu[\omega_0^n] = e^{\sum_{k=0}^{n-D} \phi(\omega_k^{k+D})} \mu[\omega_0^{D-1}] \tag{2.19}$$

Using equation (2.16), we have:

$$\begin{aligned}
\sum_{k=0}^{n-D} \phi(\omega_k^{k+D}) &= \sum_{k=0}^{n-D} \left(\mathcal{H}(\omega_k^{k+D}) - \log R(\omega_k^{k+D-1}) + \log R(\omega_{k+1}^{k+D}) - \log s\right) \\
&= \sum_{k=0}^{n-D} \mathcal{H}(\omega_k^{k+D}) - \log R(\omega_0^{D-1}) + \log R(\omega_{n-D+1}^n) - (n-D+1)\log s
\end{aligned}$$

Where the last equality is obtained by applying the sum to each term (most of the terms $\log R(\cdot)$ cancel out).

Taking exponential we get:

$$e^{\sum_{k=0}^{n-D} \phi(\omega_k^{k+D})} = e^{\left(\sum_{k=0}^{n-D} \mathcal{H}(\omega_k^{k+D}) - \log R(\omega_0^{D-1}) + \log R(\omega_{n-D+1}^n) - (n-D+1)\log s\right)}$$

using properties of the exponential function and multiplying both sides by (2.17) we get:

$$e^{\sum_{k=0}^{n-D}\phi(\omega_k^{k+D})}\mu[\omega_0^{D-1}] = \frac{e^{\sum_{k=0}^{n-D}\mathcal{H}(\omega_k^{k+D})}R(\omega_{n-D+1}^{n})R(\omega_0^{D-1})L(\omega_0^{D-1})}{R(\omega_0^{D-1})s^{n-D+1}}$$

Finally, for the left hand we use equation (2.19) and the right hand side is obtained canceling out the term $R(\omega_0^{D-1}) > 0$.

$$\mu[\omega_0^n] = \frac{e^{\sum_{k=0}^{n-D}\mathcal{H}(\omega_k^{k+D})}R(\omega_{n-D+1}^{n})L(\omega_0^{D-1})}{s^{n-D+1}}. \tag{2.20}$$

On practical grounds, when the state space is finite, we can rely on the standard results of linear algebra to study the asymptotic behavior of homogeneous Markov chains. Therefore, to characterize uniquely the spike train statistics of a neural network whose spikes are interacting following a potential $\mathcal{H}$. The Perron-Frobenius theorem is all that is needed, at least in the theory. Let us illustrate what have been explained so far with an example.

2.2.7.2 Example: How to go from a bounded potential $\mathcal{H}$ to its invariant measure μ ?

Let us consider a range-2 potential with two neurons:

$$\mathcal{H}(\omega) = h_1\omega_1(1)\omega_2(0).$$

The transfer matrix (2.15) associated to $\mathcal{H}$ is:

$$\mathcal{L}(\mathcal{H}) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & e^{h_1} & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

As this matrix is primitive by construction, thus, satisfy the hypothesis of the Perron Frobenius theorem, there exist a unique maximum eigenvalue, root of the characteristic polynomial, that in this case reads:

$$Q(s) = s^4 - 4s^3 + s^2 - e^{h_1}s^2 = s^2(s^2 - 4s + 1) - s^2e^{h_1} \tag{2.21}$$

Removing the trivial roots we get:

$$e^{h_1} - 1 = s^2 - 4s \tag{2.22}$$

so we have 2 eigenvalues equal to 0 and the others can be obtained solving the second degree equation. They are: $2 \pm \sqrt{e^{h_1} + 3}$, thus the largest eigenvalue is $s = 2 + \sqrt{e^{h_1} + 3}$. The left and right eigenvectors associated to this largest eigenvalue are respectively :

$$L\begin{pmatrix}0\\0\end{pmatrix}=1; L\begin{pmatrix}0\\1\end{pmatrix}=1; L\begin{pmatrix}1\\0\end{pmatrix}=s-3; L\begin{pmatrix}1\\1\end{pmatrix}=1$$

$$R\begin{pmatrix}0\\0\end{pmatrix}=1; R\begin{pmatrix}0\\1\end{pmatrix}=s-3; R\begin{pmatrix}1\\0\end{pmatrix}=1; R\begin{pmatrix}1\\1\end{pmatrix}=1$$

The normalization function ζ 2.16 takes therefore the following values:

$$\begin{aligned}
\zeta\begin{bmatrix}0&0\\0&0\end{bmatrix} &= \log R\begin{pmatrix}0\\0\end{pmatrix}-\log R\begin{pmatrix}0\\0\end{pmatrix}+\log s=\log s\\
\zeta\begin{bmatrix}0&0\\0&1\end{bmatrix} &= \log R\begin{pmatrix}0\\0\end{pmatrix}-\log R\begin{pmatrix}0\\1\end{pmatrix}+\log s=-\log(s-3)+\log s\\
\zeta\begin{bmatrix}0&1\\0&0\end{bmatrix} &= \log R\begin{pmatrix}0\\0\end{pmatrix}-\log R\begin{pmatrix}1\\0\end{pmatrix}+\log s=\log s\\
\zeta\begin{bmatrix}0&1\\0&1\end{bmatrix} &= \log R\begin{pmatrix}0\\0\end{pmatrix}-\log R\begin{pmatrix}1\\1\end{pmatrix}+\log s=\log s\\
\zeta\begin{bmatrix}0&0\\1&0\end{bmatrix} &= \log R\begin{pmatrix}0\\1\end{pmatrix}-\log R\begin{pmatrix}0\\0\end{pmatrix}+\log s=\log(s-3)+\log s\\
\zeta\begin{bmatrix}0&0\\1&1\end{bmatrix} &= \log R\begin{pmatrix}0\\1\end{pmatrix}-\log R\begin{pmatrix}0\\1\end{pmatrix}+\log s=\log s\\
\zeta\begin{bmatrix}0&1\\1&0\end{bmatrix} &= \log R\begin{pmatrix}0\\1\end{pmatrix}-\log R\begin{pmatrix}1\\0\end{pmatrix}+\log s=\log(s-3)+\log s\\
\zeta\begin{bmatrix}0&1\\1&1\end{bmatrix} &= \log R\begin{pmatrix}0\\1\end{pmatrix}-\log R\begin{pmatrix}1\\1\end{pmatrix}+\log s=\log(s-3)+\log s\\
\zeta\begin{bmatrix}1&0\\0&0\end{bmatrix} &= \log R\begin{pmatrix}1\\0\end{pmatrix}-\log R\begin{pmatrix}0\\0\end{pmatrix}+\log s=\log s\\
\zeta\begin{bmatrix}1&0\\0&1\end{bmatrix} &= \log R\begin{pmatrix}1\\0\end{pmatrix}-\log R\begin{pmatrix}0\\1\end{pmatrix}+\log s=-\log(s-3)+\log s\\
\zeta\begin{bmatrix}1&1\\0&0\end{bmatrix} &= \log R\begin{pmatrix}1\\0\end{pmatrix}-\log R\begin{pmatrix}1\\0\end{pmatrix}+\log s=\log s\\
\zeta\begin{bmatrix}1&1\\0&1\end{bmatrix} &= \log R\begin{pmatrix}1\\0\end{pmatrix}-\log R\begin{pmatrix}1\\1\end{pmatrix}+\log s=\log s\\
\zeta\begin{bmatrix}1&0\\1&0\end{bmatrix} &= \log R\begin{pmatrix}1\\1\end{pmatrix}-\log R\begin{pmatrix}0\\0\end{pmatrix}+\log s=\log s\\
\zeta\begin{bmatrix}1&0\\1&1\end{bmatrix} &= \log R\begin{pmatrix}1\\1\end{pmatrix}-\log R\begin{pmatrix}0\\1\end{pmatrix}+\log s=-\log(s-3)+\log s\\
\zeta\begin{bmatrix}1&1\\1&0\end{bmatrix} &= \log R\begin{pmatrix}1\\1\end{pmatrix}-\log R\begin{pmatrix}1\\0\end{pmatrix}+\log s=\log s\\
\zeta\begin{bmatrix}1&1\\1&1\end{bmatrix} &= \log R\begin{pmatrix}1\\1\end{pmatrix}-\log R\begin{pmatrix}1\\1\end{pmatrix}+\log s=\log s;
\end{aligned}$$

Taking the exponential of equation (2.16) on the r.h.s we get the values from the transfer matrix and we divide by the previously computed values of the normaliza-

tion function. Therefore the Markov chain transition matrix (2.2) reads :

$$e^{\phi} = \frac{1}{s}\begin{pmatrix} 1 & s-3 & 1 & 1 \\ \frac{1}{s-3} & 1 & \frac{e^{h_1}}{s-3} & \frac{1}{s-3} \\ 1 & s-3 & 1 & 1 \\ 1 & s-3 & 1 & 1 \end{pmatrix},$$

with $e^{h_1} - 1 = s^2 - 4s$, so that e^{ϕ} is correctly normalised (the sum of entries on each row is equal to 1).

The invariant probability of the Markov chain is given by equation (2.17) after normalization by $\langle L, R\rangle = 2s - 4$:

$$\mu\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \frac{1}{2s-4}, \quad \mu\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{s-3}{2s-4}, \quad \mu\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{s-3}{2s-4}, \quad \mu\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{2s-4}$$

Here we can use the Chapman-Kolmogorov equation (2.12) to obtain the invariant measure of blocks of length 2.

$$\mu\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mu\begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{\phi\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} = \frac{s-3}{s}\frac{s-3}{2s-4};$$

$$\mu\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \mu\begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{\phi\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} = \frac{1}{s}\frac{e^{h_1}}{2s-4};$$

With these equations, using (2.13) we can verify that *in general the detailed balance condition is not satisfied*:

$$\mu\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbb{P}\left[\begin{matrix} 0 \\ 1 \end{matrix} \middle| \begin{matrix} 1 \\ 0 \end{matrix}\right] \neq \mu\begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbb{P}\left[\begin{matrix} 1 \\ 0 \end{matrix} \middle| \begin{matrix} 0 \\ 1 \end{matrix}\right]$$

$$\Leftrightarrow$$

$$\mu\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \neq \mu\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Remark: The detailed balance condition is only satisfied in the trivial case $h_1 = 0$, in this case the maximum eigenvalue $s = 4$ and the invariant probability is the uniform assigning $\frac{1}{16}$ to each block of range 2 of the system.

2.2.8 From Markov chains to Gibbs distributions

Let us outline what we have obtained. Given a potential $\mathcal{H}$ of finite range R, written as a linear combination of monomials, with finite coefficients, it is always possible to associate a homogeneous Markov chain with memory D, and invariant distribution μ. Let us now show that μ corresponds to the concept of Gibbs distribution coming from statistical physics. To do that let us start with some definitions.

2.2.8.1 Entropy of a Markov chain

We define the *entropy rate* (or Kolmogorov-Sinai entropy) of a stationary Markov chain $\mathcal{S}[\mu]$ (Kitchens, 1998):

$$\mathcal{S}[\mu] = -\sum_{\omega_0^D} \mu[\omega_0^D] \mathbb{P}\left[\omega(D) \,\middle|\, \omega_0^{D-1}\right] \log \mathbb{P}\left[\omega(D) \,\middle|\, \omega_0^{D-1}\right], \tag{2.23}$$

When $D = 0$, $\mathcal{S}[\mu]$ reduces to the classical memory-less definition:

$$\mathcal{S}(\mu) = -\sum_{\omega(0)} \mu[\omega(0)] \log \mu[\omega(0)]. \tag{2.24}$$

Note that the time index (here, zero) does not play a role, since we hace assumed μ to be stationary (time translation invariant).

2.2.8.2 Variational principle and equilibrium states

Assume that $\mathcal{H}$ is bounded and has finite range[2]:

$$\mathcal{P}[\mathcal{H}] = \sup_{\nu \in \mathcal{M}_{inv}} (\mathcal{S}[\nu] + \nu[\mathcal{H}]) = \mathcal{S}[\mu] + \mu[\mathcal{H}], \tag{2.25}$$

where $\mathcal{P}[\mathcal{H}]$ is called the *free energy or topological presure* and its properties are listed below. $\nu[\mathcal{H}] = \sum_{\omega_0^D} \mathcal{H}(\omega_0^D)\nu(\omega_0^D)$ is the average value of $\mathcal{H}$ with respect to the probability ν.

Looking at the second equality, the variational principle (2.25) selects, among all possible probabilities ν the probability μ realizing the supremum of $\mathcal{S}[\nu] + \nu[\mathcal{H}]$.

The supremum is attained - not necessarily at a unique probability. Each probability μ for which the supremum is attained is called an *equilibrium state* for $\mathcal{H}$. As in statistical physics the word "state" is here used synonymously with "distribution".

A variant of this principle holds when the average value of observables $\mathcal{O}_k$ is constrained to a value C_k, fixed e.g. by experimental observations. For a probability measure ν we have $\nu[\mathcal{H}] = \sum_{k=1}^{K} h_k \nu[\mathcal{O}_k]$, which becomes $\sum_{k=1}^{K} h_k C_k$ if the average value of all observables $\mathcal{O}_k$ is constrained. In this case the variational principle (2.25) reduces to maximizing the entropy on the set of measures $\nu \in \mathcal{M}_{inv}$ such that $\nu[\mathcal{O}_k] = C_k$. Then, one is lead to a classical Lagrange multipliers problem where the h_ks are the Lagrange multipliers. This is the classical approach introduced by (Jaynes, 1957).

2.2.8.3 Free energy

The quantity $\mathcal{P}[\mathcal{H}]$ has the following properties:

[2] The variational principle still holds if the range is infinite and its variation decays sufficiently fast with m, typically exponentially. It can be shown that its invariant measure μ satisfy the variational principle (Ruelle, 1978, Bowen, 1975, Chazottes and Keller, 2008)

- $\mathcal{P}[\mathcal{H}]$ is a log generating function of cumulants. We have:

$$\frac{\partial \mathcal{P}[\mathcal{H}]}{\partial h_l} = \mu[m_l], \tag{2.26}$$

the average of m_l with respect to μ and:

$$\frac{\partial^2 \mathcal{P}[\mathcal{H}]}{\partial h_k \partial h_l} = \frac{\partial \mu[m_l]}{\partial h_k} = \sum_{n=-\infty}^{+\infty} C_{m_k,m_l}(n), \tag{2.27}$$

where $C_{m_k,m_l}(n)$

$$C_{\mathcal{O}_k \mathcal{O}_l}(n) = \mu[\mathcal{O}_k \mathcal{O}_l \circ \sigma^n] - \mu[\mathcal{O}_k]\mu[\mathcal{O}_l],$$

is the correlation function between the two monomials m_k and m_l at time n in the equilibrium state μ. Note that correlation functions decay exponentially fast whenever $\mathcal{H}$ has finite range and $\mathcal{H} > -\infty$, thus $\sum_{n=-\infty}^{+\infty} C_{m_k,m_l}(n) < +\infty$. Eq. (2.27) characterizes the variation in the average value of m_l when varying h_k (linear response). The corresponding matrix is a *susceptibility matrix*. It controls the Gaussian fluctuations of monomials around their mean (central limit theorem) (Bowen, 1975). When considering potential of range 1 ($D = 0$) eq (2.27) reduces to the classical fluctuation-dissipation theorem, because the corresponding process has no memory (successive times are independent thus $C_{m_k,m_l}(n) = 0$ unless $n = 0$).

- For finite range potentials $\mathcal{P}(\mathcal{H})$ is a convex function of h_l's. This ensures the uniqueness of the solution of (2.25).

Gibbs property: The invariant measure of the Markov chain μ obtained from $\mathcal{H}$ have the following properties:

1. μ obeys the variational principle (2.25) and

$$\mathcal{P}[\mathcal{H}] = \log s. \tag{2.28}$$

When considering a normalized potential ϕ, the transfer matrix becomes a stochastic transition matrix with maximal eigenvalue 1. Thus $\mathcal{P}[\phi] = 0$.

2. It follows from (2.18) that

$\exists\, A, B > 0$ such that, for any block ω_0^n the invariant distribution obeys (Bowen, 1975, Keller, 1998):

$$A \leq \frac{\mu[\omega_0^n]}{e^{-(n-D+1)\mathcal{P}(\mathcal{H})} e^{-\sum_{k=0}^{n-D} \mathcal{H}(\omega_k^{k+D})}} \leq B. \tag{2.29}$$

This equation can be verified by taking exponential to equation 2.28

$$s = e^{\mathcal{P}[\mathcal{H}]}$$

Plugging in equation 2.20 we obtain:

$$\frac{\mu[\omega_0^n]}{e^{-(n-D+1)\mathcal{P}[\mathcal{H}]}e^{-\sum_{k=0}^{n-D}\mathcal{H}(\omega_k^{k+D})}} = R(\omega_{n-D+1}^n)L(\omega_0^{D-1})$$

As $R(\omega_{n-D+1}^n)L(\omega_0^{D-1})$ are positive numbers, there exist $A, B > 0$ s.t.

$$A \leq \frac{\mu[\omega_0^n]}{e^{-(n-D+1)\mathcal{P}[\mathcal{H}]}e^{-\sum_{k=0}^{n-D}\mathcal{H}(\omega_k^{k+D})}} \leq B$$

This is actually the definition of Gibbs distributions in ergodic theory and this is the definition we use in this thesis [3].

This definition encompasses the classical definition of Gibbs distributions, $\frac{e^{\mathcal{H}}}{Z}$ found in standard textbooks of statistical physics. Let us indeed consider a potential of range $R = 1, (D = 0)$. This is a limit case in the definition of the transfer matrix where transitions between spike patterns $\omega(0) \to \omega(1)$ are considered and where all transitions are legal. $\mathcal{L}_{\omega(0),\omega(1)} = e^{\mathcal{H}(\omega(0))}$, thus each column has the form:

$$(e^{\mathcal{H}(\omega(0))}, e^{\mathcal{H}(\omega(0))}, \ldots, e^{\mathcal{H}(\omega(0))}).$$

The matrix $\mathcal{L}$ is degenerated with a maximum eigenvalue:

$$s = Z = \sum_{\omega(0)} e^{\mathcal{H}(\omega(0))}$$

and all other eigenvalues 0. The left eigenvector corresponding to $s = Z$ is:

$$L = (\frac{1}{Z}, \frac{1}{Z}, \ldots, \frac{1}{Z})$$

whereas $R(\omega(0)) = e^{\mathcal{H}(\omega(0))}$. Note that we have normalized L so that $\langle L, R\rangle = 1$. We have therefore:

$$\mu(\omega(0)) = \frac{e^{\mathcal{H}(\omega(0))}}{Z}, \tag{2.30}$$

the classical form for the Gibbs distribution. The normalized potential in the limiting case is $\phi(\omega(0)) = \mathcal{H}(\omega(0)) - \log Z$, whereas the Markov chain has no memory: successive events are independent. This last remark reflects a central weakness of memory-less MaxEnt models to describe neuron dynamics. They neither involve

[3] When considering finite range potentials equilibrium states and Gibbs distributions are equivalent notions. This equivalence requires additional assumptions for infinite range potentials (Chazottes and Keller, 2008).

memory nor time causality. Note that from equation (2.18): $\mu[\omega_0^n]$ has not the form $\frac{e^{\mathcal{H}(\omega(0))}}{Z_n}$ with $Z_n = \sum_{\omega_0^n} e^{\mathcal{H}(\omega_0^n)}$. However the following holds:

$$\mathcal{P}[\mathcal{H}] = \lim_{n\to\infty} \frac{1}{n} \log Z_n. \tag{2.31}$$

This outlines a crucial point: as soon as one introduces memory in the MaxEnt, infinite time limits have to be considered in order to fully characterize the statistics. This is *mutatis mutandis* the same procedure as taking the thermodynamic limit in spatial lattices (Georgii, 1988).

Remarks:

- While Gibbs distributions are defined by specifying certain conditional probabilities, equilibrium states are defined by a variational principle. Gibbs states are always equilibrium states, but the two notions do not coincide in general. However, for a class of sufficiently regular potentials, equilibrium states are also Gibbs states (Chazottes and Keller, 2008).
- For finite range potentials $\mathcal{H}$, one has always a unique equilibrium state which has the "Gibbs property" (2.29).
- Coming back to our example using the free energy we can push forward the analysis made in our example 2.2.7.2 and compute the derivative with respect to the parameter h_1.

$$\begin{aligned}
\frac{\partial \mathcal{P}[\mathcal{H}]}{\partial h_1} &= \frac{\partial \log(2+\sqrt{e^{h_1}+3})}{\partial h_1}, \\
&= \frac{e^{h_1}}{4\sqrt{e^{h_1}+3}+2e^{h_1}+6}, \\
&= \frac{1}{s}\frac{e^{h_1}}{2s-4}, \\
&= \mu \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}
\end{aligned}$$

Verifying equation 2.26. Taking the second derivative of the Pressure gives:

$$\begin{aligned}
\frac{\partial^2 \mathcal{P}[\mathcal{H}]}{\partial h_1^2} &= \frac{\partial \frac{e^{h_1}}{4\sqrt{e^{h_1}+3}+2e^{h_1}+6}}{\partial h_1}, \\
&= \frac{e^{h_1}(3\sqrt{e^{h_1}+3}+e^{h_1}+6)}{2\sqrt{e^{h_1}+3}(2\sqrt{e^{h_1}+3}+e^{h_1}+3)^2} \\
&= \frac{\partial \mu \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}{\partial h_1}
\end{aligned}$$

Characterizing how varies $\mu \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ with respect to h_1.

2.3 Infinite range Gibbs distributions

In this section we introduce the construction of Gibbs distributions when the state space is made of infinite elements, thus rendering the matrix-vector markovian approach exposed in the previous section inappropriate. The construction of a Gibbs measure for a system with an infinite number of interacting components is done by the idea of specifying the interdependence structure by means of a suitable class of conditional probabilities (Georgii, 1988). This construction will be suited for the stochastic processes arising from the infinite memory GLM and for the Conductance Based Integrate and Fire with and without gap junctions.

To define properly under which hypothesis we obtain a Gibbs distribution let us start with definitions.

2.3.1 Continuity with respect to a spike train

Consider an infinite spike train. For $n \in \mathbb{Z}$, we note $\mathcal{A}_{-\infty}^{n-1}$ the set of sequences $\omega_{-\infty}^{n-1}$. Assume that we are given a set of transitions probabilities, like in the previous section, possibly depending on an infinite past[4], i.e. of the form $\mathbb{P}\left[\omega(n) \,\middle|\, \omega_{-\infty}^{n-1}\right]$. We give in the next chapter an example of neural network model where such transition probabilities with an infinite memory do appear.

Even if transition probabilities involve an infinite memory $\omega_{-\infty}^{n-1}$, it is reasonable to consider situations where the effects of past spikes decreases as they are more distant in the past. This corresponds to the mathematical notion of *continuity with respect to a spike train*. We note, for $n \in \mathbb{Z}$, $m \geq 0$ and r integer:

$$\omega \overset{\mathrm{m,n}}{=} \omega' \quad \text{if} \quad \omega(r) = \omega'(r),\, \forall r \in \{\, n-m, \ldots, n \,\}\,.$$

Consider a function f depending both on discrete time n and on the spike train part of ω *anterior to* n. We write $f(n,\omega)$ instead of $f(n,\omega_{-\infty}^{n-1})$. f is *continuous with respect to the spike train* ω if its *m-variation*:

$$var_m\left[f(n,.)\right] = \sup\left\{\, |\, f(n,\omega) - f(n,\omega')\,| \,:\, \omega \overset{\mathrm{m,n}}{=} \omega' \right\}. \tag{2.32}$$

tends to 0 as $m \to +\infty$. This precisely means that the influence of a change in the spikes has less and less effect on the actual value of f at time n as these changes are more distant in the past ($n-m$ decreases). From the definition we see clearly that any finite range potential is continuous.

[4] In this case, one has to assume that (i) for every $\omega(n) \in \mathcal{A}$, $\mathbb{P}\left[\omega(n)\,|\,.\,\right]$ is measurable with respect to $\mathcal{F}_{\leq n-1}$, the sigma-algebra on $\mathcal{A}_{-\infty}^{n-1}$; (ii) for every $\omega_{-\infty}^{n-1} \in \mathcal{A}_{-\infty}^{n-1}$, $\sum_{\omega(n)\in\mathcal{A}} \mathbb{P}\left[\omega(n) \,\middle|\, \omega_{-\infty}^{n-1}\right] = 1$.

Définition 1 *A Gibbs distribution is a probability measure μ such that:*

(i) for all $n \in \mathbb{Z}$ and all $\mathcal{F}_{\leq n}$-measurable functions f:

$$\int f\left(\omega_{-\infty}^{n}\right)\mu(d\omega) = \int \sum_{\omega(n)\in\mathcal{A}} f\left(\omega_{-\infty}^{n-1}\omega(n)\right)\mathbb{P}\left[\omega(n)\,\middle|\,\omega_{-\infty}^{n-1}\right]\mu(d\omega).$$

(ii) $\forall n \in \mathbb{Z}$, $\forall \omega_{-\infty}^{n-1} \in \mathcal{A}_{-\infty}^{n-1}$, $\mathbb{P}\left[\omega(n)\,\middle|\,\omega_{-\infty}^{n-1}\right] > 0$.

(iii) For each $n \in \mathbb{Z}$, $\mathbb{P}\left[\omega(n)\,\middle|\,\omega_{-\infty}^{n-1}\right]$ is continuous with respect to ω.

The condition (i) is a natural extension of the condition defining the invariant probability of an homogeneous Markov chain (2.6). In its most general sense (i) does not require stationarity and affords the consideration of an infinite memory. It defines so-called *compatibility conditions*. This states that the average of a function $f(n,\omega)$ with respect to μ, at time n (left hand side), is equal to the average computed from transition probabilities (right hand side). This equality must hold for any time n.

The remarkable aspect of this construction is the fact that a Gibbs measure for a given type of potential may fail to be unique. In physical terms, this means that a system with this interaction can take several distinct equilibrium. The phenomenon of nonuniqueness of a Gibbs measure can thus be interpreted as a phase transition. Therefore, the conditions under which a potential leads to a unique or to several Gibbs measures turns out to be of central importance. There exist several theorems guaranteeing the existence and uniqueness of a Gibbs distribution (Georgii, 1988, Fernandez and Maillard, 2005).

2.3.2 Conditions for uniqueness of Gibbs measure for infinite chains

There are several criteria of uniqueness. Here we consider the one presented in (Fernandez and Maillard, 2005), as is more suited for our developments.

Let:

$$m(p) = \inf_{n\in\mathbb{Z}} \inf_{\omega\in\mathcal{A}_{-\infty}^{n}} \mathbb{P}[\omega(n) \mid \omega_{-\infty}^{n-1}] \tag{2.33}$$

and

$$v(p) = \sup_{m\in\mathbb{Z}} \sum_{n'>m} var_{m-n'}\mathbb{P}[\omega(n) \mid \cdot] \tag{2.34}$$

If $m(p) > 0$ and $v(p) < \infty$, then there exists a unique Gibbs measure consistent with $\mathbb{P}[\omega(n) \mid \omega_{-\infty}^{n-1}]$.

2.4 Statistical Estimation

We now focus on the problem of estimate the probability of spike patterns. In Chapter 1 we have seen how experimentalist can obtain spikes from biological neural networks as the retina using MEA. Now, having spike train data what can we do to characterize the statistics of spike trains? We present methods to give an answer to this question.

2.4.1 Maximum Entropy Method

We have seen how to obtain from transition probabilities of a Markov chain and the unique invariant probability μ (2.17). Thus, the problem of obtaining the invariant probability can be summarized as the correct estimation of the transition probabilities. Transition probabilities $\mathbb{P}[\omega(n) \mid \omega_{n-D}^{n-1}]$ can be estimated from data. Just counting how many times the pattern $\omega(n)$ have appeared after the block ω_{n-D}^{n-1}. The problem in the context of spike train statistics is the following: Consider a network of N neurons in the most simple possible scenario where the state of the system does not dependent on the past. $\omega(n)$ is a spike pattern that can take 2^N possible values thus, N does not need to be very big to arrive a situation where it is not possible to observe all possible states by doing simulations of spiking neurons or doing real data acquisition. Take for instance 100 neurons. The spike pattern can take 1267650600228229401496703205376 possible values. This fact renders the frequentist approach unsuitable.

In section (2.2.7.1) we have exposed how, given a potential, we can obtain transition probabilities without need of sample them directly from the data. Here the problem is how to choose the appropriate potential. The method is called Maximum Entropy. An important property of the solution of the MaxEnt problem is that the solution always reads as a Gibbs distribution and as we will see, this is very advantageous.

The relevant question for the spike train statistics is: What can we tell about the statistics of these states given the fact that we observe a small part of all the possible patterns? In (Jaynes, 1957) is introduced the maximum entropy method in which you can build a probability distribution] that can be used for networks of neurons to estimate $\mathbb{P}\left[\omega(n)\right]$ and $\mathbb{P}\left[\omega_{n_1}^{n_2}\right]$ without need to count how many times each pattern appear in the spike train.

2.4.1.1 Average of monomials

As we have seen, in general, we cannot obtain samples of conditional probabilities of all the states. On the other hand, there is something relatively easy to do which is count how many times a neuron has spike or how many times two neurons have fired at the same time, or delayed in time. MaxEnt method takes this information to infer conditional probabilities. Given a spike train ω of length T, we note $\pi_\omega^{(T)}[\mathcal{O}]$ the empirical average of the observable $\mathcal{O}$. For example, the empirical firing rate of

neuron k is:

$$\pi_\omega^{(T)}[\omega_k] = \frac{1}{T}\sum_{n=0}^{T-1}\omega_k(n);$$

and the empirical probability that two neurons k, j fire at the same time is:

$$\pi_\omega^{(T)}[\omega_k\omega_j] = \frac{1}{T}\sum_{n=0}^{T-1}\omega_k(n)\omega_j(n).$$

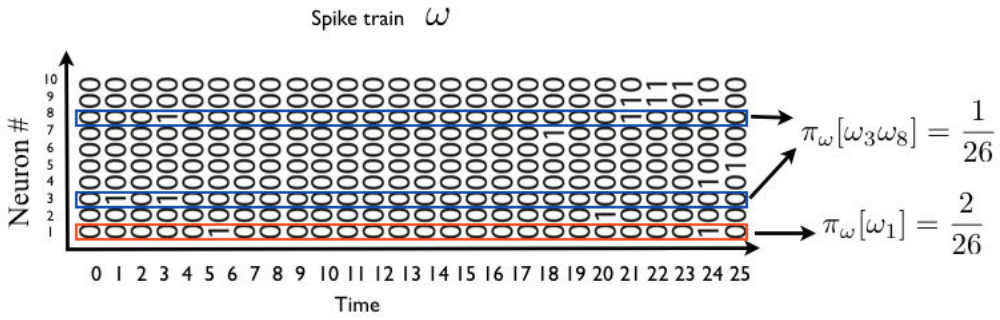


Figure 2.5: Average of monomials. In red is the sequence of spikes of neuron 1. There are two spikes (at time 5 and 24) in the 26 time steps. In blue are the sequences of spikes of neurons 3 and 8, both fire at the same time once at time 3. This is only a pedagogic example, real spike trains record up to 4 thousand neurons simultaneously for several hours.

(see fig 2.5). Considering the empirical average of monomials on the data is not enough to uniquely characterize the spike train statistics. Indeed, there are infinitely many probability distribution sharing the same empirical average of monomials. Out of those probabilities there is one particularly interesting for us, which maximizes the entropy (2.23). Note that here we are assuming ergodicity (empirical or time average is asymptotically equal to the average with respect to the invariant probability measure). In other words the Maximum entropy method in the context of spike train statistics solves the variational problem 2.2.8.2 when the empirical average of given monomials are measured from the data and constitute the constraints for the problem[5]. We give here few examples of Maximum Entropy Gibbs distributions, found in the literature.

- *Bernoulli model.* Here only firing rates of neurons are constrained. The potential has the form:

$$\mathcal{H}(\omega(0)) = \sum_{i=1}^{N}\mathsf{h}_i\omega_i(0),$$

[5]What this method do not tell us is how to choose the set of monomials (potential) to constraint the model

This is memory-less model, where transitions probabilities are given by neuron firing rates $\lambda_i = \frac{e^{\mathsf{h}_i}}{1+e^{\mathsf{h}_i}}$. The Gibbs distribution has the form:

$$\mu\,[\,\omega_m^n\,] = \prod_{l=m}^{n}\prod_{k=1}^{N} \lambda_k^{\omega_k(l)}\,(1-\lambda_k)^{1-\omega_k(l)}, \tag{2.35}$$

This is thus a product probability where neurons are independent.

- *Ising model.* This model was introduced by Schneidman et al (Schneidman et al., 2006) for retina spike trains analysis. Here, firing rates $\langle\omega_i\rangle$ and instantaneous pairwise correlations $\langle\omega_i\omega_j\rangle$ are constrained. The potential has the form:

$$\mathcal{H}(\omega(0)) = \sum_{i=1}^{N} \mathsf{h}_i\omega_i(0) + \sum_{i,j=1}^{N} \mathsf{J}_{ij}\omega_i(0)\,\omega_j(0). \tag{2.36}$$

This is memory-less model where the Gibbs distribution has the classical form (2.30).

- *Extended spatial Ising model.* A natural extension of Ising model has been proposed by Ganmor et al (Ganmor et al., 2011a), where triplets and more general synchronous spike events are considered. The potential has the form:

$$\mathcal{H}(\omega(0)) = \sum_{i=1}^{N} \mathsf{h}_i\omega_i(0) + \sum_{i,j=1}^{N} \mathsf{J}_{ij}\omega_i(0)\,\omega_j(0) + \sum_{i,j,k=1}^{N} \mathsf{J}'_{ijk}\omega_i(0)\,\omega_j(0)\,\omega_k(0) + \dots \tag{2.37}$$

This is memory-less model where the Gibbs distribution has the classical form (2.30).

- *Spatio temporal Ising model.* In (Marre et al., 2009) a spatio-temporal extension of the Ising model is considered where the potential has the form:

$$\mathcal{H}(\omega_0^1) = \sum_{i=1}^{N} \mathsf{h}_i\omega_i(0) + \sum_{i,j=1}^{N} \mathsf{J}_{ij}\omega_i(0)\,\omega_j(1). \tag{2.38}$$

In this case, spatio-temporal pairs with memory depth 1 are considered. Although the Gibbs distribution has not the form (2.30), the authors use an approximation of the exact distribution by this form, based on a detailed balance assumption. They applied this model for spike train analysis in the cat parietal cortex.

- *General Spatio temporal model.* General models of the form (2.11) have been considered in (Vasquez et al., 2012, Cessac and Palacios, 2012, Nasser et al., 2013) for the analysis of retina spike trains. A C++ implementation of methods for fitting spatio-temporal models from data is available at `http://enas.gforge.inria.fr/v3/`.

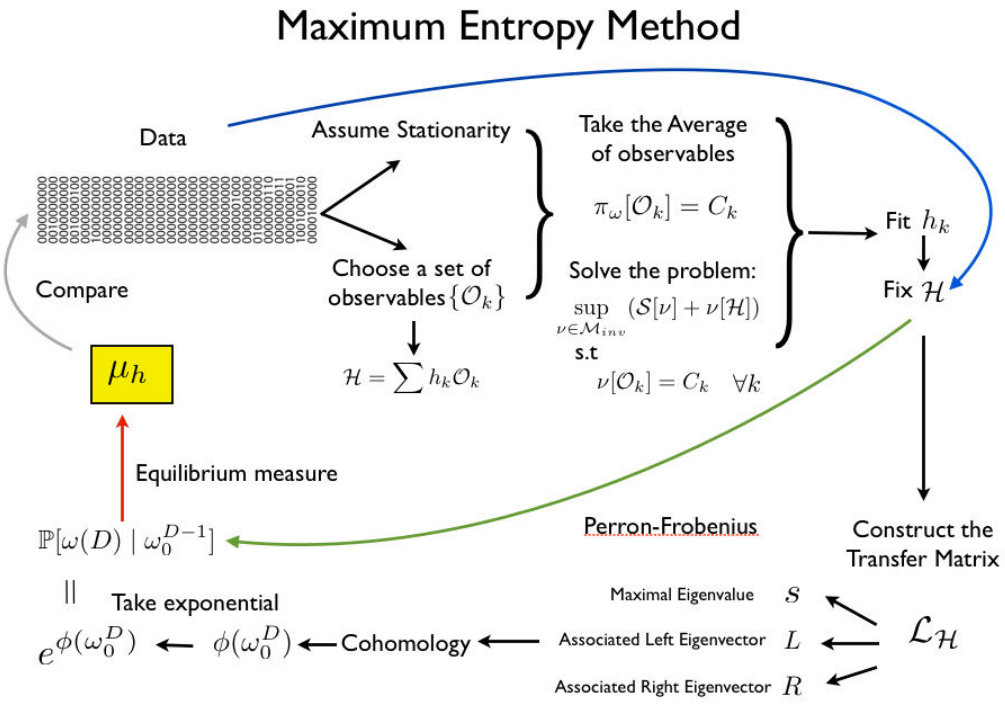


Figure 2.6: Maximum Entropy Method Diagram: The blue line indicates that from spike trains (under the hypothesis of stationarity) and after choosing a set of observables is possible to write down the MaxEnt problem and solve it using Lagrange multipliers, these multipliers are the parameters of the Maximum Entropy potential. The green line indicates that is possible to obtain transition probabilities from fixed bounded potentials. We present an example in 2.2.7.2. The red line indicates that is possible to obtain under some conditions we describe in 2.17 a unique invariant probability from conditional probabilities. This probability is compared to data to quantify how well it fits the data.

2.4.1.2 Applications of MaxEnt method to neuroscience

The maximum entropy principle has been used by several authors (Schneidman et al., 2006, Shlens et al., 2006, Tang et al., 2008, Yu et al., 2008, Ohiorhenuan et al., 2010, Tkačik et al., 2009, Ganmor et al., 2011a,b) for Multi-electrode Arrays (MEA) spike train analysis. Efficient methods have been designed to estimate the parameters of the potential, in the spatial case (Dudík et al., 2004, Broderick et al., 2007) and in the spatio-temporal case (Nasser et al., 2013).

This approach, grounded on statistical physics, attempts to find a generic model for spike statistics based on a potential of the form (2.11), where the observables and their related h parameters summarize "effective interactions" between neurons. In statistical physics language, these are parameters conjugated to the constraints, just like the inverse temperature $\beta = \frac{1}{kT}$ is conjugated to the energy, or chemical potential is conjugated with the number of particles. However, whereas inverse temperature or chemical potential have a clear interpretation thanks to the links between thermodynamics and statistical physics, the fitting parameters (Lagrange multipliers) used for spike train statistics do not benefit from such deep relations and are interpreted e.g. via loose analogies with magnetic systems. For example the Lagrange multipliers J_{ij} (2.36) conjugated to pairwise spike coincidence are interpreted as "functional interactions" (Ganmor et al., 2011a) due to their analogy with magnetic interactions in the Ising model. Likewise the parameters h_i conjugated with single spike events (whose average is the firing rate) are believed to be related with an effective stimulus received by neuron i. However, the connection between "functional" interactions J_{ij} (2.36) and real interactions (e.g synapses) in the network remains elusive, as well as the link between effective stimuli and the stimulus viewed by a neuron.

Behind this approach exists, we believe, a physicists "dream": inferring, from data analysis, the equivalent of the equation of states existing in thermodynamics; that is, summarizing the behavior of a big neuronal system by a few canonical variables (analogous e.g. to temperature, pressure, volume in a gas). To our opinion, recent exciting investigations to exhibit critical phenomena in retina spike train statistics are part of this project (Tkačik et al., 2006, 2009, 2014).

The main advantage of the MaxEnt approach is the possibility of constructing different statistical models based on a priori hypotheses on the most statistically significant events (single spikes, pairs, triplets and so on). As such, it allows to consider arbitrary forms of spatio-temporal correlations. This strength is also a weakness. Indeed, the possible forms of potentials are virtually infinite and obviously, in the setting of neuronal dynamics, one does not have the equivalent of mechanics or thermodynamics to construct the potential from general principles.

2.4.2 Generalized Linear model

We now introduce a method for statistical estimation of transition probabilities using models called Linear-Nonlinear (LN) model and Generalized Linear Model (GLM)

(Brillinger, 1988, McCullagh and Nelder, 1989, Simoncelli et al., 2004, Paninski, 2004, Truccolo et al., 2005, Pillow et al., 2005, 2008, Ahmadian et al., 2011, Pillow et al., 2011). These are methods based on the estimation of conditional probabilities. We focus here on the GLM, in particular regarding their link with Gibbs distributions. We follow the presentation found in (Ahmadian et al., 2011).

Let $x \equiv x(t)$ be a time-dependent stimulus. In response to x the network emits a spike train response ω. This response does not only depend on x, but also on the network history H. The GLM (and LN) assimilate the response ω as an inhomogeneous Poisson process: the probability that neuron k emits a spike between t and $t+dt$ is given by $\lambda_k(t|\omega)\,dt$, where $\lambda_k(t|\omega)$ is called "conditional intensity". In the GLM this function is given by:

$$\lambda_k(t|\omega) = f\left(b_k + (K_k * x)(t) + \sum_j (H_{kj} * \omega_j)(t)\right), \tag{2.39}$$

where:

- f is a non linear function (an exponential or a sigmoid);
- b_k is some constant fixing the baseline firing rate of neuron k;
- K_k is a causal, time-translation invariant, linear convolution kernel that mimics a linear receptive field of neuron k;
- $*$ is the convolution product;
- H_{kj} describes possible excitatory or inhibitory post spike effects of the j th observed neuron on the k th. As such, it depends on the past spikes, hence on ω. The diagonal components H_{kk} describe the post spike feedback of the neuron to itself, and can account for refractoriness, adaptation and burstiness depending on its shape.

2.4.2.1 Conditional independence

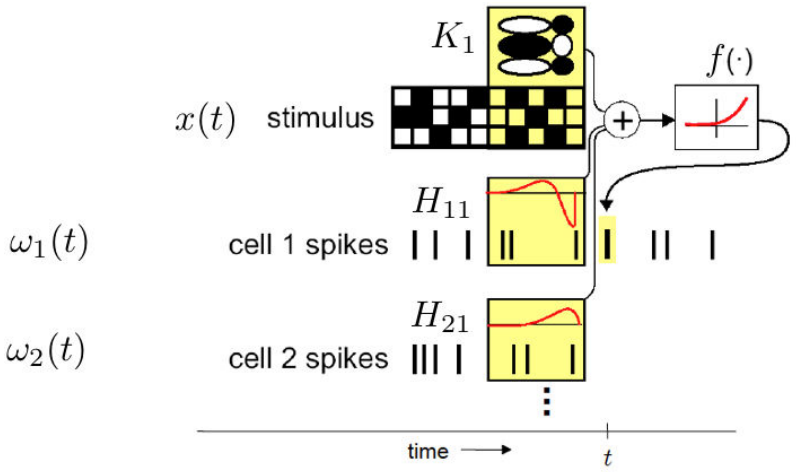
The GLM postulates that, given the history H and stimulus x, the neurons are independent (*conditional independence upon the past and stimulus*) i.e.

$$\mathbb{P}\left[\,\omega(n)\;\middle|\;\omega_{n-D}^{n-1}\,\right] = \prod_{k=1}^{N} \lambda_k(n|\omega)^{\omega_k(n)}(1-\lambda_k(n|\omega))^{1-\omega_k(n)}. \tag{2.40}$$

Where D can be $-\infty$.

2.4.2.2 Gibbs distribution from GLM

Transition probabilities are strictly positive whenever $0 < \lambda_k(n|\omega) < 1$, for all k,n. If f is e.g. a sigmoid this holds when its argument $b_i + (K_i * x)(t) + \sum_j (H_{ij} * r_j)(t)$



$$\lambda_k(t) = f\big(b_k + (K_k * x)(t) + \sum_j (H_{kj} * \omega_j)(t)\big)$$

Figure 2.7: Generalized Linear Model Diagram: The stimulus interacts with the filter H_k that mimics the receptive field of neuron k, in addition the history of spikes of the network interact with the post spike filters H_{kj}. The history can be infinite.

remains bounded in absolute value. The continuity of λ with respect to ω holds whenever f is continuous and H is continuous with respect to ω. This second condition is fulfilled in two cases:

- The memory kernel H depends on a finite past; In this case we are reduced to the framework of a Markov Chain, where we have seen in 2.2.8 the connections with Gibbs distributions.
- The memory kernel H depends on an infinite past, but the memory dependence decays sufficiently fast to ensure continuity.
- When considering the non linear function f as a sigmoid the positivity condition required by the existence and uniqueness of the Gibbs distribution is obtained under the assumption of having bounded memory kernels (physically plausible assumption). α profiles decay exponentially fast with time, so we can use the same argument used in the appendix of (Cessac, 2011b) and verify that transition probabilities match the conditions to ensure existence and uniqueness of the Gibbs distribution (1).

The Gibbs potential is:

$$\phi_n(\omega) = \sum_{k=1}^{N} \log \lambda_k(n|\omega), \tag{2.41}$$

where the dependence in ω occurs via H. It is normalized by definition.

2.4.2.3 Applications

This model has been applied in a wide variety of experimental settings (Brillinger, 1992, Chichilnisky, 2001, Theunissen et al., 2001, Brown et al., 2003, Paninski et al., 2004, Truccolo et al., 2005, Pillow et al., 2008). Efficient methods has been designed to estimate the parameters (Ahmadian et al., 2011). Note that in general, these authors consider f as an exponential function. This function has the advantage of being convex, but raises difficulties when considering the related asymptotic process.

To us, the main advantages of the GLM are:

- The transition probability is known (postulated) from the beginning and does not require the heavy normalization (2.16) imposed by potentials of the form (2.11);
- The model parameters have a neurophysiological interpretation and their number grows at most as a power law in the number of neurons.
- It has good decoding performances
- It holds for non stationary data.

Its main drawbacks are:

- It postulates an ad hoc form for the transition probability of the stochastic process;
- It uses a quite questionable assumption of conditional independence: neurons are assumed independent at time n when the past is given. On the opposite, the Maximum Entropy principle does not require this assumption.
- To us, the biophysical interpretation of the parameters H_{kj} is unclear. Do they correspond to "real" connectivity ? "functional" connectivity ?

2.4.3 Conductance based Integrate and Fire neural networks

The next section is completely devoted to analyze a particular version of this model for the estimation of transition probabilities. Comparing to previous approaches it consider a completely different point of view. It consider the neural network dynamics of membrane potentials and from this knowledge, the conditional probabilities are obtained. The main difference with the previous approaches it that it does not consider an ad-hoc form of transition probabilities, does not rely in the conditional independence assumption. The biophysical interpretation of the parameters of the model is not controversial, since are not obtained from fitting parameters. The model we analyze is an extension of the conductance based Integrate-and-Fire neuron model proposed in (Rudolph and Destexhe, 2006). Dynamics is ruled by a set of differential equations where parameters, corresponding to conductances, depend upon the action potentials emitted in the past by the neurons. In this way, the dynamical system defined here is ruled both by continuous time and discrete time dynamical variables. The relation of this model with Gibbs distributions is given in the context of infinite range potentials.

2.5 Conclusion

In this chapter we have argued that Gibbs distributions considered in a general sense may constitute generic statistical models to fit and to characterize spike trains. We have seen that Markov chains are appropriate stochastic processes to represent the spike activity of network of neurons. We have seen how Gibbs distributions appear from the MaxEnt principle. In the purely spatial case the solution of the MAxEnt problem is a Gibbs distribution in the classical sense, while in the spatio-temporal case take a different form, but is still a Gibbs distribution (in the Bowen sense). The example of GLM model suggests that such distribution could be also defined for more elaborated neural networks models (FitzHugh-Nagumo or Hodgkin- Huxley). In particular, the existence and uniqueness of a Gibbs measure holds whenever there is continuity with respect to a raster, with a sufficiently fast decay of the variations of the potential. As shown (Cessac, 2011b, Cofré and Cessac, 2013) this property is ensured when interactions between neurons decay exponentially fast. This is typically the case for chemical synapses where the α-profile (see chapter 1) decays exponentially fast with time. We left for the next chapter a complete development of

the Conductace based Integrate-and-Fire with chemical and electric synapses. The three models we describe in this chapter (MaxEnt, GLM, and Conductace based Integrate-and-Fire) introduce fall into the category of Gibbs distributions.

Chapter 3

Dynamics and spike trains statistics in neural networks with chemical and electric synapses

Overview

We present and analyze a natural and classical model of neural network with chemical and electric synapses, where the spike train statistics is described by a Gibbs distribution. The material of this chapter has been published in Cofré and Cessac (2013).

Contents

3.1 Introduction

In this chapter we develop in more detail the conductance based Integrate and Fire model introduced in chapter 2. Compared to MaxEnt approach, apart from getting rid of the stationary assumption, this approach has the advantage to characterize the spike train statistics in terms of the parameters describing the neural network. Therefore, the role in the dynamics and spike train statistics of each parameter can be tracked. Now a natural question arise: What is the appropriate neural network model to consider? It certainly depend on the questions we want to answer. As seen in Chapter 1 neural networks admit complex branching formations integrating inputs from other cells in a network, via both chemical and electrical synapses. The characterization of the spike train statistics from neural network models have been previously addressed several times in the literature . The role of chemical synapses in shaping the population spike train statistics have been already addressed in (Cessac, 2011b), but without gap junctions. In this Chapter we analyze the dynamics and spike train statistics of a conductance based integrate and fire model considering gap-junctions. Our main motivation comes from the fact that gap junctions could play a fundamental role in the retina. In this context to predict the probability of a spike pattern we need the whole past history of spikes of the entire network.

3.2 Neural Network models including gap junctions

Electric synapses transmission is mediated by gap junctions with a direct electrical communication between cells (Coombes and Zachariou, 2007), allowing faster communication than chemical synapses. Electrical coupling between cells can be found in many parts of the nervous system (Bennett and Zukin, 2004),(Connors and Long, 2004), and also outside: for example, certain cells in the heart and pancreas are connected by gap junctions (Keener and Sneyd, 1998). At the network level, electric synapses have several prominent effects such as neurons synchronization (Beierlein et al., 2000), (Galarreta and Hestrin, 1999), (Ostojic and Brunel, 2009), and the generation of neural rhythms (Hormuzdi et al., 2004), (Bennett and Zukin, 2004).

On theoretical grounds, the role of gap junctions in shaping collective dynamics has been quite less studied than the role of chemical synapses, although different models and approaches have been used to address this problem, in the context of pattern formation, using techniques such as: Poincaré map (Chow and Kopell, 2000, Gao and Holmes, 2007, Pfeuty et al., 2005), Lyapunov functions (Medvedev et al., 2003), mean-field approach and Fokker-Planck equation (Ostojic and Brunel, 2009), variance analysis (Medvedev, 2009), and phase plane analysis (Coombes, 2008, Coombes and Zachariou, 2007). The effects of gap junctions on spike trains statistics is even less known.

The goal of this chapter is to push one step forward the mathematical analysis of the join effects of chemical and electric synapses on neurons dynamics and spike statistics. The main advantage of this type of model, rendering the mathematical analysis tractable, is that the sub-threshold dynamics of membrane potential is

described by a linear system of stochastic differential equations (SDE). This system is nevertheless highly non trivial, as we show, since it is non autonomous, non homogeneous, with, additionally, a flow depending on the whole (spike) history via chemical synapses. Moreover, electric synapses introduce another mechanism of history dependence, where the membrane potential of a neuron depends on its past values, even those which are anterior to the last firing and reset of that neuron. The flow of this linear system of stochastic differential equations can be explicitly written.

From this, one can compute a family of transition probabilities, characterizing the probability of a spike pattern at a given time, given the past. These transition probabilities define a Gibbs distribution characterizing spike train statistics. The potential of this Gibbs distribution can be approached by an explicit form, as we show. We have therefore an explicit characterization of spike statistics in a model including chemical and electric synapses.

This result has several implications in the realm of biological spike trains analysis. Especially, as we show, the Gibbs potential of our model encompasses existing models for spike trains statistics analysis such as MaxEnt models (Schneidman et al., 2006, Tkačik et al., 2009, Shlens et al., 2006, 2009, Ohiorhenuan et al., 2010, Ganmor et al., 2011a,b, Vasquez et al., 2012) and Generalized-Linear Models (GLM) (Pillow et al., 2005, 2008, Ahmadian et al., 2011, Pillow et al., 2011, Macke et al., 2011). Moreover, our formalism affords the study of non stationary dynamics, while stationarity is a major assumption when using maximum entropy models. Additionally, as we show, gap junctions introduce major correlations effects ruining the hope of having conditionally-upon-the-past independent neurons, a central hypothesis in GLM. Finally, three types of effects are responsible for neuron correlations (pairwise and higher order): shared stimulus, chemical couplings, and electric couplings. The two last types of correlations persists even when the stimulus is switched-off.

3.3 Model definition

The Integrate-and-Fire model remains one of the most ubiquitous model for simulating and analyzing the dynamics of neuronal circuits. Despite its simplified nature, it captures some of the essential features of neuronal dynamics (see (Lindner and Schimansky-Geier, 2002, Lindner et al., 2004, Burkitt, 2006a,b, Lindner, 2009) for a review). In this chapter we consider an extension of the conductance based Integrate-and-Fire neuron model proposed in (Rudolph and Destexhe, 2006). The model-definition follows the presentation given in (Cessac and Viéville, 2008, Cessac, 2011b). Dynamics is ruled by a set of stochastic differential equations where parameters, corresponding to chemical conductances, depend upon a sequence of discrete variables summarizing the action potentials emitted in the past by the neurons. In this way, the dynamical system defined here is ruled both by continuous time and

discrete time dynamical variables. Let us first introduce the discrete time variables.

3.3.1 Membrane potential dynamics

Neurons are considered here as points, with neither spatial extension nor biophysical structure (axon, soma, dendrites). We note $V_k(t)$ the membrane potential of neuron $k = 1, \ldots, N$, at time t. Denote $V(t)$ the vector with entries $V_k(t)$. The continuous-time dynamics of $V(t)$ is defined as follows. Fix a real variable $\theta > 0$ called "firing threshold". For a fixed time t, we have two possibilities:

1. Either $V_k(t) < \theta$, $\forall k = 1, \ldots, N$. This corresponds to *sub-threshold dynamics*.
2. Or, $\exists k$, $V_k(t) \geq \theta$. Then, we speak of *firing dynamics*.

3.3.1.1 Sub-threshold dynamics

Let us remember what has been already exposed in chapter 2. The sub-threshold variation of the membrane potential of neuron k at time t is given by:

$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L) - \sum_j g_{kj}(t,\omega)(V_k - E_j) - \sum_j \overline{g_{kj}}\,(V_k - V_j) + I_k(t). \quad (3.1)$$

- C_k is the membrane capacity of neuron k.
- The term $I_k(t)$ is a current given by:

$$I_k(t) = i_k^{(ext)}(t) + \sigma_B \xi_k(t), \quad (3.2)$$

  where $i_k^{(ext)}(t)$ is a deterministic external current ("stimulus"). The noise term $\xi_k(t)$ mimics: (i) the random variation in the ionic flux of charges crossing the membrane per unit time at the post synaptic button, upon opening of ionic channels due to the binding of neurotransmitter, (ii) the fluctuations in the current resulting from the large number of opening and closing of ion channels (Schwalger et al., 2010); (iii) noise coming from electrical synapses. It is common to model this noise by a Wiener white noise (diffusion approximation). We use this modelling choice in the chapter. We note $\sigma_B > 0$ the amplitude (mean-square deviation) of $\xi_k(t)$.

- $g_{L,k}$ is a leak conductance, $E_L < 0$ is the leak reversal potential.
- $g_{kj}(t,\omega)$ corresponds to the conductance of the chemical synapses from the pre-synaptic neuron j to the post-synaptic neuron k, while E_j is the reversal potential associated with the neurotransmitter emitted by neuron j. In this model, the conductance $g_{kj}(t,\omega)$ at time t depends upon the spikes emitted by the pre-synaptic neuron j. The description of this dependence is made explicit in section 3.3.1.2.

- $\overline{g_{kj}}$ is the electric conductance (gap junctions) between two different neurons j and k. Although, the gap junction strength is biophysically non static (Pan et al., 2010, Hu et al., 2010), we model it here as a simple ohmic conductance, tending to equalize membrane potentials of the neurons they connect. As a consequence, $\overline{g_{kj}} = \overline{g_{jk}} \geq 0$.

- For chemical as well as electric contacts terms of eq. (3.1) the sums $\sum_j$ hold on all neurons, but we afford some conductances $g_{kj}(t,\omega)$ or $\overline{g_{kj}}$ to be 0, corresponding to no connection (chemical or electrical), between neuron j and neuron k. In this way, we can define chemical and electric network topologies. The mathematical results obtained in section 3.4 hold for any such network.

3.3.1.2 Update of chemical synapses conductances

Upon firing of the pre-synaptic neuron j at (discrete) time $t_j^{(r)}(\omega)$, the membrane conductance $g_{kj}(t)$ of the post-synaptic neuron k is modified as (Rudolph and Destexhe, 2006):

$$g_{kj}(t) = g_{kj}(t_j^{(r)}(\omega)) + G_{kj}\alpha_{kj}(t - t_j^{(r)}(\omega)), \quad t > t_j^{(r)}(\omega), \tag{3.3}$$

where $G_{kj} \geq 0$ characterizes the maximal amplitude of the conductance during a post-synaptic potential.

The function α_{kj} (called "alpha" profile (Destexhe et al., 1998)) mimics the time course of the chemical synaptic conductance upon the occurrence of the spike. We assume that the alpha profiles have the form:

$$\alpha_{kj}(t) = h(t)e^{-\frac{t}{\tau_{kj}}}H(t), \tag{3.4}$$

with $\alpha(0) = 0$, where $h(t)$ is a polynomial function continuous at 0, typically $h(t) = \frac{t}{\tau_{kj}}$, where τ_{kj} is the characteristic decay time. $H(t)$ is the Heaviside function.

As a consequence, upon the arrival of spikes at times $t_j^{(r)}(\omega)$ in the time interval $[s,t]$, eq. (3.3) becomes:

$$g_{kj}(t) = g_{kj}(s) + G_{kj} \sum_{r;s\leq t_j^{(r)}(\omega)<t} \alpha_{kj}\left(t - t_j^{(r)}(\omega)\right).$$

If we assume that this relation extends to $s \to -\infty$ and if we set $\lim_{s\to-\infty} g_{kj}(s) = 0$ (see (Cessac, 2011b) for a justification), we finally obtain:

$$g_{kj}(t,\omega) = G_{kj} \sum_{r;t_j^{(r)}(\omega)<t} \alpha_{kj}\left(t - t_j^{(r)}(\omega)\right),$$

The notation (t,ω) makes explicit the dependence of the conductance upon the (past) firing activity of the pre-synaptic neuron j. This dependence decays exponentially fast thanks to the exponential decay of α_{kj}.

3.3.1.3 Firing dynamics

If, at time t, some neuron k reaches its threshold θ, $V_k(t) = \theta$, then this neuron elicits a spike. In neurophysiology, the spike emission is a complex nonlinear mechanism involving additional physical variables (probability of opening/closing of ionic channels), as described by the Hodgkin-Huxley equations (Hodgkin and Huxley, 1952). The time course of the spike has a typical shape (depolarization / repolarization / refractory period) with a finite duration. One important simplification in Integrate-and-Fire models is to describe neuron dynamics in terms of the membrane potential only. As a consequence, the spike shape has to be simplified. In the simplest Integrate-and-Fire models (Gerstner and W.Kistler, 2002, Ermentrout and Terman, 2010) a spike is registered at time t whenever $V_k(t) = \theta$, whereas the membrane potential is instantaneously reset to 0. Moreover, the refractory period and transmission propagation delays are set to 0. Beyond the bio-physical fact that reset, refractory period and transmission propagation are not instantaneous, this modelling leads to severe mathematical problems and logical inconsistencies, such as the possibility of having uncountably many spikes within a finite time interval, or situations where the state of a neuron cannot be defined (see (Cessac and Viéville, 2008, Cessac, 2010) for a discussion).

To avoid those problems we consider that spikes emitted by a given neuron are separated by a minimal time scale $\tau_{sep} > 0$. Additionally, to conciliate the continuous time dynamics of membrane potentials and the discrete time dynamics of spikes we define the spike and reset as follows.

1. The neuron membrane potential V_k is reset to 0 at the *next integer time* (in δ units) after t, namely $[t]+1$. Between t and $[t]+1$ the membrane potential keeps on evolving according to (3.1). The main reason for this modelling choice is that it makes the mathematical analysis simpler.

2. A spike is registered at time $[t]+1$. This allows us to represent spike trains as events on a discrete time grid. It has the drawback of artificially synchronizing spikes coming from different neurons, in the deterministic case (Cessac and Viéville, 2008, Kirst and Timme, 2009). However, the presence of noise in membrane potential dynamics eliminates this synchronization effect.

3. We consider that $\tau_{sep} > 0$ is a multiple of δ (thus an integer).

4. Between $[t]+1$ and $[t]+\tau_{sep}$ the membrane potential V_k is maintained to 0 (refractory period). From time $[t]+\tau_{sep}$ on, V_k evolves according to (3.1) until the next spike.

5. When the spike occurs (at time $[t]+1$), the spike train ω as well conductances $g_{kj}(t,\omega)$ are updated.

Figure 3.1 illustrates these modelling choices.

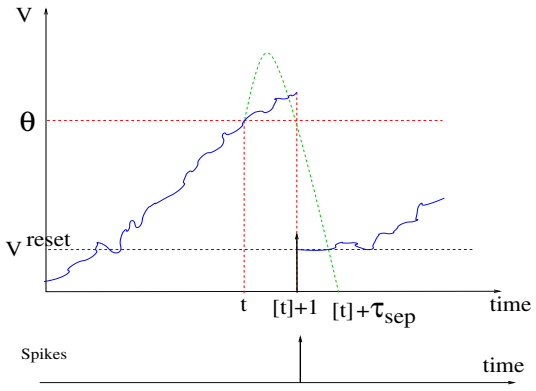


Figure 3.1: Top: Continuous time course of the membrane potential in our model. The green dashed curve illustrates the shape of a real spike, but what we model is the blue curve. Bottom: A spike is registered at integer time $[t]+1$.

3.3.1.4 Matrix-Vector representation of subthreshold dynamics

The sub-threshold dynamics can be rewritten in the form of a stochastic linear non-autonomous and non-homogeneous differential equation:

$$C\frac{dV}{dt} + \left[\, G(t,\omega) - \overline{G}\,\right] V = I(t,\omega), \tag{3.5}$$

where C is a diagonal matrix which contains the capacity of each neuron. For simplicity we assume from now on that all neurons have the same capacity c so that $C = c\mathcal{I}_N$ where $\mathcal{I}_N$ is the identity matrix of dimension N; V is the vector of membrane potentials; $G(t,\omega)$ is a diagonal matrix:

$$G_{kl}(t,\omega) = \left[\, g_{L,k} + \sum_{j=1}^{N} g_{kj}(t,\omega) \right] \delta_{kl} \stackrel{\text{def}}{=} g_k(t,\omega)\delta_{kl}. \tag{3.6}$$

$\overline{G}$ is a matrix with entries $\overline{g_{kj}}$ for $k \neq j$ and $-\sum_j \overline{g_{kj}}$ for $k = j$. It is therefore symmetric and the non diagonal part of this matrix specifies the connection topology of electric synapses in the network. Finally, $I(t,\omega)$ is the vector of currents that can be separated in 3 components.

$$I(t,\omega) = I^{(cs)}(t,\omega) + I^{(ext)}(t) + I^{(B)}(t), \tag{3.7}$$

with:

$$I_k^{(cs)}(t,\omega) = \sum_j W_{kj}\alpha_{kj}(t,\omega), \tag{3.8}$$

where:

$$W_{kj} \stackrel{\text{def}}{=} G_{kj}E_j, \tag{3.9}$$

is the synaptic weight from neuron k to neuron j;

$$I_k^{(ext)}(t) = g_{L,k}E_L + i_k^{(ext)}(t); \tag{3.10}$$

and

$$I_k^{(B)}(t) = \sigma_B \xi_k(t).$$

The first term corresponds to the current received by neuron k from chemical synapses (cs), the second term represents the external current (ext) and the last term (B) is the the noise part of the current.

Define:

$$\Phi(t,\omega) = C^{-1}\left(\overline{G} - G(t,\omega)\right), \tag{3.11}$$

which is a symmetric matrix. All entries of $\Phi(t,\omega)$ are bounded and continuous in time (alpha profiles are continuous in time and the components of $\Phi(t,\omega)$ are composition of continuous functions).

Defining:

$$f(t,\omega) = C^{-1}I^{(cs)}(t,\omega) + C^{-1}I^{(ext)}(t), \tag{3.12}$$

and using the decomposition of currents (3.7) the system (3.5) can be expressed as a system of coupled Stochastic Differential Equations (SDE) of Ornstein-Uhlembeck type (O-U) in $\mathbb{R}^N$

$$\frac{dV}{dt} = \underbrace{C^{-1}(\overline{G} - G(t,\omega))}_{\Phi(t,\omega)} V + \underbrace{C^{-1}I^{(cs)}(t,\omega) + C^{-1}I^{(ext)}(t)}_{f(t,\omega)} + C^{-1}I^{(B)}(t),$$

$$\Leftrightarrow$$

$$\begin{cases} dV &= (\Phi(t,\omega)V + f(t,\omega))dt + \frac{\sigma_B}{c}\mathcal{I}_N dW(t), \\ V(t_0) &= v, \end{cases} \tag{3.13}$$

where v is the initial condition at time t_0. Here $f(t,\omega)$ is thus a non random, measurable and locally bounded function of t; $\frac{\sigma_B}{c}$ is a constant; and $W(t)$ is a standard N-dimensional Brownian motion independent of $v \in \mathbb{R}^N$.

3.3.1.5 Remarks

- Although (5.6) is a linear system, it has a rather complex structure, due to the ω-dependence of $\Phi(t,\omega)$, $f(t,\omega)$. Indeed, these functions integrate the past spike-activity of the network from the initial time t_0 to time t. Thus, the membrane potentials at time t are determined by the past spikes-activity which, in turn, is determined by the trajectory $V(s), s \in [t_0, t]$ of the membrane potentials: the evolution of the network depends on its whole history via ω. If ω is given, the integration of (5.6) generates a flow which is explicitly computed in section 3.4.1.4.

- The discrete structure of the spike trains set Ω (discrete events and discrete times) induces a partition on the set $\mathcal{P}$ of trajectories of V. A trajectory

$\mathcal{V} \in \mathcal{P}$ belongs to the partition element $\mathcal{P}_\omega$, associated with the spike train ω, if and only if:

$$\forall k = 1, \dots N, \forall n \in \mathbb{Z}, \quad \begin{cases} \omega_k(n) = 0 \Leftrightarrow \max_{t\in]n-1,n]} V_k(t) < \theta; \\ \\ \omega_k(n) = 1 \Leftrightarrow \max_{t\in]n-1,n]} V_k(t) \geq \theta; \end{cases} \tag{3.14}$$

This constitutes *compatibility conditions* between spike trains and membrane potential trajectories.

In the absence of noise ($\sigma_B = 0$) some partition elements $\mathcal{P}_\omega$ (depending on model-parameters) are not visited by any trajectory. It has been shown in some variants of (5.6) considered in (Cessac, 2008, Cessac and Viéville, 2008), that the set of non empty $\mathcal{P}_\omega$s is finite, leading to specific, although quite rich, periodic orbit structure of the attractors. In the presence of noise, all $\mathcal{P}_\omega$s are visited by any trajectory with a positive probability.

- The Wiener process on noise trajectories induces a probability measure μ on spike trains, described in section 3.5, characterizing spike train statistics.

3.3.2 Particular cases of the model with electric synapses

The model (3.1) is quite general as it considers chemical and electric synapses, with a spike-history dependence. We don't know any analysis of this model in its most general form in the literature. However, upon simplifications it reduces to several models studied in the literature. As an illustration of the methods developed here, we refer to some of these examples all along the chapter Instead of presenting them in a section following the main results of the chapter, we introduce them step by step, in order to be as didactic as possible. As a conclusion of this section 3.3, let us here briefly present those examples.

3.3.2.1 No electric synapse

As mentioned in the introduction the model without electric synapse has been studied in (Cessac, 2011b). The sub-threshold dynamics reads:

$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L) - \sum_j g_{kj}(t,\omega)(V_k - E_j) + I_k(t).$$

In this case

$$\Phi(t,\omega) = C^{-1}G(t,\omega),$$

while $f(t,\omega)$ takes the same form (3.12) as in the general model.

3.3.2.2 Simple chemical conductances

By "simple" we mean chemical conductances that do not depend explicitly on spike history and a conductance matrix that takes the form $G(t,\omega) \equiv G(t) = \kappa(t)\mathcal{I}_N$,

where $\kappa(t)$ is a real function independent of ω. Note that in this case $G(t)$ and $\overline{G}$ commute, while this not the case if $G(t)$ is a general matrix. Classical examples are $\kappa(t) = 0$ or $\kappa(t) = g_L$, the leak conductance.

In this case the sub-threshold equation (3.5) reads:

$$C\frac{dV}{dt} = -\kappa(t)\,\mathcal{I}_N\,V + \overline{G}V + I^{(cs)}(t,\omega) + I^{(ext)}(t) + \sigma_B dW(t). \tag{3.15}$$

Let us consider two examples.

The Medvedev model

In (Medvedev, 2009), this author considers Integrate-and-Fire neurons subject to a randomly perturbed synaptic input and an electrically coupled ensemble of such neurons. He shows that in electrically coupled groups, neurons are less affected by noise than when in isolation. The magnitude of this effect depends on network size, topology of electrical coupling and the electrical coupling strength. Via a direct computation of the variances of the stochastic processes generated by the electrically coupled model, he shows that variances can be made arbitrarily small by increasing the number of neurons in the network and the strength of electrical coupling. In this way the organization of neurons in electrically coupled groups at the network level, may be involved in the filtering noise and therefore may play an important role in the information processing mechanisms.

The model is, in Medvedev notations:

$$\varepsilon\frac{dV}{dt} = -V + D(g)V + P(t) + \sqrt{\varepsilon}\sigma\mathcal{I}_N dW(t), \tag{3.16}$$

where $D(g)$ is a matrix whose components are functions of the maximal electric conductance over all neurons $g = \max_{ij} g_{ij}$ and $P(t)$ is the external current. This is as a particular case of (3.15) taking $c = \varepsilon$; $\kappa(t) = 1$; $I^{(ext)}(t) = P(t)$; $\overline{G} = D(g)$, and $\sigma_B = \sqrt{\varepsilon}\sigma$.

Upon reaching the threshold, the system generates an action potential, followed by the refractory period. After that, the evolution of the system is again governed by equation (3.16) until the next action potential. In this example the (SDE) (5.6) can be written using $\Phi = C^{-1}(\overline{G} - \mathcal{I}_N)$; $f(t) = C^{-1}(P(t))$ and diffusion term $\sqrt{\varepsilon}\sigma C^{-1}dW(t)$.

The Ostojic-Brunel-Hakim model

In (Ostojic and Brunel, 2009) these authors investigate how synchrony can be generated or induced in networks of electrically coupled Integrate-and-Fire neurons subject to noisy and heterogeneous inputs. Using a mean field approach and a Fokker-Planck analysis they find analytically the bifurcation between synchronous and asynchronous states for different input and coupling parameters. Two regimes are considered in this analysis: when the input parameters are homogeneous (equal

for all neurons) and heterogeneous (different, but with equal distribution). The model is characterized by a sub-threshold dynamics:

$$\begin{aligned} C_k \frac{dV_k(t)}{dt} &= -g_m V_k(t) + \frac{\beta c_m}{N} \sum_j \sum_{t_j^{(r)}<t} \delta(t - t_j^{(r)}(\omega)) \\ &\quad + \frac{\gamma}{N} \sum_j (V_j(t) - V_k(t)) + i_k^{ext}(t) + \sigma_B \xi_k(t). \end{aligned}$$

The correspondence between the parameters introduced by the authors (with their notations) and our model is:

$$G_{kj} = \frac{\beta}{N}; \quad \alpha_{kj}(x) = \delta(x); \quad \overline{g_{kj}} = \frac{\gamma}{N}; \quad g_{L,k} = g_m; \quad E_L = 0; \quad E^+ = 0; \quad E^- = 0.$$

The input parameters are $(\mu_{ext}, \sigma) = (\frac{i^{ext}}{g_m+\gamma}, \frac{\sigma_B}{\sqrt{g_m+\gamma}})$ and the coupling parameters $(g_c, \beta) = (\frac{\gamma_c}{g_m+\gamma}, \beta)$.

If we set $C_k = 1$ for all ks this system can be expressed in a matrix form:

$$\frac{dV}{dt} = \Phi V(t,\omega) + I(t,\omega)$$

where Φ is the symmetric matrix

$$\Phi_{kj} = (-g_m - \gamma)\delta_{kj} + \frac{\gamma}{N}.$$

Note that in the limit $N \to +\infty$ (the "mean-field" limit considered in (Ostojic and Brunel, 2009), $\Phi_{kj} \to (-g_m - \gamma)\delta_{kj}$, a diagonal matrix.

Additionally:

$$I(t,\omega) = \begin{pmatrix} \frac{\beta c_m}{N} \sum_j \sum_{t_j^{(r)}<t} \delta(t - t_j^{(r)}(\omega)) + i_1^{(ext)}(t,\omega) + \sigma_B \xi_1(t) \\ \frac{\beta c_m}{N} \sum_j \sum_{t_j^{(r)}<t} \delta(t - t_j^{(r)}(\omega)) + i_2^{(ext)}(t,\omega) + \sigma_B \xi_2(t) \\ \vdots \\ \frac{\beta c_m}{N} \sum_j \sum_{t_j^{(r)}<t} \delta(t - t_j^{(r)}(\omega)) + i_N^{(ext)}(t,\omega) + \sigma_B \xi_N(t) \end{pmatrix}.$$

In this example the (SDE) (5.6) can be written using $\Phi = (-g_m - \gamma)\delta_{kj} + \frac{\gamma}{N}$ and $f(t,\omega) = I^{det}(t,\omega)$.

In this model the reset is instantaneous and the refractory period is not considered.

Besides, the conductance term $G(t,\omega)$ of our model reduces to the leak γ, while the spike dependent history does not appear in the conductance, but in the current.

3.4 Solutions of the stochastic differential equation

We now derive several mathematical results allowing the integration of the SDE (5.6) and the consideration of firing dynamics.

3.4.1 Flow in the sub-threshold regime

We consider first the integration of (5.6) on a time interval $[t_0, t]$ in the sub-threshold regime, $V_k(s) < \theta$, $k = 1, \ldots, N$, $s \in [t_0, t]$. We assume that ω is fixed. Note that, necessary, ω obeys $\omega_k(n) = 0$, $k = 1, \ldots, N$, $t_0 < n < t$ (cf the compatibility conditions (3.14)). Nevertheless, we don't make any assumption on ω before time t_0, that is, we can have any spike history prior to t_0.

3.4.1.1 General form of the flow

We start by solving the associated homogeneous Cauchy problem

$$\left\{ \begin{array}{lcl} \frac{dV(t,\omega)}{dt} & = & \Phi(t,\omega)V(t,\omega), \\ V(t_0) & = & v, \end{array} \right. \tag{3.17}$$

The following theorem is standard and can be found e.g. in (Brockett, 1970).

Theorem 4 *If $\Phi(t,\omega)$ is a square matrix whose elements are bounded, the sequence of matrices $M_k(t_0,t,\omega)$ defined recursively by:*

$$M_0(t_0,t,\omega) = \mathcal{I}_N$$

$$M_k(t_0,t,\omega) = \mathcal{I}_N + \int_{t_0}^{t} \Phi(s,\omega) M_{k-1}(s,t) ds, \quad t \leq t_1,$$

converges uniformly in $[t_0, t_1]$. We note:

$$\Gamma(t_0,t,\omega) \stackrel{\text{def}}{=} \lim_{k\to\infty} M_k(t_0,t,\omega) \tag{3.18}$$

the limit function called "flow".

3.4.1.2 Exponential flow

If $\Phi(t,\omega)$ and $\Phi(s,\omega)$ commute, $\forall s, t$, then the flow takes the form of an exponential:

$$\Gamma(t_0,t,\omega) = \sum_{k=0}^{\infty} \frac{1}{k!} \Big(\int_{t_0}^{t} \Phi(s,\omega) ds \Big)^k = e^{\int_{t_0}^{t} \Phi(s,\omega) ds}.$$

From (5.23) the commutation condition reads $C^{-1}\left[\overline{G} - G(t,\omega)\right] C^{-1}\left[\overline{G} - G(s,\omega)\right] = C^{-1}\left[\overline{G} - G(s,\omega)\right] C^{-1}\left[\overline{G} - G(t,\omega)\right]$. Since we have assumed that all neurons have the same capacity, the commutation condition reduces to $\left[\overline{G} - G(t,\omega)\right]\left[\overline{G} - G(s,\omega)\right] = \left[\overline{G} - G(s,\omega)\right]\left[\overline{G} - G(t,\omega)\right]$, i.e. $\overline{G}$ and $G(t,\omega)$ commute for all t. Since $G(t,\omega)$ is diagonal, the commutation condition reads $g_i(t,\omega)\overline{G}_{ij} = \overline{G}_{ij} g_j(t,\omega)$, $\forall i, j$. Therefore, here are the *only cases* where the commutation property holds:

1. $\overline{G} = 0$;
2. $G(t,\omega) = \kappa(t,\omega)\mathcal{I}_N$ where $\kappa(t,\omega))$ is a real function.

These cases correspond respectively to the following examples.

No electric synapse

In this case $\Phi(t,\omega) = -\frac{1}{c}\,G(t,\omega)$ is a diagonal matrix. Thus, the flow takes the exponential form:

$$\Gamma(t_0,t,\omega) = e^{\int_{t_0}^{t}\Phi(s,\omega)ds} = e^{-\frac{1}{c}\int_{t_0}^{t}G(s,\omega)ds}, \tag{3.19}$$

which is also a diagonal matrix.

Simple chemical conductances

In this case $\Phi(t,\omega) = \frac{1}{c}\overline{G} - \frac{\kappa(t,\omega)}{c}\mathcal{I}_N$ and the flow takes an exponential form:

$$\Gamma(t_0,t,\omega) = e^{-\frac{1}{c}\int_{t_0}^{t}\kappa(s,\omega)\,ds}e^{\frac{1}{c}\overline{G}(t-t_0)}. \tag{3.20}$$

It is not diagonal in the canonical basis, but it can be diagonalized by an orthogonal variable change.

3.4.1.3 General form for the flow

In the general case, namely *in any model taking into account simultaneously chemical and electric synapses with a generic form*, $\overline{G}$ and $G(t,\omega)$ do not commute, and the flow (3.18) does not read as an exponential but as a general Dyson series:

$$\Gamma(t_0,t,\omega) = \mathcal{I}_N + \sum_{n=1}^{+\infty}\int_{t_0}^{t}\int_{t_0}^{s_1}\dots\int_{t_0}^{s_{n-1}}\Delta(s_1,\omega)\dots\Delta(s_n,\omega)\,ds_n\dots ds_1.$$

Setting $B = C^{-1}\overline{G}$ and $A(t,\omega) = -C^{-1}G(t,\omega)$, this equation reads:

$$\Gamma(t_0,t,\omega) = \mathcal{I}_N + \sum_{n=1}^{+\infty}\int_{t_0}^{t}\cdots\int_{t_0}^{s_{n-1}}\prod_{k=1}^{n}(B + A(s_k,\omega))d_{s_1}\cdots d_{s_n},,$$

where $\prod$ denotes the matrix product, hence ordered. Finally, the flow can be written:

$$\Gamma(t_0,t,\omega) = \mathcal{I}_N + \sum_{n=1}^{+\infty}\sum_{\substack{X_1=(B,A(s_1,\omega))\\ X_2=(B,A(s_2,\omega))\\ \dots\\ X_n=(B,A(s_n,\omega))}}\int_{t_0}^{t}\cdots\int_{t_0}^{s_{n-1}}\prod_{k=1}^{n}X_k\,d_{s_1}\cdots d_{s_n}. \tag{3.21}$$

This form is quite cumbersome when $A(t,\omega), B$ do not commute. It has nevertheless the following property.

3.4.1.4 Exponentially bounded flow

Definition: An exponentially bounded flow is a two parameter (t_0,t) family $\{\Gamma(t_0,t,\omega)\}_{t\le t_0}$ of flows such that, $\forall\omega\in\Omega$:

1. $\Gamma(t_0,t_0,\omega) = \mathcal{I}_N$ and $\Gamma(t_0,t,\omega)\Gamma(t,s,\omega) = \Gamma(t_0,s,\omega)$ whenever $t_0\le t\le s$;

2. For each $v \in \mathbb{R}^N$ and $\omega \in \Omega$, $(t_0, t) \to \Gamma(t_0, t, \omega)v$ is continuous for $t_0 \leq t$;

3. There is $M > 0$ and $m > 0$ such that :

$$||\Gamma(s,t,\omega)|| \leq Me^{-m(t-s)}, s \leq t. \tag{3.22}$$

Recall that a strong solution of the SDE (5.6) is a stochastic process for which the paths are right-continuous with left limits everywhere with probability one, adapted to the filtration generated by $W(t)$. From (Wooster, 2011) we have the following theorem:

Theorem 5 *If (3.18) converges to an exponentially bounded flow $\Gamma(t_0, t, \omega)$, there is a unique strong solution for $t \geq t_0$ given by:*

$$V(t_0,t,\omega) = \Gamma(t_0,t,\omega)v + \int_{t_0}^{t} \Gamma(s,t,\omega)f(s,\omega)ds + \frac{\sigma_B}{c}\int_{t_0}^{t} \Gamma(s,t,\omega)dW(s). \tag{3.23}$$

Thus, given an initial condition v at a time t_0 and a noise trajectory, this equation gives the membrane potential vector at time t by integration of the flow, provided $\max_{k=1\ldots N} \max_{u\in[t_0,t]} V_k(u) < \theta$ (sub-threshold dynamics). This is a classical form although Γ has a complex structure (3.21) and a non trivial dependence in the spike train history.

Let us now show that (3.21) converges to an exponentially bounded flow. If $\bar{G} = 0$ then $\Gamma(s,t,\omega)$ given by (3.19), is diagonal and exponentially bounded. In this case $\Gamma(s,t,\omega) = diag(e^{-\frac{1}{c}\int_s^t g_k(u,\omega)du})$ where g_k is given by (3.6). Consequently, setting $g_L = \inf_{\omega,u,k} g_k(u,\omega)$, the smallest conductance value attained when no neuron fires ever, in the absence of gap junctions, we have:

$$||\Gamma(s,t,\omega)|| \leq e^{-\frac{g_L}{c}(t-s)}. \tag{3.24}$$

This is therefore an exponentially bounded flow.

When $\bar{G} \neq 0$ we use the following perturbational result (for details see corollary 2.2.3 in (Gil, 2005)). Set $\hat{\Phi}(t,\omega) = -C^{-1}G(t,\omega)$. Considering $B = C^{-1}(\bar{G})$, as a (not necessarily small) perturbation of $\hat{\Phi}(t,\omega)$, we have:

Theorem 6 *Let the flow $\hat{\Gamma}(s,t,\omega)$ be the exponentially bounded flow of equation (3.24), obtained when $\hat{\Phi}(t,\omega) = -C^{-1}G(t,\omega)$. Then the flow $\Gamma(s,t,\omega)$ of equation (3.17), when $\bar{G} \neq 0$ and $\Phi(t,\omega) = \hat{\Phi}(t,\omega) + B$ obeys the inequality*

$$\|\Gamma(s,t,\omega)|| \leq e^{-\frac{g_L}{c}(t-s)}e^{\int_s^t ||\hat{\Phi}(r,\omega)-\Phi(r,\omega)\|dr}.$$

Moreover,

$$\|\Gamma(s,t,\omega) - \hat{\Gamma}(s,t,\omega)|| \leq e^{-\frac{g_L}{c}(t-s)}(e^{\int_s^t ||\hat{\Phi}(r,\omega)-\Phi(r,\omega)\|dr} - 1). \tag{3.25}$$

Using (3.24) the flow associated to equation (5.6) satisfies the following inequality with $||B|| = b$:

$$||\Gamma(s,t,\omega)|| \leq e^{-\frac{g_L}{c}(t-s)} e^{b(t-s)} = e^{(b-\frac{g_L}{c})(t-s)} \tag{3.26}$$

Therefore we can ensure exponentially bounded flow for equation (5.6) when $b < \frac{g_L}{c}$. Since $\bar{G}$ is symmetric and $C = c\mathcal{I}_N$ the norm $b = \|C^{-1}\bar{G}\|$ is equal to $\frac{\sigma_1}{c}$, where σ_1 is the largest eigenvalue of $\bar{G}$ and has the physical dimension of a conductance. So, theorem 6 provides a (sufficient) condition for the existence of a strong solution, given by:

$$\sigma_1 < g_L. \tag{3.27}$$

The largest eigenvalue of $\bar{G}$ has to be smaller than the leak conductance.

3.4.1.5 Remarks

- As stated e.g. in (Galarreta and Hestrin, 2001) the typical electrical conductance values are of order 1 nano-Siemens, while according e.g. to (Wohrer and Kornprobst, 2009) the leak conductance of retinal ganglion cells is of order 50 micro-Siemens. Therefore, the condition (3.27) is compatible with the biophysical values of conductances in the retina.

- Looking at the series (3.21) one may think that the exponentially bounded flow is ensured whenever $\Phi(t)$ have a negative spectrum. This property *is in general not* determined by the eigenvalues of $\Phi(t)$ in the non-autonomous case. There are examples in which the matrix $\Phi(t)$ have negative real eigenvalues $\forall t$, but the solutions of the corresponding differential equation grow in time. For a review and intuitive explanation see (Josić and Rosenbaum, 2008) and example 3.5 in (Chicone and Latushkin, 1999). For a more complete mathematical analysis see (Wooster, 2011). Therefore, in order to ensure a unique strong solution to (5.6) one needs to find conditions ensuring exponentially boundedness.

- The Dyson series (3.21) is obviously intractable on numerical and on analytical grounds. A natural idea is to truncate this series. If gap junctions conductances are small compared to chemical synapses, one can reorder the series writing first terms containing no term $\overline{G}$, then one term $\overline{G}$, two terms, and so on. The order of truncation is fixed by the norm of $\overline{G}$, which is equal to its spectral radius since this is a symmetric matrix. If this radius is small enough truncation at small order in $\overline{G}$ provide a good approximation of (3.21).

3.4.2 Flow and firing regime

The flow (3.21) characterizes the evolution of the membrane potential vector in the sub-threshold regime. Let us now extend its definition to the firing regime.

3.4.2.1 Extended flow

For a spike train ω, recall that $t_j^{(r)}(\omega)$ is the r-th time of firing of neuron j in the spike train ω. Let $\mathcal{T}^{(p)}(\omega)$ be the p-th time in which a neuron is reset, i.e. $\mathcal{T}^{(1)}(\omega) = \min_j\{t_j^{(1)}(\omega)\}$, and $\mathcal{T}^{(p)}(\omega) = \min_{j,r}\{t_j^{(r)}(\omega) > \mathcal{T}^{(p-1)}(\omega)\}$. Let $\mathcal{K}^{(p)}(\omega)$ be the set of neurons firing at time $\mathcal{T}^{(p)}(\omega)$ i.e. $\forall j \in \mathcal{K}^{(p)}(\omega)$, $\omega_j(\mathcal{T}^{(p)}(\omega)) = 1$.

Consider a membrane potential trajectory *compatible* with ω. Thus, for an initial time $t_0 < \mathcal{T}^{(1)}(\omega)$, $v_k < \theta$, $k = 1, \ldots, N$. The membrane potential vector V follows the evolution (3.23) until time $\mathcal{T}^{(1)}(\omega)$, where the membrane potential of neurons in $\mathcal{K}^{(1)}(\omega)$ is set to 0. It keeps this value during the refractory period, until time $\mathcal{T}^{(1)}(\omega) + \tau_{sep}$. During this period the neurons $\notin \mathcal{K}^{(1)}(\omega)$ keep on evolving according to (3.23), but flow $\Gamma(t_0, s, \omega)$ undergoes a variation at time $\mathcal{T}^{(1)}(\omega)$, where conductances are updated. The smoothness of this variation depends on the assumed regularity of the α function (3.4). After the refractory period, the membrane potential of all neurons follows the evolution (3.23), until another group $\mathcal{K}^{(2)}(\omega)$ fires at time $\mathcal{T}^{(2)}(\omega)$, and so on.

To take into account the refractory periods where neurons are in the rest state, we introduce a diagonal matrix $\chi \equiv \chi(s, \omega)$ with entries:

$$\chi_{kk}(s,\omega) = \begin{cases} 0, & \text{if } s \in \bigcup_r [\mathcal{T}^{(r)}(\omega), \mathcal{T}^{(r)}(\omega) + \tau_{sep}], \text{and } k \in \mathcal{K}^{(r)}(\omega); \\ 1, & \text{otherwise.} \end{cases}$$

Then we replace Γ in (3.23) by $\Gamma\chi$. This gives

$$V(t_0,t,\omega) = \Gamma(t_0,t,\omega)\,\chi(t,\omega)\,v + \int_{t_0}^{t} \Gamma(s,t,\omega)\,\chi(s,\omega)\,f(s,\omega)ds$$

$$+\frac{\sigma_B}{c}\int_{t_0}^{t} \Gamma(s,t,\omega)\,\chi(s,\omega)\,dW(s).$$

To simplify the computations made in the section 3.5, we add the following modelling choice. Denote $\tau_k(t,\omega) - \tau_{sep}$ the last time[1] before t where neuron k has been reset in the past. When $t_0 < \tau_k(t,\omega)$ we replace the integral $\int_{t_0}^{t} \Gamma(s,t,\omega)\,\chi(s,\omega)\,dW(s)$ by $\int_{\tau_k(t,\omega)}^{t} \Gamma(s,t,\omega)\,\chi(s,\omega)\,dW(s)$. In this way, the stochastic dependence upon the past is reset when the neuron's membrane potential is reset. On one hand, this modeling choice does not deeply change the phenomenology. On the other hand it allows us to formulate the determination of spike train statistics as a first passage problem (section 3.5.1). On the opposite, we keep the (deterministic) integral $\int_{t_0}^{t} \Gamma(s,t,\omega)\,\chi(s,\omega)\,f(s,\omega)ds$ unchanged. This term contains indeed a deep effect intrinsic to gap junctions, explained in details in section 3.4.2.2.

[1] By construction of the model this is an integer $\leq [t] - 1$ only determined by t, ω.

To summarize, we have thus:

$$V(t_0,t,\omega) = \Gamma(t_0,t,\omega)\,\chi(t,\omega)\,v + \int_{t_0}^{t}\Gamma(s,t,\omega)\,\chi(s,\omega)\,f(s,\omega)ds + \frac{\sigma_B}{c}\int_{\tau_k(t,\omega)}^{t}\Gamma(s,t,\omega)\,dW(s). \quad (3.28)$$

3.4.2.2 Role of gap junctions on history dependence

In the case $\overline{G}=0$, Γ is diagonal (eq. (3.19)) as well as $\Gamma\,\chi$. The reset of the membrane potential has the effect of removing the dependence of $V_k(t,\omega)$ on its past since $V_k([t]+\tau_{sep},\omega)$ is replaced by 0. As a consequence, the k component of eq. (3.28) holds, from the time $\tau_k(t,\omega)$, introduced in the previous section, up to time t. Therefore, eq. (3.28) factorizes as a set of N equations (Cessac, 2011b):

$$V_k(t,\omega) = \Gamma_{kk}(\tau_k(t,\omega),t,\omega)\,v_k + \int_{\tau_k(t,\omega)}^{t}\Gamma_{kk}(s,t,\omega)\,f_k(s,\omega)ds + \frac{\sigma_B}{c}\int_{\tau_k(t,\omega)}^{t}\Gamma_{kk}(s,t,\omega)\,dW_k(s), \quad (3.29)$$

with $\Gamma_{kk}(\tau_k(t,\omega),t,\omega)=e^{-\frac{1}{c}\int_{t_0}^{t}g_k(s,\omega)ds}$ in agreement with (3.19).

In this case, the reset has the effect of erasing the dependence of V_k on its past anterior to its last firing time. Note that, in equation (3.29), the flow is integrated from the time $\tau_k(t,\omega)$, but the total conductance defining the flow (the term $\int_{t_0}^{t}g_k(s,\omega)ds$) corresponds to an integral starting from the initial time t_0. This is because, contrarily to membrane potentials, conductances *are not reset* when a neuron fire.

The situation is quite more subtle when electric synapses are present, $\overline{G}\neq 0$. Neuron k membrane potential is still reset at firing. From this time on, its evolution depends on the whole vector $V(t)$, in particular $V_j(t)$. But $V_j(t)$ depends on $V_k(s)_{s\leq t}$ via the gap junction connection. Due to this interaction type, the evolution of V_k depends on its past before firing, via the membrane potential of the other neurons.

3.4.2.3 Initial conditions

Equation (3.28) still depends on the initial condition $V(t_0)=v$. However, we are free to choose any $t_0<\mathcal{T}^{(1)}(\omega)$. Especially, we can take $t_0\to-\infty$. This corresponds to a situation where the neural network has started to exist in a distant past (longer than all characteristic relaxation times in the system) and we observe it after transients. This corresponds to an "asymptotic" regime which not necessarily stationary, if the external current $i^{(ext)}$ depends on time.

From the exponentially bounded flow property (3.22) $\Gamma(t_0,t,\omega)v\to 0$ as $t_0\to-\infty$. Therefore upon taking $t_0\to-\infty$ we may write (3.28) as:

$$V(t,\omega)=\int_{-\infty}^{t}\Gamma(s,t,\omega)\chi(s,\omega)f(s,\omega)ds+\frac{\sigma_B}{c}\int_{\tau_k(t,\omega)}^{t}\Gamma(s,t,\omega)\,dW(s).$$

The integrals are well defined thanks to the exponentially bounded flow property. From now on we work with the limit $t_0 \to -\infty$. To alleviate notation we remove the variable t_0 in the expression of the membrane potential.

3.4.2.4 Membrane potential decomposition

The stochastic process $V(t,\omega)$ is therefore the sum:

$$V(t,\omega) = V^{(d)}(t,\omega) + V^{(noise)}(t,\omega), \tag{3.30}$$

of a deterministic part:

$$V^{(d)}(t,\omega) = \int_{-\infty}^{t} \Gamma(s,t,\omega)\,\chi(s,\omega)\,f(s,\omega)ds = V^{(cs)}(t,\omega) \,+\, V^{(ext)}(t,\omega), \tag{3.31}$$

with:

$$V^{(cs)}(t,\omega) = \frac{1}{c}\int_{-\infty}^{t} \Gamma(s,t,\omega)\,\chi(s,\omega)\,I^{(cs)}(s,\omega)ds, \tag{3.32}$$

the chemical synapses contribution to the membrane potential;

$$V^{(ext)}(t,\omega) = \frac{1}{c}\int_{-\infty}^{t} \Gamma(s,t,\omega)\,\chi(s,\omega)\,I^{(ext)}(s,\omega)ds, \tag{3.33}$$

the external current + leak term contribution, and a stochastic part:

$$V^{(noise)}(t,\omega) = \frac{\sigma_B}{c}\int_{\tau_k(t,\omega)}^{t} \Gamma(s,t,\omega)\; dW(s). \tag{3.34}$$

Thanks to the limit $t_0 \to -\infty$ which has removed the dependence in the initial condition $V^{(d)}(t,\omega)$ is uniquely determined by the spike history ω (and the time dependence of the external current $i^{(ext)}$, if any). Likewise, $V^{(noise)}(t,\omega)$ is the integral of the Wiener process with a weight $\Gamma(s,t,\omega)$ depending on the spike history.

3.5 Spike train statistics and Gibbs distribution

This section is devoted to the characterization of spike train statistics in the model. The main result establishes that spike train are distributed according to a Gibbs distribution. Note that we do not make any assumption on the *stationarity* of dynamics: the present formalism affords to consider as well a time dependent external current (stimulus). There is no explicit form of the potential determining the Gibbs distribution in the general case, but upon an assumption discussed in section 3.5.3, the potential can be approached by an analytic form. This form, is further discussed in the section 3.6 with its deep connections, on one hand, with the Generalized Linear Model (GLM) and on the other with maximum entropy models.

For the non specialized readers, we give here the main ideas behind mathematics. Spike statistics is characterized by a family of transition probabilities giving the probability to have a spike pattern $\omega(n)$ given a past spike history $\omega_{-\infty}^{n-1}$. From this

set of transitions probabilities, one defines the Gibbs distribution depending on the bio-physical parameters defining the model, as exposed in chapter 2. This is done in a similar way as homogeneous and positive Markov chains transition probabilities define the invariant distribution of the chain, although the case considered here is more general (we do not assume stationarity and we do not assume a finite memory).

3.5.1 Transition probabilities

We want to determine the probability $\mathbb{P}\left[\omega(n) \,\middle|\, \omega_{-\infty}^{n-1}\right]$ to have a spiking pattern $\omega(n)$ at time n given the spike history $\omega_{-\infty}^{n-1}$. This can be stated as a first passage problem (Burkitt, 2006a,b, Touboul and Faugeras, 2007).

Fix ω, n and $t < n$. Set:

$$\widehat{\theta}_k(t,\omega) = \theta - V_k^{(d)}(t,\omega), \tag{3.35}$$

the distance of the deterministic part of the membrane potential to the threshold. Neuron k will emit a spike at time n if there exist a time $t \in [n-1, n]$ such that $V_k^{(noise)}(t,\omega) = \widehat{\theta}_k(t,\omega)$.

Denote:

$$\sigma_k^2(t,\omega) = \frac{\sigma_B^2}{c^2} \sum_{j=1}^{N} \int_{\tau_k(t,\omega)}^{t} \Gamma_{kj}^2(s,t,\omega)\, ds.$$

Following (Touboul and Faugeras, 2007) Dubins-Schwarz' theorem can be used to change the time scale to write $V_k^{(noise)}(t,\omega)$ as a Brownian motion $V_k^{(noise)}(t,\omega) \stackrel{\text{def}}{=} W_{\sigma_k^2(t,\omega)}$ and the spiking condition at time n reads:

$$W_{\sigma_k^2(t)} = \widehat{\theta}_k(t,\omega).$$

This equation characterizes the first hitting time h_k of the Brownian motion $W_{\sigma_k^2(t)}$ to the"boundary" $\widehat{\theta}_k(t,\omega)$.

Denote $\mathbb{P}\left[h_1, \ldots, h_N\right]$ the joint law of the first hitting times of neurons $1, \ldots, N$. For a spiking pattern $\omega(n)$ divide the set of indices $\{1, \ldots, N\}$ into two subsets $\mathcal{S}^+(n,\omega) = \{k \in \{1, \ldots, N\}, \omega_k(n) = 1\}$ and $\mathcal{S}^-(n,\omega) = \{k \in \{1, \ldots, N\}, \omega_k(n) = 0\}$. We have:

$$\mathbb{P}\left[\omega(n) \,\middle|\, \omega_{-\infty}^{n-1}\right] = \mathbb{P}\left[\bigcap_{k \in \mathcal{S}^+(n,\omega)} \{h_k \in [n-1, n[\} \cap \bigcap_{l \in \mathcal{S}^-(n,\omega)} \{h_l > n\}\right]. \tag{3.36}$$

This equation can be written in terms of an integral of the join density of hitting times.

The first passage problem can be solved in simple one dimensional situations following a method developed by Lachal (Lachal, 1998). The Laplace transform of the first hitting time density can be obtained as a solution of a PDE. However, we

haven't been able to find a general form for the joint density of the N-dimensional problem in our model.

An alternative approach is to use the Fokker-Planck approach for IF models developed by several authors (Brunel and Hakim, 1999, Lindner et al., 2004, Burkitt, 2006a,b, Ostojic and Brunel, 2009). The resulting equations are seemingly not solvable unless using a mean-field approximation in the thermodynamic limit as done e.g. by Ostojic-Brunel-Hakim (Ostojic and Brunel, 2009) in the specific case where chemical and electric synapses are constant and equals. The Fokker-Planck method is a dual approach to the first passage problem but it meets the same difficulties.

In the next section, we discuss about the general conditions ensuring nevertheless that this system of transition probabilities define a Gibbs distribution. Then, in section 3.5.3 we propose an approximation allowing an explicit computation of transition probabilities.

3.5.2 The Gibbs distribution

The transition probabilities (3.36) define a stochastic process on the set of spike trains, where the probability of having a spiking pattern $\omega(n)$ at time n depends on an infinite past $\omega_{-\infty}^{n-1}$. Under suitable conditions exposed in section 2, a sequence of transition probabilities defines a unique probability distribution μ on the set of spike trains, a "Gibbs distribution". corresponding to the following Gibbs potential:

$$\phi\,(\,n,\omega\,) = \log \mathbb{P}\left[\,\omega(n)\;\middle|\;\omega_{-\infty}^{n-1}\,\right], \tag{3.37}$$

3.5.3 Approximating the Gibbs potential

The previous subsections provide mathematical results, but they do not allow explicit computations. The basic step for this would be to have an explicit form for the Gibbs potential (resp. the transition probabilities). As mentioned in section 3.5.1 this is not tractable in general. We propose here an approximation.

3.5.3.1 The stochastic term $V^{(noise)}(t,\omega)$ and compatibility conditions

$V^{(noise)}(t,\omega)$ is the integral of a Wiener process (eq. (3.34)). As a consequence this is a Gaussian process, with independent increments, mean 0, and covariance matrix:

$$\mathcal{Q}(t,\omega) \stackrel{\text{def}}{=} Cov\left[\,V^{(noise)}(t,\omega),\,V^{(noise)^T}(t,\omega)\,\right],$$

where ^{T} denotes the transpose. From standard Wiener integration we have:

$$\mathcal{Q}(t,\omega) = \frac{\sigma_B^2}{c^2}\mathbb{E}\left[\int_{\tau_k(t,\omega)}^{t}\Gamma(s,t,\omega)\,\chi(s,\omega)\,dW(s)(\int_{\tau_k(t,\omega)}^{t}\Gamma(s',t,\omega)\,\chi(s',\omega)\,dW(s'))^T\right]$$

$$= \frac{\sigma_B^2}{c^2}\int_{\tau_k(t,\omega)}^{t}\int_{\tau_k(t,\omega)}^{t}\Gamma(s,t,\omega)\,\chi(s,\omega)\chi^T(s',\omega)\Gamma^T(s',t,\omega)\,\mathbb{E}\left[\,dW(s)dW(s')^T\,\right],$$

so that:

$$\mathcal{Q}(t,\omega) = \frac{\sigma_B^2}{c^2} \int_{\tau_k(t,\omega)}^{t} \Gamma(s,t,\omega)\, \chi(s,\omega)\, \Gamma^T(s,t,\omega)\, ds, \tag{3.38}$$

where we used $\chi(s,\omega)\chi^T(s,\omega) = \chi(s,\omega)$.

However, as mentioned in section 3.3.1.5 the model definition imposes compatibility conditions between the trajectories of the membrane potential $V(t,\omega)$ and the spike train ω (see eq. (3.14)). Fixing a spike train ω, fixes the flow Γ as well as the deterministic part of the membrane potential $V^{(d)}(t,\omega)$ (eq. (3.31)). Then, $V^{(noise)}(.,\omega)$ has to obey the following compatibility conditions:

$$\forall k = 1,\dots N,\ \forall n \le t, \quad \begin{cases} \omega_k(n) = 0 & \Leftrightarrow \quad \max_{u\in]n-1,n]} \left[V_k^{(noise)}(u,\omega) - \widehat{\theta}_k(u,\omega) \right] < 0; \\ \omega_k(n) = 1 & \Leftrightarrow \quad \max_{u\in]n-1,n]} \left[V_k^{(noise)}(u,\omega) - \widehat{\theta}_k(u,\omega) \right] \ge 0; \end{cases} \tag{3.39}$$

As a consequence, when computing the law of $V_k^{(noise)}(.,\omega)$ we have to take these constraints into account. We write $\mathbb{P}_{\mathcal{C}}\left[V_k^{(noise)}(.,\omega) \right]$ the law of $V_k^{(noise)}(.,\omega)$ given these constraints. Although, $V_k^{(noise)}(t,\omega)$ is Gaussian, $\mathbb{P}_{\mathcal{C}}\left[V_k^{(noise)}(.,\omega) \right]$ is not a Gaussian distribution. For example, if $\omega_k(n) = 0$ then, necessarily, $V_k^{(noise)}(t,\omega) < \theta - V_k^{(d)}(t,\omega)$, $[\,t\,] + 1 = n$, while a Gaussian random variable takes unbounded values.

3.5.3.2 A Gaussian approximation.

A plausible approximation consists then of approximating $\mathbb{P}_{\mathcal{C}}\left[V^{(noise)}(.,\omega) \right]$ by the Gaussian law of $V^{(noise)}$, i.e. "ignoring" compatibility conditions. This approximation can be justified under the following conditions.

- **Weak noise.** Consider the compatibility conditions for $\omega_k(n) = 0$: $\forall u \in]n-1,n]$, $V_k^{(noise)}(u,\omega) < \widehat{\theta}_k(u,\omega)$. What is the probability that the Gaussian noise $V_k^{(noise)}(u,\omega)$ violates this condition for some u in the interval $]n-1,n]$?

  Denoting $\sigma_k^2(u,\omega) \stackrel{\text{def}}{=} \mathcal{Q}_{kk}(u,\omega)$, the variance of $V_k^{(noise)}(u,\omega)$ at time u (given explicitly by eq. (3.38)), this probability is given by:

  $$\mathbb{P}\left[V_k^{(noise)}(u,\omega) \ge \widehat{\theta}_k(u,\omega) \right] = \frac{1}{\sqrt{2\pi}} \int_{\frac{\widehat{\theta}_k(u,\omega)}{\sigma_k(u,\omega)}}^{+\infty} e^{-\frac{h^2}{2}}\, dh.$$

  From the explicit form of $\sigma_k^2(u,\omega)$ (eq. (3.38)) and from the exponentially bounded flow property of Γ, $\sigma_k^2(u,\omega)$ is upper bounded, uniformly in u and ω, by $\frac{\sigma_B^2}{c^2} B$ where B is some constant depending on the model parameters and where σ_B is the Wiener noise intensity. Thus, $\sigma_k^2(u,\omega) \to 0$ as $\sigma_B \to 0$.

If $\sigma_B = 0$ then the compatibility conditions for $\omega_k(n) = 0$ is $\forall u \in]n-1, n]$, $\widehat{\theta}_k(u,\omega) > 0$ (i.e. $V_k^{(d)}(u,\omega) < \theta$). Assume that $\widehat{\theta}_k(u,\omega) > \varepsilon$ for some $\varepsilon > 0$.

If $\sigma_B \neq 0$, $\frac{\widehat{\theta}_k(u,\omega)}{\sigma_k(u,\omega)} > \frac{\varepsilon c}{\sqrt{B}}\frac{1}{\sigma_B}$, so that:

$$\mathbb{P}\left[V_k^{(noise)}(u,\omega) \geq \widehat{\theta}_k(u,\omega)\right] < \frac{1}{\sqrt{2\pi}}\int_{\frac{\varepsilon c}{\sqrt{B}}\frac{1}{\sigma_B}}^{+\infty} e^{-\frac{h^2}{2}}dh = \frac{1}{2}\left(1 - erf\left(\frac{\varepsilon c}{\sqrt{B}}\frac{1}{\sigma_B}\right)\right),$$

where $erf(x)$ admits the following series expansion as $x \to +\infty$:

$$erf(x) = 1 - e^{-x^2}\frac{1}{\sqrt{\pi}}\left(\frac{1}{x} - \frac{1}{2x^3} + \frac{3}{4x^5} - \frac{15}{8x^7}\right) + \circ\left(x^{-8}e^{-x^2}\right).$$

Therefore, as $\sigma_B \to 0$, we have:

$$\mathbb{P}\left[V_k^{(noise)}(u,\omega) \geq \widehat{\theta}_k(u,\omega)\right] < \frac{1}{2}e^{-x^2}\frac{1}{\sqrt{\pi}}\left(\frac{1}{x} - \frac{1}{2x^3} + \frac{3}{4x^5} - \frac{15}{8x^7}\right) + \circ\left(x^{-8}e^{-x^2}\right), \tag{3.40}$$

with $x = \frac{\varepsilon c}{\sqrt{B}}\frac{1}{\sigma_B}$.

This shows that the probability that the noise violates the compatibility condition decreases exponentially fast as $\sigma_B \to 0$. As a consequence, when the noise is small, approximating $\mathbb{P}_{\mathcal{C}}\left[V_k^{(noise)}(.,\omega)\right]$ by the Gaussian law of $V_k^{(noise)}$ provides a reliable approximation. This amounts to considering that the spikes arising between $[n-1, n[$ are determined by the deterministic part of the membrane potential, not by the noise. If a neuron is about to fire at time n because its deterministic part crosses the threshold at a time $t \in [n-1, n[$, the (weak) noise can affect the time t where the crossing occurs, but, with high probability this time stays in the interval $[n-1, n[$.

- **Time discretization.** Beyond the compatibility conditions, an additional aspect makes the analytic computation of transition probabilities delicate: membrane potential time evolution is continuous. Since we are focusing here on spike statistics, where spike occur on discrete times, this obstacle can be raised using the following approximation. We replace the spike condition of neuron k at time n: $\exists u \in]n-1, n], V_k^{(noise)}(u,\omega) \geq \widehat{\theta}_k(u,\omega)$ by: $V_k^{(noise)}(n,\omega) \geq \widehat{\theta}_k(n,\omega)$. The argument supporting this choice is the following. Suppose that $V_k^{(noise)}(n-1,\omega) < \widehat{\theta}_k(n-1,\omega)$; if $V_k^{(noise)}(n,\omega) \geq \widehat{\theta}_k(n,\omega)$ then with probability 1, the continuous process crosses the threshold between $]n-1, n]$, therefore the spike is registered at time n. On the other hand if $V_k^{(noise)}(n,\omega) < \widehat{\theta}_k(n,\omega)$ there is nevertheless a positive probability that the continuous process crosses the threshold at some time t between $]n-1, n[$. In this case, there will be a spike that our approximation neglects since $V_k^{(noise)}(n,\omega) < \widehat{\theta}_k(n,\omega)$. The probability of occurrence of such an event can be explicitly computed for general diffusion processes (see (Baldi

and Caramellino, 2002)) and depend on the noise intensity σ_B and time step discretization δ. It is given by:

$$C_{\widehat{\theta}} \exp\left(\frac{-2(\widehat{\theta}_k(n-1,\omega) - V_k^{(noise)}(n-1,\omega))(\widehat{\theta}_k(n,\omega) - V_k^{(noise)}(n,\omega))}{\delta\sigma_B} \right),$$

where $C_{\widehat{\theta}}$ is a bounded function depending on the value of $\widehat{\theta}$ at time $n-1$. Therefore, when σ_B, δ are small (as in our case) this probability is also small.

3.5.3.3 Approached form of the potential.

Thanks to these approximation we have now the following result. If we note:

$$\mathcal{J}_k(n,\omega) = \begin{cases}]-\infty, \widehat{\theta}_k(n-1,\omega)[, & \text{if } \omega_k(n) = 0; \\ \\ [\widehat{\theta}_k(n-1,\omega), +\infty[, & \text{if } \omega_k(n) = 1; \end{cases}$$

and

$$\mathcal{J}(n,\omega) = \prod_{k=1}^{N} \mathcal{J}_k(n,\omega), \tag{3.41}$$

where $\prod$ denotes the Cartesian product of intervals, eq. (3.36) becomes, using the Gaussian approximation for $V^{(noise)}(.,\omega)$:

$$\mathbb{P}\left[\, \omega(n) \,\middle|\, \omega_{-\infty}^{n-1} \,\right] = \int_{\mathcal{J}(n,\omega)} \frac{e^{\frac{-V^T \mathcal{Q}^{-1}(n-1,\omega)V}{2}}}{(2\pi)^{\frac{N}{2}} |\mathcal{Q}(n-1,\omega)|^{\frac{1}{2}}} dv, \tag{3.42}$$

where $dv = \prod_{i=1}^{N} dV_i$.

Taking the log of (3.42) we obtain the approached form of the Gibbs potential.

3.5.3.4 Remarks

The apparently simple form (3.42) hides a real complexity.

- The integration domain $\mathcal{J}(n,\omega)$ corresponds to a product of intervals involving the variables $\widehat{\theta}_k(n-1,\omega)$, the distance of $V_k^{(d)}(n-1,\omega)$ to the threshold. Now, from eq. (5.6), (3.31), involves the chemical synapse current $I^{(cs)}(t,\omega)$, eq. (3.8), and the external current $I^{(ext)}(t)$ (3.10), integrating, via the flow Γ, the spike history of the network. As a consequence, the transitions probabilities depend on all parameters defining the system: synaptic weights W_{kj} (eq. (3.9), gap junctions, external current, and biophysical parameters such as membrane capacity, or characteristic time scale of the post-synaptic potential $\alpha_{kj}(t)$ (eq. (3.4)).

- Without electric synapses the covariance (3.38) takes a diagonal form since Γ and χ are diagonal matrices. Thus, in this case, neurons are (conditionally

upon the past), *independent*. As a consequence the transition probability $\mathbb{P}\left[\,\omega(n)\,\middle|\,\omega_{-\infty}^{n-1}\,\right]$ factorizes:

$$\mathbb{P}\left[\,\omega(n)\,\middle|\,\omega_{-\infty}^{n-1}\,\right] = \prod_{k=1}^{N} \mathbb{P}\left[\,\omega_k(n)\,\middle|\,\omega_{-\infty}^{n-1}\,\right]. \tag{3.43}$$

3.6 Consequences

In this section we adopt the following point of view. Assuming that the model presented here captures enough of the biophysics of real neurons, what are the - relevant for neuroscience- consequences of the mathematical results developed in the previous sections ? We essentially focus on spike trains analysis and argue that:

1. Spikes correlations are not only due to shared stimulus: there are correlations induced by dynamics, that persists without stimulus. Moreover, *in the absence of electric synapses* neurons are conditionally independent upon the past.

2. The Gibbs potential (3.37) or even its Gaussian approximation (3.42), includes existing models for spike trains statistics such as maximum entropy models (Schneidman et al., 2006, Tkačik et al., 2009, Shlens et al., 2006, 2009, Ohiorhenuan et al., 2010, Ganmor et al., 2011a,b, Vasquez et al., 2012) and GLM (Pillow et al., 2005, 2008, Ahmadian et al., 2011, Pillow et al., 2011, Macke et al., 2011).

3.6.1 Correlations structure

3.6.1.1 Transition probabilities do not factorize in general

This can be illustrated in the Gaussian approximation. The covariance matrix $\mathcal{Q}(n-1,\omega)$ can always be diagonalized by an orthogonal variable change $P(n-1,\omega)$ depending on n and ω:

$$\mathcal{Q}(n-1,\omega)\,P(n-1,\omega) = P(n-1,\omega)\,\Sigma(n-1,\omega),$$

where $\Sigma(n-1,\omega)$ is diagonal with real, positive eigenvalues $\sigma_k^2(n-1,\omega)$.

Upon the variable change $V = P(n-1,\omega)V'$, the transition probability reads therefore:

$$\mathbb{P}\left[\,\omega(n)\,\middle|\,\omega_{-\infty}^{n-1}\,\right] = \int_{\mathcal{J}'(n,\omega)} \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{V_k'^2}{2\sigma_k^2(n-1,\omega)}}\, dV_k',$$

where $\mathcal{J}'(n,\omega)$ is the image of $\mathcal{J}(n,\omega)$ in the variable change. As a classical result the variables change has transformed the join Gaussian density of $V^{(noise)}(n-1,\omega)$ into a product of one dimensional Gaussian with mean zero and variance $\sigma_k^2(n-1,\omega)$. However, the same variable change has transformed the domain $\mathcal{J}(n,\omega)$, reading as a product of intervals (eq. 3.41) into the domain $\mathcal{J}'(n,\omega)$ which is not a product any more, except in the case with no gap junctions.

Therefore, in general:

$$\mathbb{P}\left[\,\omega(n)\;\middle|\,\omega_{-\infty}^{n-1}\,\right] \neq \prod_{k=1}^{N}\mathbb{P}\left[\,\omega_k(n)\;\middle|\,\omega_{-\infty}^{n-1}\,\right]. \tag{3.44}$$

3.6.2 Correlations

Note first that *conditional independence upon the past* does not mean *independence*. Conditional independence upon the past means that the Gibbs potential reads:

$$\phi_n\left(\,\omega\,\right) = \sum_{k=1}^{N}\phi_{n,k}\left(\,\omega\,\right).$$

This is a sum of per-neuron potentials $\phi_{n,k}$, but each of this potential is a function of the past *network* activity ω. On the opposite, neurons independence would mean that:

$$\phi_n\left(\,\omega\,\right) = \sum_{k=1}^{N}\phi_{n,k}\left(\,\omega_k\,\right),$$

where now the per-neuron potentials $\phi_{n,k}$ depends on the past activity of neuron k only.

Now, what are the sources of general correlations (i.e. pairwise and higher order) in our model ? Namely, what makes the potential depend on the network history ? Even in the absence of electric synapses, correlations can be induced on one hand by the stimulus (the term $V^{ext}(t,\omega)$, eq (3.33)) if the external current $i^{(ext)}(t)$ in (1) has correlations between its entries ($i_k^{(ext)}(t)$ and $i_l^{(ext)}(t)$ for $k \neq l$ are correlated), and on the other hand by the chemical synapses term $V^{(cs)}(t,\omega)$, (eq. (3.32)). Even if the stimulus is zero, the chemical synapses term remains. We arrive then to the (somewhat obvious) conclusion that in the model the main source of correlations *is not* a (shared) stimulus. It is due to interactions between neurons.

A deeper question is to quantify the intensity of correlations induced by (i) chemical synapses; (ii) electric synapses; (iii) stimulus and under which conditions (parameters value) shared stimulus correlations are dominant. This is the main additional issue to investigate neuronal encoding by a population of neurons. This requires an extended investigation.

3.6.3 The Gibbs potential form includes existing models for spike trains statistics

3.6.3.1 Markovian approximation

The exponentially bounded flow property (3.22) means that the norm of the flow $\Gamma(s,t,\omega)$ decays exponentially fast as $t-s$ grows. A consequence of this is the *exponential decay of the variation of* $\Gamma(s,t,\omega)$ *with respect to* ω. Assume that we

know the spike patterns of the spike train ω from some integer time $m < t$ to time t, but we ignore the spike patterns occurring before m. What is the maximal error than we can make on the value of $\Gamma(s,t,\omega)$? As shown in the appendix of (Cofré and Cessac, 2013), thanks to the exponential bounded flow property, this error decays exponentially fast as $t-m$ grows. The same property holds for the membrane potential. It holds as well for the Gibbs potential under conditions stated in the appendix.

As a consequence the norm of the difference between the exact Gibbs potential (3.37) and an approximate potential where the past spike history $\omega_{-\infty}^{n-1}$ is replaced by ω_{n-D}^{n-1} for some integer $D > 0$ called *memory depth*, decreases exponentially fast with D.

Therefore, we may replace the infinite range potential (3.37) corresponding to a process with infinite memory, by a truncated potential, corresponding to a Markov chain with memory depth D. In this case, the Gibbs distribution is the invariant probability of the corresponding Markov chain. For a small number of neurons this probability can be characterized using transfer matrices techniques (Vasquez et al., 2012). For larger networks Monte Carlo methods can be used (Nasser and Cessac, 2014).

If ϕ has finite range D: $\phi(n,\omega) \equiv \phi(\omega_{n-D}^{n})$ then, a classical result from Hammersley and Clifford (Hammersley and Clifford, 1971) establishes that ϕ can be written as:

$$\phi(\omega_{n-D}^{n}) \;=\; \sum_{k=0}^{2^{ND}} \phi_k \, \mathcal{O}_k(\omega_{n-D}^{n}), \tag{3.45}$$

where the ϕ_ks are real parameters, (non linear) functions of the network parameters of the model.

Now, the inspection of the Gibbs potential in our model leads to several strong conclusions:

1. The Gibbs potential is definitely not Ising, and is actually quite far from an Ising model. The main reason for that is that Ising model involves only instantaneous (pairwise) events, with no memory effects. On the opposite, strong memory effects exist in our model, requiring to consider transition probabilities depending on spike history.

2. The exact expansion (3.45) involves 2^{ND} constraints, making rapidly intractable any numerical methods attempting to match all constraints. Obviously, one can argue that some ϕ_ks can be close to 0 so that the corresponding constraints can be ignored in the approximation of μ. Unfortunately, this argument does not tell us *which are* the negligible terms. It might well be that the answer depend on the network parameters as well. This aspect is investigated in the next chapter.

3. Although the expansion (3.45) of a potential into a linear combination of observables is quite natural in the realm of statistical physics and Jaynes principle, it may not be appropriate for the study of neural networks of the type

studied here. Indeed, the parameters ϕ_ks, whose number increases exponentially fast with the number of neurons, are *redundant*: they are functions of the network parameters, such as synaptic weights, whose number increases, at most, like N^2. Using approximations like the Gaussian one allows to obtain an explicit form for the potential as a function of these parameters. One obtains in the end a nonlinear function, but with quite a bit less parameters to tune; additionally those parameters have a straightforward biophysical interpretation. This constitutes an alternative strategy to estimate spike statistics from empirical data.

3.7 Conclusion

In this work we have analyzed the joint effects of spike history dependent chemical synapses conductances and gap junctions in an Integrate and Fire model. We have pointed out several technical and conceptual difficulties mainly lying in the fact that, in general, the chemical conductance matrix $G(t, \omega)$ and the gap junctions matrix $\overline{G}$ do not commute. As we showed, this has no impact on the well posedness of the model, provided that the values of chemical and electrical conductances are compatible with biophysics. There exists a strong solution of the stochastic differential equations. However, except in some specific cases, the flow reads as a Dyson series, hardly tractable.

We have also considered the statistics of spike trains in this model and showed that it is characterized by Gibbs distribution, time-dependent (non stationary) whenever the external current is time-dependent. The corresponding potential can be computed under a Gaussian approximation: it has infinite range with exponentially decaying interactions.

The main observation resulting from our analysis is that spike statistics is indecomposable. The probability of spike patterns does not factorize as a product of marginal, per-neuron, distributions. This effect is enhanced by the presence of gap junctions. As a consequence, in that model, *there is absolutely now way to defend that neurons act as independent sources.* Additionally, correlations mainly result from the chemical and electrical interactions between neurons (correlations persist even if there is no external current / stimulus). Our work suggests especially that electric synapses could have a strong influence in spike train statistics of biological neural systems, especially the retina where gap junctions connections are ubiquitous.

As mentioned in the introduction, one of the main motivation of this work was to better understand how ganglion cells in the retina supporting both chemical and electric synapses coordinate spatio-temporal spike patterns to encode information conveyed to the brain. Developing a further understanding of the regulation of gap junctions, as well as the dynamic relationship between electrical and chemical transmission, is an important challenge for the future (Bloomfield and Völgyi, 2009).

Chapter 4

Exact computation of the Maximum Entropy Potential from spiking neural networks models

Overview

Preceding chapters have left us with several distinct approaches to analyze neural network spike train statistics. We know, in particular fron chapter 2 that a MaxEnt potential of the form (2.11) is always associated to a set of transition probabilities corresponding to a Markov chain. Here we address the inverse correspondence: Having a set of transition probabilities (GLM, IF) with stationary dynamics, can one construct a MaxEnt potential leading the same Gibbs distribution? As we show, there are infinitely many such potentials. Nevertheless, there is a unique canonical one, in the sense described in this chapter. We present a method, based on equivalence between potentials, Hammersley-Clifford hierarchy and periodic orbits invariance to recover this canonical MaxEnt potential. We present an example based on a discrete time Integrate and Fire model in which we compute explicitly the "local fields" and "Ising couplings" as non linear functions of the chemical synaptic architecture and stimulus. We finally present the conclusions of this work in which we address especially the issue of the difference of dimensionality between the space of parameters of MaxEnt and neuro-mimetic models. This chapter is based on the published work (Cofré and Cessac, 2014).

Contents

4.1 Introduction

Let us introduce this chapter by contextualizing the main point in terms of the guiding lines of this thesis i.e. the transition probabilities. As we have seen in the previous chapter when considering neural network models the transition probabilities and therefore the spike train statistics can be written in terms of the parameters defining the model. This fact present an enormous advantage with respect to the maximum entropy approach in which the transition probabilities can also be obtained in terms on the maximum entropy fitting parameters, but lacking interpretation in terms of the neural network model of the tissue which has generated the data.

An alternative to the MaxEnt approach (explained in chapter 2) is based on spiking neuron models, providing a mathematical description of neural dynamics (see chapter 3 for an explicit example). These models give a probabilistic mapping between network architecture, stimuli, spiking history of the network and spiking response in terms of conditional probabilities of spike pattern given the network history. Prominent examples are the Linear-Non Linear model (LN), the Generalized Linear Model (GLM) (Brillinger, 1988, Chichilnisky, 2001) or Integrate-and-Fire models (Gerstner and W.Kistler, 2002). In all these examples the conditional probabilities that a spike pattern occurs at time t given the network spike history are explicit functions of "structural" parameters in the neural network (that can be interpreted as synaptic weights $\mathcal{W}$ matrix, and stimulus $\mathcal{I}$) (Fig. 4.1).

These conditional probabilities define a Markov process that mimics the biophysical dynamics of neurons in a network and the mechanisms that govern spike trains emission, including stimulus dependence and neurons interactions via synapses. We call them *neuro-mimetic* statistical models.

To summarize, at least two different representations can be used to analyze spike train statistics in neural networks (fig. 4.1). The goal of this chapter is to establish an explicit and exact correspondence between these two representations.

A previous result attempting to describe such a relationship can be found in (Cocco et al., 2009). Here, the authors fit a leaky Integrate and Fire model matching spike train data from a population of retinal ganglion cells. At the same time they fit a MaxEnt Ising model from this data. This allows them to compare in particular synaptic weights W_{ij} with MaxEnt Ising couplings J_{ij}. Another work in this direction can be found in (Granot-Atedgi et al., 2013) in which stimulus dependent MaxEnt is introduced based on (LN) model, attempting to include stimulus information into the "local fields" of the Ising model. Both examples are limited to the Ising model, thus do not include memory effects in the MaxEnt statistics.

We propose here a generalization which allows us to handle more general types of neuro-mimetic models as well as general spatio-temporal MaxEnt distributions

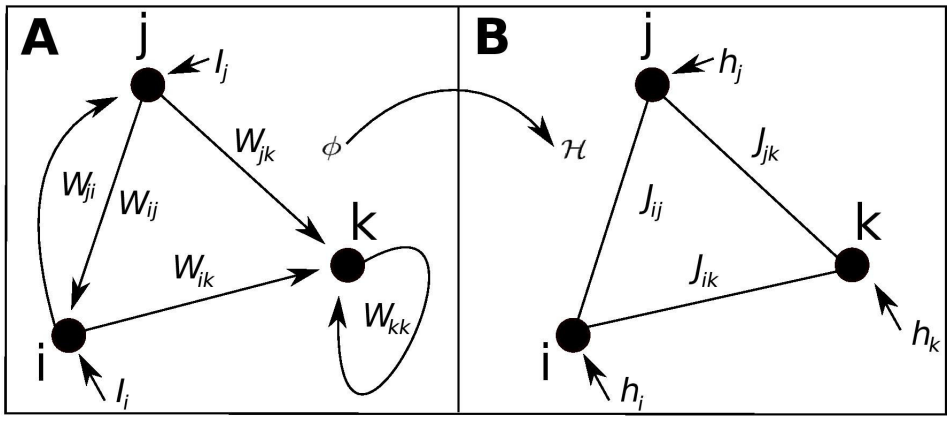


Figure 4.1: (A) Neuro-mimetic approach. Neurons are interacting via synaptic weights W_{ij} and submitted to a stimulus I. Spike probabilities are explicit functions of these parameters. (B) MaxEnt statistical approach. Here the relations between neurons are expressed by functional parameters allowing to correctly fit the correlations in the model. The graph represents the Ising model where only local fields and pairwise interactions are drawn. More general interactions are considered in the text. In Ising model pairwise interactions are symmetric (represented without arrows). We are looking for an explicit and exact correspondence between these two representations.

(including memory). The method we used is based on Hammersley-Clifford decomposition (Hammersley and Clifford, 1971) and periodic orbit invariance from ergodic theory (Pollicott and Weiss, 2003). The techniques are therefore different from (Cocco et al., 2009, Granot-Atedgi et al., 2013).

More generally, we answer the following questions:

Question 1: *Given an ergodic Markov process, where the transition probabilities are known, can we construct a MaxEnt potential, with a minimum of constraints, reproducing exactly the (spatio-temporal) statistics of this process?*

This is the most general question we answer in this chapter. It is important to notice that our results are not restricted to spike trains and neural networks, but to any ergodic Markov chain. The next question focuses on the correspondence between MaxEnt parameters and structural parameters defining spiking neural networks.

Question 2: *Given the transition probabilities from a neuro-mimetic model, is it possible to derive an analytic correspondence between MaxEnt fitting parameters (Lagrange multipliers) and neuro-mimetic structural parameters? For example can*

we establish a correspondence between the local fields h_i and the external stimulus I_i? Between Ising couplings $\mathtt{J}_{ij}$ with W_{ij}, the synaptic weights?

We show that there exists an exact and analytic correspondence revealing that Lagrange multipliers are complex non-linear functions of structural parameters. For example the local field $\mathcal{H}_i$ or the functional interactions $\mathtt{J}_{ij}$ are non-linear functions depending *generically* on all stimuli and all synaptic weights.

Additionally this correspondence raises up the question of dimensionality. A neuro-mimetic model with N neurons typically has N^2+N independent parameters (N^2 synaptic weights and N stimuli); a MaxEnt model with memory depth D may have up to $2^{N(D+1)}$ independent parameters (see text). The dimensionality of these two types of models is drastically different. When mapping a MaxEnt model to a neuro-mimetic model there is clearly a loss of dimensionality.

Question 3: *Consider a MaxEnt model equivalent to a neuro-mimetic model. Then, the difference in dimensionality between the spaces of parameters of both models suggests that either many of MaxEnt parameters are zero, or that there are relations among them, i.e. they are not independent. What is the generic situation?*

At the end of this chapter we address the issue of the difference of dimensionality between the space of parameters of MaxEnt and neuro-mimetic models.

4.2 Setting

We study a network composed by N neurons as in the introduction of this thesis. We also consider a *potential* $\mathcal{H}$ of range $R = D+1$. We assume $\mathcal{H}(\omega_0^D) > -\infty$. As we have seen in chapter 2 any such potential can be written:

$$\mathcal{H}(\omega_0^D) = \sum_{l=0}^{L} h_l \, m_l(\omega_0^D), \tag{4.1}$$

This potential defines a unique stationary probability μ.

4.2.1 Equivalent potentials

Although a potential $\mathcal{H}$ of the form (4.1) corresponds (if $\mathcal{H} > -\infty$) to a unique normalized potential ϕ and Gibbs distribution μ, this correspondence is not one to one. To a normalized potential ϕ corresponds infinitely many potentials of the form (4.1). Hence two potentials $\mathcal{H}^{(1)}, \mathcal{H}^{(2)}$ can correspond to the same Gibbs distribution (We call them equivalent).
A standard result in ergodic theory states that $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ are equivalent if and only if there exists a range $D > 0$ function f such that (Bowen, 1975):

$$\mathcal{H}^{(2)}\left(\omega_0^D\right) = \mathcal{H}^{(1)}\left(\omega_0^D\right) - f\left(\omega_0^{D-1}\right) + f\left(\omega_1^D\right) + \Delta, \tag{4.2}$$

where $\Delta = \mathcal{P}\left[\mathcal{H}^{(2)}\right] - \mathcal{P}\left[\mathcal{H}^{(1)}\right]$. This relation establishes a *strict* equivalence and does not correspond e.g. to renormalization. The validity of (4.2) can be readily seen by plugging $\mathcal{H}^{(2)}$ in the variational formula (2.25) the terms corresponding to f cancels because ν is time-translation invariant. Therefore, the supremum is reached for the same Gibbs distribution as $\mathcal{H}^{(1)}$ whereas Δ is indeed the difference of free energies. The "only if" part is more tricky.

Equation (2.16) is a particular case of equation (4.2), where $\mathcal{H}^{(2)} = \phi, \mathcal{H}^{(1)} = \mathcal{H}, f = \log R$ and $\Delta = -\log s$. This equation has the virtue to unify two very different approaches. It establishes a relation between Markov chain normalized potentials (2.13) on one hand and potentials of the form (4.1) on the other hand (the arrow $\phi \to \mathcal{H}$ in fig. 4.1). Equation (4.2) answers therefore the first part of the question (1), but, by itself does not provide a straightforward way to exploit it, due to the arbitrariness in the choice of f. Indeed, there are infinitely many potentials $\mathcal{H}$ corresponding to the same Gibbs distribution (the same normalized potential ϕ).

This arbitrariness in the choice of f raises a natural question closely related to the second part of question (1). Given a normalized potential is it possible to find, among the infinite family of equivalent potentials, a canonical form of $\mathcal{H}$ with a minimal number of terms ? The situation is a bit like normal forms in bifurcations theory where variable changes allows one to eliminate locally non resonant terms in the Taylor expansion of the vector field (Arnold, 1983). Here, the role of the variable changes is played by f. By suitable choices of f one should be able to eliminate some monomials in the expansion (4.1). An evident situation corresponds to monomials related by time translation, e.g. $\omega_i(0)$ and $\omega_i(1)$: since any $\nu \in \mathcal{M}$ is time translation invariant $\nu\left[\omega_i(0)\right] = \nu\left[\omega_i(1)\right]$, the firing rate of neuron i. Such monomials correspond to the same constraint in (2.25) and can therefore be eliminated. A potential where monomials, related by time translation, have been eliminated (the corresponding h_l vanishes) is called *canonical*. A canonical potential contains thus, in general, $2^{NR} - 2^{N(R-1)}$ terms. We now show that canonical potentials cannot be further reduced.

4.2.2 Canonical interactions cannot be eliminated using the equivalence equation (4.2)

Assume that we are given two potentials $\mathcal{H}^{(1)}, \mathcal{H}^{(2)}$ in the canonical form, where $\mathcal{H}^{(1)}$ has a zero coefficient for the canonical interaction m_l whereas $\mathcal{H}^{(2)} = \mathcal{H}^{(1)} + h_l\, m_l$, $h_l \neq 0$. Let us show that these two potentials are not equivalent. For this we need to introduce a bit of notations further used in the text.

Since a monomial is defined by a set of spike events (k_u, n_u), one can associate to each monomial a spike block or "mask" where the only bits '1' are located at (k_u, n_u), $u = 1, \ldots, r$. This mask has therefore an index. Whereas the labeling of monomials in (4.1) was arbitrary, m_l denotes from now on the monomial with mask $\omega^{(l)}$. Let us define the block *inclusion* $\sqsubseteq$, by $\omega_0^D \sqsubseteq \omega'^D_0$ if $\omega_k(n) = 1 \Rightarrow \omega'_k(n) = 1$, with the convention that the block of degree 0, $\omega^{(0)}$, is included in all blocks. Then, for two integers l, l':

$$m_{l'}(\omega^{(l)}) = 1 \text{ if and only if } \omega^{(l')} \sqsubseteq \omega^{(l)}. \tag{4.3}$$

Now, from (4.2), $\mathcal{H}^{(2)} = \mathcal{H}^{(1)} + h_l\, m_l$ and $\mathcal{H}^{(1)}$ are therefore equivalent if one can find a D-dimensional function f such that, $\forall \omega_0^D$:

$$f\left(\omega_1^D\right) - f\left(\omega_0^{D-1}\right) + \Delta + h_l \mathbb{1}_{\omega^{(l)} \sqsubseteq \omega_0^D} = 0,$$

where $\mathbb{1}_{\omega^{(l)} \sqsubseteq \omega_0^D}$ is the standard indicator function that takes value 1 when $\omega^{(l)} \sqsubseteq \omega_0^D$ and 0 otherwise. Let us consider 2 specific blocks. The block only composed by '1's contains all other blocks, and it is translation invariant so that the terms involving f cancel in the equation above. We have therefore $\Delta + h_l = 0$. The block only composed by '0's is also translation invariant and, if $l > 0$ we obtain $\Delta = 0$, so that $h_l = 0$, in contradiction with the hypothesis. Therefore, two canonical potentials are equivalent *if and only if* all their canonical their coefficients h_l, $l > 0$, are equal (Cessac and Cofré, 2013).

There is still an arbitrariness due to the term h_0 ("Gauge" invariance). One can set it equal 0 without loss of generality. In this way, there is only one canonical potential, with a minimal number of monomials, corresponding to a given stationary Markov chain.

In the next section we show how to compute the coefficients of the canonical potential $\mathcal{H}^{(2)}$ equivalent to a known $\mathcal{H}^{(1)}$.

4.3 Method

Given a spike block $\omega^{(l_0)}$, a *periodic orbit* of period κ is a sequence of spike blocks $\omega^{(l_n)}$ where $\omega^{(l_{p\kappa+n})} = \omega^{(l_n)}$, $p \geq 0$, $0 \leq n \leq \kappa - 1$. From equation (4.2) we have, for such a periodic orbit,

$$\sum_{n=0}^{\kappa-1} \mathcal{H}^{(2)}\left(\omega^{(l_n)}\right) = \sum_{n=0}^{\kappa-1} \mathcal{H}^{(1)}\left(\omega^{(l_n)}\right) + \kappa\Delta, \tag{4.4}$$

because the f-terms disappear when summed along a periodic orbit. It follows that the sum of a potential along a periodic orbit is an invariant (up to the constant term $\kappa\Delta$) in the class of equivalent potentials. This is a classical result in ergodic theory extending to infinite range potentials (Livšic, 1972). This equation is valid whatever periodic orbit is considered. It is singularly useful if one takes advantage of an existing hierarchy between blocks and between monomials, the Hammersley-Clifford (H-C) hierarchy (Hammersley and Clifford, 1971) that we explain now.

4.3.1 Hammersley-Clifford hierarchy

The construction of our method is based on the (H-C) factorization theorem, proved in the seminal although unpublished paper (Hammersley and Clifford, 1971). Later simpler proofs were given in (Grimmett, 1973, Besag, 1974). This result establishes the equivalence between Markov random fields and Gibbs distributions. It was proved in the context of undirected graphs where the clique structures provide the factorization of the potential. Our result is based on a decomposition of the potential over inclusions $\sqsubseteq$ of (spatio-temporal) blocks defined previously. Inclusions provide a hierarchical structure similar to the blackening algebra of (H-C). However (H-C) theorem does not provide by itself an explicit method to obtain from a Markov chain the corresponding canonical MaxEnt potentials. On the opposite, our method of periodic orbits allows to perform this computation.

We can express $\mathcal{H}^{(2)}$ in the form (4.1), then using (4.3) it follows that (4.4) becomes:

$$\sum_{n=0}^{\kappa-1} \sum_{l \sqsubseteq l_n} h_l^{(2)}\, m_l(\omega^{(l_n)}) = \sum_{n=0}^{\kappa-1} \mathcal{H}^{(1)}(\omega^{(l_n)}) + \kappa\Delta \tag{4.5}$$

where, with a slight abuse of notations $l \sqsubseteq l_n$ stands for $\omega^{(l)} \sqsubseteq \omega^{(l_n)}$.

The interesting fact about this representation is that the l.h.s of this equation is written entirely in terms of blocks included in the blocks considered in the periodic orbits. Therefore, in order to compute all the coefficients h_l's that characterize the canonical MaxEnt potential we can proceed by first obtaining the coefficient of degree 0, then the coefficients of degree 1,2, and so on. We use equation (4.5) to compute from a known $\mathcal{H}^{(1)}$ potential its associated canonical potential $\mathcal{H}^{(2)}$. From now on we focus in the particular case when $\mathcal{H}^1 = \phi$ is a normalized potential. To alleviate notation we note $\mathcal{H}^{(2)} = \mathcal{H}$.

Degree 0: Free energy
Start from the first mask in hierarchy, the mask $\omega^{(0)}$ containing only 0's, whose corresponding monomial is $m_0 = 1$ and consider its periodic orbit, of period $\kappa = 1$, $\{\, \omega^{(0)} \,\}$. The application of (4.5) gives $h_0 = \phi(\omega^{(0)}) + \mathcal{P}\,[\,\mathcal{H}\,]$ and since we choose $h_0 = 0$ for the canonical potential we obtain a direct way to compute the free energy of $\mathcal{H}$.

$$\mathcal{P}\,[\,\mathcal{H}\,] = -\phi(\omega^{(0)}) \tag{4.6}$$

Degree 1: Local Fields
Let us now consider masks of degree 1:

$$\omega^{(l_0)} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

(where dots correspond to 0) corresponding to the monomial $\omega_i(D)$. Also consider the periodic orbit obtained by a R-circular shift of this block ($\kappa = R$):

$$\omega^{(l_0)} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \to \omega^{(l_1)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

$$\cdots \to \omega^{(l_D)} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Since the corresponding monomials of the blocks in the orbit $\omega^{(l_0)}, \omega^{(l_1)}, \ldots, \omega^{(l_D)}$ are related by time translation they correspond to the same constraint in (2.25). The coefficient of all but one of these monomials is therefore set to 0 in the canonical potential $\mathcal{H}$. We use the convention to keep the monomial m_{l_0} whose mask contains a 1 in the right most column. This convention extends to the monomials considered below. The block $\omega^{(l)}$ considered to generate this periodic orbit has one spike corresponding to neuron i. To make this explicit we note $h_l \equiv \mathsf{h}_i$: Then, applying (4.5) to this periodic orbit we obtain:

$$\begin{aligned} \mathsf{h}_i = {} & \phi(\omega^{(l_0)}) + \phi(\omega^{(l_1)}) + \cdots + \phi(\omega^{(l_D)}) \\ & + R\phi(\omega^{(0)}) \end{aligned} \tag{4.7}$$

We have thus obtained the coefficient corresponding to the monomial $\omega_i(0)$ which is precisely h_i in Ising model (2.36).

Considering a different $\omega^{(l_0)}$ of degree 1 and the periodic orbit generated by its R-circular shift we get another local field term. We do the same for the N neurons.

Degree 2: Instantaneous pairwise interactions

Let us now consider instantaneous pairwise interactions. We consider masks of the form :

$$\omega^{(l_0)} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.,$$

corresponding to the monomial $\omega_i(D)\omega_j(D)$, the procedure is the same as above i.e. take the periodic orbit:

$$\omega^{(l_0)} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \to \omega^{(l_1)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

$$\cdots \to \omega^{(l_D)} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

The coefficients corresponding to this monomials are J_{ij} in the Ising model (2.36). We have, from (4.5):

$$\mathsf{J}_{ij} = \sum_{n=0}^{R-1} \phi\left(\omega^{(\sigma^n l)}\right) + R\phi(\omega^{(0)}) - \sum_{n=0}^{R-1} \sum_{l'_n \sqsubset \sigma^n l} h_{l'_n}. \tag{4.8}$$

For blocks $l'_n \sqsubset \sigma^n l$ of degree 1 the spike is either on neuron i or neuron j. The contribution of these blocks is $h_i + h_j$. In the blocks $l'_n \sqsubset \sigma^n l$ there is also the block $\omega^{(0)}$, whose contribution is $h_0 = 0$. Therefore, we finally have:

$$\mathsf{J}_{ij} = \sum_{n=0}^{R-1} \phi\left(\omega^{(\sigma^n l)}\right) + R\phi(\omega^{(0)}) - h_i - h_j. \tag{4.9}$$

Degree 2: (1 time-step memory):

For the one step of memory pairwise coefficients (e.g. $\omega_j(0)\omega_i(1)$) the situation is slightly different;

$$\omega^{(l_0)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix},$$

Here the periodic orbit generated by the R-circular shift of $\omega^{(l_0)}$ is:

$$\omega^{(l_0)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix} . \to \omega^{(l_1)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

$$\omega^{(l_2)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix} \to \cdots \to \omega^{(l_D)} = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots \end{pmatrix}$$

This orbit is not sufficient because it contains 2 unknowns in eq (4.5), namely the first and second blocks $\omega^{(l_0)}$ and $\omega^{(l_1)}$ correspond to monomials $\omega_j(D-1)\omega_i(D)$ and $\omega_j(D)\omega_i(0)$ which are not related by time translation, so correspond to different canonical constraints both having degree 2. Therefore it is not possible to solve (4.5) just generating one circular periodic orbit. Fortunately is possible to circumvent this problem by generating additional periodic orbits.

Let us now consider the following periodic orbit:

$$\omega^{(l_0)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix} \rightarrow \omega^{(l_1)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

$$\omega^{(l_2)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix} \rightarrow \omega^{(l_3)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

$$\omega^{(l_4)} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix} \cdots \rightarrow \omega^{(l_{R+2})} = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ 0 & 1 & 0 & \vdots \\ \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 1 & \vdots \\ \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & \cdots \end{pmatrix} \tag{4.10}$$

We have generated a periodic orbit, where (4.5) has only one unknown, the coefficient associated to the first block. All the other blocks in the orbit are either of lower degree, thus we have already computed them; or are time translations of the first block, thus their coefficient is set to zero.

This is a particular example of a general procedure that we describe now. It allows to compute hierarchically any h_l. The procedure is general, but we illustrate it with:

$$\omega^{(l)} = \left[\begin{array}{cccccc} 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right]$$

- Step 1. Shift circularly $\omega^{(l)}$ until the left-most spiking pattern has at least a 1. Each of the circular shifts generate a mask, which corresponds to the same constraint in (2.25) so the corresponding h_l coefficient is set to zero.

$$\left[\begin{array}{cccccc} 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right] \rightarrow \left[\begin{array}{cccccc} 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{array} \right] \rightarrow \left[\begin{array}{cccccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \end{array} \right]$$

- Step 2. Continue circularly left shifting but, before shifting, remove the 1 with the lower neuron index, on the left most spike pattern. Tag the 1s that has been removed. Do this until the total number of left shifts including step 1 and 2 is R.

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

- Step 3. Same as step 1. All the masks generated at this step correspond to the same constraint and thus have a zero coefficient.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

- Step 4. Do the opposite of what was done in step 2: Restore the 1's that has been removed on the left most spike pattern while circularly shifting. In this way we finally regenerate $\omega^{(l)}$.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\to \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

As claimed we have generated a periodic orbit where all monomials, but $\omega^{(l)}$, have either a coefficient 0 or have a degree smaller than $\omega^{(l)}$ and have therefore been already computed. Obviously, when getting to larger and larger degrees the method becomes rapidly intractable because of the exponential increase in the number of terms. The hope is that the influence of monomials decays rapidly with their degree. Additionally, applying it to real data where transition probabilities are not exactly known leads to severe difficulties. Our goal here was to answer the first question asked in the introduction. This goal is now achieved.
We now switch to the second question.

4.4 Example: The discrete time Leaky Integrate and Fire model

In this section we illustrate our result in a stochastic leaky Integrate-and-Fire model with noise and stimulus (Soula et al., 2006) analyzed rigorously in (Cessac, 2011a).

This model is a discretization of the usual leaky Integrate-and-Fire model. Its dynamics reads:

$$V(t+1) = F(V(t)) + \sigma_B B(t), \tag{4.11}$$

where $V(t) = (\, V_i(t)\,)_{i=1}^N$ is the vector of neuron's membrane potential at time t; $F(V)$ is a vector-valued function with entries:

$$F_i(V) = \gamma V_i(1 - S\,[\,V_i\,]) + \sum_{j=1}^{N} W_{ij} S\,[\,V_j\,] + I_i, \quad i = 1 \dots N$$

where $\gamma \in [0, 1[$, is the (discrete-time) "leak rate [1]"; S is a function characterizing the neuron's firing: for a firing threshold $\theta > 0$, $S(x) = 1$ whenever $x \geq \theta$ and $S(x) = 0$ otherwise; I_i is an external current. In the most general version of this model, I_i depends on time. Here, we focus on the case where I_i is constant, ensuring the stationarity of dynamics.

Finally, in (4.11), $\sigma_B > 0$ is a variable controlling the noise intensity, where the vector $B(t) = (B_i(t))_{i=1}^N$ is an additive noise. It has Gaussian independent and identically distributed entries with zero mean and variance 1.

4.4.1 The normalized potential

The normalized potential of the model (4.11) has infinite range. Indeed, a neuron has memory only back to the last time when it has fired. But this time is unbounded (although the probability that the last firing time arises before time m decays exponentially fast as $m \to -\infty$). Nevertheless, the exact potential can be approximated by the finite range potential (Cessac, 2011a).

$$\begin{aligned} \phi(\omega_0^D) \;=\; & \sum_{k=1}^{N} \Big[\omega_k(D)\, \log \pi \left(X_k \left(\omega_0^{D-1} \right) \right) \\ & + (\, 1 - \omega_k(D)\,) \log \left(1 - \pi(X_k \left(\omega_0^{D-1} \right) \right) \Big], \end{aligned} \tag{4.12}$$

where the function π is:

$$\pi(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{\frac{-u^2}{2}} du.$$

All functions appearing below depend on the spike block ω_0^{D-1} and make explicit the dependence of the network state (membrane potentials) on the spike history of the network.

The term:

$$X_k \left(\omega_0^{D-1} \right) = \frac{\theta - V_k^{(d)} \omega_0^{D-1}}{\sigma_k(\omega_0^{D-1})}, \tag{4.13}$$

contains the network spike history dependence of the neuron k at time D. More precisely, the term $V_k^{(d)} \omega_0^{D-1}$ contains the deterministic part of the membrane potential of neuron k at time D, given the network spike history ω_0^{D-1}, whereas $\sigma_k(\omega_0^{D-1})$

[1] Thus, it corresponds to $\gamma = 1 - \frac{dt}{RC}$ in the continuous-time LIF model.

characterizes the variance of the integrated noise in the neuron k's membrane potential. We have:

$$V_k^{(d)}\omega_0^{D-1} = \sum_{j=1}^{N} W_{kj}\,\eta_{kj}\left(\omega_0^{D-1}\right) + I_k\frac{1-\gamma^{D-\tau_k\left(\omega_0^{D-1}\right)}}{1-\gamma}.$$

The first term is the network contribution to the neuron k's membrane potential, where:

$$\eta_{kj}\left(\omega_0^{D-1}\right) = \sum_{l=\tau_k\left(\omega_0^{D-1}\right)}^{D-1} \gamma^{D-1-l}\omega_j(l),$$

is the sum of spikes emitted by j in the past, with a weight γ^{D-1-l} corresponding to the leak decay of the spike influence as time goes on. The notation $\tau_k\left(\omega_0^{D-1}\right)$ means the last time before $D-1$ where neuron k has fired, with the convention that this time is 0 if neuron k didn't fire between 0 and $D-1$ in the block ω_0^{D-1}. In the definition of $\eta_{kj}\left(\omega_0^{D-1}\right)$ we sum from $\tau_k\left(\omega_0^{D-1}\right)$: this is because the membrane potential of neuron k is reset whenever k fires, hence loosing the memory of its past. Finally, in (5.28), we have:

$$\sigma_k^2(\omega_0^{D-1}) = \sigma_B^2\frac{1-\gamma^{2\left(D-\tau_k\left(\omega_0^{D-1}\right)\right)}}{1-\gamma^2}.$$

(see (Cessac, 2011a) for details)

4.4.2 Explicit calculation of the canonical Maximum Entropy Potential

The goal now is to derive from (4.12) a canonical potential $\mathcal{H}$ of the form (4.1) whose spike interactions terms h_l's are functions of the network parameters: the synaptic weight matrix $\mathcal{W}$ and the external stimulus $\mathcal{I}$, $h_l \equiv h_l(\mathcal{W},\mathcal{I})$.

Equation (4.5) gives a relation between the normalized potential and an equivalent non-normalized potential. From this equation, after considering the elimination of equivalent interactions is it possible to compute explicitly the values of the interaction terms h_l's.

Free energy:
From (4.6) and (4.12) we get the free energy:

$$-\phi(\omega^{(0)}) = \mathcal{P}[\mathcal{H}] = -\sum_{k=1}^{N}\log\left[1-\pi\left(\frac{\theta - I_k\frac{1-\gamma^D}{1-\gamma}}{\sigma_B\sqrt{\frac{1-\gamma^{2D}}{1-\gamma^2}}}\right)\right].$$

Local fields:
They are computed using equation (4.7). We consider the periodic orbit obtained by the R-circular shift of the block corresponding to the monomial $\omega_i(D)$. We have

therefore to compute $\phi(\omega^{(l_0)}) + \phi(\omega^{(l_1)}) + \cdots + \phi(\omega^{(l_D)})$ using equation (4.12). To obtain this quantity we have compute X_k (5.28) for all the blocks in the periodic orbit. Note that X_k does not depend on the last column of the blocks in the orbit. We abuse the notation by writing $X_k\left(\omega^{(\sigma^n l)}\right)$ instead of $X_k\left(\omega_0^{(\sigma^n l)D-1}\right)$. The same holds for $\eta_{kj}\left(\omega^{(\sigma^n l)}\right)$ and $\sigma_k(\omega^{(\sigma^n l)})$. We obtain:

$$X_k\left(\omega^{(\sigma^n l)}\right) = \begin{cases} \frac{\theta - W_{ki}\,\gamma^{n-1} - I_k \frac{1-\gamma^D}{1-\gamma}}{\sigma_B \sqrt{\frac{1-\gamma^{2D}}{1-\gamma^2}}}, 1 \leq n \leq R-1, k \neq i; \\ \frac{\theta - W_{kk}\,\gamma^{n-1} - I_k \frac{1-\gamma^n}{1-\gamma}}{\sigma_B \sqrt{\frac{1-\gamma^{2n}}{1-\gamma^2}}}, 1 \leq n \leq R-1, k = i; \\ \frac{\theta - I_k \frac{1-\gamma^D}{1-\gamma}}{\sigma_B \sqrt{\frac{1-\gamma^{2D}}{1-\gamma^2}}}, \forall k,\, n = 0. \end{cases} \tag{4.14}$$

Combining equations (4.7) , (4.12) and (4.14) we obtain:

$$\begin{array}{c} \mathcal{H}_i = \sum_{n=1}^{R-1} \sum_{k=1}^{N} \log\left[1 - \pi\left(X_k\left(\omega^{(\sigma^n l)}\right)\right)\right] + \\ \sum_{k \neq i} \log\left[1 - \pi\left(X_k\left(\omega^{(\sigma^0 l)}\right)\right)\right] + \\ \log\left[\pi\left(X_i\left(\omega^{(\sigma^0 l)}\right)\right)\right] - R\phi(\omega^{(0)}). \end{array} \tag{4.15}$$

which is an explicit function of synaptic weights and stimuli. Clearly:

- The "local field" of a neuron i depends non linearly on *all* stimuli (not only I_i).
- It depends non linearly on the incoming synaptic weights connected to i.

Pairwise interactions (instantaneous):
We get:

$$X_k\Big(\omega^{(\sigma^n l)}\Big)=\begin{cases}\frac{\theta-(W_{ki}+W_{kj})\,\gamma^{n-1}-I_k\frac{1-\gamma^D}{1-\gamma}}{\sigma_B\sqrt{\frac{1-\gamma^{2D}}{1-\gamma^2}}}, 1\leq n\leq R-1, k\neq i,j;\\ \frac{\theta-(W_{kk}+W_{kj})\,\gamma^{n-1}-I_k\frac{1-\gamma^n}{1-\gamma}}{\sigma_B\sqrt{\frac{1-\gamma^{2n}}{1-\gamma^2}}}, 1\leq n\leq R-1, k=i;\\ \frac{\theta-(W_{kk}+W_{ki})\,\gamma^{n-1}-I_k\frac{1-\gamma^n}{1-\gamma}}{\sigma_B\sqrt{\frac{1-\gamma^{2n}}{1-\gamma^2}}}, 1\leq n\leq R-1, k=j;\\ \frac{\theta-I_k\frac{1-\gamma^D}{1-\gamma}}{\sigma_B\sqrt{\frac{1-\gamma^{2D}}{1-\gamma^2}}}, \forall k, n=0.\end{cases} \tag{4.16}$$

Plugging (4.16) in (4.12) and using (4.9), one finally obtains J_{ij} as a explicit function of synaptic weights and stimulus.

Remarks:

- The "instantaneous pairwise" interaction J_{ij} depends not only on W_{ij}, but in all synaptic weights of neurons connected with i or j.
- It also depends in the stimulus of all neurons in the network.

Pairwise interactions (1 time-step):
As mentioned in the previous section, in order to compute this term, the periodic orbit obtained by the R-circular shift of the block $\omega^{(l_0)}$ corresponding to the monomial $\omega_i(1)\omega_j(0)$ is not sufficient. We have therefore to use the periodic orbit obtained by our procedure (4.10). From $\omega^{(l_1)}$ to $\omega^{(l_4)}$ we have already computed their corresponding value $X_k\left(\omega^{(\sigma^n l)}\right)$ when computing the Local fields. From $\omega^{(l_4)}$ to $\omega^{(l_{R+2})}$ we just circularly shift $\omega^{(l_4)}$. We compute the corresponding $X_k\left(\omega^{(\sigma^n l)}\right)$:

$$X_k\Big(\omega^{(\sigma^n l)}\Big)=\begin{cases}\frac{\theta-W_{ki}\gamma^{n-4}-W_{kj}\gamma^{n-3}-I_k\frac{1-\gamma^D}{1-\gamma}}{\sigma_B\sqrt{\frac{1-\gamma^{2D}}{1-\gamma^2}}}, 4\leq n\leq R+2, k\neq i,j;\\ \frac{\theta-W_{kk}\gamma^{n-4}-W_{kj}\gamma^{n-3}-I_k\frac{1-\gamma^{n-3}}{1-\gamma}}{\sigma_B\sqrt{\frac{1-\gamma^{2(n-3)}}{1-\gamma^2}}}, 4\leq n\leq R+2, k=i;\\ \frac{\theta-W_{kk}\gamma^{n-4}-W_{ki}\gamma^{n-3}-I_k\frac{1-\gamma^{n-2}}{1-\gamma}}{\sigma_B\sqrt{\frac{1-\gamma^{2(n-2)}}{1-\gamma^2}}}, 4\leq n\leq R+2, k=j;\end{cases} \tag{4.17}$$

We then apply equation (4.5) to obtain the desired term, from previously computed interaction terms.

A numerical illustration of our method is presented in figure (4.2). We start from the normalized potential (4.12) and construct the canonical equivalent potential. We then compare the conditional probability of patterns predicted by $\mathcal{H}$ with the empirical probabilities inferred from a spike train generated by (4.12). This is just an illustration, and not a systematic study. Note that this numerical analysis is limited to small N, R since the number of terms in $\mathcal{H}$ grows exponentially fast, rendering intractable the method for $NR \geq 20$.

4.5 Conclusion

We have presented a method capable to recover explicitly the MaxEnt potential associated to a set of transition probabilities of a Markov chain. In other words, we have found a way to revert the well known mapping (presented in detail in chapter 2, see also the example 2.2.7.2) from MaxEnt potentials (bounded and finite range) to transition probabilities corresponding to a Markov chain. When the normalized potential ϕ is derived from a neuro-mimetic model (e.g. eq. (4.12)), it follows that the "local fields" h_i depends non linearly on the complete stimulus $\mathcal{I}$ (not only the stimulus applied to neuron i), and the synaptic weights matrix $\mathcal{W}$. This is not that surprising. Even considering an Ising model of two neurons with no memory, a strong favorable pairwise interaction between the two neurons will increase the average firing rate of both neurons, even in the absence of an external field. Likewise, J_{ij} depends on the whole synaptic weights matrix $\mathcal{W}$ and not only on the connection between i and j. This example clearly shows that there is no straightforward relation between the so-called "functional connectivity" in Ising model J_{ij} and the neural synaptic connectivity (W_{ij}).

Our method allows a mechanistic and causal understanding of the origin of correlations observed in recordings on retinal ganglion cells using the MaxEnt approach, in consequence, opens up new possibilities allowing a better understanding of the role of different circuit topologies and stimulus on the collective spike train statistics.

Our result is not limited to spike trains and could also impact different areas of scientific knowledge where binary time series are considered.

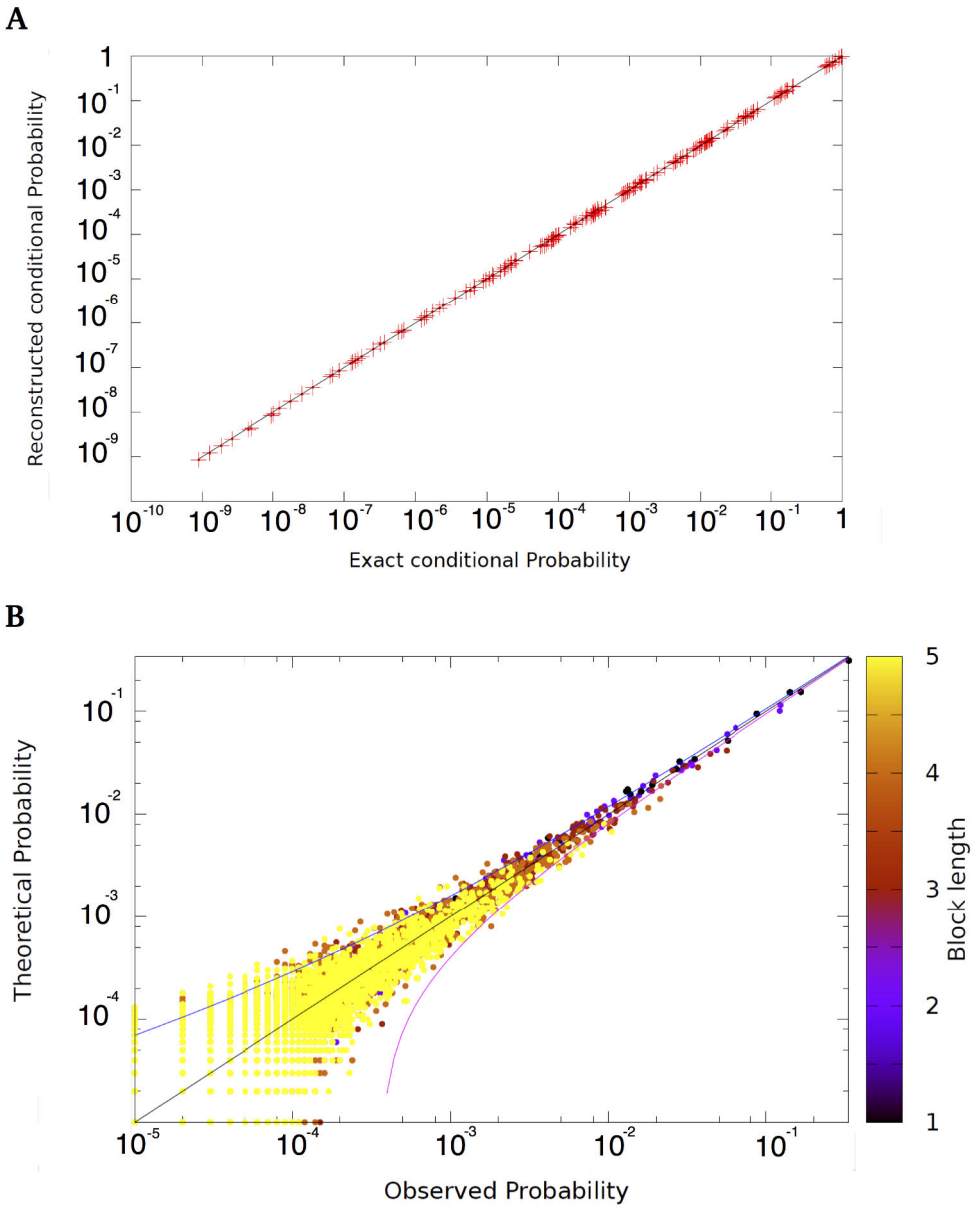


Figure 4.2: (A). Exact conditional probabilities for blocks of range R obtained from the normalized potential (4.12), vs exact conditional probabilities associated with the potential (4.1). (B) Empirical probabilities of blocks ω_0^k, $k = 1, \dots 5$ (darker lower length), obtained from a discrete leaky integrate and fire spike train of size $T = 10^5$ vs the probabilities of the same blocks predicted by the Gibbs ditribution with potential $\mathcal{H}$ (4.1). Each dot stands for one of the 2^{Nk} spatio-temporal patterns, where k is the block length. Diagonal shows equality. Confidence bounds (blue and red lines) correspond to fluctuations ruled by Central Limit Thorem. Plot is in log scale. This figure corresponds to $N = 5, R = 3, \gamma = 0.2, \sigma_B = 0.2, \theta = 1, I_k = 0.7$, $k = 1, \dots 5$. The synaptic weights are random and sparse. Each neuron was randomly connected to other 2 neurons whose weights were drawn from a gaussian 0 mean and variance $\frac{J^2}{N}$. In this example $J = 3$.

Chapter 5

Characterizing the collective response of a neural network model to weak stimuli

In this chapter we take advantage of the formalism developed in the previous chapters to propose an extension of the standard approach of modeling the impact of the external stimulus on neural responses considering the network structure and its dynamics in a neural network model. While the current knowledge allows to assess the impact of external stimulus application on single neuron firing rates, we show how this can (formally) be extended to spatio-temporal correlations (more generally observables) in the spike trains they produce. Our theoretical approach is based on Gibbs distributions and linear response theory. We exploit properties of the Gibbs potential associated to a neural network model to recover the convolution term characterizing the single neuron response. The convolution is written in terms of a correlation matrix, computed with respect to the spontaneous probability. This matrix can also be used to propose a new definition of "preferred stimulus". We mention at the end of this chapter other possible uses of this matrix in the realm of retinal spike train data analysis.

Contents

5.1 Introduction

A central question in computational neuroscience is whether neurons within a population encode stimuli independently, by modulating their firing rates, or collectively using spatio-temporal patterns of spikes (Haslinger et al., 2013). Classical models to predict firing rates are based on the estimation of kernel functions that shape the firing rates of single neurons. The kernel function that interacts with the stimulus is often referred as the "receptive field" of the neuron (already introduced in chapter 1 and 2) and is usually derived from the mean of the spike-triggered stimulus ensemble (STA) (Pillow et al., 2005, Truccolo et al., 2005). This approach considers each spike as an independent message, rather than considering that the information might be conveyed through spatio-temporal patterns of neural activity distributed across the space and time. At the level of sensory neurons, there is however, clear evidence that information is encoded not by single neurons, but instead, by populations or networks of neurons responding collectively to characteristics of the stimulus. The spatio-temporal patterns of spikes produced in the retina reflect indeed both the intrinsic dynamics of the structured network and the temporal characteristics of the stimulus. The retina provides clear evidence that the stimulus is encoded by spatio-temporal patterns. This encoding strategy is known as *population coding*, and turns out to be common throughout the nervous system (Pouget et al., 2000).

While the importance of studying populations of neurons rather than just single neurons has been recognized for decades, the experimental and theoretical tools to empirically investigate the computational properties of neural populations is still lacking (Macke et al., 2008). Nowadays, thanks notably to the MEA technology and theoretical work done in the last years, it is possible to characterize the spike train statistics of populations of neurons responding to stimuli.

In this chapter we propose a framework to characterize, in the context of a neural network model, the impact of a weak time dependent stimulus application in the spatio-temporal correlations of the spike trains produced by the model.

Although our approach is general, we base our result considering the conductance based Integrate and Fire model introduced in chapter 3, but without gap junctions, the evoked time dependent potential is explicitly computed in terms of the neural network parameters.

5.1.1 Characterizing the single neuron response to sensory stimuli: The experimental approach

A simple approach to characterize the single neuron response to a "sensory" stimulus is called "Linear model". In this approach, the firing rate of a neuron at time t depends on the behavior of the stimulus over a period of time. The influence of the rest of the population here is somewhat hidden in the shape of the response model. In this approach the spikes are supposed to obey a non homogeneous Poisson distribution with rate $r_k(t)$. The basic problem is to construct an estimate $\hat{r}_k(t)$ of the firing rate evoked by a stimulus $I(t)$ (in the case of retinal ganglion cells correspond to images projected on the photoreceptors). The simplest way to construct an estimate is to assume that the firing rate at any given time can be expressed as a weighted sum of the values taken by the stimulus at earlier times. This effect is represented by a convolution. The estimate is Dayan and Abbott (2005):

$$\hat{r_k}(t) = r_0 + D_k * I(t), \tag{5.1}$$

where the term r_0 accounts for a background firing occurring without stimulus $I = 0$. Here D_k is a filter that characterize the receptive field of neuron k, determining how strongly, and with what sign, the value of the stimulus affects the firing rate at time t. Note that the convolution in this equation corresponds to linear filtering.
Different methods can be used to estimate the D_k filters, notably the Spike-Triggered average. A generalization of the linear model is the Linear non linear model (LNL), which assumes that spikes are generated by non homogeneous Poisson process with a time varying rate $\hat{r_k}(t)$, which depends only on the stimulus vector. Subsequently, a monotonic nonlinearity f transforms the real-valued output of the linear filtering into a nonnegative instantaneous firing rate (Simoncelli et al., 2004).

$$\hat{r_k}(t) = f(r_0 + D_k * I(t)). \tag{5.2}$$

Another extension of the previous approach, where firing rate depend not only on the stimulus but also on the history of spikes generated by the neuron is the Generalized linear model (GLM), presented in chapter 2.

5.1.2 Problem setting

In this section, we explain how the convolution form (5.1) can be obtained considering a neural network model via the Gibbs distributions setting.

We suppose that the conductance based neural network model has evolved until time t_0 without external current i.e., we fix all the parameters of the synaptic weights, threshold, etc.. to given values and set the external stimulus to 0. As shown in (Cessac, 2011b) to this set of parameters is associated a unique normalized potential and a unique Gibbs distribution characterizing the "spontaneous" spike train statistics (note that neurons can fire without external stimulation). Considering

now one perturbs the model by including a time dependent external stimulus at time t_0, what is the new "evoked" spike train statistics? This situation is illustrated in figure 5.1.

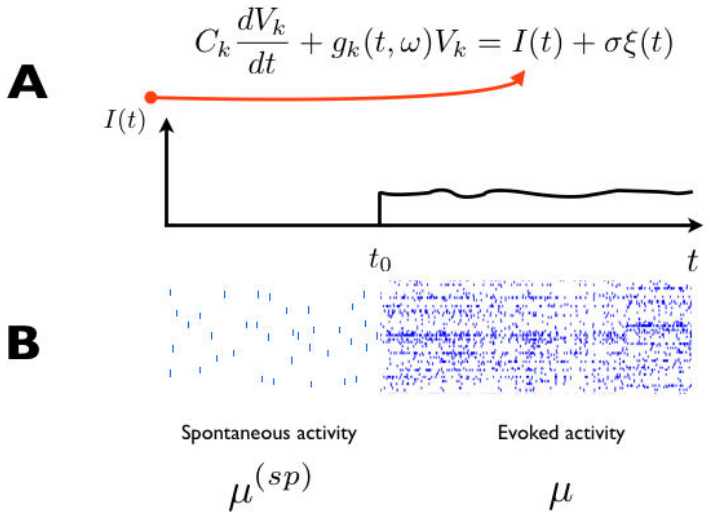


Figure 5.1: A) We consider the Conductance based Integrate and Fire without gap junctions and without current before t_0. B) We assess the impact of the time dependent current injection to the equation in the difference between the spontaneous and evoked probability distributions.

Indeed, this question have been already addressed before in this thesis. What is new in this chapter is the fact that the normalized potential obtained from the conditional probabilities considering a model with time dependent external stimulus can be divided in two parts: one independent of the external stimulus, thus time independent corresponding to spontaneous activity, and the other characterizing the stimulus, time dependent. This fact allows to easily assess the influence of the time dependent stimulus on the statistics of spikes.

An important result in this direction is presented in Brunel and Fourcaud (2002), where the response of the firing rate to weak current injection in an integrate and fire model, is expressed as a convolution, whose kernel is computed analytically in the frequency domain. The analysis is done on the LIF model which does not consider the structure of connectivity between neurons in the network. The approach is based on a small perturbation of the steady state of the Fokker-Planck equation..

An important advantage of our framework is that is that allows us to consider models that takes into consideration any connectivity structure of the network.

Special attention should be taken when comparing our approach with the experimental approach. While in the experimental approach one attempts to fit a

kernel or "Receptive field" from the spike trains and the sensory stimulus (belonging to the external world, in the case of retinal ganglion cells: visual images), in the neural network modeling approach the external stimulus is of different nature and correspond to a time dependent current injection that affect the fluctuations in the membrane potential of the neurons. In the case of retinal ganglion cells, they receive directly current coming from bipolar and amacrine cells, which integrate currents from photoreceptors and horizontal cells.

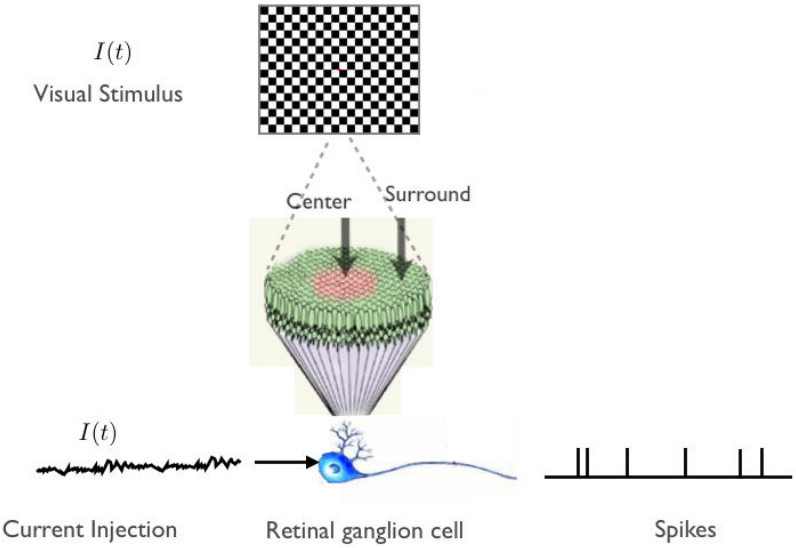


Figure 5.2: Receptive fields of sensory neurons (for instance of retinal ganglion cells) are usually represented as a filter that respond to images presented to the retina. However a retinal ganglion cell does not "see" directly the image, it receives currents from the outer plexiform layer (OPL) of the retina causing variations in its membrane potential due to the network that has been stimulated previously by the visual stimulus..

Our aim is to characterize the average response of an arbitrary observable f to a time dependent perturbation on its membrane potential.

The response of the system to this perturbation can be characterized thanks to the knowledge of the spontaneous Gibbs distribution $\mu^{(sp)}$. This represent a great advantage, because this distribution can be obtained from the MaxEnt principle using the Perron-Frobenius theorem as explained in chapter 2.

In this chapter we look to bridge the effects of a weak external stimulus perturbation to the concepts of receptive fields, convolution kernel, preferred stimulus. In particular we are interested in a more general characterization of these concepts and investigate potential application to real recordings of retinal ganglion cells under different stimuli scenarios. This work is still under development.

5.1.3 Ansatz of the Linear response to weak stimulus

How does the paradigm of filters of a stimulus, obtained from real spiking data characterizing receptive fields of a neuron, connects with the paradigm of Gibbs distributions? Are these two visions coherent?

Given a normalized Gibbs potential ϕ obtained from a neural network model it is possible to compute the linear response to a stimulus considered as a weak perturbation of the spontaneous dynamics (when there is no stimulus). This linear response is characterized by a convolution kernel depending on the observables which could be compared to the models of receptive field filters used in the literature (Cessac and Palacios, 2012). However, one should take care about one point. While the concept of receptive field applies to sensory neurons receiving the external stimulus from the external world (i.e for visual receptive in the cortex V1 an angled bar or a natural image for retinal ganglion cells) in neural network model framework, modeling retinal ganglion cells for instance, we assess the influence of the injection of current affecting the membrane potentials in the spike train statistics . While sensory neurons only feels the external stimulus through current coming from other neurons (see image 5.2).

Under this setting we ask the following: Can we infer the impact of a small time dependent current injection in the correlations of spike trains from the parameters of the Gibbs potential?

We call $\delta\mu\,[\,f(t,.)\,]$ the difference in the mean value of the observable f between the evoked and spontaneous regime.

$$\mu[f(t,\omega)] - \mu^{(sp)}\,[\,f(t,\omega)\,] \;=\; \delta\mu\,[\,f(t,.)\,] \tag{5.3}$$

We assume that the current injection $I^{(ext)}(t) = \varepsilon\,\iota(t)$, where $\varepsilon > 0$ is a small parameter and ι contains the time-dependence of this current. For small ε the variation $\delta\mu\,[\,f(t,.)\,]$ is proportional to ε, provided the exponential decay of the variations of the normalized potential ϕ characterizing the transition probabilities of the neural network (see Cessac (2011b)). We call *linear response of f to the external current $I^{(ext)}$ at time t* the limit (Ruelle, 1999) :

$$\mathcal{R}_t\,[\,f\,] = \lim_{\varepsilon\to 0} \frac{\delta\mu\,[\,f(t,.)\,]}{\varepsilon}.$$

An alternative formulation of $\mathcal{R}_t\,[\,f\,]$, allowing explicit computation, can be formally obtained from derivatives of the free energy. This computation is indeed formal, because the free energy is defined in the context of time-translation invariant dynamics (stationarity) whereas, we consider here a time dependent perturbation. For a rigorous derivation in the stationary dynamics see (Hanus et al., 2002). The situation is the same as considering, in statistical mechanics, say a magnetic system with a non spatially uniform magnetic field. As a consequence of $I^{(ext)}(t) = \varepsilon\,\iota(t)$, the perturbation of the normalized potential $\delta\phi(l,\omega)$ is proportional to ε. In this section

we write $\delta\phi(l,\omega) = \varepsilon\psi(l,\omega)$. From the definition of the derivative we have:

$$\left.\frac{\partial\mathcal{P}\left(\phi^{(sp)}+\varepsilon\psi(n,.)+\lambda f(t,.)\right)}{\partial\varepsilon}\right|_{\varepsilon=0}$$

$$=\lim_{\varepsilon\to 0}\frac{\mathcal{P}\left(\phi^{(sp)}+\varepsilon\psi(n,.)+\lambda f(t,.)\right)-\mathcal{P}\left(\phi^{(sp)}+\lambda f(t,.)\right)}{\varepsilon},$$

Therefore,

$$\mathcal{R}_t[f]=\left.\frac{\partial^2\mathcal{P}\left(\phi^{(sp)}+\varepsilon\psi(n,.)+\lambda f(t,.)\right)}{\partial\lambda\partial\varepsilon}\right|_{\lambda=0,\varepsilon=0}. \tag{5.4}$$

From chapter 2 eq (2.27), where $\psi(n,\omega)=0$ if $n\leq 0$ (no external current for $t\leq 0$), we finally obtain the discrete convolution:

$$\mathcal{R}_t[f]=\sum_{l=1}^{+\infty}\mathcal{C}^{(sp)}[\psi(n-l,.),f(t,.)]=\sum_{l=1}^{n}\mathcal{C}^{(sp)}[\psi(l,.),f(t,.)] \tag{5.5}$$

where the finite sum of correlations is obtained after a change of index and considering the fact that $\psi(n,\omega)=0$ if $n\leq 0$. This equation relates the variation in the average of f, upon application of an external current, to correlation functions measured with respect to the spontaneous Gibbs distribution. This is the *fluctuation-response* relation, classical in non equilibrium statistical physics. The power of eq. (5.4), (5.5) is that the linear response is directly given as the derivative of the free energy, that can be numerically approximated. This computation is general and does not depend on the underlying neural network model.

5.2 The strategy and the model

We focus in the Conductance based Integrate and Fire model introduced in section 3, but without gap junctions. We chose this model because, in this case, we have a simple form to approximate the potential and as shown in (Cessac, 2011b) for any choice of parameters characterizing the sub-threshold dynamics always exist a unique Gibbs measure. We suppose that the system has evolved from $-\infty$ to time 0^-, spontaneously, i.e. the external stimulus is switched off in eq. (5.7) for $t=-\infty$ to $t=0$. Then, at time $t_0=0^+$ one applies a time dependent external stimulus which, for instance, can be a random current (the situation is illustrated in figure 5.1).

Thanks to a Taylor expansion in the normalized potential obtained from the neural network model, around the spontaneous regime, this potential can be separated in two parts: one time and stimulus independent and other time and stimulus dependent. This procedure provides a very convenient representation of the evoked probability measure allowing to assess the impact of the current injection in any correlation caused in the spike trains.

Let us first recall the form of the model without gap junctions.

We may rewrite eq. (3.1) as:

$$C_k \frac{dV_k}{dt} + g_k\,(\,t,\omega\,)\,V_k = i_k(t,\omega), \tag{5.6}$$

Without gap junctions, the term $i_k(t,\omega)$ reads:

$$i_k(t,\omega) = g_{L,k}\,E_L + \sum_{j=1}^{N} W_{kj}\,\alpha_{kj}(t,\omega) \;+\; i_k^{(ext)}(t) \;+\; \sigma_B \xi_k(t), \tag{5.7}$$

where W_{kj} is the synaptic weight:

$$\begin{cases} W_{kj} = E^+ G_{kj}, & \text{if} \qquad j \in \mathcal{E}, \\ W_{kj} = E^- G_{kj}, & \text{if} \qquad j \in \mathcal{I}. \end{cases}$$

We use the same notations as in chapter 3. Assume that the spike train ω is given. Then, it is straightforward to solve the linear equation (5.6). In this case the flow reads 3.19:

$$\Gamma_k(t_1,t_2,\omega) = e^{-\frac{1}{C_k}\int_{t_1}^{t_2} g_k(\,u,\omega\,)\,du}. \tag{5.8}$$

In the case without gap junctions the reset condition introduced in (3.3.1) has the consequence of removing the dependence of neuron k on the past anterior to $\tau_k(t,\omega)$. In this setting, the membrane potential of neuron k at time t is a function of the spike occuring in the network before t. We can decompose the solution of equation (5.6) as done in (Cessac, 2011b):

$$V_k(t,\omega) = V_k^{(det)}(t,\omega) + V_k^{(noise)}(t,\omega). \tag{5.9}$$

5.2.1 Deterministic part of the integrated membrane potential

This is :

$$V_k^{(det)}(t,\omega) = V_k^{(syn)}(t,\omega) + V_k^{(ext)}(t,\omega), \tag{5.10}$$

where:

$$V_k^{(syn)}(t,\omega) \;=\; \frac{1}{C_k}\sum_{j=1}^{N} W_{kj} \int_{\tau_k(t,\omega)}^{t} \Gamma_k(t_1,t,\omega)\alpha_{kj}(t_1,\omega)dt_1, \tag{5.11}$$

is the synaptic contribution to the membrane potential at time t. Moreover,

$$V_k^{(ext)}(t,\omega) \;=\; \frac{E_L}{\tau_{L,k}} \int_{\tau_k(t,\omega)}^{t} \Gamma_k(t_1,t,\omega)dt_1 + \frac{1}{C_k} \int_{\tau_k(t,\omega)}^{t} i_k^{(ext)}(t_1)\Gamma_k(t_1,t,\omega)dt_1, \tag{5.12}$$

where we set:

$$\tau_{L,k} \stackrel{\text{def}}{=} \frac{C_k}{g_{L,k}}, \tag{5.13}$$

the characteristic leak time of neuron k. We have included the leak reversal potential term in this "external" term for convenience. Even if there is no external current this term is non zero.

5.2.2 Stochastic part of the integrated membrane potential

The term:

$$V_k^{(noise)}(t,\omega) = \Gamma_k(\tau_k(t,\omega),t,\omega)V_{reset} + V_k^{(B)}(t,\omega), \tag{5.14}$$

is the stochastic part of the membrane potential. The first term in the left-hand side comes from the reset of the membrane potential to a random value after reset. The second one is due the integration of synaptic noise from $\tau_k(t,\omega)$ to t:

$$V_k^{(B)}(t,\omega) \stackrel{\text{def}}{=} \frac{\sigma_B}{C_k}\int_{\tau_k(t,\omega)}^{t}\Gamma_k(t_1,t,\omega)dB_k(t_1). \tag{5.15}$$

As a consequence, for a fixed ω, $V_k^{(noise)}(t,\omega)$ is the k-th component, at time t, of a N-dimensional Gaussian process $V^{(noise)}(.,\omega)$ with mean zero and covariance:

$$Cov\left[V_k^{(noise)}(t,\omega),\, V_l^{(noise)}(u,\omega)\right] =$$

$$\delta_{kl}\left[\left(\tfrac{\sigma_B}{C_k}\right)^2\int_{\max(\,\tau_k(t,\omega),\tau_k(u,\omega)\,)}^{\min(\,t,u\,)}\Gamma_k(t_1,t,\omega)\,\Gamma_k(t_1,u,\omega)\,dt_1 \;+\; \sigma_R^2\,\Gamma_k^2(\tau_k(t,\omega),t,\omega)\,\delta(t-u)\right].$$

In particular, $V_k^{(noise)}(t,\omega)$ is Gaussian, centered, with variance

$$\sigma_k^2(t,\omega) \stackrel{\text{def}}{=} \left(\frac{\sigma_B}{C_k}\right)^2\int_{\tau_k(t,\omega)}^{t}\Gamma_k^2(t_1,t,\omega)\,dt_1 \;+\; \sigma_R^2\,\Gamma_k^2(\tau_k(t,\omega),t,\omega). \tag{5.16}$$

Under the approximation discussed in section 3, the probability of $\omega(n)$, given the past sequence $\omega_{-\infty}^{n-1}$ is given by:

$$\mathbb{P}_n[\omega(n)|\omega_{-\infty}^{n-1}] = e^{\phi(\,n,\omega\,)}, \tag{5.17}$$

with:

$$\phi\,(\,n,\omega\,) = \sum_{k=1}^{N}\phi_k\,(\,n,\omega\,)\,, \tag{5.18}$$

$$\phi_k\,(\,n,\omega\,) = \omega_k(n)\,\log\pi\,(\,X_k(n-1,\omega)\,) + (\,1-\omega_k(n)\,)\,\log\,(\,1-\pi\,(\,X_k(n-1,\omega)\,)\,)\,, \tag{5.19}$$

$$X_k(n-1,\omega) = \frac{\theta - V_k^{(det)}(n-1,\omega)}{\sigma_k(n-1,\omega)}, \tag{5.20}$$

and

$$\pi(x) = \frac{1}{\sqrt{2\pi}}\int_x^{+\infty}e^{-\frac{u^2}{2}}\,du. \tag{5.21}$$

5.2.3 Gibbs distribution

As shown in chapter 3 there is a unique Gibbs distribution associated with ϕ, where for all $m < n \in \mathbb{Z}$, the probability of a spike block ω_m^n, given the past $\omega_{-\infty}^{m-1}$ is:

$$\mathbb{P}[\omega_m^n \mid \omega_{-\infty}^{m-1}] \;=\; e^{\Phi(m,n,\omega)}, \tag{5.22}$$

$$\Phi(m,n,\omega) = \sum_{l=m}^{n}\phi\,(\,l,\omega\,)\,. \tag{5.23}$$

5.2.3.1 Probability of spike blocks and expected values of observables

As a consequence of (5.22),

$$\mu\left[\,\omega_m^n\,\right] = \int_{\mathcal{A}_{-\infty}^{m-1}} e^{\Phi(m,n,\omega)}\mu(d\omega). \tag{5.24}$$

In this equation the integral holds on the set $\mathcal{A}_{-\infty}^{m-1}$ of sequences $\omega_{-\infty}^{m-1}$.

As another consequence of (5.22), for an observable $f(t,\omega)$, setting $n = [\,t\,]$ the integer part of t, the expectation of $f(t,\omega_{-\infty}^n)$ with respect to μ obeys:

$$\mu\left[\,f(t,.)\,\right] = \int f\left(\,t,\omega\,\right)\mu(d\omega) = \sum_{\omega(n)\in\mathcal{A}} \int f\left(\,t,\omega_{-\infty}^{n-1}\omega(n)\,\right) e^{\phi(\,n,\omega\,)}\mu(d\omega), \tag{5.25}$$

and, by induction, for any integer $m < n$:

$$\mu\left[\,f(t,.)\,\right] = \sum_{\omega_m^n\in\mathcal{A}^{n-m+1}} \int f\left(\,t,\omega_{-\infty}^{m-1}\omega_m^n\,\right) e^{\Phi(m,n,\omega)}\mu(d\omega). \tag{5.26}$$

5.2.4 Spontaneous Gibbs potential.

Let us now consider the situation where the system (5.6) has evolved spontaneously, i.e. $i_k^{(ext)}(t)$ is switched off. More generally, it could be a time constant external current. But, there is no loss of generality in assuming that this current vanishes, since a time constant current can be included in (5.12) by e.g. a modification in the term $\frac{E_L}{\tau_{L,k}}\int_{\tau_k(t,\omega)}^t \Gamma_k(t_1,t,\omega)dt_1$.

We denote the corresponding potential $\phi^{(sp)}$ where:

$$\phi_k^{(sp)}\left(\,l,\omega\,\right) = \omega_k(l)\,\log\pi\left(\,X_k^{(sp)}\left(\,l-1,\omega\,\right)\right) + \left(\,1-\omega_k(l)\,\right)\log\left(\,1-\pi\left(\,X_k^{(sp)}\left(\,l-1,\omega\,\right)\right)\right), \tag{5.27}$$

with

$$X_k^{(sp)}\left(\,l-1,\omega\,\right) = \frac{\theta - V_k^{(syn)}(l-1,\omega) - \frac{E_L}{\tau_{L,k}}\int_{\tau_k(l-1,\omega)}^{l-1}\Gamma_k(t_1,l-1,\omega)dt_1}{\sigma_k(l-1,\omega)}. \tag{5.28}$$

The Gibbs distribution corresponding to this potential (statistics in spontaneous activity) is denoted $\mu^{(sp)}$.

5.2.5 Perturbative expansion of the potential

We write $X_k(t,\omega) = X_k^{(sp)}\left(\,t,\omega\,\right) + \delta X_k\left(\,t,\omega\,\right)$ where $X_k^{(sp)}\left(\,t,\omega\,\right)$ is given by (5.28) while,

$$\delta X_k\left(\,t,\omega\,\right) = -\frac{1}{C_k}\frac{\int_0^t i_k^{(ext)}(t_1)\Gamma_k(t_1,t,\omega)dt_1}{\sigma_k(t,\omega)}, \tag{5.29}$$

the time-dependent term, for $t \geq 0$. We have thus

$$\begin{aligned}\phi_k\,(\,l,\omega\,) = \omega_k(l)\,\log\pi\,\Big(\,X_k^{(sp)}\,(\,l-1,\omega\,)+\delta X_k\,(\,l-1,\omega\,)\,\Big)\\ +\,(\,1-\omega_k(l)\,)\,\log\Big(\,1-\pi\,\Big(\,X_k^{(sp)}\,(\,l-1,\omega\,)+\delta X_k\,(\,l-1,\omega\,)\,\Big)\,\Big)\,.\end{aligned}\tag{5.30}$$

In the reminder of this section we write $X_k^{(sp)}, \delta X_k$ instead of $X_k^{(sp)}\,(\,l-1,\omega\,), \delta X_k\,(\,l-1,\omega\,)$ whenever it causes no confusion, to alleviate notations.

$\pi\,\Big(\,X_k^{(sp)}+\delta X_k\,\Big)$ is bounded away from 0, $\forall\omega$, provided $X_k^{(sp)}+\delta X_k > A > -\infty$ with probability $1-\varepsilon$ where $\varepsilon\to 0$ as $A\to-\infty$. We can make a series expansion of $\log\pi(X_k^{(sp)}+\delta X_k)$ at $X_k^{(sp)}$.

$$\log\pi(X_k^{(sp)}+\delta X_k) = \log\pi(X_k^{(sp)}) + \sum_{r=1}^{+\infty} a_r(X_k^{(sp)})\;(\,\delta X_k\,)^r\,,\tag{5.31}$$

$$\log\Big(\,1-\pi(X_k^{(sp)}+\delta X_k)\,\Big) = \log\Big(\,1-\pi(X_k^{(sp)})\,\Big) + \sum_{r=1}^{+\infty} b_r(X_k^{(sp)})\;(\,\delta X_k\,)^r\tag{5.32}$$

where:

$$a_1(x) = \frac{\pi'(x)}{\pi(x)} = -\frac{1}{\sqrt{2\pi}}\frac{e^{-\frac{x^2}{2}}}{\pi(x)},\quad b_1(x) = -\frac{\pi'(x)}{1-\pi(x)},\tag{5.33}$$

a_r and b_r are the $r-th$ derivative of $\log\pi(x)$ and $\log(1-\pi(x))$ respectively.

Therefore,

$$\phi_k\,(\,l,\omega\,) = \phi_k^{(sp)}\,(\,l,\omega\,) + \delta\phi_k\,(\,l,\omega\,),$$

where $\phi_k^{(sp)}\,(\,l,\omega\,)$ is given by (5.27) and

$$\delta\phi_k\,(\,l,\omega\,) = \sum_{r=1}^{+\infty}\delta\phi_k^{(r)}\,(\,l,\omega\,),\tag{5.34}$$

the time dependent perturbation. We have set:

$$\delta\phi_k^{(r)}\,(\,l,\omega\,) = \mathcal{H}_k^{(r)}\,(\,l,\omega\,)\;(\,\delta X_k(l-1,\omega)\,)^r\,,\tag{5.35}$$

with:

$$\mathcal{H}_k^{(r)}\,(\,l,\omega\,)\stackrel{\text{def}}{=}\omega_k(l)\,a_r\,\Big(\,X_k^{(sp)}\,(\,l-1,\omega\,)\Big)\,+\,(\,1-\omega_k(l)\,)\,b_r\,\Big(\,X_k^{(sp)}\,(\,l-1,\omega\,)\Big)\,.\tag{5.36}$$

Those quantities are time-translation invariant. Note that both $\phi_k^{(sp)}\,(\,l,\omega\,)$ and $\phi_k\,(\,l,\omega\,)$ are normalised (log of conditional probabilities).

5.2.6 Influence of the current injection in the average value of an observable f

Setting:

$$\delta\phi\,(\,l,\omega\,)\,=\,\sum_{k=1}^{N}\delta\phi_k\,(\,l,\omega\,)\,, \tag{5.37}$$

and

$$\delta\Phi(1,n,\omega)\,=\,\sum_{l=1}^{n}\delta\phi\,(\,l,\omega\,)\,, \tag{5.38}$$

for $n \geq 1$, we have :

$$\begin{aligned} e^{\Phi(\,1,n,\omega\,)} &= e^{\Phi^{(sp)}(\,1,n,\omega\,)+\delta\Phi(\,n,\omega\,)} \\ &= e^{\Phi^{(sp)}(\,1,n,\omega\,)}\,\left[\sum_{p=0}^{+\infty}\frac{\delta\Phi(1,n,\omega)^p}{p!}\right] \end{aligned} \tag{5.39}$$

Where the last equation is obtained taking the Taylor expansion of $e^{\delta\Phi(\,n,\omega\,)}$. Taking the first order approximation in $\delta\Phi(1,n,\omega)$ from equation (5.39) we get:

$$e^{\Phi(\,1,n,\omega\,)} \sim e^{\Phi^{(sp)}(\,1,n,\omega\,)}\,[\,1\,+\,\delta\Phi(1,n,\omega)\,]\,, \tag{5.40}$$

While $\Phi\,(\,1,n,\omega\,)$ and $\Phi^{(sp)}\,(\,1,n,\omega\,)$ are normalized potentials, using the first order approximation, $e^{\Phi^{(sp)}(\,1,n,\omega\,)}\,[\,1\,+\,\delta\Phi(1,n,\omega)\,]$ is not anymore a normalized potential. We introduce therefore the conditional probability given the past as follows:

$$e^{\Phi(\,1,n,\omega\,)} = \frac{e^{\Phi^{(sp)}(\,1,n,\omega\,)}\,[\,1\,+\,\delta\Phi(1,n,\omega)\,]}{Z[\omega_{-\infty}^{0}]}, \tag{5.41}$$

and,

$$\begin{aligned} Z_n[\omega_{-\infty}^{0}] &= \sum_{\omega_1^n} e^{\Phi^{(sp)}(1,n,\omega)}[1+\delta\Phi(1,n,\omega)] \\ &= 1+\sum_{\omega_1^n}\mathbb{P}[\omega_1^n\mid\omega_{-\infty}^{0}]\delta\Phi(1,n,\omega) \end{aligned} \tag{5.42}$$

Therefore,

$$Z_n[\omega_{-\infty}^{0}] \sim 1+\mathbb{E}^{(sp)}[\delta\Phi(1,n,\omega)\mid\omega_{-\infty}^{0}],$$

where $\mathbb{E}^{(sp)}$ stands for the expected value under the spontaneous probability distribution.

Therefore:

$$e^{\Phi(\,n,\omega\,)} \sim e^{\Phi^{(sp)}(1,n,\omega)}[1+\delta\Phi(1,n,\omega)][1-\mathbb{E}^{(sp)}[\delta\Phi(1,n,\omega)\mid\omega^0_{-\infty}] \tag{5.43}$$

At first order in $\delta\Phi(1,n,\omega)$:

$$\mathbb{P}(\omega_1^n\mid\omega^0_{-\infty}) = e^{\Phi(\,n,\omega\,)} \sim e^{\Phi^{(sp)}(1,n,\omega)}\left(1+\delta\Phi(1,n,\omega)-\mathbb{E}^{(sp)}[\delta\Phi(1,n,\omega))\mid\omega^0_{-\infty})]\right) \tag{5.44}$$

Now, as we are interested in the average of an observable after a small perturbation, we consider equation (5.25), which reads now:

$$\mu[f(t,\omega)] = \int \sum_{\omega_1^n\in\mathcal{A}_1^n} f(t,\omega)\mathbb{P}(\omega_1^n\mid\omega^0_{-\infty})\mu(d\omega)$$

and replace the conditional probability by r.h.s of equation (5.44):

$$\mu[f(t,\omega)] = \mu^{(sp)}[f(t,\omega)]+\mu^{(sp)}[f(t,\omega)\delta\Phi(1,n,\omega)]-\mu^{(sp)}[f(t,\omega)\mathbb{E}[\delta\Phi(1,n,\omega)\mid\omega^0_{-\infty}]]$$

Assuming that $f(t,\omega)$ and $\mathbb{E}^{(sp)}[\delta\Phi(1,n,\omega)\mid\omega^0_{-\infty}]$ are independent, we obtain:

$$\mu^{(sp)}[f(t,\omega)]\mu^{(sp)}\left(\mathbb{E}^{(sp)}[\delta\Phi(1,n,\omega)\mid\omega^0_{-\infty}]\right) = \mu^{(sp)}[f(t,\omega)]\mu^{(sp)}[\delta\Phi(1,n,\omega)]$$

So far, we have not been able to find the conditions under which is natural to consider that $f(t,\omega)$ and $\mathbb{E}^{(sp)}[\delta\Phi(1,n,\omega)\mid\omega^0_{-\infty}]$ are independent

Finally,

$$\mu[f(t,\omega)] = \mu^{(sp)}[f(t,\omega)] + \mu^{(sp)}[f(t,\omega)\delta\Phi(1,n,\omega)] - \mu^{(sp)}[f(t,\omega)]\mu^{(sp)}[\delta\Phi(1,n,\omega)]$$

$$\mu[f(t,\omega)] = \mu^{(sp)}[f(t,\omega)] + C^{(sp)}[f(t,\omega),\delta\Phi(1,n,\omega)] \tag{5.45}$$

where $C^{(sp)}$ stands for the correlation function under the spontaneous probability distribution. Plugging equation (5.38) in the correlation we obtain:

$$\mu[f(t,\omega)] = \mu^{(sp)}[f(t,\omega)] + C^{(sp)}[f(t,\omega),\sum_{l=1}^{n}\delta\phi(l,\omega)] \tag{5.46}$$

As the correlation is a bilinear operator we get:

$$\mu[f(t,\omega)] = \mu^{(sp)}[f(t,\omega)] + \sum_{l=1}^{n} C^{(sp)}[f(t,\omega),\delta\phi(l,\omega)] \tag{5.47}$$

One obtains the stimulus-dependent averages from averages with respect to $\mu^{(sp)}$, i.e. averages in the spontaneous regime.

5.2.7 Explicit form of the convolution kernel

The linear response is characterized by a convolution kernel κ_f, called the *response function*, depending on f:

$$\mathcal{R}_t\,[\,f\,] = \int_0^t \kappa_f\,(\,t - t_1\,)\;\iota(t_1)\,dt_1 = [\,\kappa_f * \iota\,]\,(\,t\,)\,, \tag{5.48}$$

where $\kappa_f\,(\,u\,) = 0$ if $u < 0$ (causality), while here $\iota(t') = 0$ if $t' \leq 0$. Note that the response kernel depends on the observable f.

Let us now write κ_f in a more explicit form. From (5.29) we have:

$$\delta X_k\,(\,t,\omega\,) = -\frac{\int_0^t \iota_k(t_1)\Gamma_k(t_1,t,\omega)dt_1}{C_k\,\sigma_k(t,\omega)}.$$

As $\varepsilon \to 0$, all terms $\delta\phi_k^{(r)}\,(\,l,\omega\,)$ in (5.35) with $r > 1$ can be neglected with respect to the first order term $\delta\phi_k^{(1)}\,(\,l,\omega\,) = \mathcal{H}_k^{(1)}\,(\,l,\omega\,)\;\delta X_k(l-1,\omega)$. Therefore, $\delta\phi(l,\omega) \sim \sum_{k=1}^N \delta\phi_k^{(1)}\,(\,l,\omega\,)$, and since $\delta\phi(l,\omega) = \varepsilon\psi(l,\omega)$,

$$\psi(l,\omega) \sim -\sum_{k=1}^N \int_0^t \frac{\mathcal{H}_k^{(1)}\,(\,l,\omega\,)\;\Gamma_k(t_1,t,\omega)}{C_k\,\sigma_k(t,\omega)}\;\iota_k(t_1)dt_1 \tag{5.49}$$

as $\varepsilon \to 0$.

We have thus $\mathcal{C}^{(sp)}\,[\,\psi(l,.),f(t,.)\,] = -\int_0^t \sum_{k=1}^N \mathcal{C}^{(sp)}\left[\frac{\mathcal{H}_k^{(1)}(\,l,.\,)\,\Gamma_k(t_1,t,.)}{C_k\,\sigma_k(t,.)}, f(t,.)\right]\iota_k(t_1)dt_1$ and, from (5.5),

$$\mathcal{R}_t\,[\,f\,] \;=\; -\sum_{k=1}^N \int_0^t \sum_{l=1}^n \mathcal{C}^{(sp)}\left[\frac{\mathcal{H}_k^{(1)}\,(\,l,.\,)\;\Gamma_k(t_1,t,.)}{C_k\,\sigma_k(t,.)}, f(t,.)\right]\iota_k(t_1)dt_1.$$

By identification we obtain the linear response kernel as a vector valued operator with components:

$$\kappa_{k,f}\,(\,t-t_1\,) \;=\; -\sum_{l=1}^n \mathcal{C}^{(sp)}\left[\frac{\mathcal{H}_k^{(1)}\,(\,l,.\,)\;\Gamma_k(t_1,t,.)}{C_k\,\sigma_k(t,.)}, f(t,.)\right].$$

Finally, setting

$$h_k(n,\omega) \;=\; \sum_{l=1}^n \mathcal{H}_k^{(1)}\,(\,l,\omega\,)$$

and using (5.33) so that:

$$h_k(n,\omega) \;=\; \sum_{l=1}^n \pi'\left(X_k^{(sp)}\,(\,l-1,\omega\,)\right)\left[\frac{\omega_k(l)}{\pi\left(X_k^{(sp)}\,(\,l-1,\omega\,)\right)} + \frac{1-\omega_k(l)}{1-\pi\left(X_k^{(sp)}\,(\,l-1,\omega\,)\right)}\right], \tag{5.50}$$

we obtain:

$$\kappa_{k,f}\,(\,t-t_1\,) \;=\; -\mathcal{C}^{(sp)}\left[\frac{h_k(n,.)\,\Gamma_k(t_1,t,.)}{C_k\,\sigma_k(t,.)}, f(t,.)\right].$$

Since $\mu^{(sp)}$, $\Gamma_k(t_1, t, .)$ and $\sigma_k(t, .)$ are invariant under time translation we have:

$$\kappa_{k,f}\,(\,t - t_1\,) \;=\; -\mathcal{C}^{(sp)}\left[\frac{h_k(n - t_1, .)\,\Gamma_k(0, t - t_1, .)}{C_k\,\sigma_k(t - t_1, .)}, f(t - t_1, .)\right]. \tag{5.51}$$

Thus, $\kappa_{k,f}$ is a function of $t - t_1$. This is the general form of the response kernel for the Conductance based Integrate and Fire model, and depends on the system parameters via h_k, Γ_k, on the noise via σ_k and on the observable f

5.2.8 Remarks:

- The response function $\mathcal{R}_t[f]$ is the convolution of the perturbation of the potential and the observable f. It can be interpreted as the response associated to the observable f.
- The response function is written in terms of correlation functions, computed according to the Gibbs distribution associated with the unperturbed system. Although this is a well known result in statistical mechanics, up to our knowledge, it has not been explored previously in the context of spiking neural networks. These correlations can be directly obtained as the derivative of the free energy of the evoked potential, that can be numerically approximated.

5.3 Example

In order to give a concrete example of the use of our results, we consider here a simplified model: the leaky Integrate and Fire model (LIF), compared to the conductance based model, here we consider constant conductances and replace the exponentially decaying alpha profiles by delta functions. In the simplified model, as conductances are constant, the evolution operator does not depend any more on ω:

$$\Gamma_k(t_1, t, \omega) = e^{\frac{-1}{C_k}\int_{t_1}^{t} g_k(u,\omega)du} = e^{-\frac{g_k}{C_k}(t - t_1)}$$

Thus, the deterministic part of the solution of the dynamic equation:

$$V_k^{(det)}(t, \omega) = V_k^{(syn)}(t, \omega) + V_k^{(ext)}(t, \omega)$$

Now reads:

$$\begin{aligned} V_k^{(syn)}(t, \omega) &= \frac{1}{C_k}\sum_{j=1}^{N} W_{kj} \int_{\tau_k(t,\omega)}^{t} \Gamma_k(t_1, t)\delta(t_1 - \hat{t}_j)dt_1 \\ &= \frac{1}{C_k}\sum_{j=1}^{N} W_{kj} \sum_{\tau_k(t,\omega)<\hat{t}_j<t} \Gamma_k(\hat{t}_j, t) \\ &= \frac{1}{C_k}\sum_{j=1}^{N} W_{kj} \sum_{\tau_k(t,\omega)<\hat{t}_j<t} e^{-\frac{g_k}{C_k}(t - \hat{t}_j)} \end{aligned} \tag{5.52}$$

where $\hat{t}_j$ is the time at which neuron j has fired, and

$$V_k^{(ext)}(t,\omega) = \frac{E_L}{\tau_{L,k}} \int_{\tau_k(t,\omega)}^{t} e^{-\frac{g_k}{C_k}(t-t_1)} dt_1 + \frac{1}{C_k} \int_{\tau_k(t,\omega)}^{t} i^{(ext)}(t) e^{-\frac{g_k}{C_k}(t-t_1)} dt_1$$

When the external current is constant ($i^{(ext)}(t) = I$) :

$$V_k^{(ext)}(t,\omega) = (\frac{E_L}{\tau_{L,k}} + \frac{I}{C_k}) \int_{\tau_k(t,\omega)}^{t} e^{-\frac{g_k}{C_k}(t-t_1)} dt_1 = (\frac{E_L}{\tau_{L,k}} + \frac{I}{C_k}) \frac{(1 - e^{-\frac{g_k}{C_k}(t-\tau_k(t-t_1,\omega))})}{\frac{g_k}{C_k}} \tag{5.53}$$

Therefore:

$$X_k^{(sp)}(t,\omega) = \frac{\theta - \frac{1}{C_k}\sum_{j=1}^{N} W_{kj} \sum_{\tau_k(t,\omega)<\hat{t}_j<t} e^{-\frac{g_k}{C_k}(t-\hat{t}_j)} - (\frac{E_L}{\tau_{L,k}} + \frac{I}{C_k}) \frac{(1-e^{-\frac{g_k}{C_k}(t-\tau_k(t-t_1,\omega))})}{\frac{g_k}{C_k}}}{\sigma_k(t,\omega)} \tag{5.54}$$

Moreover,

$$\begin{aligned} \sigma_k^2(t-t_1,\omega) &= (\frac{\sigma_B}{C_k})^2 \int_{\tau_k(t-t_1,\omega)}^{t} \Gamma_k^2(0, t-t_1) dt_1 \\ &= \frac{\sigma_B^2}{C_k^2} \frac{(1 - e^{\frac{-2g_k}{C_k}(t-\tau_k(t-t_1,\omega))})}{\frac{2g_k}{C_k}} \end{aligned} \tag{5.55}$$

Knowing the last firing time of neuron k at time $(t - t_1)$, this random function becomes deterministic.

5.3.1 General response of the observable f

The convolution kernel $\kappa_{k,f}$ equation (5.51) in the simplified model reads:

$$\kappa_{k,f}(t-t_1) = -e^{-\frac{g_k}{C_k}(t-t_1)} \mu^{(sp)} [\frac{\sum_{l=1}^{t-t_1} \pi'(X_k^{(sp)}(l-1,\omega)) [\frac{\omega_k(l) f(t-t_1,\omega)}{\pi(X_k^{(sp)}(l-1,\omega))} + \frac{(1-\omega_k(l)) f(t-t_1,\omega)}{1-\pi(X_k^{(sp)}(l-1,\omega))}]}{C_k \sigma_k(t-t_1,\omega)}]$$

$$+e^{-\frac{g_k}{C_k}(t-t_1)} \mu^{(sp)} [\frac{\sum_{l=1}^{t-t_1} \pi'(X_k^{(sp)}(l-1,\omega)) [\frac{\omega_k(l)}{\pi(X_k^{(sp)}(l-1,\omega))} + \frac{(1-\omega_k(l))}{1-\pi(X_k^{(sp)}(l-1,\omega))}]}{C_k \sigma_k(t-t_1,\omega)}] \mu^{(sp)}[f(t-t_1,\omega)] \tag{5.56}$$

We note the exponential decay with respect to the characteristic time of the neuron $\frac{g_k}{C_k}$. This result is (at first sight) in agreement with Brunel and Fourcaud (2002). This last equation can be simplified. From the integral term appearing in

$X_k^{(sp)}(t,\omega)$ in equation (5.54), we note that necessarily $t > \tau_k(t,\omega)$, thus in equation (5.56) for $X_k^{(sp)}(l-1,\omega)$ we must consider $l-1 > \tau_k(t-t_1,\omega) \to l > \tau_k(t-t_1,\omega)+1$. In that case, the first part of the braket is zero (because, by definition $\omega_k(l) = 0, \forall l > \tau_k(t-t_1,\omega)+1$). Therefore,

$$\kappa_{k,f}(t-t_1) = -e^{-\frac{g_k}{C_k}(t-t_1)}\mu^{(sp)}[\frac{\sum_{l=\tau_k(t-t_1,\omega)+1}^{t-t_1}\frac{\pi'(X_k^{(sp)}(l-1,\omega))}{1-\pi(X_k^{(sp)}(l-1,\omega))}f(t-t_1,\omega)}{C_k\sigma_k(t-t_1,\omega)}] + e^{-\frac{g_k}{C_k}(t-t_1)}\mu^{(sp)}[\frac{\sum_{l=\tau_k(t-t_1,\omega)+1}^{t-t_1}\frac{\pi'(X_k^{(sp)}(l-1,\omega))}{1-\pi(X_k^{(sp)}(l-1,\omega))}}{C_k\sigma_k(t-t_1,\omega)}]\mu^{(sp)}[f(t-t_1,\omega)] \tag{5.57}$$

Taking out of the sum the terms independent of l we get finally:

$$\kappa_{k,f}(t-t_1) = -e^{-\frac{g_k}{C_k}(t-t_1)}\mu^{(sp)}[\frac{f(t-t_1,\omega)\sum_{l=\tau_k(t-t_1,\omega)+1}^{t-t_1}\frac{\pi'(X_k^{(sp)}(l-1,\omega))}{1-\pi(X_k^{(sp)}(l-1,\omega))}}{C_k\sigma_k(t-t_1,\omega)}] + \frac{e^{-\frac{g_k}{C_k}(t-t_1)}\mu^{(sp)}[\sum_{l=\tau_k(t-t_1,\omega)+1}^{t-t_1}\frac{\pi'(X_k^{(sp)}(l-1,\omega))}{1-\pi(X_k^{(sp)}(l-1,\omega))}}{C_k\sigma_k(t-t_1,\omega)}]\mu^{(sp)}[f(t-t_1,\omega)] \tag{5.58}$$

Considering the observable $f(t-t_1,\omega)$ as the firing rate of neuron p i.e. $f(t-t_1,\omega) = \omega_p(t-t_1)$ the convolution kernel (5.58) reads:

$$\kappa_{k,\omega_p}(t-t_1) = -e^{-\frac{g_k}{C_k}(t-t_1)}\mu^{(sp)}[\frac{\omega_p(t-t_1)\sum_{l=\tau_k(t-t_1,\omega)+1}^{t-t_1}\frac{\pi'(X_k^{(sp)}(l-1,\omega))}{1-\pi(X_k^{(sp)}(l-1,\omega))}}{C_k\sigma_k(t-t_1,\omega)}] + e^{-\frac{g_k}{C_k}(t-t_1)}\mu^{(sp)}[\frac{\sum_{l=\tau_k(t-t_1,\omega)+1}^{t-t_1}\frac{\pi'(X_k^{(sp)}(l-1,\omega))}{1-\pi(X_k^{(sp)}(l-1,\omega))}}{C_k\sigma_k(t-t_1,\omega)}]\mu^{(sp)}[\omega_p(t-t_1)] \tag{5.59}$$

This result must be confronted to numerical simulations (in progress).

5.4 Linear Response obtained from the correlation matrix between monomials

In the previous section we have focused on a neural network model whose potential has infinite range and is time dependent. In this section we focus is a much more simple case: finite range and time independent. Then, we can go further in the consequences of the Gibbs distribution setting and linear response. As we have seen

in the previous chapter we can decompose any (finite length) observable in the base of monomials. We decompose here ψ representing the perturbation of the potential $\phi^{(sp)}$

$$\psi = \sum_l \psi_l m_l \tag{5.60}$$

We also decompose the observable f:

$$f = \sum_l f_l m_l$$

Taking equation (5.5) decomposed using the new representation of ψ and develop in the base of monomials, we obtain:

$$\begin{aligned} C_{\psi,f}(n) &= \mu^{(sp)}[\psi \circ \sigma^n f] - \mu^{(sp)}[\psi]\mu[f] \\ &= \sum_{l,l'} \psi_l f_{l'} \mu^{(sp)}[m_l \circ \sigma^n m_{l'}] - \sum_{l,l'} \psi_l f_{l'} \mu^{(sp)}[m_l]\mu^{(sp)}[m_{l'}] \\ &= \sum_{l,l'} \psi_l f_{l'} C_{ll'}(n) \end{aligned} \tag{5.61}$$

Where:

$$C_{ll'}(n) = \mu^{(sp)}[m_l \circ \sigma^n m_{l'}] - \mu^{(sp)}[m_l]\mu[m_{l'}]$$

Summing on n we obtain the following matrix (indexed by integers representing monomials):

$$\chi_{ll'} = \sum_{-\infty}^{\infty} C_{ll'}(n),$$

Therefore, we can write the linear response formula in terms of a quadratic form as follows:

$$\mathcal{R}[f] = \sum_{ll'} \psi_l \chi_{ll'} f_{l'} = \langle \psi, \chi f \rangle \tag{5.62}$$

This is the linear response giving the variation of the average of f when there is a variation ψ of the potential $\phi^{(sp)}$

Remarks

- Note that $C_{ll'}(n)$, and thus, the correlation matrix is obtained with respect to the spontaneous measure $\mu^{(sp)}$.
- The matrix χ can be obtained numerically from simulated spiking data or real recording of retinal ganglion cells (see image 5.3 obtained with the software EnaS *http://enas.gforge.inria.fr/v3/*).

- This formula allows to characterize the receptive field of the observable f, i.e. measure the change in the statistics in the spontaneous and evoked regime.

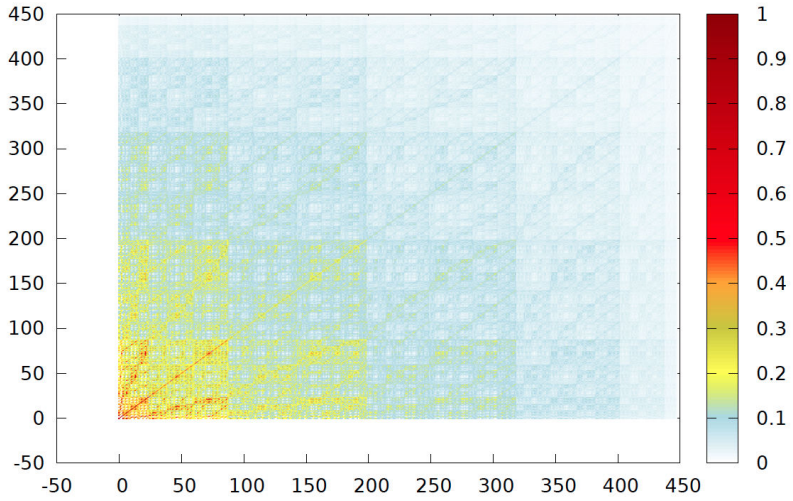


Figure 5.3: Correlation matrix between monomials obtained from simulated data. Each number (in the axes) represent one monomial (this correspondence was explained in chapter 2). For each pair of monomials the temporal correlation is computed. The singular structure appears due to the arbitrary choice of coding blocks to numbers

5.4.1 Eigenvalue decomposition of χ

Noting $|v_k\rangle$ the eigenvectors of χ associated to the real eigenvalues σ_k we can decompose χ as follows:

$$\chi = \sum_k \sigma_k |v_k\rangle\langle v_k| \tag{5.63}$$

Where $\langle v_k|$ represent the transpose of the eigenvector $|v_k\rangle$. Therefore taking equation 5.62:

$$\langle \psi|\chi f\rangle = \sum_k \sigma_k \langle \psi|v_k\rangle\langle v_k|f\rangle \tag{5.64}$$

For a given observable f, $\langle v_k|f\rangle$ is constant.

5.4.2 Consequences and perspectives

Apart from the application concerning the impact of a perturbation in the stimulus in the average value of a given observable, it has been shown recently (Panas et al.,

2014) a novel application of the matrix χ concerning the identification of insensitive regions in parameter space of pairwise maximum entropy models, where the global network statistics is slightly altered. This study only consider a simplified version of χ (only considering spatial monomials) Regions of high sensitivity are also identified. This work is done considering a purely spatial pairwise MaxEnt model. The authors argue that this form of degeneracy endows neuronal networks with the flexibility to continuously remodel and explore large regions of parameter space without compromising stability and function. Using tools exposed in this chapter we can extend this analysis to the spatio-temporal case. Indeed, we are able to compute the matrix χ for spatio-temporal observables. From this matrix the Fisher information matrix can be computed and the same analysis can reveal interesting mechanism taking place when time in taken into account. As the correlation matrix χ can be obtained also from neural network models, a more ambitious project is to link this two approaches.

Another important application comes from the fact that is it possible to derive formally the "preferred stimulus" of an observable from the normalized potential associated to a structured model of spiking neurons? We ask if exist a vector $\vec{\mathcal{I}}$ that maximizes $\langle\vec{\mathcal{I}}|\chi f\rangle$. This vector can be interpreted as the "preferred stimulus of the observable f". In the case when f is a monomial corresponding to a single spike event (rate), this notion should coincide with the receptive field of the associated neuron.

We note $u_k = \langle\vec{\mathcal{I}}|v_k\rangle$, thus:

$$\langle\vec{\mathcal{I}}|\chi f\rangle = \mathcal{G}_f(u) = \sum_k \sigma_k \langle v_k|f\rangle u_k$$

If the solution exist it depends on the observable and the perturbation, as it is expected. Then one looks under which conditions the function $\mathcal{G}_f(u)$ have a maximum. These conditions are under current investigation.

5.5 Conclusion

We have presented a closed formula characterizing the statistical response of a structured neural network model to a small perturbation on the stimulus. Our results allow to assess the impact of this perturbation on any observable defined in the spike train this model produce. One interesting application is found in the generalization of the concept of receptive fields. Our result is obtained from the correlation matrix, computed with respect to the invariant Gibbs distribution corresponding to the neural network model under the unperturbed regime. An interesting point is that this matrix can be computed directly from spiking data, is symmetric and real, thus is possible to obtain a singular value decomposition. From this representation

it is natural to find the principal components, allowing to investigate what is the "favorite stimulus" for each observable. The results of this chapter form part of a more ambitious research project attempting to obtain known results in receptive field models as particular cases of our characterization. This certainly needs further investigation. In particular we are interested to link our results with those obtained in (Sharpee et al., 2004), where using tools of information theory they obtain the maximally informative dimensions of the stimulus for firing rates.

Chapter 6

General conclusion, discussion and future research

6.1 General conclusion

Experimental recordings from MEA provide an unprecedented amount of spiking data, opening up the possibility of studying the statistical structure of neural activity in large populations of neurons responding simultaneously to external stimuli. A more complete understanding of "the code" that neurons use to send spatio-temporal spike patterns to the brain requires not only powerful statistical methods allowing to predict responses, it also requires an "explanatory" idea of the mechanistic and causal structure behind these responses, going in this way, beyond the strictly predictive ability. So far, computational neuroscience lacks a general theory capable to unify and explain the diversity of phenomena observed in spiking neurons under different experimental settings: "neuroscience is data rich and theory poor" (Churchland and Sejnowski, 2002).

The aim of this thesis has been to answer the four very focused questions asked in the introduction. The goal was certainly much more modest than to build a theory of neural circuits or retinal responses, but we believe that is a step forward in that direction.

We repeat the four questions here. Below each question we stress the results and general conclusions. More general discussions and consequences of our results are left to the next section.

Question 1: *Is it possible to characterize the network dynamics and the population spike train statistics in a neural network model?*

We have build a biologically realistic model including both chemical and electric synapses. We have characterized the dynamics and spike train statistics arising from this model. We show that these probabilities are Gibbs distributions depending on the parameters shaping the neural network model. These Gibbs distributions are non-stationary and depend mathematically on an infinite past i.e the spike statistics have an infinite memory and as a consequence is non-Markovian (although Markovian approximation is possible). The fact that the stochastic process is non-Markovian is a consequence of the modeling choice and may not represent a biophysical fact. A question arising from our result is: How far should we consider the past? This question has been addressed mainly in chapter 3.

Question 2: *Are Gibbs distributions good candidates to analyze the spike train statistics from experimental data?*

This question is present throughout this thesis, but is mainly addressed in chapter 2, where we concluded that all canonical models used for spike train statistics we analyzed lead to Gibbs distributions. In chapter 2 we provide a complete theoretical treatment for the generalization of the Maximum Entropy approach to the characterization of the spike train statistics (including spatio-temporal constraints) and show how Gibbs distributions arise from this approach. We have shown in chapter 2 that the introduction of memory requires to define the Gibbs distribution in a more general setting than is usual in statistical physics. We present the Generalized Linear model and compute explicitly its Gibbs potential, we address the issue of conditions ensuring the existence and uniqueness of the Gibbs distributions. In chapter 3 we present a Conductance based Integrate-and-Fire model with electric synapses, we also show an approximation of its Gibbs potential. Gibbs distributions introduce a unifying framework that integrates many aspects considered in different models of spike train statistics.

Question 3: *When attempting to characterize spike train statistics, Gibbs distributions arise from data-based approaches (Maximum entropy models) and from neuro-mimetic model-based approaches (Generalized linear models and conductance based integrate and fire). Is it possible to take advantage of properties of Gibbs distributions to link both approaches?*

To answer question 3 we took advantage of properties of (stationary) Gibbs distributions. Starting from the conditional probabilities of a discrete time Integrate-and-Fire neural network we provide an explicit example for the 1-time step spatio-temporal extension of the Ising model, where we compute explicitly the Maximum Entropy parameters in terms of neural-network parameters. Our method extends to potentials with general spatio-temporal constraints. This method can be further exploited to investigate in more detail the mechanistic origins of spatio-temporal correlations observed in spike train data.

Question 4: *Can we derive a more general notion of receptive fields looking at the difference between the spontaneous and evoked spiking response of a network of neurons using properties of Gibbs distributions?*

Using a neural network model and perturbing the stimulus of this model, we analyze using linear response theory the statistical response of the spiking network to the perturbation. We get a closed formula based on correlations with respect to the unperturbed regime, that represent the average behavior of any observable, generalizing in this way the concept of receptive fields to general spatio-temporal events (not just firing rates). This chapter is under development.

Results of this thesis have contributed to the general understanding of spike train statistics. Our results set a solid ground for future studies attempting to unveil the mechanistic origins from data observed spatio-temporal correlations. Our results have been used in the software EnaS (a C++ library that allow users to manage spike data, perform empirical statistics, modeling and visualizing results) *http://enas.gforge.inria.fr/v3/*.

6.2 Conclusion générale

Les enregistrements expérimentaux obtenus à partir du MEA fournissent une quantité sans précédent de données de potentiel d'action, offrant la possibilité d'étudier la structure statistique de l'activité neuronale dans de grandes populations de neurones qui répondent simultanément à des stimuli externes. Une compréhension plus complète du "code neural" que les neurones utilisent pour envoyer des motifs spatio-temporels au cerveau, exige non seulement des méthodes statistiques puissantes permettant de prédire les réponses, il faut aussi avoir une idée "explicative" de la structure mécanique et causale derrière ces réponses, allant dans ce sens, au-delà de la capacité strictement prédictive.

Jusqu'à présent, dans les neurosciences computationnelles il manque une théorie générale apte à unifier et expliquer la diversité des phénomènes observés dans les neurones pour différents paramètres expérimentaux: "les neurosciences sont riches en données et pauvres en la théorie" (Churchland and Sejnowski, 2002). L'objectif de cette thèse a été de répondre aux quatre questions très ciblées posées dans l'introduction. Le but était certainement beaucoup plus modeste que de construire une théorie des circuits neuronaux de la rétine ou de ses réponses aux stimuli, mais nous croyons que c'est un pas en avant dans cette direction.

Nous répétons les quatre questions ici. En dessous de chaque question, nous soulignons les résultats et les conclusions générales. Des discussions plus générales et les conséquences de nos résultats sont laissées à la section suivante.

Question 1: *Est-il possible de caractériser la dynamique du réseau et les statistiques de potentiel d'action dans un modèle de réseau de neurones?*

Nous avons construit un modèle biologiquement réaliste comprenant à la fois des synapses chimiques et électriques. Nous avons caractérisé la dynamique et les statistiques de potentiel d'action résultant de ce modèle. Nous montrons que ces probabilités sont des distributions de Gibbs en fonction des paramètres du modèle de réseau neuronal. Ces distributions de Gibbs sont non stationnaires et dépendent mathematiquement, sur un passé infini. C'est-à-dire que les statistiques de potentiel d'action ont une mémoire infinie et par conséquent non-markovien (Néanmoins l'approximation Markovienne est possible). Le fait que le processus

stochastique est non Markovien est une conséquence du choix de modélisation et peut ne pas représenter un effet biophysique. Une question qui se pose à partir de notre résultat est: jusqu'où devrions-nous considérer le passé? Cette question a été abordée principalement dans le chapitre 3.

Question 2: *Sont-ils les distributions de Gibbs de bons candidats pour analyser les statistiques de potentiel d'action pour données expérimentales?*

Cette question est présente tout au long de cette thèse, mais s'adresse principalement dans chapitre 2, où nous avons conclu que tous les modèles canoniques utilisées pour les statistiques de potentiel d'action que nous avons analysé donnent distributions de Gibbs. Dans le chapitre 2, nous offrons un traitement théorique complet pour la généralisation de l'approche d'entropie maximale pour la caractérisation des statistiques de potentiel d'action (y compris avec des contraintes spatio-temporelles) et montrons comment les distributions de Gibbs découlent de cette approche. Nous avons montré au chapitre 2 que l'introduction de mémoire nécessite de définir les distribution de Gibbs dans un cadre plus général qu'habituellement en physique statistique. Nous présentons le modèle linéaire généralisé et calculons explicitement son potentiel de Gibbs, nous abordons la question des conditions garantissant l'existence et l'unicité des distributions de Gibbs. Dans le chapitre 3, nous présentons un modèle base en conductance de type Intégre-et-tir avec les synapses électriques, nous montrons également une approximation de son potentiel de Gibbs. Les distributions de Gibbs introduisent un cadre unificateur qui intègre de nombreux aspects pris en compte dans les différents modèles de statistiques de potentiel d'action.

Question 3: *En essayant de caractériser les statistiques de potentiel d'action, les distributions de Gibbs apparaissent naturellement tant dans des approches basées sur des données (modèles d'entropie maximale) que dans des approches basées sur des modèles neuro-mimétiques (modèles linéaires généralisés et Intégre-et-tire). Est-il possible de profiter des propriétés des de Gibbs approches pour relier les différentes approches?*

Pour répondre à la question 3, nous avons utilisé des propriétés des distributions de Gibbs (stationnaires). A partir des probabilités conditionnelles à temps discret obtenus à partir du modèle Intégre-et-tir dans un réseau de neurones, nous fournissons un exemple explicite pour le modèle d'Ising spatio-temporel (avec 1 pas de temps), où nous calculons explicitement les paramètres du potentiel d'entropie maximale canonique en termes de paramètres de réseaux de neurones. Notre méthode s'étend à des potentiels avec des contraintes spatio-temporelles plus générales. Notre méthode peut être exploitée pour étudier plus en détail les origines mécaniques de corrélations spatio-temporelles observées dans les données expérimentales.

Question 4: *Peut-on tirer une notion plus générale de champs réceptifs à partir de la différence entre le regime spontané et celui évoqué par le stimulus dans une modèle de réseau de neurones en utilisant les propriétés des distributions de Gibbs?*

En utilisant un modèle de réseau neuronal et en le perturbant, nous analysons en utilisant la théorie de la réponse linéaire, la réponse statistique du réseau de neurones induite par la perturbation. Nous obtenons une formule basée sur des corrélations obtenues par rapport au régime non perturbé, qui représentent le comportement moyen d'une observable, généralisant ainsi la notion de champs récepteurs à des événements spatio-temporels généraux? (non seulement les taux de tir). Ce chapitre est en cours de développement.

Les résultats de cette thèse ont contribué à la compréhension générale de la statistiques de potentiel d'action. Nos résultats donnent une base solide pour de futures études tentant de dévoiler les origines mécaniques à partir des données observées corrélations spatio-temporelles. Nos résultats ont été utilisés dans le logiciel ENAS (une bibliothèque C ++ qui permettent aux utilisateurs de gérer les données de potentiel d'action, calculer des statistiques empiriques, la modélisation et la visualisation des résultats) http://enas.gforge.inria.fr/v3/.

6.3 Discussion

In this section we discuss some issues appearing in the development of this thesis. We focus in particular on those concerning current approaches to model spike trains of retinal ganglion cells.

6.3.0.1 Difficulty of direct comparison of neural network approach and MaxEnt approach due to binning

There are ongoing debates on the sources of correlations observed between spikes of ganglion cells in the retina: are ganglion cells independent encoders or do they act in a correlated way? (Nirenberg and Latham, 1998, 2003, Roudi et al., 2009, Mastronarde, 1983, Schneidman et al., 2006, Ganmor et al., 2011a,b) What are the sources of correlations (Panzeri and Schultz, 2001, Schneidman et al., 2003) ? Are they only due to a shared stimulus or shared noise, or do they also result from interactions between neurons (Tyrcha et al., 2012) ? We have shown that we can answer these questions in the context of a neural network model. However, it is important to highlight that the conductance based integrate and fire model that we have considered in chapter 3 constitutes an inaccurate model of the retina: it involves only firing cells while most cell types in the retina (amacrine, horizontal, bipolar) do not "spike". Nevertheless, our analysis raises questions about retinal spike trains: if dynamical correlations are so important in this simple model, how can they become weaker in a more complex neural system

like the retina ? If such correlations exist why are they so difficult to exhibit in experiments, without controversy (Schneidman et al., 2006, Roudi et al., 2009) ? Concerning this last question, note that the spike train we obtained is not *binned*. In binning data one makes strong assumptions about the way information is encoded: assuming a fixed temporal resolution and a fixed internal clock defining the bin boundaries (Staude et al., 2010). Note that, by binning, information about the exact spike times is lost. Therefore, any posterior analysis will have limited accuracy (Park et al., 2008). The price to pay in leaving the data on a rather fine temporal resolution is a huge quantity of sparse data. Binning data at $10-20$ ms generates a time scale larger than the characteristic time scale of gap junctions and of the order of the time decay of chemical synapses. Due to this huge difference in relevant time scales binned data may not reveal the presence of gap junctions. The impact of the binning transformation must be taken into account. The effect of binning in our model needs however to be further investigated.

6.3.0.2 The "Ising miracle"

We have pointed out that the Gibbs potential obtained from the conductance-based model we consider in chapter 3 (in the stationary case) is quite a lot more complex than the Ising model or similar models used in retinal spike train analysis (Schneidman et al., 2006, Tkačik et al., 2010, Ganmor et al., 2011a,b, Tyrcha et al., 2012). In particular, it involves spatio-temporal spike patterns. A neuro-mimetic model with N neurons has $O(N^2)$ parameters, whereas a MaxEnt model with N neurons and memory depth D has $O(2^{NR})$ parameters h_l (canonical potential). Given the fact that the Ising model characterizes quite well spike recordings of retinal ganglion cells, we are presented with either a paradox or a miracle. We can only speculate how the Ising model arise from a neuro-mimetic approach, a more complete clarification requires further investigation. We have two possibilities:

(i) If a large number of MaxEnt parameters h_l's vanish, in particular higher order terms.

(ii) If the coefficients h_l's are related among them and the higher order terms do not add additional information to the statistical model.

The two possibilities are actually not exclusive. Let us first address this question from the mathematical (dynamical systems) viewpoint which was the line followed in this thesis.
Consider a neuro-mimetic model with a well defined dynamics (e.g. (4.11)) and the associated normalized potential $\phi = \phi(\mathcal{W}, \mathcal{I})$ (e.g. (4.12)). We may view a normalized potential as a point in a space: the coordinates of this point are fixed by $\mathcal{W}, \mathcal{I}$. A neuro-mimetic model corresponds therefore to the space of normalized potential of dimension $O(N^2)$. Using the same representation MaxEnt models with memory depth D span a space of dimension $O(2^{NR})$, but MaxEnt models

equivalent to our neuro-mimetic model span a space of dimension $O(N^2)$. There is therefore a huge projection effect. Now eq (4.8), or, more generally, eq. (4.5) show that (ii) always functionally holds: h_l are non linear functions of $\mathcal{W}, \mathcal{I}$ and are related to each other. This has a dramatic effect. Assume that we want to fit (exactly) a neuro-mimetic model with a MaxEnt. We will need 2^{NR} terms whereas $O(N^2)$ are sufficient. This is exactly what happens in Fig 2. Now, one may hope that many h_l's are zero or close to zero. This is actually where MaxEnt models could make a breakthrough; showing that, in real spike trains many h_l's (almost) cancel would reveal a hidden law of nature.
What is the hope for this? If we address this question from the dynamical systems viewpoint, there is almost no hope. Indeed in this context, one has to look for generic conditions under which the h_l's vanish (case (i)). But it results from our analysis that the h_l's of a canonical potential corresponding to a neuro-mimetic model are *generically* non zero: considering e.g. *random* synaptic weights W_{ij}, the probability that some h_l's in (4.5) vanish is indeed zero[1].

However, real neural networks are non generic: synaptic weights are not drawn at random but result from a long phylogenetic and ontogenetic evolution, also they are subject to change in time due to plasticity. Also the statistics of the stimulus modify the MaxEnt parameters, and thus could play a role in eliminating some parameters. When trying to "explain" spike statistics of real neural networks with the Maximum Entropy Principle, one is seeking a general law which has to be expressed with relatively few phenomenological parameters in the potential (4.1). The hope is that many coefficients coming from real data are 0 or close to 0. This could explain the efficiency of pairwise MaxEnt models (Bialek and Ranganathan, 2007) for spike trains analysis (although this effect could also arise due e.g. binning).

6.4 Directions on future research

Our contributions suggest specific open questions and directions for future research. The remainder of this section will lay out some of these more immediately accessible avenues for future research.

6.4.1 Model with Gap junctions in terms of the model without them

We have seen in Chapter 3 that, the model including gap junctions, leads to a much more complicated dynamics than in the model without them (see chapter 4). One possible direction to explore is to consider the model without gap junctions and "perturb" this model with small conductances corresponding to gap junctions.

[1] There are two notions of generic properties: topological in which the generic property holds on a dense open set; and metric in which the generic property holds almost everywhere. In our case, the coefficients are generically non null in both senses

After doing this one can consider first order approximations and look how the perturbed model (with gap junctions) can be written in terms of the model without gap junctions. This representation could be very advantageous.

6.4.2 Criticality in retinal spike train recordings

One of the biggest ambitions in the field of spike train statistics is that techniques and ideas from statistical mechanics and thermodynamics will help us understand the collective "macroscopic" behavior of big populations spiking neurons with relatively few parameters. A recent paper (Tkačik et al., 2014), discuss about signatures of criticality found in retinal spike trains. The main idea is to use concepts and procedures from statistical mechanics considering networks with larger and larger numbers of neurons, expecting to see the emergence of a "thermodynamic limit" providing simpler universal behavior for the network as a whole, independent of microscopic details. The authors discover signatures of criticality by changing the parameters of the Maximum Entropy potential along one axis in parameter space (scaling). The fact that they observe a peak in the specific heat suggests that the real network is poised in the parameter space very close to a maximum in the variance of log(probability), which constitutes the dynamic range of surprise that can be represented by the network. This peak is a signature of criticality (second order phase transition) in statistical physics. An important point discussed in this paper is that systems near critical points are maximally responsive to certain external signals, and this sensitivity may be functionally useful for the retina (Tkačik et al., 2014). The authors finish the paper suggesting further tests of criticality. In statistical mechanics, scale invariance in the correlation length is a feature of second order phase transitions. Near a phase transition or critical point, fluctuations occur at all length scales, and thus one should look for an explicitly scale invariant theory to describe the phenomena. Finite size scaling techniques help to know how we can extract the correct values for the interesting quantities of the infinite system out of the finite system.

With the tools developed in this thesis we can answer the following questions. Does the correlation matrix χ have scale invariance properties for real retinal data? Can we say something in the case of retina recordings ? Do this properties translate to the spectra? Note that in comparison to (Tkačik et al., 2014), our matrix χ also handles spatio-temporal correlations. This is an interesting avenue of research that will be developed in the near future.

One point that is not discussed and that, we believe deserves further investigation is the role of binning in this result. Clearly, changing the binning size one changes the data set and thus different parameters for the Maximum Entropy model would be found. We believe that this result would be more robust if is stable under a different choose of time binning, i.e., if the divergence of the specific heat is independent of the time binning.

There is still work to be done in order to make sense in a more general picture of what is going on behind phase transitions. We leave the question of alterna-

tive neuro-mimetic approaches to spike train distribution modeling open for future research.

6.4.3 Can we give an example of Conductance based Integrate and Fire model exhibiting critical behavior?

Taking advantage of the results of this thesis another alternative to find critical behavior is to identify a neural network model capable to exhibit second order phase transitions. Identifying the underlying dynamical system would help producing a more "complete" global picture of what is going on. In particular, would help to answer "how" and eventually "why" the system is poised near the critical point. This can not be done using the methods exposed in (Tkačik et al., 2014). Note that such property (criticality) is again non generic, unless a so specific mechanism (self organized criticality) drives the system to a critical point. But, which biophysical mechanism would correspond to self organized criticality?

Conductance based Integrate and Fire generate spatio-temporal spike patterns whose temporal correlations can be measured. If these correlations decay as power law, is a signature of critical behavior. Under which conditions in the parameters of the model or in the model this could happen? In order to give a sense of previous work done in MaxEnt models (under the hypothesis of stationarity and ergodicity) we are looking for a system that is ergodic, but not mixing, because mixing implies exponential decay of temporal correlations, which implies no second order phase transitions. In particular we look for a dynamical system whose spectral gap (in the transfer matrix) disappear when the number of neurons goes to infinity ("thermodynamic limit").

6.4.3.1 Example: From MaxEnt to Neural network model

In chapter 4 we have shown how to go from transition probabilities to MaxEnt potentials. A natural extension of our work is the following: Given a spatio-temporal MaxEnt potential and a neuro-mimetic model can we identify the parameters of the neuro-mimetic model leading the Gibbs distribution obtained from the given spatio-temporal MaxEnt potential. Particularly important is to identify (if possible) a neural network model "equivalent" to the Ising model. Here the main difficulty is that while considering neural network models is natural to include memory, the Ising model consider successive spikes as independent events.

Bibliography

L. F. Abbott. Lapicque's introduction of the integrate-and-fire model neuron (1907). *Brain Research Bulletin*, 150:303–304, 1999. (Cited on pages 13 and 14.)

Y. Ahmadian, J. Pillow, and L. Paninski. Efficient Markov Chain Monte Carlo Methods for Decoding Neural Spike Trains. *Neural Computation*, 23(1):46–96, 2011. ISSN 1530-888X. doi: 10.1162/NECO_a_00059. URL http://dx.doi.org/10.1162/NECO_a_00059. (Cited on pages 46, 48, 53 and 74.)

V. I. Arnold. *Geometrical Methods in the Theory of Ordinary Differential Equations*. Springer, New York, 1983. (Cited on page 83.)

P. Baldi and L. Caramellino. Asymptotics of hitting probabilities for general one-dimensional pinned diffusions. *The Annals of Applied Probability.*, 12(3):1071–1095, 2002. (Cited on page 72.)

M. Beierlein, J. R. Gibson, and B. W. Connors. A network of electrically coupled interneurons drives synchronized inhibition in neocortex. *Nature Neuroscience*, 3(9):904–910, 2000. URL http://www.ncbi.nlm.nih.gov/pubmed/10966621. (Cited on pages 3 and 52.)

M. Bennett and R. Zukin. Electrical coupling and neuronal synchronization in the mammalian brain. *Neuron*, 41(4):495 – 511, 2004. ISSN 0896-6273. doi: 10.1016/S0896-6273(04)00043-1. URL http://www.sciencedirect.com/science/article/pii/S0896627304000431. (Cited on pages 3 and 52.)

J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*, 36:192–236, 1974. (Cited on page 85.)

W. Bialek and R. Ranganathan. Rediscovering the power of pairwise interactions. *arXiv.org:0712.4397 [q–bio.QM]*, 2007. (Cited on page 125.)

S. Bloomfield and B. Völgyi. The diverse functional roles and regulation of neuronal gap junctions in the retina. *Nature Reviews Neuroscience*, 10(7):495–506, 2009. URL http://www.ncbi.nlm.nih.gov/pubmed/19491906. (Cited on page 77.)

R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lect. Notes.in Math.* Springer-Verlag, New York, 1975. (Cited on pages 35, 36 and 82.)

P. Bremaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues.* Springer, 1999. (Cited on pages 22, 27 and 28.)

D. R. Brillinger. Maximum likelihood analysis of spike trains of interacting nerve cells. *Biol Cybern*, 59(3):189–200, 1988. (Cited on pages 46 and 80.)

D. R. Brillinger. Nerve Cell Spike Train Data Analysis - a Progression of Technique. *J Amer Statist Assn*, 87(418):260–271, 1992. (Cited on page 48.)

R. Brockett. *Finite Dimensional Linear Systems*. John Wiley and Sons, 1970. (Cited on page 62.)

T. Broderick, M. Dudik, G. Tkačik, R. E. Schapire, and W. Bialek. Faster solutions of the inverse pairwise ising problem. *Submitted (see http://arxiv.org/abs/0712.2437)*, 2007. URL `http://arxiv.org/abs/0712.2437`. (Cited on page 45.)

E. N. Brown, R. Barbieri, U. T. Eden, and L. M. Frank. Likelihood methods for neural spike train data analysis. *Computational Neuroscience: A Comprehensive Approach*, 2003. (Cited on page 48.)

N. Brunel and N. Fourcaud. Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural Computation*, 14(9):2057–2110, 2002. (Cited on pages 100 and 112.)

N. Brunel and V. Hakim. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Computation*, 11:1621–1671, 1999. (Cited on page 70.)

A. N. Burkitt. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biol Cybern*, 95:1–19, 2006a. (Cited on pages 53, 69 and 70.)

A. N. Burkitt. A review of the integrate-and-fire neuron model: Ii. inhomogeneous synaptic input and network properties. *Biol Cybern*, 95:97–112, 2006b. (Cited on pages 53, 69 and 70.)

A. Calabrese, J. Schumacher, D. Schneider, L. Paninski, and S. Woolley. A generalized linear model for estimating spectrotemporal receptive fields from responses to natural sounds. *PLoS ONE*, 6(1):470–484, 2011. (Cited on page 13.)

B. Cessac. A discrete time neural network model with spiking neurons. rigorous results on the spontaneous dynamics. *J. Math. Biol.*, 56:311–345, 2008. (Cited on page 59.)

B. Cessac. A view of neural networks as dynamical systems. *International Journal of Bifurcations and Chaos*, 20(6):1585–1629, 2010. doi: 10.1142/S0218127410026721. URL `http://lanl.arxiv.org/abs/0901.2203`. (Cited on page 56.)

B. Cessac. A discrete time neural network model with spiking neurons ii. dynamics with noise. *J. Math. Biol.*, 62:863–900, 2011a. doi: 10.1007/s00285-010-0358-4. URL `http://lanl.arxiv.org/abs/1002.3275`. (Cited on pages 90, 91 and 92.)

B. Cessac. Statistics of spike trains in conductance-based neural networks: Rigorous results. *Journal of Mathematical Neuroscience*, 1(8), 2011b. doi: doi:10.1186/

2190-8567-1-8. (Cited on pages 15, 17, 48, 49, 52, 53, 55, 59, 67, 99, 102, 103 and 104.)

B. Cessac and R. Cofré. Estimating maximum entropy distributions from periodic orbits in spike trains. *INRIA Research Report Number 8329*, 2013. (Cited on page 84.)

B. Cessac and A. Palacios. *Spike train statistics from empirical facts to theory: the case of the retina*, volume in "Current Mathematical Problems in Computational Biology and Biomedicine". Springer, 2012. (Cited on pages 43 and 102.)

B. Cessac and T. Viéville. On dynamics of integrate-and-fire neural networks with adaptive conductances. *Frontiers in neuroscience*, 2(2), 2008. (Cited on pages 53, 56 and 59.)

J. Chazottes and G. Keller. Pressure and equilibrium states in ergodic theory. *Israel Journal of Mathematics*, 131(1), 2008. (Cited on pages 35, 37 and 38.)

E. Chichilnisky. A simple white noise analysis of neuronal light responses. *Network Computational Neural Systems*, 12, 2001. (Cited on pages 48 and 80.)

C. Chicone and Y. Latushkin. *Evolution semigroups in dynamical systems*, page 61. American Mathematical Society, 1999. URL `http://books.google.fr/books/about/Evolution_semigroups_in_dynamical_system.html`. (Cited on page 65.)

C. C. Chow and N. Kopell. Dynamics of spiking neurons with electrical coupling. *Neural Computation*, 12:1643–1678, 2000. (Cited on page 52.)

P. Churchland and T. Sejnowski. *The Computational Brain*. The MIT Press, 2002. (Cited on pages 119 and 121.)

S. Cocco, N. Leibler, and R. Monasson. Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. *PNAS*, 106(33): 14058–14062, 2009. doi: 10.1073/pnas.0906705106. URL `www.pnas.org/cgi/doi/10.1073/pnas.0906705106`. (Cited on pages 80 and 81.)

R. Cofré and B. Cessac. Dynamics and spike trains statistics in conductance-based integrate-and-fire neural networks with chemical and electric synapses. *Chaos, Solitons and Fractals*, 50(8):13–31, 2013. (Cited on pages 49, 51 and 76.)

R. Cofré and B. Cessac. Exact computation of the maximum entropy potential of spiking neural networks models. *Physical Review E*, 107(5):368–368, 2014. (Cited on page 79.)

B. Connors and M. Long. Electrical synapses in the mammalian brain. *Annual Review Neuroscience*, 27:393–418, 2004. (Cited on page 52.)

S. Coombes. Neuronal networks with gap junctions: A study of piece-wise linear planar neuron models. *SIAM Journal on Applied Dynamical Systems*, 7(3):1101–1129, 2008. URL `http://epubs.siam.org/SIADS/`. (Cited on page 52.)

S. Coombes and M. Zachariou. *Gap junctions and emergent rhythms*, pages 77–94. Springer, 2007. URL http://eprints.nottingham.ac.uk/894/. (Cited on page 52.)

P. Dayan and L. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press, 2005. (Cited on page 99.)

A. Destexhe, Z. M. ZF, and T. Sejnowski. *Kinetic models of synaptic transmission*. Cambridge, MA: MIT Press, 1998. (Cited on page 55.)

M. Dudík, S. Phillips, and R. Schapire. Performance guarantees for regularized maximum entropy density estimation. In *Proceedings of the 17th Annual Conference on Computational Learning Theory*, 2004. (Cited on page 45.)

G. B. Ermentrout and D. H. Terman. *Mathematical Foundations of Neuroscience*. Springer, 1st edition. edition, 2010. ISBN 038787707X. (Cited on page 56.)

R. Fernandez and G. Maillard. Chains with complete connections : General theory, uniqueness, loss of memory and mixing properties. *J. Stat. Phys.*, 118(3-4):555–588, 2005. (Cited on pages 22, 29 and 40.)

M. Galarreta and S. Hestrin. A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature*, 402(6757):72–75, 1999. URL http://www.ncbi.nlm.nih.gov/pubmed/10573418. (Cited on pages 3 and 52.)

M. Galarreta and S. Hestrin. Electrical synapses between gaba-releasing interneurons. *Nature Reviews Neuroscience*, 2:425–433, 2001. (Cited on page 65.)

A. Galves and E. Löcherbach. Infinite systems of interacting chains with memory of variable length-a stochastic model for biological neural nets. *Journal of Statistical Physics*, 151:896–921, 2013. (Cited on page 21.)

E. Ganmor, R. Segev, and E. Schneidman. The architecture of functional interaction networks in the retina. *The journal of neuroscience*, 31(8):3044–3054, 2011a. (Cited on pages 43, 45, 53, 74, 123 and 124.)

E. Ganmor, R. Segev, and E. Schneidman. Sparse low-order interaction network underlies a highly correlated and learnable neural population code. *PNAS*, 108 (23):9679–9684, 2011b. (Cited on pages 45, 53, 74, 123 and 124.)

F. R. Gantmacher. *The theory of matrices*. AMS Chelsea Publishing, Providence, RI, 1998. (Cited on page 30.)

J. Gao and P. Holmes. On the dynamics of electrically-coupled neurons with inhibitory synapses. *Journal of Computational Neuroscience*, 22(1):39–61, 2007. URL http://www.ncbi.nlm.nih.gov/pubmed/16998640. (Cited on page 52.)

J. Gauthier, G. Field, A. Sher, M. Greschner, J. Shlens, A. Litke, and E. J. Chichilnisky. Receptive fields in primate retina are coordinated to sample visual space more uniformly. *PLoS Biol*, 7(4), 2009. (Cited on page 8.)

H.-O. Georgii. *Gibbs measures and phase transitions*. De Gruyter Studies in Mathematics:9. Berlin; New York, 1988. (Cited on pages 30, 38, 39 and 40.)

W. Gerstner and W.Kistler. *Spiking Neuron Models*. Cambridge University Press, 2002. (Cited on pages 1, 56 and 80.)

M. Gil. *Explicit Stability Conditions for Continuous Systems: A Functional Analytic Approach*. Lecture Notes in Control and Information Sciences. Springer, 2005. (Cited on page 64.)

T. Gollisch and M. Meister. Eye smarter than scientists believed: neural computations in circuits of the retina. *Neuron*, 65(2):150–164, 2010. doi: doi:10.1016/j.neuron.2009.12.009. URL http://dx.doi.org/doi:10.1016/j.neuron.2009.12.009. (Cited on page 6.)

E. Granot-Atedgi, G.Tkačik, R.Segev, and E.Schneidman. Stimulus-dependent maximum entropy models of neural population codes. *Plos Computational Biology*, 9, 2013. (Cited on pages 80 and 81.)

G. Grimmett. A theorem about random fields. *Bulletin of the London Mathematical Society*, 5:81–84, 1973. (Cited on page 85.)

J. M. Hammersley and P. Clifford. Markov fields on finite graphs and lattices. *unpublished*, 1971. URL http://www.statslab.cam.ac.uk/~grg/books/hammfest/hamm-cliff.pdf. (Cited on pages 76, 81, 84 and 85.)

P. Hanus, D. Mauldin, and M. Urbanski. Thermodynamic formalism, multifractal analysis of conformal infinite iterated function. *Acta Math. Hungarica*, 96:27–98, 2002. (Cited on page 102.)

R. Haslinger, G.Pipa, L.Gordon, D. Laura, D. Nikolić, Z. Williamsc, and E. Brown. Encoding through patterns: Regression tree-based neuronal population models. *Neural Comput.*, 25(8):1953–1993, 2013. ISSN 0899-7667. doi: 10.1162/NECO_a_00464. URL http://dx.doi.org/10.1162/NECO_a_00464. (Cited on page 98.)

D. Hegger. *Perception lecture Notes*. New York University, 2006. (Cited on page 5.)

A. Hodgkin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve cells. *Journal of Physiology*, 117:500–544, 1952. (Cited on page 56.)

S. Hormuzdi, M. Filippov, G. Mitropoulou, and R. Monyer, H.and Bruzzone. Electrical synapses: a dynamic signaling system that shapes the activity of neuronal networks. *Biochimica et Biophysica Acta*, 1662(1-2):113–137, 2004. URL http://www.ncbi.nlm.nih.gov/pubmed/15033583. (Cited on pages 3 and 52.)

E. Hu, F. Pan, B. Völgyi, and S. Bloomfield. Light increases the gap junctional coupling of retinal ganglion cells. *The Journal of Physiology*, 588(Pt 21):4145–4163, 2010. (Cited on page 55.)

E. Izhikevich. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT Press, Cambridge, MA, 2007. (Cited on page 2.)

E. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106(620), 1957. (Cited on pages 13, 35 and 41.)

K. Josić and R. Rosenbaum. Unstable solutions of nonautonomous linear differential equations. *SIAM Rev.*, 50(3):570–584, 2008. ISSN 0036-1445. doi: 10.1137/060677057. URL http://dx.doi.org/10.1137/060677057. (Cited on page 65.)

J. Keener and J. Sneyd. *Mathematical Physiology*, volume 8 of *Interdisciplinary Applied Mathematics*. Springer, New York, 1998. (Cited on page 52.)

G. Keller. *Equilibrium States in Ergodic Theory*. Cambridge University Press, 1998. (Cited on pages 30 and 36.)

C. Kirst and M. Timme. How precise is the timing of action potentials ? *Front. Neurosci.*, 3(1):2–3, 2009. (Cited on page 56.)

B. P. Kitchens. *Symbolic Dynamics: One-sided, Two-sided and Countable State Markov Shifts*. Springer-Verlag, 1998. (Cited on page 35.)

N. Kriegeskorte and G. Kreiman. *Understanding visual population codes - towards a common multivariate framework for cell recording and functional imaging*. MIT Press, Cambridge, MA, 2010. (Cited on page 5.)

A. Lachal. Some martingales related to the integral of brownian motion. applications to the passage times and transience. *Journal of Theoretical Probability*, 11:127–156, 1998. ISSN 0894-9840. URL http://dx.doi.org/10.1023/A:1021646925303. 10.1023/A:1021646925303. (Cited on page 69.)

L. Lapicque. Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. *J. Physiol. Pathol. Gen.*, 9:620–635, 1907. (Cited on page 13.)

B. Lindner. *Stochastic Methods in Neuroscience.*, chapter A brief introduction to some simple stochastic processes. G. Lord and C. Laing Eds, pages 1–28. Oxford University Press, 2009. (Cited on page 53.)

B. Lindner and L. Schimansky-Geier. Maximizing spike train coherence or incoherence in the leaky integrate-and-fire model. *Physical Review E*, 66:031916, 2002. (Cited on page 53.)

B. Lindner, J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geiere. Effects of noise in excitable systems. *Physics Reports*, 392:321–424, 2004. (Cited on pages 53 and 70.)

A. Livšic. Cohomology of dynamical systems. *Math. U.S.S.R. Izvestija*, 6:1278–1371, 1972. (Cited on page 84.)

J. Macke, G. Zeck, and M. Bethge. Receptive fields without spike-triggering. In *21th Neural Information Processing Systems Conference*, volume 12, pages 280–286. The MIT Press, 2008. (Cited on page 98.)

J. Macke, L. Buesing, J. Cunningham, B. Yu, K. Shenoy, and M. Sahani. Empirical models of spiking in neural populations. In *Advances in Neural Information Processing Systems 24*, pages 1350–1358. NIPS, 2011. (Cited on pages 53 and 74.)

O. Marre, S. El Boustani, Y. Frégnac, and A. Destexhe. Prediction of spatiotemporal patterns of neural activity from pairwise correlations. *Physical review letters*, 102 (13), 2009. ISSN 0031-9007. URL `http://view.ncbi.nlm.nih.gov/pubmed/19392405`. (Cited on pages 9, 13 and 43.)

D. N. Mastronarde. Interactions between ganglion cells in cat retina. *J Neurophysiol*, 49(2):350–65, 1983. (Cited on page 123.)

P. McCullagh and J. A. Nelder. *Generalized linear models (Second edition)*. London: Chapman and Hall, 1989. (Cited on page 46.)

G. Medvedev, C. J. Wilson, J. Callaway, and N. Kopell. Dendritic synchrony and transient dynamics in a coupled oscillator model of the dopaminergic neuron. *Journal of Neuroscience*, 15:53–69, 2003. (Cited on page 52.)

G. S. Medvedev. Electrical coupling promotes fidelity of responses in the networks of model neurons. *Neural Comput.*, 21(11):3057–3078, 2009. ISSN 0899-7667. (Cited on pages 52 and 60.)

M. Meister and M. Berry II. The neural code of the retina. *Neuron*, 22(3):435–450, 1999. (Cited on page 5.)

H. Nasser and B. Cessac. Parameter estimation for spatio-temporal maximum entropy distributions: Application to neural spike trains. *Entropy*, 16(4), 2014. (Cited on page 76.)

H. Nasser, O. Marre, and B. Cessac. Spatio-temporal spike trains analysis for large scale networks using maximum entropy principle and monte-carlo method. *Journal Of Statistical Mechanics*, 2013. (Cited on pages 43 and 45.)

S. Nirenberg and P. Latham. Population coding in the retina. *Current Opinion in Neurobiology*, 8:488—493, 1998. (Cited on page 123.)

S. Nirenberg and P. Latham. Decoding neuronal spike trains: how important are correlations. *Proceeding of the Natural Academy of Science*, 100(12):7348–7353, 2003. (Cited on page 123.)

I. E. Ohiorhenuan, F. Mechler, K. P. Purpura, A. M. Schmid, Q. Hu, and J. D. Victor. Sparse coding and high-order correlations in fine-scale cortical networks. *Nature*, 466(7):617–621, 2010. (Cited on pages 45, 53 and 74.)

S. Ostojic and V. Brunel, N.and Hakim. Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities. *Journal of computational neuroscience*, 26(3):369–392, 2009. (Cited on pages 52, 60, 61 and 70.)

F. Pan, D. Paul, S. Bloomfield, and B. Völgyi. Connexin36 is required for gap junctional coupling of most ganglion cell subtypes in the mouse retina. *Journal of Comparative Neurology*, 518(6):911–927, 2010. (Cited on page 55.)

D. Panas, H. Amin, A. Maccione, O. Muthmann, M. van Rossum, L. Berdondini, and M. Hennig. Spontaneous neuronal network remodeling takes place along sloppy parameter dimensions. *preprint*, 2014. (Cited on page 115.)

L. Paninski. Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Comput. Neural Syst.*, 15(04):243–262, 2004. doi: 10.1088/0954-898X/15/4/002. URL http://www.iop.org/EJ/abstract/0954-898X/15/4/002. (Cited on page 46.)

L. Paninski, M. Fellows, S. Shoham, N. Hatsopoulos, and J. Donoghue. Superlinear population encoding of dynamic hand trajectory in primary motor cortex. *J. Neurosci.*, 24:8551–8561, 2004. doi: 10.1523/JNEUROSCI.0919-04.2004. URL http://dx.doi.org/10.1523/JNEUROSCI.0919-04.2004. (Cited on page 48.)

S. Panzeri and S. Schultz. A unified approach to the study of temporal, correlational, and rate coding. *Neural Comput*, 13:1311–1349, 2001. (Cited on page 123.)

I. Park, A. Parva, T. DeMarse, and J. Principe. An efficient algorithm for continuous-time cross correlation spike trains. *J. Neurosci. Methods*, 128(2), 2008. (Cited on page 124.)

B. Pfeuty, G. Mato, D. Golomb, and D. Hansel. The combined effects of inhibitory and electrical synapses in synchrony. *Neural Comput.*, 17(3):633–670, 2005. ISSN 0899-7667. doi: 10.1162/0899766053019917. URL http://dx.doi.org/10.1162/0899766053019917. (Cited on page 52.)

J. Pillow, L. Paninski, V. Uzzell, E. Simoncelli, and E. Chichilnisky. Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *J. Neurosci*, 25:11003–11013, 2005. (Cited on pages 46, 53, 74 and 98.)

J. Pillow, Y. Ahmadianr, and L. Paninski. Model-based decoding, information estimation, and change-point detection techniques for multineuron spike trains. *Neural Comput.*, 23(1):1–45, 2011. ISSN 0899-7667. doi: 10.1162/NECO_a_00058. URL http://dx.doi.org/10.1162/NECO_a_00058. (Cited on pages 46, 53 and 74.)

J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. Litke, E. J. Chichilnisky, and E. Simoncelli. Spatio-temporal correlations and visual signaling in a complete

neuronal population. *Nature*, 454(7206):995–999, 2008. (Cited on pages 46, 48, 53 and 74.)

M. Pollicott and H. Weiss. Free energy as a dynamical invariant (or can you hear the shape of a potential?). *Communications in Mathematical Physics*, 240:457–482, 2003. (Cited on page 81.)

A. Pouget, R. Zemel, and P. Dayan. Information processing with population codes. *Nature Review Neuroscience*, 1(2):125–132, 2000. (Cited on page 98.)

F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek. *Spikes, Exploring the Neural Code*. The M.I.T. Press, 1996. (Cited on pages 8 and 9.)

J. Roederer. *Information and Its Role in Nature*. Springer., 2005. (Cited on page 5.)

R. Rojas. *Neural Networks: A Systematic Introduction*. Springer, 1996. (Cited on page 5.)

Y. Roudi, S. Nirenberg, and P. Latham. Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't. *PLOS Computational Biology*, 5(5), 2009. (Cited on pages 123 and 124.)

M. Rudolph and A. Destexhe. Analytical integrate and fire neuron models with conductance-based dynamics for event driven simulation strategies. *Neural Computation*, 18:2146–2210, 2006. URL `http://www.mitpressjournals.org/doi/abs/10.1162neco.2006.18.9.2146`. (Cited on pages 14, 15, 17, 49, 53 and 55.)

D. Ruelle. *Thermodynamic formalism*. Addison-Wesley,Reading, Massachusetts, 1978. (Cited on page 35.)

D. Ruelle. Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics. *J. Statist. Phys.*, 95:393–468, 1999. (Cited on page 102.)

E. Schneidman, W. Bialek, and M. Berry II. Synergy, redundancy, and independence in population codes. *J Neurosci*, 23(37):11539–53, 2003. (Cited on page 123.)

E. Schneidman, M. Berry II, R. Segev, and W. Bialek. Weak pairwise correlations imply string correlated network states in a neural population. *Nature*, 440:1007–1012, 2006. (Cited on pages 13, 43, 45, 53, 74, 123 and 124.)

T. Schwalger, K. Fisch, J. Benda, and B. Lindner. How noisy adaptation of neurons shapes interspike interval histograms and correlations. *PLoS Comput Biol*, 6(12): e1001026, 2010. doi: 10.1371/journal.pcbi.1001026. (Cited on page 54.)

M. Shadlen and W. Newsome. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci*, 18(10): 3870–3896, 1998. (Cited on page 9.)

L. T. Sharpee and A. Stockman. Two rod pathways: the importance of seeing nothing. *Trends in Neurosciences*, 22:497–504, 1999. (Cited on page 7.)

T. Sharpee, N. Rust, and W.Bialek. Analyzing neural responses to natural signals: maximally informative dimensions. *Neural Comput*, 16:223–250, 2004. (Cited on page 117.)

J. Shlens, G. D. Field, J. L. Gauthier, M. I. Grivich, D. Petrusca, A. Sher, A. M. Litke, and E. J. Chichilnisky. The structure of multi-neuron firing patterns in primate retina. *J Neurosci*, 26(32):8254–66, 2006. (Cited on pages 45, 53 and 74.)

J. Shlens, G. Field, J. Gauthier, M. Greschner, A. Sher, A. Litke, and E. J. Chichilnisky. The structure of large-scale synchronized firing in primate retina. *The Journal of Neuroscience*, 29(15):5022–5031, 2009. doi: 10.1523/JNEUROSCI.5187-08.2009. (Cited on pages 53 and 74.)

E. P. Simoncelli, J. P. Paninski, J. Pillow, and O. Schwartz. Characterization of Neural Responses with Stochastic Stimuli. *The cognitive neurosciences*, 2004. (Cited on pages 46 and 99.)

H. Soula, G. Beslon, and O. Mazet. Spontaneous dynamics of asymmetric random recurrent spiking neural networks. *Neural Computation*, 18(1), 2006. (Cited on page 90.)

B. Staude, S. Grun, and S. Rotter. Higher-order correlations in non-stationary parallel spike trains: statistical modeling and inference. *Frontiers in computational neuroscience*, 4(16), 2010. (Cited on page 124.)

A. Tang, D. Jackson, J. Hobbs, W. Chen, J. Smith, H. Patel, A. Prieto, D. Petrusca, M. Grivich, A. Sher, P. Hottowy, W.Dabrowski, A. Litke, and J. Beggs. A maximum entropy model applied to spatial and temporal correlations from cortical networks *In Vitro*. *The Journal of Neuroscience*, 28(2):505–518, 2008. doi: 10.1523/JNEUROSCI.3359-07.2008. (Cited on page 45.)

F. E. Theunissen, S. V. David, N. C. Singh, A. Hsu, W. E. Vinje, and J. L. Gallant. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network*, 12(3):289–316, 2001. doi: 10.1080/net.12.3.289.316. URL `http://dx.doi.org/10.1080/net.12.3.289.316`. (Cited on page 48.)

G. Tkačik, E. Schneidman, M. Berry II, and W. Bialek. Ising models for networks of real neurons. *arXiv*, q-bio/0611072, 2006. (Cited on page 45.)

G. Tkačik, E. Schneidman, M. J. Berry II, and W. Bialek. Spin glass models for a network of real neurons. *arXiv: 0912.5409v1*, 2009. URL `arXiv:0912.5409v1`. (Cited on pages 45, 53 and 74.)

G. Tkačik, J. Prentice, V. Balasubramanian, and E. Schneidman. Optimal population coding by noisy spiking neurons. *PNAS*, 107(32):14419–14424, 2010. (Cited on page 124.)

G. Tkačik, T. Mora, O. Marre, D. Amodei, M. Berry II, and W. Bialek. Thermodynamics for a network of neurons: Signatures of criticality. *arXiv:1407.5946 [q-bio.NC]*, 2014. (Cited on pages 45, 126 and 127.)

J. Touboul and O. Faugeras. The spikes trains probability distributions: a stochastic calculus approach. *Journal of Physiology-Paris*, 101:78–98, 2007. (Cited on page 69.)

W. Truccolo, U. Eden, M. Fellows, J. Donoghue, and E. Brown. A point process framework for relating neural spiking activity to spiking history, neural ensemble and extrinsic covariate effects. *J Neurophysiol*, 93:1074–1089, 2005. (Cited on pages 46, 48 and 98.)

J. Tyrcha, Y. Roudi, M. Marsili, and J. Hertz. Effect of nonstationarity on models inferred from neural data. *preprint http://arxiv.org/abs/1203.5673*, 2012. (Cited on pages 123 and 124.)

J. Vasquez, O. Marre, A. Palacios, M. Berry II, and B. Cessac. Gibbs distribution analysis of temporal correlation structure on multicell spike trains from retina ganglion cells. *J. Physiol. Paris*, 2012. in press. (Cited on pages 13, 43, 53, 74 and 76.)

A. Wohrer and P. Kornprobst. Virtual retina: a biological retina model and simulator, with contrast gain control. *Journal of Computational Neuroscience*, 26(2): 219–249, 2009. (Cited on page 65.)

R. Wooster. Evolution systems of measures for non-autonomous stochastic differential equations with levy noise. *Communications on Stochastic Analysis*, 5:353–370, 2011. (Cited on pages 64 and 65.)

S. Yu, D. Huang, W. Singer, and D. Nikolic. A small world of neuronal synchrony. *Cereb. Cortex*, 2008. (Cited on page 45.)

Neuronal Networks, Spike Trains Statistics and Gibbs Distributions

Keywords: Gibbs distributions, Maximum entropy Principle, Spike train statistics.