Large deviations for the dynamics of heterogeneous neural networks

Tanguy Cabana

PhD THESIS

prepared at
Collège de France Paris

and presented at the
Université Pierre et Marie Curie
Laboratoire de Probabilités et Modèles Aléatoires

A dissertation submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF MATHEMATICS
Specialized in Neuroscience

Large deviations for the dynamics of heterogeneous neural networks

Tanguy CABANA

Advisors	Pr. Raphaël Krikorian	Université de Cergy-Pontoise, France
	Dr. Jonathan Touboul	Collège de France, France
Reviewers		
Examiners		

THESE DE DOCTORAT

preparée à
Collège de France Paris

et présentée à
Université Pierre et Marie Curie
Laboratoire de Probabilités et Modèles Aléatoires.

Thèse présentée en vue d'obtenir le titre de
DOCTEUR EN MATHEMATIQUES
Spécialité: Neurosciences

Grandes déviations pour la dynamique de réseaux de neurones hétérogènes

Tanguy CABANA

Directeurs:	Pr. Raphaël Krikorian	Université de Cergy-Pontoise, France
	Dr. Jonathan Touboul	Collège de France, France

Rapporteurs

Examinateurs

ABSTRACT

English Abstract

This thesis addresses the rigorous derivation of mean-field results for the continuous time dynamics of heterogeneous large neural networks. In our models, we consider firing-rate neurons subject to additive noise. The network is fully connected, with highly random connectivity weights. Their variance scales as the inverse of the network size, and thus conserves a non-trivial role in the thermodynamic limit. Moreover, another heterogeneity is considered at the level of each neuron. It is interpreted as a spatial location. For biological relevance, a model considered includes delays, mean and variance of connections depending on the distance between cells. A second model considers interactions depending on the states of both neurons at play. This last case notably applies to Kuramoto's model of coupled oscillators. When the weights are independent Gaussian random variables, we show that the empirical measure of the neurons' states satisfies a large deviations principle, with a good rate function achieving its minimum at a unique probability measure, implying averaged convergence of the empirical measure and propagation of chaos. In certain cases, we also obtained quenched results. The limit is characterized through a complex non Markovian implicit equation in which the network interaction term is replaced by a non-local Gaussian process whose statistics depend on the solution over the whole neural field. We further demonstrate the universality of this limit, in the sense that neuronal networks with non-Gaussian interconnections but sub-Gaussian tails converge towards it. Moreover, we present a few numerical applications, and discuss possible perspectives.

French Abstract

Cette thèse porte sur l'obtention rigoureuse de limites de champ moyen pour la dynamique continue de grands réseaux de neurones hétérogènes. Nous considérons des neurones à taux de décharge, et sujets à un bruit Brownien additif. Le réseau est entièrement connecté, avec des poids de connections dont la variance décroît comme l'inverse du nombre de neurones conservant un effet non trivial dans la limite thermodynamique. Un second type d'hétrogénéité, interprété comme une position spatiale, est considéré au niveau de chaque cellule. Pour la pertinence biologique, nos modèles incluent ou bien des délais, ainsi que des moyennes et variances de connections, dépendants de la distance entre les cellules, ou bien des synapses dépendantes de l'état des deux neurones post- et présynaptique. Ce dernier cas s'applique au modèle de Kuramoto pour les oscillateurs couplés. Quand les poids synaptiques sont Gaussiens et indépendants, nous prouvons un principe de grandes déviations pour la mesure empirique de l'état des neurones. La bonne fonction de taux associée atteint son minimum en une unique mesure de probabilité, impliquant convergence et propagation du chaos sous la loi "averaged". Dans certains cas, des résultats "quenched" sont obtenus. La limite est solution d'une équation implicite, non Markovienne, dans laquelle le terme d'interactions est remplacé par un processus Gaussien qui dépend de la loi de la solution du réseau entier. Une universalité de cette limite est prouvée, dans le cas de poids synaptiques non-Gaussiens avec queues sous-Gaussiennes. Enfin, quelques résultats numérique sur les réseau aléatoires sont présentés, et des

perspectives discutées.

ACKNOWLEDGEMENT

I would firstly like to express my sincere gratitude to my two PhD advisors Raphaël Krikorian and Jonathan Touboul, who gave me their trust and approbation for this project. I am very grateful to Jonathan Touboul for these three years of collaborations. Despite his many responsibilities, he always found the time to lend me his entire attention and share with me his advised intuitions. His inspiring conversations, his creativity, and the numerous opportunities he gave me to travel across the world have been as many elements that made these three years a wonderful experience. I am very thankful to Gilles Wainrib with who I had a great pleasure working with, notably for his patience, his inventive perspicacity and structuring pedagogy. I am also indebted to Giambattista Giacomin and Francis Comets who kindly accepted to receive me in their office to discuss my work.

I benefited from a very nice and stimulating research environment in the *Quartier Latin* in Paris, and I warmly thank the researchers and administrative staff of both Collège de France and UPMC. I am particularly grateful to the PhD students of Paris 6-7 organizers of the *Groupe de Travail des Thésards* for their useful and generous work. I also want to address my warmest thanks to the members of the Mathematical Neuroscience Team of Collège de France for these pleasant years: Justyna Signerska, Alberto Romagnoni, Jérôme Ribot, Cristobal Quiñinao, Luis Carlos Garcia del Molino, Richard Bailleul, Yi Cui, Gaëtan Vignoud, Antoine Sauvage, Valentin Figué, Guillaume Hubert, and Guillaume Penent; and also Anabelle Planques, Camille Curantz, Leticia Gomes and Marion Ruinart de Brimont. Thanks for sharing so many coffee breaks and cheerful discussions!

Finally, I dearly thanks my numerous family, my friends and Gabriela for their unfailing and priceless support.

Contents

Introduction

Neuroscience is the field of research concerned by the study of the development, organization and functioning of the nervous system. First emerged from neurobiology at the far end of the nineteenth century, its importance has grown all along the first half of the twentieth century, to become recognized as a discipline in its own right by the scientific community. Because of the tremendous complexity of its object of interest, and due to its potential repercussions on both medicine and artificial intelligence (AI), it is today a booming, highly interdisciplinary field, involving no less than chemistry, physics, linguistic, computer science, medicine, psychology, philosophy, and mathematics.

The present work pertains to mathematical neuroscience. Beyond providing a reliable formalism as well as technical tools from a number of theories such as the dynamical systems, stochastic calculus, or partial differential equations theories, and ergodic theory, this hybrid area aims at enriching the neuroscientific investigation with the originality of its viewpoint. In order to help elucidate the mysteries of cognition, it endeavors at tailoring simple tractable models equipped with predictive strength that would account for biological observations. Such an approach can for example be illustrated by the contribution of P. L. Buono and M. Golubitsky and colleagues [45]: through symmetry assumptions made on the architecture of the central pattern generator, the brain's area involved with the generation of locomotive pace in animals, the authors successfully predicted the existence of a rare gait that was then observed in nature. Moreover, by bringing into light original mathematical problems and equations, this transverse dialog has revealed itself as fruitful as mathematical physics have been for pure and applied mathematics, and might result in new mathematical concepts and theories. For its tremendous complexity and puzzling efficiency, it indeed seems that the brain has something to tell us about our mathematical world.

The object of the present thesis is to address, through probabilistic tools, a mean-field problem: that of the asymptotic dynamics emerging from a class of neurobiologically based mathematical model describing, at the microscopic level, the time evolution of networks of neurons, and in the limit where the number of neurons tends toward infinity. These models take a particular account of both the spatial extension of mammalian nervous system, as well as the highly heterogeneous and noisy properties they display at the microscopic level.

The thesis is organized as follow. Part I presents a general overview of the mathematical approach we will develop to tackle the dynamics of randomly connected neural networks. Chapter 1 will briefly describe the characteristics of the biological system at study, and present the classical mathematical models at use. Chapter

2 will introduce the basic concepts of the mean-field theory that is tailored to address the dynamics of large systems of particles. Chapter 3 will give a rapid overview on the emergence of rigorous mean-field approaches in neuroscience, and present two different strategies in order to cope with networks presenting two different kind of random interactions: *weakly random interactions* and *strongly random interactions* whose variance respectively scale as $\frac{1}{N^2}$ and $\frac{1}{N}$, N being the size o the network. In the rest of the manuscript, we shall insist on the second type of interactions which are of central importance in this work. Part II is the core of this thesis. In chapter 4 and 5, we will undertake the rigorous analysis of two different spatially extended neural networks. The first one involves spatially-dependent delays, as well as spatially-dependent mean and variance of interactions. It notably extends its scope to obtain results for the case of non-Gaussian interactions weights relevant for biology. The second model analyzed in chapter 5 involves synapses depending on both the postsynaptic and presynaptic neurons, extending the general firing-rate approach. Part III consists in a rapid presentation of some numerical results accompanied by a theoretic analysis. Chapter 6 presents the phenomenology of randomly connected neural networks, and analyze the effect of delays as well as variance and mean of connections onto the dynamics. Chapter 7 applies the results of Chapter 5 to investigate the effect of heterogeneous connections on the dynamics of the Kuramoto model of coupled oscillators. Finally, the General Appendix gives a rapid overview of some tools from probability theory to Gaussian calculus. It notably furnishes a substantial introduction on large deviations theory covered by the chapter 9. Note also that a section referencing the General Notations is also present at the end of the manuscript.

Part I

General Introduction

CHAPTER 1

PRINCIPLES OF NEURAL SCIENCE AND MODELIZATIONS BASICS

1.1 HISTORICAL NOTES

He who cannot draw on three thousand years is living from hand to mouth.

— Johann Wolfgang von Goethe

"Know thyself". Found on the forecourt of Apollo's Delphic temple, this antique aphorism - traditionally and mistakingly attributed to the Greek philosopher Socrates - could perfectly have been carved on the pediment of any neuroscientific institution. In fact, what better answer than the brain could humans - both as a species and individuals - possibly find when philosophically questioning the essence of their identity: What makes us unique on Earth? What makes me who I am?

Back in pre-Socratic times, people had a utterly different picture of the mind. For it was easily observed by the naked eye, and quite apparently extended its vessels to the entire body, the vascular system was the only convincing candidate for the carriage of both motions and sensations. It was thought to convey a life-giving air-like substance called *pneuma*, extracted from the atmosphere by the pumping of our lungs. At the center of this system lied the heart viewed as the source of human emotions and mental abilities. Nerves, in comparison, were very hard to see. They were even harder to distinguish from sinews and ligaments. The brain thus appeared relatively isolated from the rest of the body. Speculations made it a disregarded mucus carrier.

A testimony of this ancient belief lies in the embalming ritual performed by the Egyptian civilization. To prepare the passage of Pharaoh into eternal life, some of its organs were meticulously cared for, and put into jars near its body. The liver, the lungs and stomach all benefited this sacred privilege, while the heart was regarded as the most precious of them all: it was to be weighted by the god Anubis as a token of the merit of the dead sovereign. As for the brain, it experienced a far gloomier fate: extracted from the skull with an iron hook, it was thrown away as a mere garbage. It is worth noting that Egyptian were among the most skilled physician of the ancient

Mediterranean world, and have certainly greatly influenced their neighbors on medical matters. The Old Testament and Homeric epics are good witnesses of this possible transmission: they contain many references to the heart as the siege for bravery and intelligence.

Figure 1.1: Drawing of a silver-stained pigeon Purkinje cells by Santiago Ramón y Cajal. Cell body, dendritic bush, and axon are clearly distinguishable.

For centuries now, science has well established that the seat of our intelligence and memory, of our consciousness and sensations, of our emotions and communicative skills is to be found in our brain. We have also widely accepted that our experience of the world can roughly be summarized by the electrical activity of this vital organ - up to a possible immaterial and complementary element we call soul. One of the most important discovery of the modern history of neuroscience is that of the *neuron*, made by the Spanish histologist Santiago Ramón y Cajal. Using the silver-staining method invented by Camillo Golgi in 1873, he challenged the contemporary consensus that depicted the nervous system as a single continuous fibrous network. In fact, he observed that it was in fact made of many distinct elements, connecting with each other through tiny structure called synapse (meaning fasten in Greek), that constitute the siege of transmission. His theory, the *Neuron's doctrine*, can be stated as follow:
1) neurons are discrete autonomous cells able to interact through electrical impulses
2) synapses are gaps that separate neurons
3) information is directional, i.e. transmitted by the neuron in one direction: from dendrites to axon.
At first very controversial, it progressively gained into credibility to become the starting point of a new understanding of the mind.

1.2 Basic notions

This section found support in [80, 99].

1.2.1 Neurons and action potentials

What is real? How do you define ”real”? If you're talking about what you can feel, what you can smell, what you can taste and see then ”real” is simply electrical signals interpreted by your brain.

— Morpheus, Matrix

In this section we will describe the global organization of the neuron and explain how it conveys and integrates electrical signals. As will be emphasized in section 1.4.3, there exists a great diversity of neurons, dramatically differing in size function and shape. The following description does not intend to present them all, but gives a stereotyped picture underlying neurons functioning. Moreover, we will mainly overlook important supportive cells, called *glial cells*.

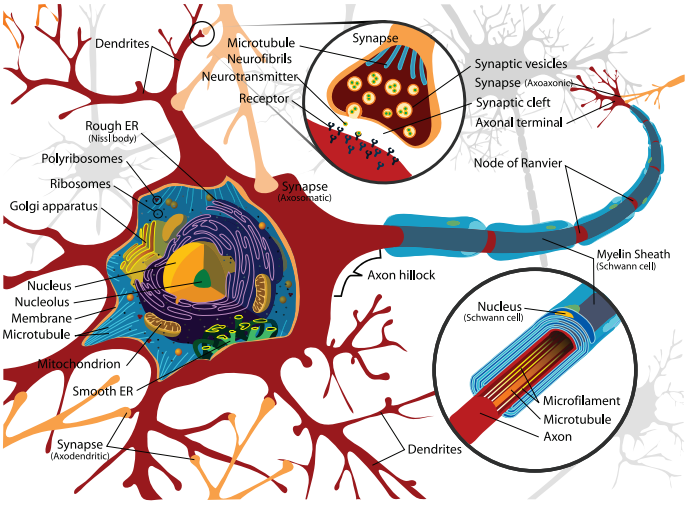


Figure 1.2: Graphic overview of a typical neuron. Reproduced from Wikipedia Commons.

Neurons are cells characterized by their aptitude to transmit information through the emission of transient electrical signals, called *action potentials* or *spikes*, that travel along their membrane. In order to establish connections with possibly remote interlocutors, they display a fibrous tree-like shape that can be divided in three distinct parts: the *dendrites*, *cell body*, and *axon*. Interestingly, these evoke, by both their shape and function, the roots, seed, and stem of a growing plant. First, the cell body, also called *soma*, is the core of the neuron. It contains the nucleus and DNA, and shelters the classical cellular machinery that notably produces the proteins and energy of the cell. It is also the pivot that links the two other parts of the neuron, both made of filamentous extensions. On one side, the bushy and highly ramified dendrites - ”tree-like” in Greek - extend in the neighborhood of the soma, and feed it with the neural raw material: information. They are the ears of the neuron that collect the electrical activity incoming from the network, to integrate it at the level of the soma. On the other side lies the solitary axon. It is a far-reaching electrical highway, able to extend at the order of the meter before branching to connect, through bud-like structures of its membrane called *synapses*, the dendrites of other neurons. Its basis,

the *axon hillock*, displays the highest density of voltage-gated sodium channels (see below) making it the most excitable part of the neuron. In order to accurately transmit possibly complex spiking patterns to far away areas (outside the cortex), axon are wrapped in insulating sheets of *myelin* (actually, these sheets are made of supportive glial cells: Schwann cells in the *Peripheral Nervous System* and of oligodendrocites in the *Central Nervous System*) which are steadily spaced by *Ranvier nodes*, a kind of exit-entrance for this biological highway. Note that a neuron can possess more than one axon, that axon-to-axon, dendrite-to-dendrite, and dendrite-to-axon connections also exist, but these ”exceptions” are relatively scarce in mammalian nervous systems.

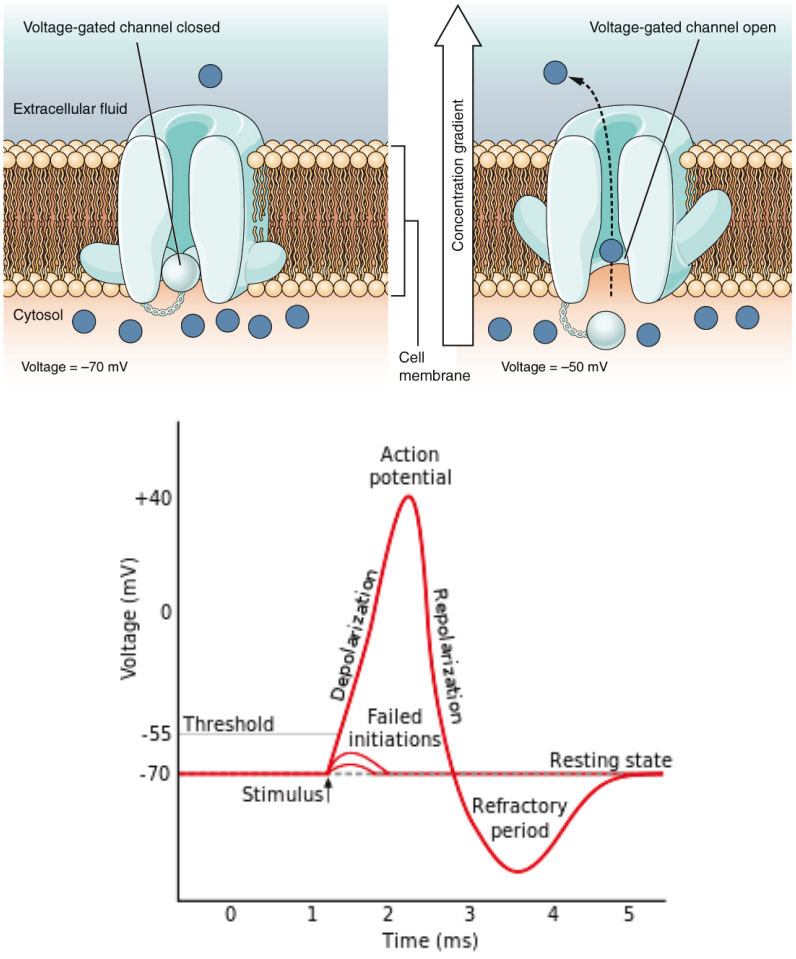


Figure 1.3: Voltage-gated ionic channel, reproduced from http://cnx.org/contents/QBrzNCkw@5/The-Action-Potential (left). Temporal profile of an action potential, reproduced from wikipedia (right).

When the membrane potential of the soma reaches a given *threshold*, a spike is triggered at the level of the axon hillock, and propagates all along its membrane to reach the synaptic terminals. Locally, this electrical wave lasts 2 milliseconds. It is governed by the opening and closure of *voltage-gated ionic channels* covering the surface of the neuron. When the neuron is at rest, some of these transmembrane proteins actively thwart the natural diffusion of ions in order to maintain a high concentration of potassium and a low concentration of sodium inside the cell. This

pumping requires ATP (energy of the cell), and represents approximatively seventy percent of the neuron's consumption in energy. This results in a polarized resting membrane potential whose typical value lies around -70mV. When a positive current is locally applied on the membrane, this prompts the opening of sodium channels that passively and massively let these ions rush into the cytoplasm for the membrane potential to attain +40mV within a millisecond. This local depolarization in turn yields both the closing of sodium channels and the opening of potassium ones. This causes a rapid outward flow of potassium ions, allowing the neuron to return to its resting state. It also impacts the potential of the directly adjacent portions of membrane, propagating the spike along the axon. This short-lasting event, is generally followed by a *refractory period* lasting around two milliseconds, and during which the ion gradient is rebuild.

1.2.2 The synapse

At the level of a synapse, two neurons are interacting: the *presynaptic neuron* sends information, whereas the *postsynaptic neuron* is listening. Remark that a neuron is perfectly capable of connecting its own dendrites. There exists two great categories of synapses. On the on hand, *electric synapses* or *gap junctions* directly connects the cytoplasm of the two neurons, allowing the bidirectional transit of various ions, molecules. This transit is made possible by the binding of two transmembrane proteins, *hemichannels*, that tightly link the two membrane, and create a small local aperture. These electrical synapses are characterized by a high speed of transmission between neurons, a very useful feature for escaping predation through stereotypical reflexes.

More sophisticated and versatile are the *chemical synapses*. They rely on the diffusion of molecular messengers called *neurotransmitters* throughout the *synaptic cleft* - a small space that is set up and maintained between the two neurons. When an action potential crosses the axon of a presynaptic neuron, and reaches the synaptic terminal, voltage-gated calcium channels covering the membrane of the *synaptic terminal* open and let this ion flow into the cellular medium. Calcium ions then bind to specific receptors attached to arm-like proteins linking both the membrane of the presynaptic neuron, and vesicles filled with neurotransmitters. This activates a spring-loaded fusion between these vesicles and the synaptic membrane through a bending of the protein, enabling a massive release of neurotransmitters into the synaptic cleft. Neurotransmitters then diffuse to bind specific receptors on the membrane of the postsynaptic neuron, yielding one of the two following scenari. In the first case, receptors are directly controlling the aperture of ion-channels, enabling an immediate effect on the postsynaptic potential. Alternatively, receptors are located on intermediary transmembrane protein that triggers a slower long-lasting response in the intracellular medium of the postsynaptic neuron. Interestingly, this kind of synaptic transmission is seemingly crucial for learning and memory. In both cases, within a 1ms [66], the synapse progressively deactivates through the clearance of neurotransmitter achieved through diverse mechanisms (antagonist molecules binding the receptors, destruction, diffusion outside the synaptic cleft, etc.).

Despite their relative slowness, chemical synapses present the advantage of the sophistication: there exists more than a hundred different kind of neurotransmitters displaying various characteristic times. Hence, they are able to induce a wide range of possible interactions, and notably enable inhibition. Moreover, the location of the

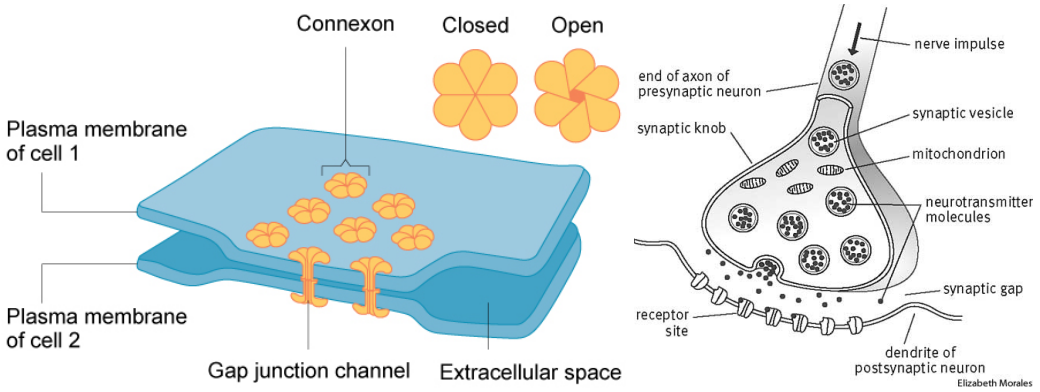


Figure 1.4: Gap junctions, composed of the binding of two hemichannels. Image courtesy Mariana Ruiz Villarreal (republished under open access license from Wikimedia Commons) (left). Synaptic terminal for a chemical synapses with direct binding of the neurotransmitters onto postsynaptic ionic channels. (Copyright © 2002. Published by Houghton Mifflin. All rights reserved.) (right).

synapse on the postsynaptic neuron's dendrite modulate its transmission: it is all the more effective than it is close from the soma. Hence, an inhibitory synapse located near the soma can short-cut several excitatory synapses located farther on the same dendrite. Furthermore, the size of the synapse is directly linked with the number of vesicles it contains, and thus, to its impact on the postsynaptic neuron. Both these properties shape the *synaptic efficiency*, which corresponds to the maximal transmission of a given terminal.

Many interesting and intricate phenomena can arise at the level of the synapse, making them particularly difficult to model. For example, synaptic transmission depends on the geometry of the synapse. Furthermore, the astrocytes that support the synapse, and notably create the synaptic cleft, might actually play an important role into the transmission [78, 187]. More importantly, the wiring of the brain is a dynamics process, as synapses are created and destroyed on a daily basis, while existing synapses adjust their efficiency along time. This phenomenon, called *synaptic plasticity*, was formalized by Donald O. Hebb [133], and has today become one of the most important fields of research in neuroscience. It is of crucial importance for the good functioning of the brain, in particular for learning and memory. All these complex and interesting phenomena widely outrange the scope of this manuscript.

1.3 Spatial organization of the brain

1.3.1 General overview

The mammalian nervous system is a three-dimensional object presenting a strong spatial organization at every scale. It can be divided into two major parts. On one hand, the *Central Nervous System* (CNS) - on which we will further insist - contains the *brain* and *spinal cord* and shelters the higher cognitive functions. On the other hand, the *Peripheral Nervous System* (PNS) is made of all the nerve cells and fibers lying outside the CNS. It is essentially involved in conveying sensory messages (touch, pain, temperature), and motor commands (voluntary movements) between the spinal

cord and the different parts of the body. Moreover, the PNS ensures the good functioning of the internal organs. This vital function, *homeostasis*, is achieved through the careful regulation of the body constants (blood pressure, temperature, etc.).

In the CNS, the brain and spinal cords continuously join through the *foramen magnum*. As vital organs, they display several level of protection. Firstly, they are encapsulated into bone structures: the skull and spine. Secondly, they are wrapped into three successive layers called *meninges*. The most external and toughest one is the *dura*, that notably carries a venous network extracting the blood out of the CNS. The second layer, *arachnoid*, contains the *cerebrospinal fluid* in which the CNS is immersed. This fluid has both a shock-absorbing and a nutritive role. Eventually, the *pia* is a delicate impermeable tissue tightly enclosing the brain and spinal cord.

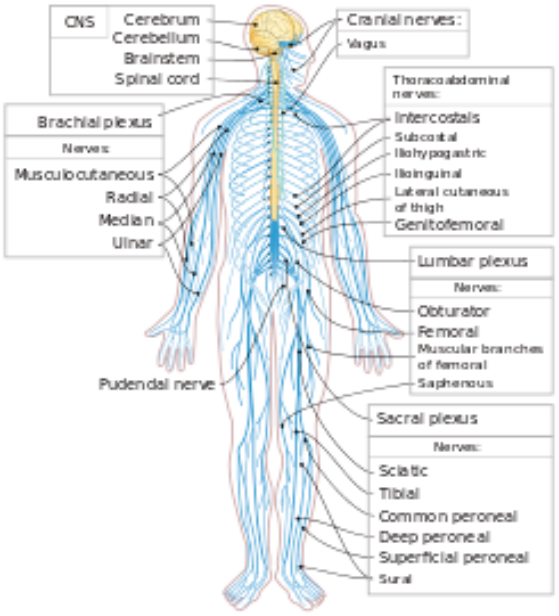


Figure 1.5: General overview of the human nervous system, with PNS (blue) and CNS (yellow). Reproduced from wikipedia Common, credit: William Crochot.

The brain can be divided into three main parts. At its base, the *brainstem* (medulla, pons and midbrain) is the continuation of the spinal cord. Besides being an unavoidable pathway for corticospinal communications, it also notably controls respiration. Behind it, at the rear of the head, lies the peach-size *cerebellum involved* in equilibrium and motor coordination. As the human cortex (see below) it displays a folded surface, characterized by very deep sulci, and condenses fifty percents of the neurons of the brain, justifying its name: ”little brain”. The *forebrain* is the rostral-most part of the nervous system, connecting the brainstem at its extremity. It is composed of the *diencephalon* and *telencephalon* or *cerebrum*. The former notably regroups the *thalamus* - a relay station for incoming pathways to sensory and motor areas of the cerebral cortex -, the *hypothalamus* regulating the autonomic nervous system, as well as the *retina*. The cerebrum is composed of two symmetric hemispheres - right and left - gathering the *cortex* as well as some subcortical areas: the *olfactory bulb*, the *hippocampus* where lies memory, the almond-shaped *amygdala* controlling emotions, and *basal ganglia* involved in procedural learning. These two hemispheres are linked through a flat bundle of nerve fibers: the *corpus callosum*, ensuring their good com-

munication.

The different part composing the nervous systems are complexly dependent, and are intricately linked by tracts of nerve fibers. This wires are so intertwined that tracking them is a difficult task. Let us insist on the presence of two different circuits: a local short-connections, far-reaching long-connections regrouping into thick myelinated cables. The latter compose the *white matter* of the brain, responsible for the good communication between distant areas, while the *gray matter* is characterized by a high density of cell bodies. Certain large fibers can be seen by the naked eye and tracked on considerable distance, giving many insights on the function of the connected regions. Nevertheless, while it was undertook for the entire nervous system of C. Elegans composed of three hundred neurons, mapping the entire *connectome* of the human brain is still out of technological reach [214].

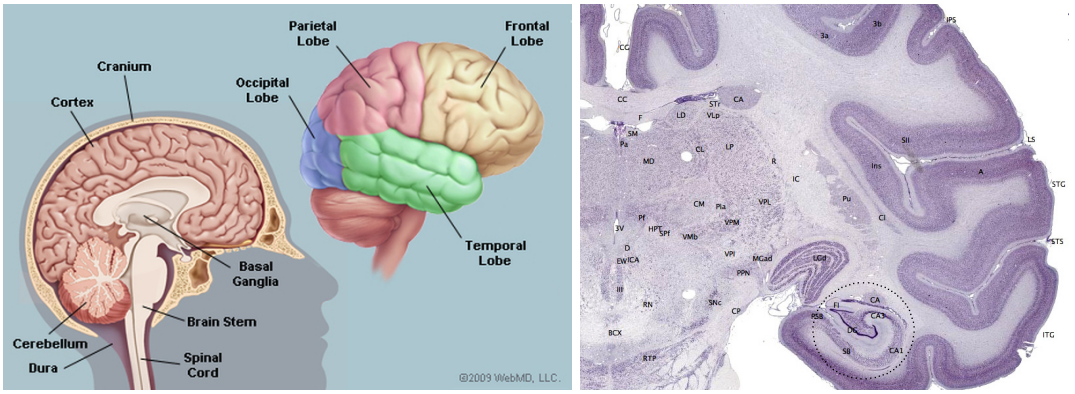


Figure 1.6: General structure of the brain. © 2014 WebMD, LLC. All rights reserved. (left). White and gray matter. Reproduced from Wikipedia Commons (right).

1.3.2 The mammalian cortex: from functional localization to microscopic organization.

The cerebral cortex is a thin extended sheet of tissue covering the outer surface of the cerebrum. While almost inexistent for fishes and amphibians and very rudimentary for reptiles, it displays for all mammals a stratified organization composed of six distinct layers. In contrast with rats or mice presenting smooth hemispheres, the human cortex is convoluted, forming cavities (sulci) and ridges (gyri), allowing its surface to triple within the skull. The most important sulci divide each hemisphere into four distinct regions: the frontal, parietal, temporal and occipital lobes, respectively associated with the planning and organization of future actions, the processing of sensory information, hearing and other aspects of language and memory, and vision. The human cortex contains up to 28 billion neurons, is approximately 2600 square centimeters wide and 3-4 mm thick. As the siege of the highest cognitive functions, it is undoubtedly the most fascinating region of the human brain.

Its important functional role has been overlooked until the beginning of the nineteenth century. In fact, since the European Renaissance, only a few scientist were keen to show the cortex any interest, and their work were largely ignored [127]. It is not a detail that the meaning of cortex is bark. Ironically, it was the dubious *phrenology*, developed by the German physician Franz Joseph Gall, that conceptually revolutionized our vision of the cortex by spatially dividing the cortex into dis-

tinct functional areas. Still impregnated with dualist considerations, the scientific community fiercely withstood this new depiction of human nature. This initiated one of the most roaring controversies of the nineteenth century known as the holism versus localism debate. It ended with the prevailing of the latter. Indeed, in 1861, Paul Broca, demonstrated the link between speech impairments with brain damage to the left hemisphere. In 1870, German researchers Fritsch and Hitzig located the motor cortex at the rear of the frontal lobes, and notably reopened an experimental highway: cortex electrical stimulation. This technique was then thoroughly put at use by the Scottish neurologist David Ferrier to dramatically strengthen localism theory. In 1875, Ferrier discovered the auditory cortex in the temporal lobe. This same year, following the pioneer work of the Italian Bartholomeo Panizza, the German physician Munk positioned the visual cortex in the occipital lobe. Furthermore, complementing Broca's research, the German Carl Wernicke found in 1874 another cortical region associated with a new form of aphasia: unintelligible speech pattern with incapability of comprehension. All these advances suggested that the cortex was divided into distinct functional area, but also lateralized as speech was mainly present in the left-hemisphere. Meanwhile these picture of the cortex should be mitigated in certain regards, it has nevertheless proven largely valid until today.

From the anatomical viewpoint, the mammalian cortex is an homogeneous medium organized both horizontally and vertically. At the horizontal level, it displays six distinct layers (laminae I to VI, the first being the most extern) that constitute the frame of cortical pathways. Each layer has its own characteristic composition, and its specific set of connections with the other layers as well as the other cortical cortical and subcortical areas. On top of this stratified organization is a vertical one whose smallest anatomical and functional unit is the cortical *microcolumn* (or *minicolumn*).

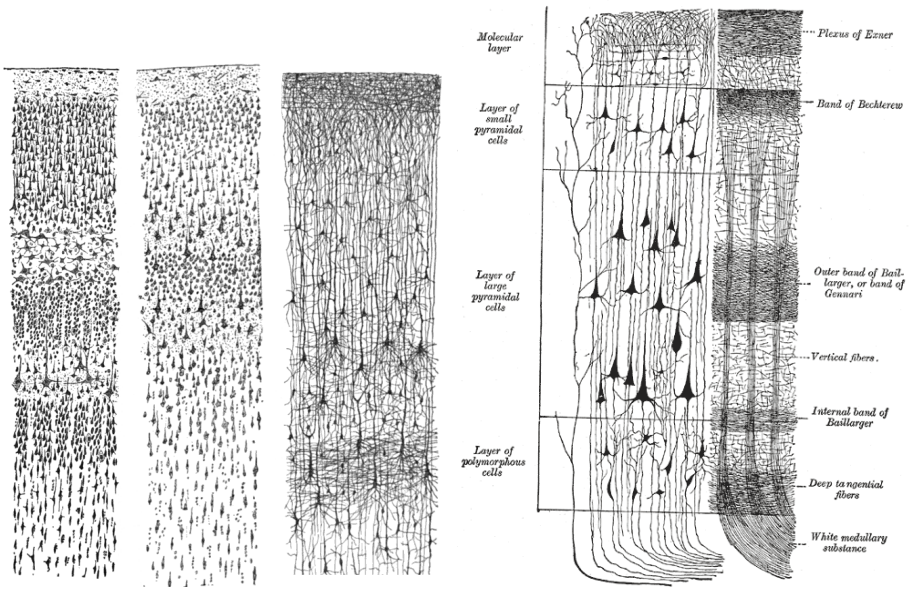


Figure 1.7: Layered structure of the cortex. Each layer has its own composition and typical connectivity. Left: drawing by Santiago Ramón y Cajal. Right: reproduced from [125]

This microscopic structure consists of around a hundred preferentially interconnected neurons, organized as column 20-60 micrometers wide, that vertically traverse

laminae II-VI. Within a microcolumn, neurons display an homogeneous level of activity. The diversity of the neurons in each column is representative of the cortex composition, with around 20 percents of inhibitory neurons playing a central role for its good functioning. In various cortical regions, microcolumns gather into larger functional unit called *cortical column* or *macrocolumn*, each specific of a given region and whose variety has been reviewed in [177]. While these macro structures are quite versatile, all have in common an approximate diameter of 0.5-1 mm. They are seemingly a well preserved organizing principle of mammalian cortex. These columns have specific functions and spatial locations resulting in the presence of delays in their interactions due to the transport of information through axons and to the typical time the synaptic machinery needs to transmit it. These delays have a clear role in shaping the neuronal activity, as established by different authors (see e.g. [70, 213]). Moreover, neurons and microcolumns typically have two different scales of connections in the cortex (see 1.8). At the microscopic level, they project many connections toward their most proximate siblings, while they send a few far-away synapses to determined areas.

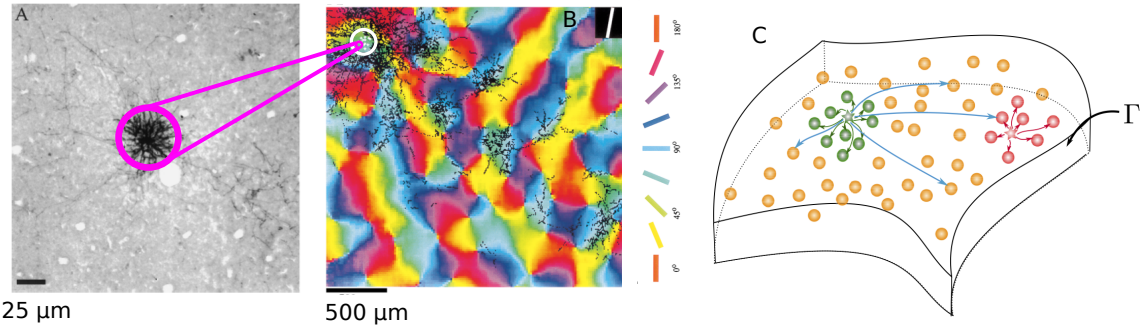


Figure 1.8: Two different scales of connection in the primary visual cortex of a cat. At the microscopic level, neurons connects many of their neighbors in a random fashion (A). At a higher level, a patch of neurons will send connections to other patches processing the same task. Colors correspond to preferred orientation of neurons (B). Sketch of the two scales of connections for an abstract representations of microcolumns (C). Modified from [31]

Lorente de Nó, a former student of Ramón y Cajal, was the first to envision this possibility of such a columnar organization [215]. It was latter evidenced electrophysiologically in the somato-sensory cortex of cat by Vernon MountCastle in [176]. Albeit controversial [138, 145], microcolumns have been accounted for by diverse techniques ranging from Nissl-staining method, optical density measurement, to metabolic 2-deoxy-D-glucose methods, and might replace single neurons as the functional units of the cortex. Several fact support this theory. First, while the synaptic transmission in the cortex is generally of 1-5 ms and might have an important impact on its function, the latency in one column is very small making it an acceptable indivisible structure [49, 177]. Furthermore, there is not enough myelinated connections in the cortex for single neurons to form one to one long-range synaptic connections. For example, in the primary visual cortex, callosal termination seems to correspond to the output of an entire orientation hypercolumns. Lastly, microcolumns can display more sophisticated behaviors than single neurons, making them very adaptable structures.

1.4 Mathematical models in Neuroscience

When taken alone, a neuron already displays a complex dynamics. It was carefully described in the seminal work of Hodgkin and Huxley in 1952 on the giant squid [135]. This animal presents uncommonly large axon fibers, yielding the possibility to wire it and record its excitable properties. Hodgkin and Huxley translated them into equations, and tailored the canonical model - still widely studied - of the single neuron dynamics: a set of four coupled non-linear ordinary differential equations, taking into account both the membrane potential, the opening and closing dynamics of the channels covering the surface (sodium and potassium), and the leak of the membrane. Because these equations were quite intricate, some reductions have been proposed in the next decades, such as Morris-Lecar model [175] and FitzHugh-Nagumo model [109, 140]. These simpler models all presented the advantage of being tractable and of preserving the excitability properties of neurons. As the present manuscript is exclusively interested in the dynamics of neural networks, we refer to [60] for further details on the intrinsic dynamics of single neurons.

Neural networks certainly fall into the scope of the theory of "complex systems". These systems are generally composed of many elementary units (spins, oscillators, neurons, etc.) interacting with each other through complex non-linear interactions and possibly subjected to transmission delays. In fact, the human brain is made of around 80 billiards neurons, each projecting up to 10 thousands synapses toward other neurons. Besides the possible intricacy of the intrinsic behavior of the components at play, the aim of complex systems theory is to discover, describe and categorize the possible dynamics of such large networks. One of its main insights is that, due to the non-linearities in the model, the system is not equal to the sum of its parts. It might indeed displays unexpected emergent behaviors, not deducible from the properties of its elements taken alone. The flocking of bats, in which millions fly coherently, is a striking example of such natural phenomena. Very often, the macroscopic behavior of the system is tuned by some parameters of the model, possibly associated with bifurcations or phase transitions. The system is thus reducible, and many details of the microscopic models become irrelevant. This can concern both the intrinsic dynamics, microscopic interactions, and some heterogeneity parameter in the model: a change in their value will not make a change! This conclusion - along with mathematical tractability - justifies, in some sense, the poor biological relevance of the intrinsic dynamics we will consider in our network equations. It also clarifies what is at stake: the aim is not to perfectly describe a biological neural network, but to capture something of its dynamics in the limit where the number of neurons is very large (known as the *thermodynamic limit*). Before giving the microscopic dynamics of the class of neural networks we will study, let us precise a few hypothesis concerning the neural code and synaptic transmission.

1.4.1 The neural code and the firing-rate hypothesis

A very intricate and interesting issue in neurobiology concerns the *neural code*. In fact, knowing that neurons communicate by electrical means does not tell us how, exactly, the information is encoded. To gain some insights, neurobiologists have carefully recorded and categorized the temporal spiking patterns of neurons. It turns out that these *spike trains* can take many forms: individual spikes, periodic firing, square-waves bursting, periodic spikes intertwined with small oscillations, etc...

There is still no dictionary to translate this bestiary, but it certainly has an important role in the encoding of neural information. Especially if one considers long-range connections, with well identified source and target. If the precise temporal spike pattern, we speak of a temporal code. We can compare this possible code to Morse, where the interval between two flashes carry information.

The problem is that, if one wants to model a network composed of many neurons, taking into account the spike trains of every one of them is an Herculean task, if possible. Even simulating the network's trajectory with a computer seems prohibited. Nevertheless, if one considers a local area of the brain containing only a few million neurons (V1 is thought to contain around 140 millions in each hemisphere [157]), each sending and receiving many connections, you might expect that some averaging effect occurs, and that individual spike trains loose their importance. This is why a common assumption to reduce this difficulty is to consider that in such integrative networks, interactions between neurons can roughly be summarized as a non-linear function of the instantaneous spiking frequency of the presynaptic neuron (integrated on a short time window), also called firing-rate. This non-linearity is described in the next subsection.

While a rough assumption, the firing-rate hypothesis actually makes sense in a number of scenarios. First, it is compatible with modern non-invasive monitoring techniques whose resolution is not precise enough to go beyond the local activity of brain's regions, containing many hundred neurons (EEG techniques, functional MRI, etc.). In fact, in most experiments where a stimulus is presented to an animal, it is possible to find a group of neurons whose firing rate will increase compared with the background activity. Moreover, in certain precise case, the firing-rate seems to condense all the information. For example, it has been demonstrated that the firing-rate of a stretch receptor neuron associated to a muscle fiber is a function of the stretching force applied to the fiber [3, 4]. In these cases, the information is contained in the firing-are of a set of neurons.

Without surprise, the significance of firing-rate hypothesis is quite limited: neglecting (by definition) important temporal features of neuronal activity, it has been challenged by many authors [1, 202, 219]. The real surprise comes from the successes it did earn, providing insights into the mechanisms underlying both hallucinations patterns, binocular rivalry, working memory or the emergence of up-down states in the cortex [94, 99]. These achievements justify the respectability this model as earned among the neuroscientific community, and explain why it is still at use today.

1.4.2 Synaptic models in Neural Networks

Synaptic transmission models are of chief importance in this manuscript. We will thus describe their evolution in details. Nonetheless, the history of mathematical modeling of neural network is a long one, and at least goes back to the 30s [130]. Interestingly, it shares a lot with the development of Artificial Intelligence (AI) and Cybernetics, and it is notable that many early neuroscientists had made contributions to both disciplines. Let us begin in 1943, when the neurophysiologist Warren S. McCulloch and logician Walter Pitts proposed one of the first formal equations for neural integration [169]. Of logical inspiration, their article would become the start-

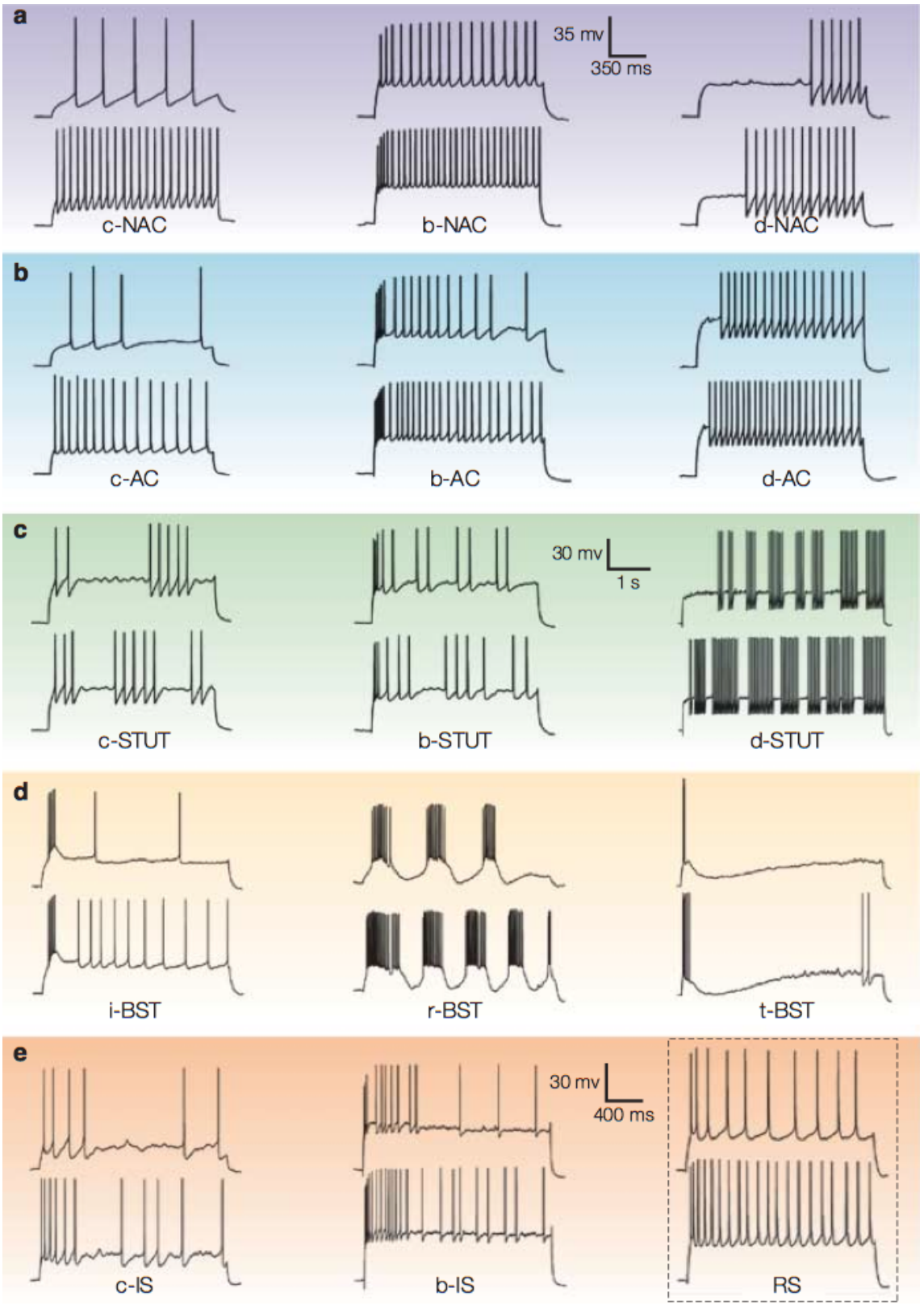


Figure 1.9: Different electrophysiological classes of inhibitory neurons produce different trains of spikes. Reproduced from [166].

ing point of the theory of neural networks, as it opened the possibility to study their discrete time evolution. At the time, it was well known that neurons presented a spiking behavior in order to implement complex representations of the world. Inspired by this biological fact, McCulloch and Pitts suggested *binary neurons*, or *formal neurons*, either excitatory or inhibitory, evolving in a discrete time, and whose *activity* would be described at time $t \in \mathbb{N}$ by a boolean 0 or 1 (or -1 and 1). Moreover, at each time step, every neuron of the network updates its state through a nonlinear function of the afferent activities with a threshold condition, unless inhibition occurs. For a given neuron i receiving inputs from neurons $j \in [\![1,N]\!]$, this writes

$$\begin{cases} u^i(t+1) = H\Big(\sum_{j=1}^N x^j(t) - \theta_i\Big) & \text{if no inhibition,} \\ u^i(t+1) = 0 & \text{if inhibition.} \end{cases}$$

where H is the Heaviside step function, $H(x) = 0, \forall x \in \mathbb{R}_-^*$ (or alternatively -1) and $H(x) = 1, \forall x \in \mathbb{R}_+$, and θ_i is the activation threshold of neuron i. We emphasize that this nonlinear behavior is of chief importance in neuroscience, as the brain is definitely not the sum of its parts. In the 50s, an improved version of this synaptic integration was then invented. Frank Rosenblatt notably put is at use for the implementation of its *perceptron* [206], one of the first learning machines, composed of one layer of neuron and designed for image recognition. In his model, the inhibition was no more absolute, and the afferent connections were weighted, so that a postsynaptic neuron could listen preferentially to a few incoming signals (neurons or inputs) and neglects others:

$$u^i(t+1) = H\Big(\sum_{j=1}^N J_{ij} u^j(t) - \theta_i\Big).$$

Here, J_{ij} is the *synaptic weight* from j to i. It is positive for excitation, and negative for inhibition. This weighted integration thus permits the neuron to integrate a great amount of contradictory messages in order to spike only when needed. This was a very successful innovation that enabled Rosenblatt's machine to discriminate linear features of images. It gave rise to several generations of learning machines, more and more sophisticated: from multi-layers feed-forward perceptrons [10], until today deep learning algorithms. All these models preserved weighted integration, but the discontinuity of the Heaviside step function was sometimes a problem for both the updating of synaptic connections, as well as for theoretical study of the dynamics. A natural way to tackle this issue was to smoothen H into a sigmoid shape function S taking value in $[0,1]$, infinitely differentiable, increasing and odd (up to an additive constant). Popular examples of sigmoid functions are e.g. $x \to \frac{1}{1+e^{-gx}}$, $x \to \mathrm{erf}(gx)$, or $\frac{\tanh(gx)+1}{2}$, where $g \in \mathbb{R}_+$ is the gain that controls the slope at the origin.
This new ingredient made the activity u^i a continuous state variable. More interestingly, the sigmoid being an invertible function, the latter somewhat abstract quantity could now be unequivocally related to the most natural description of the state of a neuron, being its real valued membrane potential x^i:

$$x^i(t) = S^{-1}(u^i(t)), \qquad u^i(t) = S(x^i(t)). \tag{1.1}$$

In addition, the discrete time dynamics of the network could be advantageously de-

scribed in term of the membrane potential:

$$u^i(t+1) = S\big(x^i(t+1)\big) = S\Big(\sum_{j=1}^{N} J_{ij}u^j(t) - \theta_i\Big),$$
$$x^i(t+1) = \sum_{j=1}^{N} J_{ij}u^j(t) - \theta_i = \sum_{j=1}^{N} J_{ij}S\big(x^j(t)\big) - \theta_i,$$

giving the continuous time dynamics

$$\frac{\mathrm{d}x^i(t)}{\mathrm{d}t} = -x^i(t) + \sum_{j=1}^{N} J_{ij}S\big(x^j(t)\big) - \theta_i. \tag{1.2}$$

This is the typical equation for the evolution of a large network. Remark that the intrinsic dynamics is an exponential discharge of the neuron, that can be accounted by a leaky membrane. You might want to replace this simplistic behavior by a more realistic $f(x^i(t),t)$. Equation (1.2) is able to describe both the time evolution of interconnected neurons, or alternatively that of a family of interconnected subnetworks. In the latter case, i is the index of a group of neuron, and x^i represents the averaged activity of the subnetwork i. In both cases, $S(x^j(t))$ needs an interpretation, and the refractory period had to be taken into account for this continuous time dynamics. These both found a solution in the firing-rate hypothesis that has progressively imposed itself among the neuroscientific community. It has at least two interpretations.

When i represents a single neuron, $x^i(t)$ is seen as a smooth average on a short time window of the otherwise spiky membrane potential of neuron i. If the neuron is very excited, it will exactly emit one spike within a time τ_r, where τ_r is the refractory period. Thus, $S(x^i(t))$ represents either the probability that the neuron spike within a time window of width τ_r, either its normalized frequency of spike. A heuristic derivation of the latter interpretation can, for example, be found in [5], [60, section V.A] or [99, Chapter. 12]. In contrast, when i represents a subnetwork of many neurons, $x^i(t)$ may represents the averaged membrane potential of its components, and $S(x^i(t))$ the instantaneous frequency of spike it generates. This view is especially natural when considering a neural field, that is a continuous spatially extended neural network, for which $x^i(t)$ accounts for the local membrane potential on a neighborhood containing infinitely many neurons (see e.g. [245]). Note also that generally $0 \leq S \leq 1$ but, as for the Heaviside function, this sigmoid some time takes value in the interval $[-1,1]$.

Of course, these models are far from embracing the whole complexity of synaptic transmission. Richer microscopic models take, for example, into account the reversal potential of the synapse, the concentration of neurotransmitters within the synaptic cleft, or the aperture's dynamics of the channels covering the membrane of the postsynaptic neuron [99]. As already discussed, these omission are consistent with the fact that we are dealing with a complex systems, and do not want to enter too much into details. Nevertheless, a biological fact that should be taken into account is that the synaptic transmission actually depends on the membrane potential of the postsynaptic neuron x_t^i [99]. In this manuscript, we propose to address A synaptic transmission of the form $b\big(x^i(t),x^j(t)\big)$ through a mean-field analysis performed in chapter 5.

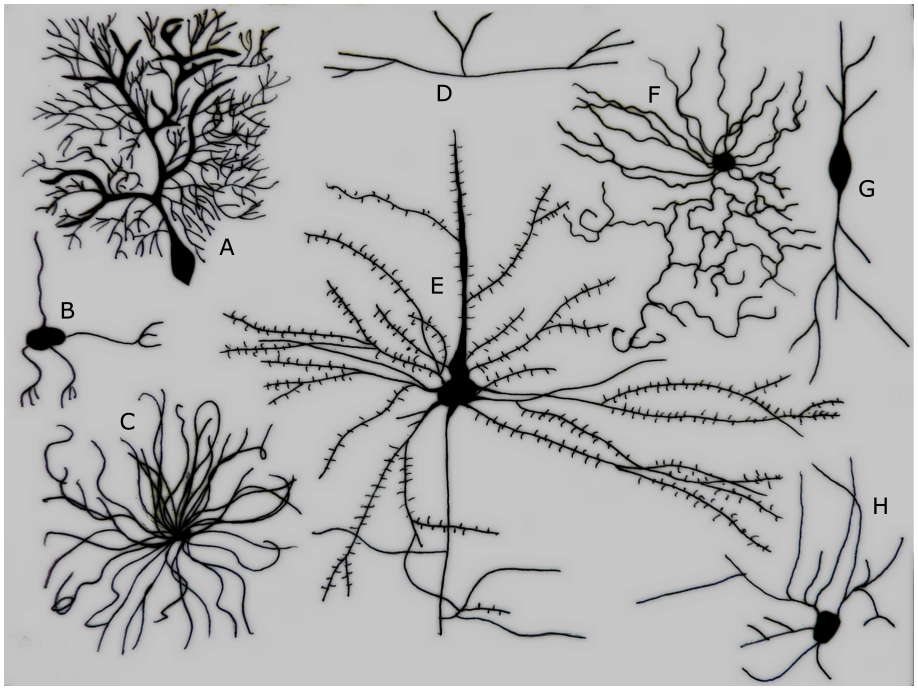


Figure 1.10: Different types of neurons. Drawing by Ramón y Cajal. A. Purkinje cell (a), B. granule cell (b) C. motor neuron (c), tripolar neuron (d), pyramidal cell (e), chandelier cell (f), spindle neuron (g), and stellate cell (h). (Credit: Ferris Jabr; based on reconstructions and drawings by Santiago Ramón y Cajal)

1.4.3 Heterogeneities and noise

The neural network equation (1.2) we have built so far is missing important aspects of real nervous system: variability. In fact, the brain is a very disordered systems, as it is made of highly heterogeneous components - from neurons to synapses, channels, and minicolumns of the neocortex - and displays a noisy dynamics at every level. These features seem to play an important role in the functioning principle underlying cognitive abilities, as well as in the emergence of pathological state. In order to fully account for the biological observations, mathematical representations of neural networks has to include these important traits.

Different neurons, different channels, etc.

The diversity of neurons composing the brain is one of its most striking feature. Some of them are well known by neuroscientists, as the excitatory pyramidal cell of the cortex or the inhibitory Purkinje cell of the cerebellum. Though, its doubtful that any neuroscientist know them all, as more than 10.000 different kinds of neurons have already been referenced [99].

Neurons can distinguish themselves in many ways. Two of their fundamental characteristics certainly are the type of channels and synapses they involve. On the one hand, channels are themselves very diverse, and can take up to a hundred forms [99]. They are specific to a given ion (calcium, potassium, sodium, chloride, etc.), can be voltage-gated, ion-gated, neurotransmitter-gated (NMDA, AMPA/kinase for glu-

tamate) or passive. They may have activation or deactivation pathways, and can be fast or slow. Along with the porosity of the membrane and the diameter of the neural fiber, they impact, through their type and density, the excitability properties of the neuron: speed of conduction, threshold, refractory period, as well as the resting membrane potential. Let us remark that a neuron generally involves more than two type of channels.

On the other hand, synapses offer many possibilities. While electrical synapses are the simplest and most stereotyped, chemical ones can involve a few neurotransmitters among dozens of candidates, display various geometries, and adopt different strategies to enhance deactivation (through pumping or destruction of neurotransmitters). Their size relates to their number of vesicles and, along with their location on the postsynaptic dendrite, importantly impact the synaptic efficiency. Moreover, their neurotransmitters can either induce excitation or inhibition, directly bind to a postsynaptic channel or trigger the G-protein, and are associated with different time characteristics.

Note also that, while microcolumns present a clone-like composition and though interconnections in the mammalian cortex seems to follow a robust pattern prescribed by neurodevelopment, it appears that these structures and the connections they display are also subject to variability [49, 151].

This microscopic diversity has certainly been a driving force in the evolution of the animal reign, and has enabled the shaping of highly specialized neural components such as the respiratory pacemakers of the brainstem, able to generate synchronized oscillations for breathing control through persistent channels. For this reason, it seems to be an unavoidable feature of nervous systems, that must necessarily impact its functioning, even at the macroscopic level. Supporting this view, it was shown that neurons of rats subject to febrile seizure displayed a similar mean resting potential than ones of normal rats, but a much greater variance [13]. Hence, a problematic is to identify which microscopic parameters are of importance in the emergence of functions and pathologies of the brain, and which are not.

This is an intricate question, as it was for example shown that different networks composed of very distinct types of neurons where nonetheless able to display very similar behavior [193]. Among the many details present at the microscopic level, the importance of multiple time scale for the emergence of rich circuit dynamics were pointed out [120]. Moreover, in a different dynamical approach, Izhikevich very coherently addressed a similar issue at the level of the single neuron. The author proposed to categorize them according to their dynamical properties and bifurcation diagrams only. He was thus able to differentiate four classes of neurons through their character monostable or bistable, as well as oscillator or resonator [139]. From our mean-field viewpoint - presented in the next chapter 2 - the criterium is again different, as the parameters of importance are those that conserve an influence on the macroscopic behavior of the system, when the number of neurons is very large. Even more important are the microscopic characteristics associated with bifurcations, the crossing of which causes drastic changes in the networks dynamics.

Such a phase transition was famously discovered by Sompolinsky, Crisanti and

Sommers [221] in the context of a homogeneous population of firing-rate neurons with leaky membrane, and displaying heterogeneous synaptic connections modeled as centered independent Gaussian variables:

$$\dot{x^i}(t) = -x^i(t) + \sum_{j=1}^{N} J_{ij} S(x^j(t)), \qquad i \in \{1,\dots,N\}, \tag{1.3}$$

with $J_{ij} \overset{\mathscr{L}}{=} \mathscr{N}\left(0, \frac{\sigma^2}{N}\right)$, and $S = tanh$.

Importantly in this model, the variance of the disordered connections scaled as $\frac{1}{N}$ (we will come back on such scalings in chapter 3). The authors found that, in the limit of an infinite number of neurons, the system displayed a sharp phase transition in the level of the variance parameter σ. When σ was smaller that unity, all the individual neural trajectories were attracted to the stable equilibrium 0. In comparison, when σ was greater than 1, the mean activity of the network was still 0, but the individual trajectories appeared to sustainably fluctuates around the origin in a very incoherent way, a sign of deterministic chaos. This discovery was a shock to the neuroscientific community, as chaos had already been observed in the rabbit olfactory bulb through EEG techniques [18, 113]. Hence, despite some biologically irrelevant aspects such as the all-to-all connectivity, this work strongly demonstrated the importance of taking into account the heterogeneous character of nervous systems in order to unravel its functioning.

Noise

It has been long known that the nervous system is an unreliable system, incapable to accurately reproduce a given behavior. Around the 50s, trial-to-trial variability had already been observed in frogs: the same stimulus repeatedly applied to the same neuron evoked different time of first spike, and small shifts in the inter-spike interval [38, 44]. This variability was attributed to noise, that was classically seen as an *unpredictable perturbation corrupting the signal*. In comparison, machines of the next decades appeared much sounder to accomplish orders.

Nevertheless, this negative understanding must be challenged in the context of neuroscience. In fact, in the last few decades, noise has progressively been accepted as one of the most important trait of biological system in general, and of the nervous system in particular. If noise is indeed unpredictable, there is a gaining attention of the community that it might also be of some benefits [96, 101, 170, 223], and constitute a significant part of the neural code [223]. This could be explained by the fact that nervous systems has seemingly adapted to noise and turned its unavoidable presence to its advantage. These new insights might resolve the apparent paradox formulated by John von Neuman in 1956 [242]: "how can a reliable nervous system be made out of unreliable elements?", and are reminiscent of the cost effective "optimal control theory" proposed by Harris an Wolpert [131] to explain the unreliability of motor control.

There are several source of noise in the brain [96, 101]. First, *sensory noise* accounts for the stochastic character inherent to every stimulus received from the out-

side world. In fact, when experiencing smell, taste, or visual stimulus, the precise number of molecules captured by your olfactory receptors is subject to chemical noise, while the number of photons absorbed by your retina is of quantal nature. *Membrane noise* results from both the perpetual motion of ions, electrons, molecules and protein due to thermal agitation within each cell, as well as the stochastic opening and closing of ion channels on the surface of the neuron, resulting in transient current through the membrane. These latter random events follow a Poisson law, and induce local depolarizations of the membrane affecting the transmission and propagation of action potential. *Synaptic noise* condensed the effects of chemical diffusion and binding of neurotransmitter within the synaptic-cleft, and those of the random spontaneous releases of vesicles inducing well accounted for postsynaptic current. Eventually, *calcium noise* should also be taken into account [99] because of the combination of a low inner concentration of calcium(10^{-4} millimolar), and a high sensitivity of calcium-gated channels to its presence.

Perhaps one of the most well-understood and documented potential benefit of noise to neural computation is the so called *stochastic resonance*. In fact, membrane noise is thought to improve the acuity of neural integration by enabling subthreshold stimulus to prompt an action potential with a positive probability. Stochastic resonance was notably discovered in the visual cortex of cats [159], and has since attracted much attention (see [170] and reference there in). Mathematically, this problem fall into the study of first-passage time of stochastic processes. Another possible noise-related advantage would be an increased adaptability of nervous system: by permitting assembly of neurons to explore more thoroughly their state space, they might overcome changing of the environment with more talent [101]. Noise might also be an important component of a cost-effective strategy, improving the trade-off of the nervous system which might result in an improved fitness. Strikingly, noise might sometimes enhance the reliability of response of the system: in the case of a frozen noise input, it was shown that the trial-to-trial variability of the response of neural units was significantly reduced [96, 164]. Moreover, theoretical models suggest that noise might also better the regularity of spikes, a phenomenon known as *stochastic coherence* [155, 191]. For example, in the case of intrinsically oscillatory neurons, Ermentrout argued that a correlated noise could facilitate the emergence of synchronization [96, 99]. Similar conclusions were drawn in another theoretic work [239], in which the effect of noise induced stabilization or destabilization of both fix points, and limit cycles. Note, however, that noise might also be involved in pathological pathways, as noised-induced epilepsy was theoretically described [150].

As advocated by various authors [101, 170], understanding the role of noise in neural computation and pathologies will require innovative stochastic mathematical models. This motivates the incorporation of independent Brownian noise at the level of each cell. Both the independence, and white noise hypothesis are convenient, as they facilitate the mathematical tractability. They will be central in our proof. Nevertheless they are somewhat crude, as we have seen that thermal noise is not the only source of stochasticity in the brain, and as there might exist correlations at least in the *sensory noise* when group of neurons are processing the same stimulus. We point out that some authors have considered models with more intricate colored Ornstein-Uhlenbeck noise, displaying autocorrelation functions with exponentially decay, but this choice did not drastically impact their conclusions [158].

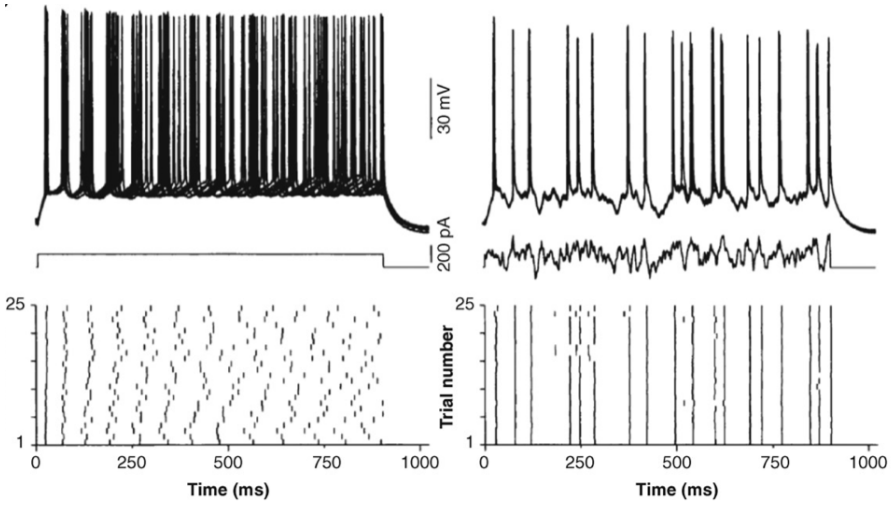


Figure 1.11: Noisy inputs cause reliable spiking in vitro and in vivo: steady-state current injection into a cortical pyramidal cell in vitro results in trains of action potentials shown as voltage traces (top) or spike rasters (bottom). On different trials the first spike is evoked at the same time on each trial, but subsequent spikes are unreliable. Reproduced from [96]

1.4.4 Neural fields

As accounted in the previous section, spatial architecture is a crucial feature of biological neuronal networks at both the microscopic and macroscopic level. Dynamically speaking, this topological extension offers the possibility for a wide repertoire of spatial patterns of activity. Early physiological evidence accounted for their existence in the nervous system [11, 47, 48, 87, 130, 132, 132, 199], including alpha, beta, and delta rhythms as well as more complex spatio-temporal recording such as wave propagations [11, 46], and thalamic undamped oscillations thought to encode external stimuli [11, 91, 110, 192].

Around the same period, technological progresses and the emergence of computers allowed to simulate the dynamics of discrete spatially extended randomly connected networks of excitatory neurons. The first such contribution was made by Rochester et al. [204], shortly followed by that of Farley and Clark [102] for a planar network. Transient diffusive reverberations, oscillatory patterns, and traveling waves were observed, but these networks were not able to display sustained activity: after a period they were either saturated, or quiescent. This latter phenomenon, poorly representative of the biological brain, was known as the *switching effect*, and has notably highlighted the tremendous importance of inhibition for the brain functioning [130].

In order to account for these interesting phenomena, a spatially continuous mathematical model of the nervous system - also called *neural field* - was needed. In 1956, a pioneering contribution was made by Beurle [28] that derived a neural field equation for a network composed of excitatory synapses, and considered the firing-rates of local populations as the quantities of interest. Moreover, in 1963, Griffith showed sustained oscillations for a network including inhibition, in an attempt to solve the switching effect [126]. Inspired by these works [69], HG Wilson and JD Cowan made

the real breakthrough in 1972 [245, 246]. They proposed a set of two coupled non-linear integro-differential equations describing the dynamics of interconnected excitatory and inhibitory populations of neurons:

$$\begin{aligned}\tau_E \dot{E}(t) &= -E(t) + (k_E - rE(t))S_E\big(J_{EE}E(t) - J_{EI}I(t) + k_E P(t)\big)\\ \tau_I \dot{I}(t) &= -I(t) + (k_I - rI(t))S_I\big(J_{IE}E(t) - J_{II}I(t) + k_I Q(t)\big).\end{aligned}$$

A spatially extended two-dimensional version of these equation followed in 1973 [246]. While mathematically intricate, their model successfully accounted for many spatio-temporal patterns as it displayed multiple spatially inhomogeneous steady states, hysteresis phenomena, sustained oscillations, traveling waves, and predicted some of their features such as the wave velocity etc. Note also that at the same period, Amari independently proposed similar equations for a two-populations inhibitory-excitatory neural networks that also displayed sustained oscillations [5, 6] but without spatial extension. The resolution of Wilson and Cowan's model was then undertaken in the context of a unidimensional neural field presenting lateral inhibition [7].

Recently, research on *neural fields* has attracted great endeavors that have built upon these precursory work in order to mathematically describe the dynamics of large spatially extended neural network. Some of them notably includes long-range connections, bi-dimensional spatial extension, and transportation delays. They exhibit a number of spatio-temporal dynamics such as solitary traveling fronts and pulses, stationary pulses, spatially localized oscillations (breathers), spiral waves, Turing-like patterns, bumps [67, 68, 93]. Many of these spatio-temporal phenomena has been physiologically accounted for. In vivo, traveling waves have been observed in the cat's primary and secondary visual cortex (V1 and V2), along with compressions and reflexions phenomena [250], while analysis of spatio-temporal patterns in the somatosensory cortex of rabbits [111], and rodents [65, 107, 190] has been performed. In vitro, electrically evoked traveling waves are observed [61, 201, 249]. Furthermore, these mathematical models have successfully accounted for a wide range of neurobiological phenomena including hallucination patterns [37, 94], orientation tuning in V1 [25, 220], short-term memory [55, 156], control of head direction [251], and motion perception [119].

Nevertheless, if these patterns are thought to be related with both normal functioning, development of the brain along with pathological states, their computational utility are, for their great majority, yet to be settled. As advocated by Paul Bressloff in [32]:
"one of the major challenges in neurobiology is understanding the relationship between spatially structured activity states and the underlying neural circuitry that supports them."

In the brain's development, calcium waves in glial cells might coordinate neuronal division and migration, spontaneous retinal waves are seemingly related to an activity-dependent development in retina, whilst traveling waves seem involved in the maturation of neurons [33].Furthermore, we know that epilepsy induces cortical waves as well as spatially localized oscillations [172]. All these phenomena are still

waiting a full resolution.

In this manuscript, we will address the mean-field dynamics of a microscopic spatially extended neural network (contained in an open set of $\mathbb{R}^d$). It presents random connections, transportation delays and is subject to Brownian noise at the level of each cell (see 4). Importantly, the variance of the random connections scale as $\frac{1}{N}$, where N is the number of neurons in the network. As a consequence, the effect of these heterogeneities is still felt in the mean-field equation. In contract, classical random connections with variance scaling as $\frac{1}{N^2}$ have no impact on the limit (this will be discussed in Chapter 3 3). Moreover, the latter also accounts for the effect of noise. Our motivations are two-folded. Our first objective is to rigorously address the mean-field limit of a spatially extended network. Our second motivation is that the obtained limit neural field equation takes into account both the heterogeneity parameter of the connections, and the amplitude of noise. In fact, these two important aspects were notably lacking in the Wilson-Cowan equation and many of its derivations. Interestingly, heterogeneities seem to importantly impact the dynamics of neural fields [34, 98, 149].

All the microscopic models of neural networks we have built so far are hard to address. In fact, they constitute high dimensional nonlinear differential equations involving random connections, possibly other sources of heterogeneities and noise. Such a system can certainly be seen as a interacting particle system, for which mean-field theory is very adapted. We now introduce this theory.

CHAPTER 2

THE MEAN-FIELD APPROACH

The anatomical picture of the cortex suggests that we may usefully employ statistical methods in the study of its function.

— Norbert Wiener, 1948

Since Ancient Greece until today, human species have strained to describe nature and to understand its laws. This thousands-year-old quest, of which Thales of Miletus has been one of the great initiators, has encountered radiant successes. The Gravitation theory of Isaac Newton, the Electro-Magnetism of James Clark Maxwell, and the Relativity theory of Albert Einstein are a few examples.

Still many systems provided by nature resist a full human understanding. In particular those composed of a great number of interacting particles with nonlinear interaction have constituted a longstanding challenge, as well as an important key issue. In fact, for the Universe is discrete and hierarchically ordered, these systems of many bodies are virtually present at every scale, in both physical and biological systems. The Milky Way with its several hundreds of billiard of stars, the brain composed of around 80 billiards neurons are only a few examples of such puzzling. Without mentioning the dynamics of fluids, and the vertigo caused by Avogadro number. Especially if you take into account the interactions.

In these models, the microscopic laws governing the individual evolution of any given particle are generally simple and well understood. For example, we can exactly describe the influence of a star onto another, and perfectly solve the equation of movement when the number of stars remains small. Nevertheless, when the number of particles and thus interactions is so huge, the macroscopic behavior displayed by the system is far from obvious, and might contain a bizarre phenomenology. In a microscopic context, the underlying dynamics is generally inaccessible for scientists. It is unobservable as the systems is composed of very tiny particles, whose state (position, velocity) are changing fast, and very hard to measure. We could only observe macroscopic quantity, that are mean quantity averaged over a great number of particles, the study of which has precisely been the object of classical thermodynamic.

Historically, the first known attempt to derive macroscopic properties from a microscopic model is due to Van der Walls, through its early work on the dynamics of gas. Nevertheless, mean-field theory, as well as the emergence of statistical mechanics, are more tightly linked with the equation of Ludwig Boltzmann for the description of the dynamics of rarefied gases. In this work, Boltzmann famously formulated a *pre-collisional independence* hypothesis between two particles just before they collide. Under this assumption, also known as *molecular chaos* or *Stosszahlansatz*, he

Figure 2.1: Ludwig Eduard Boltzmann (February 20, 1844 – September 5, 1906).

computed the collision term of two particles, and reduced the dynamics of the gas into its celebrated equation. A few decades later, its equivalent with Coulomb interaction - the Vlasov equation - appeared to describe the dynamics of both plasmas and galaxies with an astonishing accuracy. Justifying these two models have been a permanent endeavor during the twentieth century, but only partial solutions were found [141]. They still constitute the two main problems of statistical mechanics out of equilibrium [173].

Mean-field theory has now found applications in many discipline. It could be defined as the domain of Statistical Mechanics and Probability theory involved in the rigorous derivation of mesoscopic descriptions for many-body systems. While these latter are generally very difficult to solve exactly, the general approach is to reduce the microscopic dynamics into a unidimensional one that drives a "typical" particle, and keeps tracks of the statistical properties of the system. The basic idea of mean-field theory is to integrate the numerous interactions influencing the trajectory of a given particle into an effective macroscopic term, that can be seen as a (mean)-field at the level of each particle. Before putting this theory at use for our biological purpose, let us now introduce the basic framework of the theory, and explain its key concepts.

2.1 MEAN-FIELD FORMALISM

In this section, we introduce the general framework of mean-field theory. Formally speaking, the latter deals with the behavior of large system in which every particle is feeling, at all time, the contribution of a large number of other particles in the system. It also reduces its scope to cases in which pairwise interactions vanish in the *thermodynamic limit*, i.e. when the number of particles tends toward infinity. Under these hypothesis, the effective interaction term driving each particle can be described as a field accounting for the infinitesimal influences of all the other particles. This definition discards, in particular, the rarefied gas of hard spheres studied by Boltzmann as, in his model, interactions only occur locally in time and space, during the collision of two particles. Before giving our mathematical framework, let us define three different scales at which a many-body system can be studied:

- the microscopic scale is attached to perfectly describe the state of the system,

by taking into account the time evolution of all the particles in presence.

- the macroscopic scale describes emergent observable properties of the system, when they exist. It mostly consists in a collection of local mean quantities, such as magnetization for a collection of spins, or temperature for a gas.
- Between these two scale, is the mesoscopic scale. It keeps track of the statistics on the particles, and gives access to the probability that a given event occurs. It might be able to study the deviation from "standard behaviors".

We will mostly be interested into the third scale, as the first one is generally too cumbersome - if even possible - and certainly unnecessary, and as the second is too partial, and somewhat arbitrary as we actually choose the quantities of interest among those that are observable.

Let us then consider a system of N interacting particles whose states at time $t \in \mathbb{R}_+$ are described by the state-vector $X^N(t) := (x_t^1, \dots, x_t^N) \in E^N$. Here E denotes the space of all possible states accessible to a particle. We will generically assume that $E \subset \mathbb{R}^d$ for some $d \in \mathbb{N}^*$.

Among the wide variety of mean-field problems, we will for simplicity restrict our attention to systems that evolve accordingly to the following class of first order microscopic dynamics, possibly noisy, and with additive nonlinear interactions:

$$dx_t^i = \Big(f(x_t^i, t) + \sum_{j=1}^{N} b(x_t^i, x_t^j) \Big) dt + \lambda \, dW_t^i, \quad 1 \leq i \leq N. \tag{2.1}$$

In this equation, $f : E \times \mathbb{R} \to \mathbb{R}^d$ condenses the behavior of the free particle, $b : E^2 \to \mathbb{R}^d$ describes the structure of interactions, and W^i accounts for the eventual sources of randomness with $\lambda \geq 0$. We also suppose good properties for the functionals f and b in order to have existence and uniqueness of solutions for (2.1) for any initial conditions.

In the context of a large system with many particles, a natural question arises from this equation: knowing the initial states of every particle, are we able to describe the time evolution of the systems in the limit where the number of particles becomes arbitrarily large? We would be particularly interested in the behavior of the interaction term $\sum_{j=1}^{N} b(x_t^i, x_t^j)$ when $N \to \infty$. In fact, deriving its limit might lead to a reduction of the dynamics: that is a *mean-field equation*. Moreover, studying the solutions of this equation would inform us the behavior of a typical particle in the thermodynamic behavior. Of course, the real system of interest is finite. We simply hope that the thermodynamic limit satisfactorily approximates the system when N is very large. Let us suppose that it does.

To be able to take the limit $N \to +\infty$, we have to cope with several issues. The first issue is the scaling of equation (2.1): if the real system at study presents some kind of observed stability, we want to make sure that the different terms in equation (2.1) do not explode. To this purpose, scaling the equation - considering for example the quantity $y^i := \frac{x^i}{N}$ - might be necessary. We refer to [141, 173, 241] for these kind of questions, as they will not be a problem in our analysis.

A second issue is that the state-vector $X^N(t)$ is living in E^N which grows with the number of particles N. If we are to describe an asymptotic behavior for the system when N goes to infinity, we need to find a proper space independent of N as well as an alternative description of the state-vector $X^N(t)$ that would take value in this space. A natural quantity could be the arithmetic mean: $\frac{\sum_{i=1}^N x_t^i}{N} \in E$, or more generally quantities of the form $\left(\frac{1}{N}\sum_{i=1}^N \phi(x_t^i)\right)_{t\geq 0}$ for any relevant function $\phi : E \to \mathbb{R}$. These objects might converge in some sense, and provide information on the limit system. Nevertheless, such an approach would not be able to account for a typical particle, or quantify the probability that the systems deviates form its mean behavior. To obtain such a mesoscopic description, we would like to condense all these partial descriptions - also named statistics - into a unique object. It is, therefore very convenient to introduce the empirical measure of the system that constitutes the "true" density of particles:

$$\hat{\mu}_N(t) := \frac{1}{N}\sum_{i=1}^N \delta_{x_t^i} \in \mathscr{M}_1^+(E) \tag{2.2}$$

where $\mathscr{M}_1^+(E)$ is the set of probability measure on E, and $\forall x \in E, \delta_x$ denotes the element of $\mathscr{M}_1^+(E)$ singular x (For any doubt, please refer to the General Notations 7.4.2 at the end of the manuscript). In order to avoid coping with a family of empirical measure $\left(\hat{\mu}_N(t)\right)_{t\in\mathbb{R}_+}$, we will prefer the alternative empirical measure taking into account the whole trajectory of the system:

$$\hat{\mu}_N := \frac{1}{N}\sum_{i=1}^N \delta_{x^i} \in \mathscr{M}_1^+\left(\mathscr{C}(\mathbb{R}_+, E)\right) \tag{2.3}$$

where $x^i := \left(x_t^i\right)_{t\in\mathbb{R}_+}$ denotes the whole trajectory of particle i, $\mathscr{C}(\mathbb{R}_+, E)$ is the space of continuous function from $\mathbb{R}_+$ to E.

If the system at study displays some kind of macroscopic behavior, we can hope that there exists a probability measure $\mu \in \mathscr{M}_1^+\left(\mathscr{C}(\mathbb{R}_+, E)\right)$ such that the sequence of empirical measure $\left(\hat{\mu}_N\right)_{N\in\mathbb{N}}$ would converge in some sense toward μ when $N \to \infty$. There are two approaches to characterize such a measure μ. The first one, favored by probabilists, is to describe it as the solution of a mean-field stochastic differential equation (SDE in the sequel) or equivalently a martingale problem. The second one, more pleasant for specialists of partial differential equations, involves the evolution of the densities of particles through a *Fokker-Planck equation*. Deriving and studying these latter kind of equations is the object of Kinetic theory of gases that does not enter the scope of this thesis.

Going back to the empirical measure, remark that it does not tell you anything on the correlations of two, or more particles. In particular, assuming for example that μ is the law of a Gaussian process with stationary mean and variance does not tell you that the particles are also in a stationary state. They might, as well asynchronously oscillates with a common frequency (so that correlations remain), or evolve in total incoherence with each other. If one is interested in fixing this issue, it is always possible to introduce another kind of empirical measure:

$$\hat{\mu}^{2,N} = \frac{1}{N(N-1)}\sum_{i\neq j} \delta_{x^i, x^j},$$

and more generally, $\hat{\mu}^{k,N}$ for k approaching N. Of course we never go to that point. Moreover, if one wants to take into account other features of the particles into, for example to know the distribution of their velocity, one has to extend the state space and make the velocity appears in the empirical measure.

Furthermore, to be able to speak of convergence of the empirical measure toward a probability measure μ, it will be necessary to introduce proper distance on probability spaces. That will be the object of the next subsection.

In order to obtain convergence results for the empirical measure, we will use the family of p-Vasserstein distance on the space $\mathscr{M}_1^+\big(\mathscr{C}(\mathbb{R}_+,E)\big)$. These distances are compatible with the weak topology, that is with the convergence in law. They are defined for any $\mu,\nu\in\mathscr{M}_1^+\big(\mathscr{C}(\mathbb{R}_+,E)\big)$ by:

$$d^V_{\mathbb{R}_+,p}(\mu,\nu):=\inf_{\xi\in\mathscr{C}_{\mu,\nu}}\left\{\int_{\mathscr{C}(\mathbb{R}_+,E)^2}d(x,y)^p d\xi(x,y)\right\}^{\frac{1}{p}},$$

where $\mathscr{C}_{\mu,\nu}$ denotes the set of couplings of measures μ,ν, and where d is a proper distance on $\mathscr{C}(\mathbb{R}_+,E)$ that we will generically chose to be the supremum norm. For the particular case of the 1-Vasserstein distance we also have the formula

$$d^V_{\mathbb{R}_+,1}(\mu,\nu)=\sup_{f\in\mathscr{C}_b(\mathscr{C}(\mathbb{R}_+,E),\mathbb{R}),1-\mathrm{lip}}\left|\int_{\mathscr{C}(\mathbb{R}_+,E)}f(x)(d\mu-d\nu)(x)\right|.$$

More generally, for $T>0$, $\mu,\nu\in\mathscr{M}_1^+\big(\mathscr{C}([0,T],E)\big)$ and $t\in[0,T]$, we will denote:

$$d^V_{t,p}(\mu,\nu):=\inf_{\xi\in\mathscr{C}_{\mu,\nu}}\left\{\int_{\mathscr{C}([0,T],E)^2}\sup_{0\le s\le t}|x_s-y_s|^p d\xi(x,y)\right\}^{\frac{1}{p}}. \tag{2.4}$$

We will now discuss the suitable choices for initial conditions to equation (2.1), we present another simpler criteria to obtain similar convergence results in the case of symmetric particles.

2.2 Initial conditions and propagation of chaos

To solve the dynamics of the microscopic equation (2.1), we have to specify what are the initial conditions for our system. This is not a trivial matter, as mean-field approach is not classical in two ways.

First, our goal is not to confront the microscopic equation - already assumed valid - with the experiment, but to understand what mesoscopic behavior will emerge from it. Second, even if the experimentalist might want to test its model for the latter, he/she would not be able to ”prepare” the system at the microscopic level, neither to

access the approximate states of all the particles in a short time window for their number is too large, and their evolutions too fast.

Hence, the question is more to understand what are the most probable initial conditions than to actually "choose" them - only nature does. On the one hand, it easy to see why initial configuration is crucial for the convergence toward a mean-field equation. Consider for example N hard spheres in a cubic box. You can always suppose that their initial velocities are all aligned and perpendicular to two of the six faces of the box, and that the initial positions of the spheres are such that they will never collide. In this setting, it is pretty unlikely that the system will converge toward Boltzmann's equation. On the other hand, this kind of pathological initial configurations is so unlikely that you might never encounter it in nature, especially when N is large.

As particles states cannot be known with precision due to the incertitude principle of Heisenberg, it makes sense for such problems to describe initial conditions in a statistical fashion. That can be done through a probability measure μ_0^N over the state space E^N. If we are dealing with homogeneous particle, it is also natural to assume that this initial joint density is symmetric. Furthermore, as the system is "prepared" by nature, it is sensible to suppose that it is near thermodynamic equilibrium, that is near maximal entropy. As a consequence, we often assume that the particles are almost independent at initial time, that is μ_0^N to be approximately equal to a tensor product, configuration that makes the entropy of the system maximal [241].

Let us introduce the appropriate vocabulary:

Definition 2.2.1. A sequence of probability measure $\mu^N \in \mathscr{M}_1^+(E^N)$ is chaotic if exists $(\mu_i)_{1\leq i\in\mathbb{N}^*} \in \mathscr{M}_1^+(E)^{\mathbb{N}^*}$ such that for all $k \in \mathbb{N}^*$, for all bounded continuous $\varphi_1,\dots,\varphi_k : E \to \mathbb{R}$, for all distinct $i_1,\dots,i_k \in \mathbb{N}^*$

$$\int_{E^N} \Big(\prod_{j=1}^k \varphi_j(x^{i_j})\Big) d\mu^N(x^1,\dots,x^N) \to_{N\to+\infty} \prod_{j=1}^k \Big(\int_E \varphi_j(x) d\mu_i(x)\Big).$$

Moreover, if all the μ_i are equal to μ, then μ^N is μ-chaotic.

This notion was famously introduced by Mark Kac in 1956 in an attempt to derive Boltzmann's equation from the rarefied gas problem [146]. He simplified the problem into collision-less model where all the particles interact - making it a mean-field problem - and propose the so-called master equation: in these, when two particles interact, the parameter of the interactions are chosen at random. With this simple model, he was able to precise Boltzmann's concept of pre-collisional independence on mathematical grounds. In this definition, it is important to remark that k is fixed when N goes to $+\infty$. Hence, μ^N being chaotic basically means that particles become more and more independent when taken in sufficiently small number (for an improved propagation of chaos, see [24]).

For example, $(\mu^{\otimes N})_N$ is the purest μ-chaotic sequence, as the property is already valid for any finite N. Remark that it is possible to quantify how chaotic a distribution is by controlling the decay of the difference between each sides in definition 2.2.1.

Moreover, it turns out that it is equivalent to prove the property for $k=2$ only. This is a consequence of the following famous result by Alain-Sol Sznitman [231], that also crucially highlights the link between chaos and convergence in the symmetric case:

Theorem 2.2.1. *Let $(\mu^N)_N$ be a sequence of symmetric probability measures on E^N respectively, and $\mu \in \mathscr{M}_1^+(E)$. Let also $(X^1,\dots,X^N)$ have law μ^N. The following statements are equivalent:*

(i) *$(\mu^N)_N$ is μ-chaotic,*

(ii) *for any continuous bounded $\varphi_1,\varphi_2 : E \to \mathbb{R}$,*

$$\int_{E^2} \varphi_1(x_1)\varphi_2(x_2) d\mu^N(X^N) \to \Big(\int_E \varphi_1(x) d\mu(x)\Big)\Big(\int_E \varphi_2(x) d\mu(x)\Big),$$

(iii) $\mu^N \circ \hat{\mu}_N^{-1} = \mu^N\big(\hat{\mu}_N \in \cdot\big) \overset{\mathscr{L}}{\to} \delta_\mu$.

This can be seen as an illustration of the Law of Large Numbers, as it suggests that, for the empirical measure to converge toward its mean, particles should be asymptotically independent. This theorem also justifies Boltzmann pre-collisional chaos: it is indeed required to obtain the emergence of a macroscopic behavior.

As our goal is to obtain such an emergence for dynamics (2.1) at all time, we might want to show that, starting from the chaotic initial condition $\mu_0^{\otimes N}$ for a given $\mu_0 \in \mathscr{M}_1^+(E)$, the chaotic property remains at all time $t>0$. Hence, we would like the dynamics to *propagate chaos*:

Definition 2.2.2. Let $Q^N \in \mathscr{M}_1^+\big(\mathscr{C}(R_+,E)^N\big)$ be the solution of the dynamics (2.1) with initial conditions μ_0^N. We say that the dynamics propagates chaos if:

$$\mu_0^N \text{ is chaotic } \implies Q^N \text{ is chaotic},$$

or alternatively if

$$\mu_0^N \text{ is chaotic } \implies \forall t>0,\ Q^N \circ \pi_t^{-1} \text{ is chaotic},$$

where $\pi_t : \mathscr{C}\big(\mathbb{R}_+,E\big) \to E$, and $\forall x \in \mathscr{C}\big(\mathbb{R}_+,E\big), \pi_t(x) = x_t$.

As a summary, we seek the propagation of chaos property in the symmetric case for the following reason: - the empirical mean will become very close to the distribution of a typical particle with overwhelming probability in the thermodynamic limit. Reciprocally, it turns out that asymptotic correlations often discard any possibility of convergence toward a macroscopic description.
- it ensures that the independence is preserved asymptotically, furnishing statistical information on the dynamics,
- it ensures that the empirical measure stays deterministic, so that a macroscopic equation describing the evolution of this empirical measure hopefully exists.

Unfortunately, no such results exists when the symmetry is broken.

2.3 HETEROGENEOUS SYSTEMS: AVERAGED AND QUENCHED RESULTS

Suppose now every particle composing the system is unique and non-interchangeable. This is for example, the case of neurons, as each of them displays a unique set of connections in the network. For these disordered systems, Alain-Sol Sznitman theorem 2.2.1 no more holds as the fundamental hypothesis of symmetry is not satisfied. In order to simplify the analysis, one is then interested in symmetrizing the system in some way by averaging over the disorder, allowing the derivation of *averaged results*. Nevertheless, the latter results might not always apply to biologically or physically relevant system presenting a given realization of the disorder: we would speak of *quenched results*.

Let us precise the definition between these two kind of results with more details. In our models, we shall distinguish between two kinds of heterogeneities. The first type involves a family of random quantities $(r_i)_{1\leq i\leq N}$ representing individual characteristics of the particles. They are in number $O(N)$. We will assume that the r_i are independently drawn from $D\subset\mathbb{R}^d$ for some fixed $d\in\mathbb{N}^*$ with same law $\pi\in\mathscr{M}_1^+(D)$. The second type of heterogeneity depends on two particles and will be associated with the strength of their connections $(J_{ij})_{1\leq i,j\leq N}$. They are i.i.d. random variables, with mean and variance scaling as $\frac{1}{N}$ (so that the equation is properly scaled). We also suppose that the J_{ij}, r_i and Brownian paths are independent. Note that we speak of *quenched disorder*, as the r_i and J_{ij} are frozen in time. We thus deals (2.1) must be modified into the N-dimensional SDE in random environment:

$$\begin{cases} dx_t^i = \Big(f(r_i,x_t^i,t)+\sum_{j=1}^N J_{ij}b(x_t^i,x_t^j,r_i,r_j)\Big)dt+\lambda(r_i)dW_t^i, \quad 1\leq i\leq N \\ (x_0^i)_{1\leq i\leq N}\overset{\mathscr{L}}{=}\bigotimes_{i=1}^N\mu_0(r_i), \end{cases} \tag{2.5}$$

where we have chosen heterogeneous chaotic initial conditions with $\mu_0(r_i)\in\mathscr{M}_1^+(E)$. Under suitable hypothesis, this equation admits a unique solution $Q_{\mathbf{r}}^N(J)\in\mathscr{M}_1^+\big(\mathscr{C}(\mathbb{R}_+,E)^N\big)$ that depends on the realization of both J_{ij} and r_i for $1\leq i,j\leq N$. Results under $Q_{\mathbf{r}}^N(J)$ are called *quenched*.

Has these results are hard to obtain, it is often useful to put some symmetry in the system. The usual trick to deal with the heterogeneity in number $O(N)$ is to make them appear in an extended two-layer empirical measure [74]:

$$\hat{\mu}_N := \frac{1}{N}\sum_{i=1}^N \delta_{x^i,r_i}\in\mathscr{M}_1^+\big(\mathscr{C}(R_+,E)\times D\big).$$

Moreover, as it is not obvious how to extend such a trick to the heterogeneities in number $O(N^2)$, we introduce an averaged version of $Q_{\mathbf{r}}^N(J)$:

$$Q_{\mathbf{r}}^N := \mathscr{E}_J\big(Q_{\mathbf{r}}^N(J)\big),$$

where $\mathscr{E}_J$ denote the expectation over the J_{ij} only. We further define $dQ^N := dQ_{\mathbf{r}}^N(\mathbf{x})d\pi^{\otimes N}(\mathbf{r})$ provided that it exists, and observe that, under $Q^N\in\mathscr{M}_1^+\big(\mathscr{C}(\mathbb{R}_+,E)\times D\big)$, every random couple (x^i,r_i) has now the same law. Hence, we are dealing with a symmetric system for which theorem 2.2.1, and many other useful tools hold. Results under Q^N

are *averaged*.

To summarize:

- *Quenched results* directly deals with the probability measure $Q^N_{\mathbf{r}}(J)$, that only considers random the Brownian paths and initial conditions whereas the *quenched disorder* $(\mathbf{r}, J)$ remains fixed.

- In comparison, *Average results* are obtain under the symmetric probability Q^N, that considers random both the Brownian paths, initial conditions and the disorder $(\mathbf{r}, J)$.

CHAPTER 3

MEAN-FIELD THEORY FOR NEUROSCIENCE

3.1 A SHORT HISTORY OF RIGOROUS MEAN-FIELD APPROACHES FOR RANDOMLY CONNECTED NEURAL NETWORKS

This section present a short review of the mean-field approach in neuroscience. For the purpose of this thesis, it mainly deals with the endeavor to build rigorous derivation of mean-field equations from the microscopic network. It also emphasizes the importance of a particular scaling for the random interactions present in the network, as it is tightly linked with the emergence of a chaotic flow. We are well aware of the partiality of such an approach, and apologize dearly for the piece of works we have omitted, willingly or not.

The genesis of the mean-field approach in mathematical neuroscience must certainly be linked to the so called "local chaos hypothesis". It was seemingly formulated for the first time in 1969 by the Russian scholar Rozonoer [207, 208, 209] and then by Amari [5] in 1972. Though its precise origin in time remains somewhat blurry, as Rozonoer's series of work are found in original version only. Nonetheless, it is not hard to infer the necessity of such an assumption. It doubtlessly emanates from unsuccessful attempts to derive a macroscopic evolution from the microscopic dynamics of random neural network. In fact, it states that, in the thermodynamic limit, the complex correlations among both neurons and random synaptic efficiencies should vanish, allowing to compute the effective limit of the interaction term $\sum_{i=1}^{N} J_{ij} S(x_t^j)$ by a simple use of the LLN (or CLT). In this sense, the "local chaos hypothesis" can be seen as the neuroscientific expression of Boltzmann's molecular chaos.

At the time, the need to fill the gap between existing microscopic models of neural networks with observed macroscopic behavior of both neurophysiological recording and computer simulations had been expressed [72]. The beginning of the 70s gave rise to important endeavors to put the mean-field approach at the heart of mathematical neuroscience [5, 6, 8, 9, 71, 245, 246], unequivocally emphasizing the need to unravel the microscopic parameters responsible for these observed macroscopic behavior [5, 9]. Moreover, properties that would be satisfied by random networks for almost all realizations of their random parameters (quenched results), rather than the properties valid on the average (averaged results), as well as convenient macro-

scopic variables where also sought [8]. Though heuristic, the "local chaos hypothesis" importantly permitted the derivation of mean-field equation in a number of settings involving fully connected network with asymmetrical random synapses, including random thresholds, possibly multiple populations (in order to include both excitation and inhibition), either in the case of analog or formal neurons. These limit equations condensed a rich phenomenology, displaying multi-stable states, as well as oscillations in setting involving both an inhibitory and an excitatory population [5, 6, 245].

In an effort to put the "local chaos hypothesis" on firm grounds, a notable contribution was made in 1982 by S. Geman [115]. Extending his own results applicable to linear systems [116], the author relied on a coupling argument to rigorously prove - under a moment condition on the J_{ij} - that in a firing-rate setting, the macroscopic equation almost surely governed every individual neuron in the thermodynamic limit. The result only concerned the case of what we shall call *weakly random* interactions, presenting mean and variance scaling respectively as $\frac{1}{N}$ and $\frac{1}{N^2}$, N being the size of the network. In fact, most of the above neuroscientific studies used this scaling. In particular, the heterogeneity of connections was not felt in the thermodynamic limit for these networks. Geman also gave an intuition the asymptotic behavior (valid for the linear case [116]) for another scaling of synaptic connection, with both mean and variance scaling as $\frac{1}{N}$, that we shall call *strongly random* in this manuscript.

As mentioned above, this latter scaling, a breaking discovery was made by Sompolinsky, Crisanti and Sommers [221] that applied heuristic mean-field methods from physics literature (the so called dynamical mean-field theory introduced in [167], and also developed in subsequent works [81, 82, 83, 143, 222]), in order to investigate the thermodynamic limit of a homogeneous firing-rate network of neurons with leaky membranes:

$$\frac{\mathrm{d}x_t^i}{\mathrm{d}t} = -x_t^i + \sum_{j=1}^{N} J_{ij} S(x_t^j),$$

where $\forall 1 \leq i, \neq j \leq N$, $J_{ij} \sim \mathcal{N}(0, \frac{\sigma^2}{N})$, $J_{ii} = 0$, and $S = tanh(g\cdot)$, with $\sigma, g > 0$. In fact, in the thermodynamics limit, they described a sharp transition from a stationary state attracting every trajectories, toward a chaotic [1] flow in which neurons fluctuated incoherently around the origin, as the gain parameter $g\sigma$ becomes greater than 1. The authors showed that the mean field equation could be reduced to the movement of a particle in a potential. They carefully studied the solutions after the transition, and showed - by an argument involving the maximal Lyapunov exponent - that they were infinitely many, including fixed points or limit cycle, but that only the solution associated with a zero energy was stable. For this solution, the local-auto correlation function displayed an exponential decay in time, symptomatic of chaos. Moreover, simulations highlighted the fact that in the finite-size networks, existed an intermediary regime between the global attractor and chaotic flow, for whose width shrank to zero in the thermodynamic limit. In this regime, progressively appear non zero stationary solutions, and limit cycles, becoming increasingly more complex as $g\sigma$ increased.

[1] Let us emphasize that "chaos" here stands for deterministic chaos in the sense of Lorentz, while the "local chaos hypothesis" basically means independence of particles in the sense of Boltzmann.

This article launched a considerable field of research, as such a chaotic behavior had already been observed in the nervous system in sleep, epileptic seizure and sensory-pattern retrieval [18, 19, 112, 218]. In [218], through EEG recordings, the authors witnessed a spontaneous chaotic activity when the rabbit was not solicited. Interestingly, when the rabbit was presented with a learned odor, the dynamics reduced from chaos toward attractor an of lower dimension. The authors argued that, by a perpetual and thorough exploration of the state space, chaos acted as an effective way to rapidly retrieve previously learned patterns, and to learn new ones.

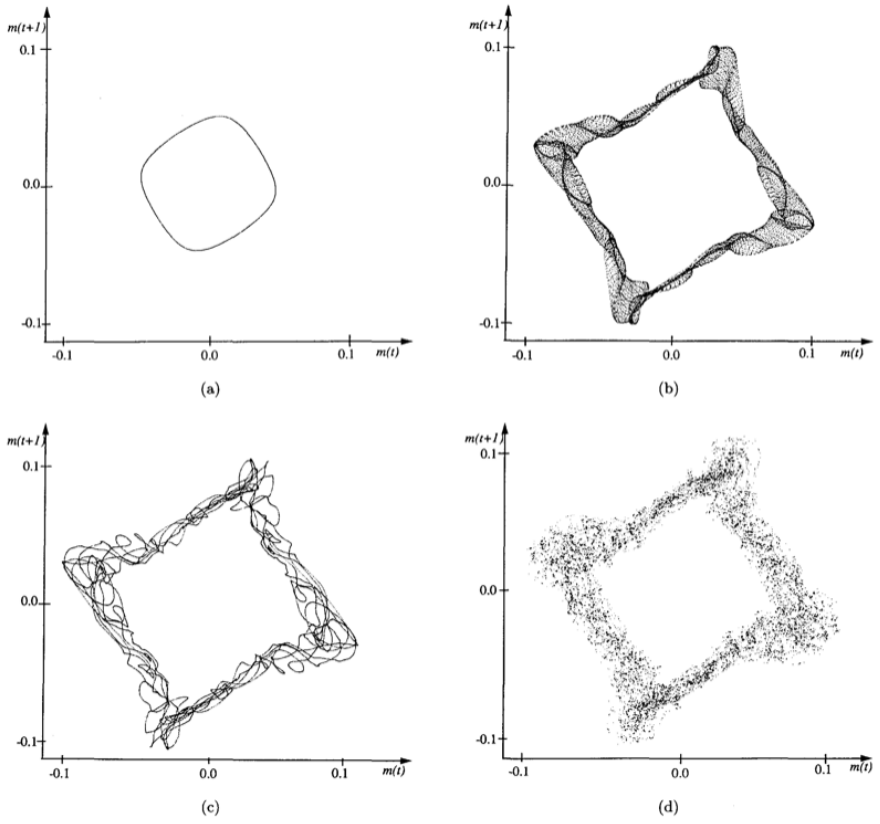


Figure 3.1: Example of a route to chaos for a discrete-time neural network as the gain parameter g is increased. The first bifurcation is a Hopf (N=128,K=16). (a) After the first bifurcation, the zero fixed point has lost its stability. The series of points $(m(t), m(t+1))$ densely covers a cycle ($g = 1$). (b) After the second Hopf bifurcation: projection of a T^2 torus ($g = 1.23$). (c) Frequency locking on the T^2 torus ($g = 1.247$). (d) Chaos ($g = 1.6$). Reproduced from [89]

In an effort to better understand Sompolinsky's transition and challenge its generalization to more biologically relevant neural networks displaying sparse non-Gaussian connections, numerical studies were realized on a ring with nearest neighbor connections [21], and for a randomly connected network [89]. Both article involved discrete time with K input connections per neuron, and confirmed the existence of the chaotic phase, using a criteria of sensitivity to initial condition. Moreover, they evidenced the fact that, for finite N, the route to chaos was a quasi-periodic one, involving several successive Hopf bifurcation has the gain parameter was increased. In [89], a theoreti-

cal analysis was also provided to account for these bifurcations, and relied on random matrix theory to discuss the possibility of pitch-fork and flip bifurcation occurrence in the route to chaos. These results where then extended in the case of a fully connected network with strongly random interactions and Gaussian random thresholds breaking the reversal symmetry $x \to -x$ [58]. A bifurcation diagram was numerically inferred in the gain and threshold parameter. More importantly, a "local chaos hypothesis" was adopted to derive mean-field Gaussian limit, as no rigorous derivation existed for the case of strongly random interactions. This limit was them studied through the dynamical equation for its mean and variance, and possible stationary regime were investigated. Moreover, in a very similar setting, a strong analogy with the Sherrington-Kirkpatrick (SK) spin glass model from physics was outlined by Cessac [57]. In fact, he proposed a quadratic distance criteria between distinct initial conditions to determine the entering into the chaotic phase, and discovered that, in a particular setting, it was determined by the Almeida-Thouless line specific of SK model. The author also showed that the two models shared some properties such as the breaking of ergodicity, ultrametricity, and a sharp increase in complexity (the expected number of equilibria. See also [243]).

It is was well known since the notable work of Hopfield on the long-term memory capacity of neural network presenting symmetric couplings [137], that there existed intimate connections between neural networks and spin glasses. In fact, in the 80s, many works involved symmetric synaptic connections for which the system presented an Hamiltonian. These systems experienced an exponential increased of stable equilibria. Nevertheless, a Lyapunov function could be constructed, and thus systematically ensured a relaxation of the system toward a stable equilibria (see for example [26, 60]). However, for biological neural networks, stable equilibrium are observed in two cases: anesthesia or death. This observation, along with the discovery of the chaotic phase created a regain of interest in asymmetrically connected model. Interestingly, the "local chaos hypothesis" has proven to be wrong in the case of symmetric interactions, and in particular for the Sherrington-Kirkpatrick model in low temperature. Indeed, the ill-obtained mean-field equation fail to capture a term that the network induced on a given particle, in reaction to its effect on the network [57].

It was again from physics literature that the next breakthrough came from. In the mid 90s, Gerard Ben Arous and Alice Guionnet rigorously addressed the mean-field derivation of the continuous time dynamics of the soft spins SK model, in both the asymmetric and symmetric case [16, 22, 23, 129]. They developed an original approach based on large deviations theory, stochastic and Gaussian calculus to derive powerful results, going beyond that of Geman, in both the averaged and the quenched case (see section 3.3). Though this quite technical proof only addressed the case of linear interactions and heavily relied on the presence of Brownian noise and Gaussian interactions, this precursory work opened a paved alley to - at last - render a solid foundations for the mean-field approach in neuroscience.

This generalization to the case of non-linear interactions was tackled in 2001 by Moynot and Samuelides for the discrete (and finite) time dynamics of a firing-rate neural network with asymmetric connections, and in the presence of white noise and Gaussian thresholds [178]. Interestingly for biological relevance, their interactions were not necessarily centered nor Gaussian, as they only necessitated to satisfy a

sub-Gaussian tail condition. In particular, this condition allowed to consider family of neurons almost exclusively inhibitory or excitatory [211], that is interactions charging negative (or positive) value with arbitrarily small probability. Without any short time hypothesis, a full large deviation principle (LDP) was obtained for the averaged network in the case of Gaussian interactions only, inducing, in an averaged sense, a convergence of the empirical measure toward the unique minimum of the good rate function, and a propagation of chaos. While the LDP was lost for non-Gaussian interactions, the two latter results remained. The authors also derived an almost sure quenched convergence through Borel-Cantelli lemma. This important contribution thus validated the "local chaos hypothesis" in the case of asymmetric interactions, and supported several subsequent works involving sparse connections and several populations [79], binary and spiking neurons [59].

Nevertheless, the large deviations approach of [178] did not apply to continuous time setting. Moreover, rigorous mean-field results in neuroscience were still very scarce. Building on [231], coupling methods were used to address the dynamics of a time continuous networks of neurons involving multiple populations, Brownian noise, and weakly random interactions [239]. Existence and uniqueness of the mean-field equation were proved relying on a simple contraction argument à la Picard. Moreover, a convergence results was obtained for the whole network uniformly in time, as well as a propagation of chaos result. A similar analysis was also undertaken to rigorously derive a mean-field limit for a continuous time dynamics involving heterogeneous transmission delays [234], and spatial extension [237]. In [234], it was importantly shown that the mean and variance of the heterogeneous delays impacted the dynamics of the network.

In [51], we extended the large deviation approach of [22] to tackle the nonlinear continuous-time dynamics of a multi-population setting with strongly random Gaussian interactions. The dynamics also included deterministic delays - putting the problem in infinite dimension - as well as a non-zero mean for the interactions. Averaged convergence of the empirical measure and propagation of chaos were obtained. Moreover, provided the initial conditions were Gaussian and the intrinsic dynamics linear, the solution of the mean field equation was also Gaussian, enabling its study through a set of implicit equations coupling its mean and variance. Moreover, the mean of the Gaussian interactions was shown to be involved with a pitch-fork bifurcation. In part III, we will present various numerical results, including a transition toward chaos, delay-induced oscillation, and mean-induced up-and-down states. Notably, large deviations were also used in a setting involving correlated strong synaptic weights [105]. The authors considered a microscopic spatially extended network on the torus, in which the neurons were regularly spaced. The system was invariant by translation, and the correlation between synaptic weights was described as a function of the distance. The authors derived a unique mean-field equation (with no more spatial extension) preserving the correlations in the thermodynamics limit.

We conclude this short overview with the continuous-time dynamics of spiking networks. In fact, virtually all the above mentioned analysis were down under the firing-rate hypothesis. Though, addressing the dynamics of large spiking networks is promising, and has been the aim of several recent endeavors including numerical results on the dynamics of spiking networks [147, 186], a heuristic mean-field

derivation for the case of strongly random interactions [147], and a rigorous one for the case of positively charged neurons described as Poisson process, and displaying weakly random interactions [203]. Importantly, it was shown that the dynamics of spiking neurons with strong interactions also experienced a transition to chaos.

3.2 COUPLING METHODS FOR WEAK INTERACTIONS

In this section, we revisit powerful probabilistic *coupling methods* reviewed by Alain-Sol Sznitman in his famous Saint-Flour lecture [231] and used in [234, 237, 239]. These will be useful to derive and prove - without much effort - the convergence toward the mean-field equations in the case of a random fully connected neural networks presenting *weak interactions* (defined below).

The idea underlying these methods is to facilitate the comparison between the solution of two SDEs by choosing the same realizations for the Brownian paths and initial conditions. With this approach, proving the convergence of the network toward the mean-field equation follows three steps:

(i) intuiting the mean-field equation by relying on the "local chaos hypothesis",

(ii) proving existence and uniqueness of solution for this heuristic mean-field equation,

(iii) ensuring that the networks converge to this solution in a suitable sense.

Remark also that similar techniques were used in [115] to obtain the convergence of the network toward the mean-field equation in the context of ODEs in random environment. The author compared individual trajectories with an averaged over the whole network. Here, we will rely on a different strategy, as we will compare directly the network to N independent copies of the mean-field solution.

Let us now discriminate between two types of interactions. In all the manuscript, we will speak of *weakly random interactions*, or *weak interactions*, connections J_{ij} whose variance decays faster than $\frac{1}{N}$, N being the number of components in the network. In contrast, *strongly random interactions*, also called *strong interactions*, scale exactly as $\frac{1}{N}$. The reason to make such a difference comes from the CLT. In fact, let the connections $(J_{ij})_{1\leq i,j\leq N}$ be a family of independent copies of the same distribution with mean and variance given by:

$$\mathbb{E}\big[J_{ij}\big] = \frac{\bar{J}}{N}, \qquad Var\big(J_{ij}\big) = \frac{\sigma^2}{a_N},$$

where $a_N \in \mathbb{R}$, and $a_N \to_{N\to+\infty} +\infty$. Set $\tilde{J}_{ij} = \frac{\sqrt{a_N}}{\sigma}\big(J_{ij} - \frac{\bar{J}}{N}\big)$ and consider the following neural network:

$$\begin{cases} dx_t^i = \sum_{j=1}^N J_{ij} S(x_t^j)dt + \lambda dW_t^i \\ (x_0^i)_{i=1,\dots,N} \overset{\mathscr{L}}{=} \mu_0^{\otimes N}. \end{cases} \tag{3.1}$$

where $\lambda \in \mathbb{R}+$ (possibly 0 to include the case without noise). Here, neurons present

no intrinsic dynamics for simplicity. Nevertheless, the following analysis is straightforward by adding a Lipschitz-continuous intrinsic dynamics. In this section, we will informally work on two concatenated probability spaces $\big(\Omega,\mathscr{F},\mathbb{P}\big)$, and $\big(\Omega,\mathscr{F}_J,\mathscr{P}_J\big)$. The first one will include all the different randomness in presence, whereas the last one only deals with the J_{ij}, $i,j\in[\![1,N]\!]$: the Brownian trajectories and initial conditions are seen as deterministic under $\mathscr{P}_J$. Moreover, we will denote by $\mathbb{E}$ the expectation over $\mathbb{P}$, and $\mathscr{E}_J$ that over $\mathscr{P}_J$. Importantly, note that this hierarchy only holds in this section, as the two space will be strictly separated in the core of our proofs. Let us denote by $Q^N(J)$ the solution of (3.1) up to a given time $T>0$, and by $Q^N:=\mathscr{E}_J\big(Q^N(J)\big)$ its averaged version (that we assume well-defined). To find the intuition about the limit equation, we have to understand the asymptotic behavior of the interaction term. It can be decomposed as follow:

$$\sum_{j=1}^N J_{ij}S(x_t^j)=\underbrace{\frac{\bar{J}}{N}\sum_{j=1}^N S(x_t^j)}_{\overset{LLN}{\to_N}\bar{J}\mathbb{E}[S(\bar{x}_t)]}+\underbrace{\frac{\sigma}{\sqrt{a_N}}\sum_{j=1}^N \tilde{J}_{ij}S(x_t^j)}_{\overset{CLT}{\to_N}\sigma\mathscr{N}\left(0,\mathbb{E}[S(\bar{x}_t)^2]\right),\text{ when }a_N=N.}}.$$

Here the limit are heuristic. Observe that the variance term vanishes as soon as $N=o(a_N)$, and explodes when $a_N=o(N)$. Hence, the only scaling preserving a finite non-zero contribution of the variance term in the mean-field equation is when $a_N=N$, that is for *strong interactions*. Hence, in the case of *weak interactions*, the limit equation seems to be

$$d\bar{x}_t^i=\bar{J}\mathbb{E}\big[S(\bar{x}_t)\big]dt+\lambda dW_t^i \tag{3.2}$$

whereas for strong interactions, it is given by

$$d\bar{x}_t^i=G_t^{\bar{x}}dt+\lambda dW_t^i, \tag{3.3}$$

where $G^{\bar{x}}$ is a Gaussian process with mean and covariance given by $\bar{J}\mathbb{E}\big[S(\bar{x}_t)\big]$ and $\sigma^2\mathbb{E}\big[S(\bar{x}_t)S(\bar{x}_s)\big]$. A first observations is that these limits does not depends on the precises distribution of the J_{ij}. Underlying this universality is the CLT. Moreover, both equations are implicit, as the effective interaction term depends on the law of the solutions (provided it exists). Consequently, we need to ensure that these equations are well-posed.

In what follows, we will try to push the coupling methods to their limits. In the case of weak interactions, we will consider the classical case $S(x_t^j)$ with S bounded and K_S Lipschitz-continuous. We will denote $\mathscr{C}:=\mathscr{C}([0,T],\mathbb{R})$, and $d_{T,p}^V$ the p-Vasserstein distance on $\mathscr{M}_1^+(\mathscr{C})$, and let $\mu_0\in\mathscr{M}_1^+(\mathscr{C})$.

We now show existence and uniqueness of solutions for the (heuristic) mean-field equation (3.2):

Theorem 3.2.1. *The following SDE is well-posed and admits a unique weak solution in* $\mathscr{M}_1^+(\mathscr{C})$*:*

$$\begin{cases} d\bar{x}_t = \bar{J}\mathbb{E}_{\bar{z}}\big[S(\bar{z}_t)\big]dt + \lambda dW_t \\ \bar{x}_0 \overset{\mathscr{L}}{=} \mu_0. \end{cases} \tag{3.4}$$

where $(W_t)_{0\le t\le T}$ *is a Brownian motion, and* $\bar{z}$ *is an independent copy of* $\bar{x}$*.*

The proof relies on a contraction principle involving the Vasserstein distance (see equation (2.4)).

Proof. Let $\mu,\nu \in \mathscr{M}_1^+(\mathscr{C})$ and $\xi \in \mathscr{C}_{\mu,\nu}$ be a coupling of these measures. Let $x^\mu \in \mathscr{C}$ be the strong solution of the following SDE

$$\begin{cases} dx_t = \bar{J}\Big(\int_{\mathscr{C}} S(y_t)d\mu(y)\Big)dt + \lambda dW_t \\ x_0 = x^0. \end{cases}$$

with $\mathscr{L}(x^0) = \mu_0$, and define similarly $x^\nu \in \mathscr{C}$ with same Brownian motion and initial condition. We denote their law by $L(\mu) \in \mathscr{M}_1^+(\mathscr{C})$ and $L(\nu) \in \mathscr{M}_1^+(\mathscr{C})$ respectively. Then

$$\begin{aligned} \big|x_t^\mu - x_t^\nu\big| &= \left|\bar{J}\int_0^t \left(\int_{\mathscr{C}} S(y_s)d\mu(y) - \int_{\mathscr{C}} S(z_s)d\nu(z)\right)ds\right| \\ \big|x_t^\mu - x_t^\nu\big| &\le |\bar{J}|\int_0^t \left(\int_{\mathscr{C}^2} \big|S(y_s) - S(z_s)\big|d\xi(y,z)\right)ds \le |\bar{J}|K_S\int_0^t d_{s,1}^V(\mu,\nu)ds \end{aligned}$$

where we have taken, in the right-hand side, the supremum in time and the infimum in ξ, and the Vasserstein distance is define in 2.4. Now, taking the supremum in time, and the expectation on the left-hand side yields

$$d_t^V\big(L(\mu),L(\nu)\big) \le \mathbb{E}\big[\sup_{0\le s\le t}\big|x_s^\mu - x_s^\nu\big|\big] \le |\bar{J}|K_S\int_0^t d_s^V(\mu,\nu)ds,$$

so that contraction of the map L^p with respect to the 1-Vasserstein distance and for p big enough, follow by classical arguments. In particular, L admits a unique fix point, solution of (3.4). □

The third line of the proof is to demonstrate the convergence of the whole network toward the solution of the mean-field equation. Let $Q \in \mathscr{M}_1^+(\mathscr{C})$ denote this solution. Here, we give a proof for weakly random interactions admitting exponential moments.

Theorem 3.2.2. *Suppose that exists* $p \in \mathbb{N}^*$, $\delta_0 > 0, C > 0$ *such that*

$$\forall \delta \le \delta_0, \quad \mathscr{E}_J\Big[e^{\delta|\tilde{J}_{ij}|^p}\Big] \le e^{\delta C}. \tag{3.5}$$

Then, for $\frac{1}{a_N} = O\big(N^{-2}\big)$ *the probability measure* $\hat{\mu}_N \overset{\mathscr{L}}{\to} Q$ *in probability under* Q^N *and for the p-Vasserstein distance. In other words:*

$$\forall \varepsilon > 0, \quad Q^N\Big(d_{T,p}\big(\hat{\mu}_N, Q\big) > \varepsilon\Big) \to_N 0.$$

Furthermore, if $p > 2$*, the convergence is almost sure.*

Remark 1. In particular, if the J_{ij} are

- Gaussian or bounded random variables, we obtain an almost sure convergence in law of the empirical measure.
- the deterministic case can be treated taking $a_N = +\infty$.

Define $\xi_i^{N,p} = \sum_{j=1}^N |\tilde{J}_{ij}|^p$, and $\bar{\xi}^{N,p} := \max_{i=1,\dots,N} \xi_i^{N,p}$. Before proving the result, we will need a control on exponential moments of $\bar{\xi}^{N,p}$, that will naturally arise in the proof of the theorem:

Lemma 3.2.3. Suppose condition (3.5) holds. Then,

$$\forall a \in \mathbb{R}_+, \quad \mathbb{E}\Big[e^{a\bar{\xi}^{N,p}N^{-1}}\Big] = O(1).$$

Proof.

$$\begin{aligned}\mathbb{E}\Big[e^{a\bar{\xi}^{N,p}N^{-1}}\Big] &= \int_0^{+\infty} e^{ax} dP_{\bar{\xi}^{N,p}N^{-1}}(x) = \int_0^{+\infty} \Big(\int_{-\infty}^{x} ae^{ay}dy\Big) dP_{\bar{\xi}^{N,p}N^{-1}}(x) \\ &= \int_{-\infty}^{+\infty} ae^{ay}\mathbb{P}\Big(\frac{\bar{\xi}^{N,p}}{N} > y\Big)dy\end{aligned}$$

$$\begin{aligned}\mathbb{P}\Big(\bar{\xi}^{N,p} > yN\Big) &= 1 - \Big\{1 - \mathbb{P}\Big(\xi_i^{N,p} > yN\Big)\Big\}^N \overset{\text{Cheb.}}{\leq} 1 - \Big\{1 - \mathbb{E}\Big(e^{\delta_0 \Sigma_{j=1}^N |\tilde{J}_{ij}|^p}\Big)e^{-\delta_0 Ny}\Big\}^N \\ &\overset{(3.5)}{\leq} 1 - \Big\{1 - e^{-N\delta_0(y-C)}\Big\}^N = Ne^{-N\delta_0(y-C)}(1+o(1)).\end{aligned}$$

Hence,

$$\begin{aligned}\mathbb{E}\Big[e^{a\bar{\xi}^{N,p}N^{-1}}\Big] &\leq \int_{-\infty}^{C} ae^{ay}dy + N(1+o(1))\int_C^{+\infty} ae^{ay - N\delta_0(y-C)}dy \\ &\leq e^{aC}\Big\{1 + N(1+o(1))\int_0^{+\infty} ae^{-(N\delta_0 - a)y}dy\Big\} \leq e^{aC}\Big\{1 + \frac{aN}{N\delta_0 - 1}(1+o(1))\Big\}.\end{aligned}$$

□

Proof. of the theorem: Let $(x^{i,0})_{1\leq i\leq N}$ be N independent realization of μ_0, and let $\big(X_t^i, 0 \leq t \leq T\big)_{1\leq i\leq N}$ be the strong solution of the network equation:

$$\begin{cases} dx_t^i = \sum_{j=1}^N J_{ij} S(x_t^j)dt + \lambda dW_t^i \\ (x_0^i)_{i=1,\dots,N} = (x^{i,0})_{i=1,\dots,N}, \end{cases} \tag{3.6}$$

where $\big(W_t^i, 0 \leq t \leq T\big)_{1\leq i\leq N}$ is a family of independent Brownian motions. Hence, it has the same law as the original network equation (3.1). Let also, for any $i \in \{1,\dots,N\}$, $\bar{X}^i$ be the strong solution of the SDE

$$\begin{cases} d\bar{x}_t^i = \bar{J}\mathbb{E}\big[S(\bar{z}_t)\big] + \lambda dW_t^i \\ \bar{x}_0^i = x^{i,0}, \end{cases} \tag{3.7}$$

with same initial conditions and Brownian motions. Remark that the $\bar{x}^i$ are independent. Now, observe that:

$$X_t^i - \bar{X}_t^i = \int_0^t \sum_{j=1}^N J_{ij}\big(S(X_s^j) - S(\bar{X}_s^j)\big)ds + \int_0^t \sum_{j=1}^N \Big(J_{ij}S(\bar{X}_s^j) - \frac{\bar{J}}{N}\mathbb{E}\big[S(\bar{X}_s)\big]\Big)ds$$
$$= \int_0^t \frac{1}{N}\sum_{j=1}^N \bar{J}\big(S(X_s^j) - S(\bar{X}_s^j)\big)ds + \int_0^t \frac{\sigma}{\sqrt{a_N}}\sum_{j=1}^N \tilde{J}_{ij}\big(S(X_s^j) - S(\bar{X}_s^j)\big)ds + \int_0^t \sum_{j=1}^N \Big(J_{ij}S(\bar{X}_s^j) - \frac{\bar{J}}{N}\mathbb{E}\big[S(\bar{X}_s)\big]\Big)ds.$$

As $(\bar{X}^i)_{1\le i\le N}$ is independent from $(J_{ij})_{1\le i,j\le N}$, the $\Big(J_{ij}S(\bar{X}_s^j) - \frac{\bar{J}}{N}\mathbb{E}\big[S(\bar{x}_s)\big]\Big)_{1\le j\le N}$ are i.i.d. centered random variables with variance $\frac{\sigma^2}{a_N}\mathbb{E}\big[S(\bar{X}_s)^2\big]$. Hence

$$\zeta_i^N(s) := \sqrt{\frac{a_N}{N}}\sum_{j=1}^N \Big(J_{ij}S(\bar{X}_s^j) - \frac{\bar{J}}{N}\mathbb{E}\big[S(\bar{X}_s)\big]\Big) \overset{\mathscr{L}}{\to} \mathcal{N}\Big(0, \sigma^2\mathbb{E}\big[S(\bar{X}_s)^2\big]\Big),$$

under $\mathbb{E}_J$. They are also i.i.d. random variables under $\mathscr{E}_J$. Thus

$$\big|X_t^i - \bar{X}_t^i\big|^p \le C_{p,t}\Big\{\int_0^t \frac{1}{N}\sum_{j=1}^N \big|X_s^j - \bar{X}_s^j\big|^p ds + \int_0^t \frac{N^{p-1}}{\sqrt{a_N}^p}\sum_{j=1}^N |\tilde{J}_{ij}|^p\big|X_s^j - \bar{X}_s^j\big|^p ds + \int_0^t (Na_N^{-1})^{\frac{p}{2}}|\zeta_i^N(s)|^p ds\Big\}.$$

$$\underbrace{\frac{1}{N}\sum_{i=1}^N \sup_{s\le t}\big|X_s^i - \bar{X}_s^i\big|^p}_{=:g(t)} \le C_{p,t}\Big\{\int_0^t g(s)ds + \int_0^t \frac{1}{N}\sum_{j=1}^N \Big\{\sup_{u\le s}\big|X_u^j - \bar{X}_u^j\big|^p \frac{N^{p-1}}{\sqrt{a_N}^p}\sum_{i=1}^N |\tilde{J}_{ij}|^p\Big\}ds$$
$$+ (Na_N^{-1})^{\frac{p}{2}}\int_0^t \frac{1}{N}\sum_{i=1}^N |\zeta_i^N(s)|^p ds\Big\}.$$

As $\frac{1}{N}\sum_{j=1}^N \Big\{\sup_{u\le s}\big|X_u^j - \bar{X}_u^j\big|^p \frac{N^{p-1}}{\sqrt{a_N}^p}\sum_{i=1}^N |\tilde{J}_{ij}|^p\Big\} \le g(s)\frac{N^{p-1}}{\sqrt{a_N}^p}\bar{\xi}^{N,p}$, we can use Gronwall lemma to obtain:

$$g(t) \le C_{p,t}\exp\Big\{C_{p,t}\Big(1 + \frac{N^p}{\sqrt{a_N}^p}\frac{\bar{\xi}^{N,p}}{N}\Big)\Big\}(Na_N^{-1})^{\frac{p}{2}}\int_0^t \frac{1}{N}\sum_{i=1}^N |\zeta_i^N(s)|^p ds,$$
$$\le C_{p,t}\exp\big\{C_{p,t}\big\}(Na_N^{-1})^{\frac{p}{2}}\int_0^t \frac{1}{N}\sum_{i=1}^N |\zeta_i^N(s)|^p \exp\Big\{C_{p,t}\frac{N^p}{\sqrt{a_N}^p}\frac{\bar{\xi}^{N,p}}{N}\Big\}ds. \tag{3.8}$$

Let $s_N(\mathbf{x}) := \frac{1}{N}\sum_{i=1}^N \delta_{x^i}$, where $x^i \in \mathscr{C}, \forall i \in \{1,\dots,N\}$ and with the notation $\mathbf{x} := (x^i)_{1\le i\le N}$. We thus have:

$$d_{t,p}^V\big(s_N(\mathbf{X}), s_N(\bar{\mathbf{X}})\big)^p \le g(t) \le C_{p,t}\exp\big\{C_{p,t}\big\}(Na_N^{-1})^{\frac{p}{2}}\int_0^t \frac{1}{N}\sum_{i=1}^N |\zeta_i^N(s)|^p \exp\Big\{C_{p,t}\frac{N^p}{\sqrt{a_N}^p}\frac{\bar{\xi}^{N,p}}{N}\Big\}ds.$$

Now, observe that

$$\mathbb{P}\big(d_{t,p}^V\big(s_N(\mathbf{X}), Q\big) > \varepsilon\big) \le \mathbb{P}\big(d_{t,p}^V\big(s_N(\bar{\mathbf{X}}), Q\big) > \frac{\varepsilon}{2}\big) + \mathbb{P}\big(d_{t,p}^V\big(s_N(\mathbf{X}), s_N(\bar{\mathbf{X}})\big) > \frac{\varepsilon}{2}\big)$$
$$\overset{\text{Markov}}{\le} Q^{\otimes N}\Big(\hat{\mu}_N \in B_{d_{t,p}^V}\big(Q, \frac{\varepsilon}{2}\big)^c\Big) + (2\varepsilon^{-1})^p\mathbb{E}\Big[d_{t,p}^V\big(s_N(\mathbf{X}), s_N(\bar{\mathbf{X}})\big)^p\Big].$$

Where $B_{d_{t,p}^V}\big(Q, \frac{\varepsilon}{2}\big)$ is the open ball centered at Q with radius $\frac{\varepsilon}{2}$ for the distance $d_{t,p}^V$, and A^c denote the complement of subset of a Polish space $A \subset \Sigma$. On the one hand, the first

term of the right-hand side is summable by Sanov's Theorem 9.5.1 and Borel-Cantelli lemma. On the other hand, as $C_{p,t}\frac{N^p}{\sqrt{a_N}^p} = O(1)$, lemma 3.2.3 ensures that:

$$\mathscr{E}_J\left[\exp\left\{2C_{p,t}\frac{N^p}{\sqrt{a_N}^p}\frac{\bar{\xi}^{N,p}}{N}\right\}\right]^{\frac{1}{2}} = O(1),$$

so that

$$\begin{aligned}\mathscr{E}_J\big(d_{t,p}^V\big(s_N(\mathbf{X}),s_N(\bar{\mathbf{X}})\big)^p\big) &\overset{C.S.}{\leq} \tilde{C}_{p,t}(Na_N^{-1})^{\frac{p}{2}}\int_0^t\frac{1}{N}\sum_{i=1}^N\mathscr{E}_J\big(|\zeta_i^N(s)|^{2p}\big)^{\frac{1}{2}}ds\\ &\overset{C.S.}{\leq} \tilde{C}_{p,t}\int_0^t(Na_N^{-1})^{\frac{p}{2}}\mathscr{E}_J\Big(|\zeta_1^N(s)|^{4p}e^{-|\zeta_1^N(s)|}\Big)^{\frac{1}{4}}\mathscr{E}_J\Big(e^{|\zeta_1^N(s)|}\Big)^{\frac{1}{4}}ds.\end{aligned}$$

As $x\to x^{4p}e^{-x}$ is a bounded function, and as $|\zeta_1^N(s)| \overset{\mathscr{L}}{\to} \mathscr{N}\big(0,\sigma^2\mathbb{E}\big[S(\bar{x}_s^2)\big]\big)$, taking the expectation over the Brownian paths and initial conditions yields:

$$\mathbb{E}\Big(d_{t,p}^V\big(s_N(\mathbf{X}),s_N(\bar{\mathbf{X}})\big)^p\Big) \leq \tilde{C}_{p,t}\int_0^t\underbrace{(Na_N^{-1})^{\frac{p}{2}}}_{O(N^{-p})}\underbrace{\mathbb{E}\Big[e^{|\zeta_1^N(s)|}\Big]^{\frac{1}{4}}}_{O(1)}ds = O(N^{-p})$$

In particular $\mathbb{P}\big(d_{t,p}^V\big(s_N(\mathbf{X}),Q\big)$ tends toward 0 when N gos toward infinity. Moreover, it is summable as soon as $p>1$. □

Note that these proofs can be easily extended to the case of bi-dependent interactions of the form $b(x,y)$ with b bounded and K_b Lipschitz-continuous in both variables, by relying on a conditional CLT. Moreover, in the case of *strongly random interactions*, coupling methods cannot be adapted for the convergence of the network equation toward the mean-field limit. Nevertheless, we are able to show existence and uniqueness of the equation (3.3) in an averaged sense. Such results will be derived in part II.

3.3 Large deviations techniques for Spin Glass dynamics

We now shall introduce large deviations techniques developed by G. Ben Arous and A. Guionnet in the context of spin glasses [16, 22, 129]. These techniques constitute a paradigm to rigorously derive limit theorems in the case of *strongly random interactions*, and have been notably adapted for neuroscience in the context of fully connected neural networks evolving in discrete times [178]. In this manuscript, we will build on such techniques to study continuous time dynamics of fully connected networks presenting delays, spatial extension, and bi-dependent synapses.

3.3.1 Setting

In their publication of the year 1995 [22], Gerard Ben Arous and Alice Guionnet were interested in the asymptotic behavior of an asymmetrical spin glass dynamics as the number of spins tended toward infinity. For a glass composed of N soft spins taking values in $[-A,A]$, they addressed the following dynamics:

$$\begin{cases} dx_t^i = \Big(-\nabla U(x_t^i) + \frac{\beta}{\sqrt{N}}\sum_{j=1}^N J_{ij}x_t^j\Big)dt + dB_t^i \\ (x_0^i)_{1\leq i\leq N} \overset{\mathscr{L}}{=} \mu_0^{\otimes N}. \end{cases} \tag{3.9}$$

Here U is a confining potential that tends to infinity sufficiently fast when $|x| \to A$ in order to ensure that the spins stay in $[-A,A]$. Moreover, β is the inverse of the temperature, the B^is are independent Brownian motions driving each spin, and μ_0 is a probability measure on $\mathscr{C}([0,T],[-A,A])$ characterizing the initial conditions supposed symmetric and chaotic. Furthermore, the J_{ij} are centered i.i.d. Gaussian variable with variance $\frac{\sigma^2}{N}$ that account for the heterogeneous interactions between spins. Remark that the variance of the J_{ij} is scaled as $\frac{1}{N}$, so that we are dealing with *strongly random connections*.

3.3.2 Frame of the analysis

We now present the authors analysis that mainly relies on large deviations theory, stochastic and Gaussian calculus. It importantly relies on:

- *Sanov's theorem* for the network without interactions,
- *Girsanov's theorem* allowing to compare the coupled network with the uncoupled one,
- An *ad hoc* version of *Varadhan's lemma*, that furnishes a LDP for the coupled network, and constitute the main part of the analysis.

Let us enter a bit more into the mathematical formalism. The authors work on a finite time window $[0,T]$, with $T > 0$, and on two distinct probability spaces. The first is a filtered complete probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\leq t\leq T}, \mathbb{P})$ satisfying the usual conditions, and that accounts for the randomness of both the Brownian paths $\left(B_t^i, 0 \leq t \leq T\right)_{1\leq i\leq N}$ and initial conditions. In contrast, the random Gaussian interactions J_{ij} are drawn out of the complete probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \mathscr{P})$. Hence, there are two distinct sources of randomness:

- *A random environment*: interactions $J = (J_{ij})_{1\leq i,j\ leqN}$ are random variables of a probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \mathscr{P})$. They define the structure of the network, are independent and frozen in time. Moreover, expectation on the interactions is denoted $\mathscr{E}$.
- *A stochastic dynamics*: states of spins are stochastic variables, solutions of a SDE driven by a collection of independent $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\leq t\leq T}, \mathbb{P})$-Brownian motions $\left(B_t^i, 0 \leq t \leq T\right)_{1\leq i\leq N}$.

They denote by $P_\beta^N(J)$ the unique weak solution in $\mathscr{M}_1^+(\mathscr{C}([0,T],[-A,A])^N)$ for the dynamics (3.9). It obviously depends on the realization of the matrix J. The main problem to tackle its analysis is the non-exchangeability of the different spins because of the asymmetry which emanates from J. To cope with this difficulty, the authors adopt an averaging strategy. They introduce the uncoupled networks dynamics for which $J = 0$:

$$\begin{cases} dx_t^i = -\nabla U(x_t^i)dt + dB_t^i \\ (x_0^i)_{1\leq i\leq N} \overset{\mathscr{L}}{=} \mu_0^{\otimes N}. \end{cases} \tag{3.10}$$

As particles are independent and identically distributed in this network, the weak solution is a tensorial product $P^{\otimes N}$. As we are working on a Polish space, these good

property of independence and symmetry allow to rely on Sanov's theorem 9.5.1: under $P^{\otimes N}$ the empirical measure of the spins states satisfies a full LDP principle with good rate function $I(\cdot|P)$, the relative entropy with respect to P. This also induces a convergence of the empirical measure toward P with exponential speed for the uncoupled network.

The strategy of the authors is now to compare the coupled network with the uncoupled network to derive another LDP for the former. This step is two-folded. First, Girsanov's theorem ensure that $P_\beta^N(J) \ll P^{\otimes N}$ with density:

$$\frac{\mathrm{d}P_\beta^N(J)}{\mathrm{d}P^{\otimes N}} = \exp\left\{ \sum_{i=1}^N \int_0^T \beta\Big(\frac{1}{\sqrt{N}}\sum_{j=1}^N J_{ij}x_t^j\Big)dB_t^i - \frac{\beta^2}{2}\int_0^T \Big(\frac{1}{\sqrt{N}}\sum_{j=1}^N J_{ij}x_t^j\Big)^2 dt \right\}.$$

allowing to define the averaged solutions of the networks (3.9) by integrating on the Gaussian realizations of the interaction matrix:

$$Q_\beta^N = \int P_\beta^N(J(\omega))d\gamma(\omega).$$

In particular, Fubini's theorem ensures that $Q_\beta^N \ll P^{\otimes N}$ with density

$$\frac{\mathrm{d}Q_\beta^N}{\mathrm{d}P^{\otimes N}} = \mathscr{E}\left(\exp\left\{ \sum_{i=1}^N \int_0^T \beta\Big(\frac{1}{\sqrt{N}}\sum_{j=1}^N J_{ij}x_t^j\Big)dB_t^i - \frac{\beta^2}{2}\int_0^T \Big(\frac{1}{\sqrt{N}}\sum_{j=1}^N J_{ij}x_t^j\Big)^2 dt \right\}\right),$$

giving, by independence of the J_{ij}

$$\frac{\mathrm{d}Q_\beta^N}{\mathrm{d}P^{\otimes N}} = \exp\left\{ \sum_{i=1}^N \log\left(\mathscr{E}\left[\exp\left\{ \int_0^T \beta\Big(\frac{1}{\sqrt{N}}\sum_{j=1}^N J_{ij}x_t^j\Big)dB_t^i - \frac{\beta^2}{2}\int_0^T \Big(\frac{1}{\sqrt{N}}\sum_{j=1}^N J_{ij}x_t^j\Big)^2 dt \right\}\right]\right)\right\}.$$

Because of the stability of Gaussian law, this density can be put into the form

$$\frac{\mathrm{d}Q_\beta^N}{\mathrm{d}P^{\otimes N}} = \exp\big\{N\Gamma(\hat{\mu}_N)\big\},$$

with

$$\Gamma(\mu) := \int \log\left(\int \exp\left\{ \beta\int_0^T G_t^\mu dW_t(x) - \frac{\beta^2}{2}\int_0^T {G_t^\mu}^2 dt \right\} d\gamma\right) d\mu(x),$$

where the $(G^\mu)_{\mu\in\mathscr{M}_1^+(\mathscr{C})}$ is a family of centered Gaussian processes of the space $(\tilde{\Omega},\tilde{\mathscr{F}},\tilde{\mathscr{P}})$ (actually, this formalism with μ as an exponent is borrowed from [129], but is equivalent to the original one), whose covariance structure is given by:

$$K_\mu(s,t) := \sigma^2 \int_{\mathscr{C}} x_t x_s d\mu(x).$$

Thus the proof heavily relies on the Gaussian nature of the interactions, as well as the noisy nature of the dynamics.
Second, this form for the density suggests the possible application of Varadhan's lemma 9.6.3. Unfortunately, Γ is neither bounded nor continuous, so that the lemma does not readily applies. Nevertheless, the authors were able to demonstrate a LDP for the sequence of probabilities $\big(Q_\beta^N(\hat{\mu}_N \in .)\big)_N$ with good rate function $H := I(\cdot|P) - \Gamma$, and obtain convergence results for the empirical measure. The proof follows three important lines:

- Proving that H is indeed a good rate function.
- Proving a LDP for the sequence $\left(Q_\beta^N(\hat{\mu}_N \in .)\right)_N$ with good rate function H.
- Characterizing the minima of H through a variational analysis, and proving that this characterization admits a unique minimum Q.

The two first steps of the demonstrations were initially done through a time discretization, before being extended to the continuous case. We now state the main results of [22].

3.3.3 Main results

The main result that G. Ben Arous and A. Guionnet obtained in [22] are valid under the short time, or high temperature hypothesis $\beta^2 A^2 T < 1$. It ensures the exponential tightness of $Q_\beta^N(\hat{\mu} \in \cdot)$, so that a full LDP follows. This full LPD, in turn induces a convergence of the empirical measure and a propagation of chaos in the averaged case, as well as a full LDP for the quenched sequence $P_\beta^N(J)$, yielding a quenched convergence. Moreover, it appears that H achieves its minimum at a unique non-markovian probability measure Q that satisfies an intricate implicit stochastic differential system. We now state the mains results of the authors:

Theorem 3.3.1. *There exists a good rate function H, such that if $\beta^2 A^2 T < 1$, $Q_\beta^N(\hat{\mu} \in \cdot)$ satisfies a full LDP with good rate function H.*

Theorem 3.3.2. *H achieves its minimal value at a unique probability measure Q on $\mathscr{C}([0,T],[-A,A])$ which is implicitly given by the following procedure:*
Let $P(h)$ be the law of the diffusion on $\mathscr{C}([0,T],[-A,A])$

$$\begin{cases} dx_t = -\nabla U(x_t)dt + dB_t + \beta h_t dt, \\ x_0 \overset{\mathscr{L}}{=} \mu_0 \end{cases} \tag{3.11}$$

for a deterministic process h in $L^2\left([0,T],dt\right)$.
Then, Q satisfy the non-linear equation

$$Q = \int P(G^Q(\omega))d\mathscr{P}(\omega),$$

where G^Q is, under $\mathscr{P}$, a centered Gaussian process with covariance

$$\int G_t^Q G_s^Q d\mathscr{P} = \int_{\mathscr{C}([0,T],[-A,A])} x_t x_s dQ(x).$$

Theorem 3.3.3. (i) *If $\beta^2 A^2 T < 1$, then $Q_\beta^N(\hat{\mu} \in \cdot)$ converges weakly to δ_Q, i.e.*

$$\forall F \in \mathscr{C}_b\Big(\mathscr{M}_1^+\big(\mathscr{C}([0,T],[-A,A])\big)\Big),\ \lim_{N\to+\infty} \mathscr{E}\left(\int_{\mathscr{C}([0,T],[-A,A])^N} F\left(\frac{1}{N}\sum_{i=1}^N \delta_{x^i}\right) dP_\beta^N(J)(\mathbf{x})\right) = F(Q).$$

In particular, if $f \in \mathscr{C}_b\big(\mathscr{C}([0,T],[-A,A])\big)$,

$$\lim_{N\to+\infty} \mathscr{E}\left(\int_{\mathscr{C}([0,T],[-A,A])^N} \frac{1}{N}\sum_{i=1}^N f(x^i) dP_\beta^N(J)(\mathbf{x})\right) = \int_{\mathscr{C}([0,T],[-A,A])} f(x)dQ(x).$$

(ii) *As a consequence, if* $\beta^2 A^2 T < 1$, $\forall k \in \mathbb{N}, \forall \varphi_1, \dots, \varphi_k \in \mathscr{C}_b\big(\mathscr{C}([0,T],[-A,A])\big)$,

$$\lim_{N\to+\infty} \mathscr{E}\left(\int_{\mathscr{C}([0,T],[-A,A])^N} \Big(\prod_{j=1}^{k} \varphi_j(x^j) \Big) dP_\beta^N(J)(\mathbf{x}) \right) = \prod_{j=1}^{k} \left(\int_{\mathscr{C}([0,T],[-A,A])} \varphi_j(x) dQ(x) \right).$$

Theorem 3.3.4. *There exists a good rate function* H *such that if* $\beta^2 A^2 T < 1$, *for any closed subset* F *of* $\mathscr{M}_1^+\big(\mathscr{C}([0,T],[-A,A])\big)$, *for almost all* J,

$$\limsup_{N\to+\infty} \frac{1}{N} \log P_\beta^N(J)\big(\hat{\mu}_N \in F\big) \leq -\inf_F H.$$

Theorem 3.3.5. *If* $\beta^2 A^2 T < 1$,

(i) *For any bounded continuous function* $F \in \mathscr{C}_b\Big(\mathscr{M}_1^+\big(\mathscr{C}([0,T],[-A,A])\big)\Big)$, *and for almost all* J,

$$\lim_{N\to+\infty} \int_{\mathscr{C}([0,T],[-A,A])^N} F\left(\frac{1}{N} \sum_{i=1}^{N} \delta_{x^i} \right) dP_\beta^N(J)(\mathbf{x}) = F(Q).$$

(ii) *For any* $f \in \mathscr{C}_b\big(\mathscr{C}([0,T],[-A,A])\big)$, *for almost all* J, *for almost all* $\big((x^i)_{1\leq i\leq N}\big)_{N\in\mathbb{N}^*}$,

$$\lim_{N\to+\infty} \mathscr{E}\left(\int_{\mathscr{C}([0,T],[-A,A])^N} \frac{1}{N} \sum_{i=1}^{N} f(x^i) dP_\beta^N(J)(\mathbf{x}) \right) = \int_{\mathscr{C}([0,T],[-A,A])} f(x) dQ(x).$$

Let us further precise that, relying on a replica argument, the authors show a quenched propagation of chaos, that provides statistical properties on the quenched single spin in the thermodynamic limit. Nevertheless, this analysis does not apply to our setting, as it relies on the hypothesis that

$$\int x_t dQ(x) = 0, \ \ \forall t \in [0,T].$$

In our model, this hypothesis would write:

$$\int S(x_t) dQ(x) = 0, \ \ \forall t \in [0,T],$$

which is not likely to happen as $S > 0$.

In a more recent publication of Alice Guionnet [129], the approach is slightly different. The author is still interested in convergence and propagation of chaos results, but intends to get rid of the high temperature and short time hypothesis. As a consequence the full LDP is discarded, but the author manages to conserve the upper-bound result for compact sets (see 9.4.3), and to obtain a tension result on Q_β^N. This again yields an averaged and quenched convergence of the empirical measure, as well as an averaged propagation of chaos results. Moreover, a quenched propagation of chaos result is obtained under further assumptions of symmetry for the potential U and law μ_0.

In the present manuscript, we have tried to capitalize on theses two publications in order to prove similar results for our neural networks. In many ways, our demonstration is a mere adaptation of their results to our biological framework, so that the

thread of the proof remains unchanged. As in [129] we shall not systematically restrict our results to a short time hypothesis, and thus shall not establish a full LDP but an upper-bound for compact along with a tension result. This approach will be adopted in Chapter 4. Nevertheless, it is always possible to prove that the empirical law satisfies a full LDP, provided a short-time hypothesis as in [22]. We will prove such a full LDP, along with quenched results in Chapter 5, as our analysis strongly relies on short times.

Even so, we have succeeded to substantially extend the scope of G. Ben Arous and A. Guionnet analysis. First of all, our networks are no longer homogeneous has they admit a spatial extension that impacts the dynamics. Alternatively, we could say that the network is composed of infinitely many different population (labeled by a continuous parameter), each satisfying its own dynamics, even though they appear to be quite similar. Remark that the case of a finite number of populations has already been addressed in [79] in the case of a discrete time dynamics. In addition, the voltage is here only almost surely bounded, the interactions are nonlinear and we introduce characteristic transportation delays depending on the position of the two neurons at play. This sets the problem in an infinite-dimensional space, as initial conditions are probability measure on the space of continuous function from a compact interval to $\mathbb{R}$. In addition, we consider synapses interactions depending on the state of both particles at play in chapter 5 5. Eventually, the synaptic weights are still i.i.d Gaussian variables, but are not centered. This was already addressed for strongly random synapses in [178]). In Chapter 4 we go a bit further, their mean and variance will be continuous maps of the heterogeneous positions of both neurons at play.

Part II

Large deviations for heterogeneous neural networks

CHAPTER 4

SPATIALLY EXTENDED NETWORKS

We investigate the asymptotic behavior of a spatially extended stochastic neural networks dynamics in random environment with highly random connectivity weights. These systems model the spatiotemporal activity of the brain, thus feature (i) communication delays depending on the distance between cells and (ii) heterogeneous synapses: connectivity coefficients are random variables whose law depends on the neurons positions and whose variance scales as the inverse of the network size. When the weights are independent Gaussian random variables, we show that the empirical measure satisfies a large-deviation principle. The associated good rate function achieves its minimum at a unique spatially extended probability measure, implying convergence of the empirical measure and propagation of chaos. The limit is characterized through complex non Markovian implicit equation in which the network interaction term is replaced by a non-local Gaussian process whose statistics depend on the solution over the whole neural field. We further demonstrate the universality of this limit, in the sense that neuronal networks with non-Gaussian interconnection weights converge towards it provided that synaptic weights have a sufficiently fast decay. This universality results holds under a technical condition on time horizon, noise and heterogeneity.

4.1 INTRODUCTION

We study the asymptotic behavior of spatially-extended neuronal networks with heterogeneous interconnections at a mesoscopic scale in which averaging effects occur but where one can still resolve fine spatial structures. In detail, we will consider a stochastic network equation of size N in random environment, in which

- neurons have random locations on a compact set $D \subset \mathbb{R}^d$,
- the amplitude of the interaction between two cells are heterogeneous. Their statistics depend on the cells positions and have a mean and a variance scaling as $1/N$,
- neurons communicate after a delay, also depending on the cells locations, associated with transport and transmission of information.

Each network is characterized by a random *configuration* that does not evolve in time. Within this fixed network configuration, the state of each neuron is described by a stochastic nonlinear process. The motivation for developing this model lies in the understanding of spatio-temporal patterns of activity of the cortex, as we

review in section 4.1.1. For spatially extended networks with "weak" interaction heterogeneities (variance of interconnection weights scaled as $1/N^2$), coupling methods have been used to derive a non-local McKean-Vlasov thermodynamic limit [236, 237] where the effective interaction term involves a non-local integral in space. This limit depends explicitly on the averaged spatial structure of the brain, thus preserving important information on spatiotemporal patterns of activity [235]. In the context of networks on lattices with non-random synapses and no delay, compactness methods were used to show a convergence result towards a nonlinear Fokker-Planck equation [161]. For interacting heterogeneous diffusions with non-random interconnections, large-deviations techniques were developed [74] and convergence of double-layer empirical distributions including state variance and heterogeneity was proved. In all these cases, the heterogeneity of the interconnections was not sufficient to affect the asymptotic behavior.

Strongly stochastic connections have been the object of intense studies in the domain of mathematical physics. Sophisticated techniques were developed in the context of spin glasses (see e.g. the reference books [232, 233]). Of particular relevance to our purposes, and in the same context, large-deviations techniques were devised for randomly connected networks with strongly heterogeneous interactions [14, 16, 22, 129]. The methods were then adapted for biological neural networks in discrete time settings in a number of models [60, 79, 104], and were recently extended to continuous-time diffusions with multiple populations and delays [51]. In all these contributions, synaptic weights were considered Gaussian and the limit found involved an implicit effective interaction term that has a Gaussian law. Although methods of proof use Gaussian calculus, the Gaussian nature of the limit process does not require weights to be Gaussian, but similarly to the central limit theorem, is valid for a broad class of couplings. This was rigorously addressed in [178] in the case of discrete-time dynamics for weights with sub-Gaussian tails. It proves surprisingly complex to generalize their approach in a continuous-time setting.

We undertake in this manuscript the characterization of spatially extended networks with continuous-time dynamics and strongly heterogeneous synapses as motivated by the study of the spatio-temporal cortical patterns of activity. To this purpose, we combine large-deviations estimates and the methods developed for spatially extended particle systems to demonstrate the thermodynamic convergence of the network equation and identify their non-Markovian limit, for Gaussian and non-Gaussian synaptic weights. Before we proceed to the exposition of the setting and main results, we briefly review our motivations and model.

4.1.1 Biological background

It has been widely shown that mammalian brain displays precise spatiotemporal patterns of activity that correlate with brain states and cognitive processes. Classical examples include transient and local activation of specific regions in the cortex while recalling a memory (see e.g. [114]), visual illusions [142] or the propagation of a localized stimulus [180]. A popular and very efficient approach to describe these phenomena is the Wilson and Cowan neural field equation [245, 246], characterizing the spatiotemporal evolution of the activity $u(r,t)$ of cells at location r on the neural field D through a simple integro-differential equation of type:

$$\frac{\partial u}{\partial t} = -u(r,t) + \int_D J(r,r')S(u(r',t))\,dr' + I(r,t) \tag{4.1}$$

where $I(r,t)$ represents the input to the population at location r, $J(r,r')$ is the averaged interconnection weight from neurons at location r' onto neurons at location r and the non-decreasing map S associates to a level of activity u the resulting spiking rate. This equation has been very successful in reproducing a number of biological phenomena, in particular working memory [148] and visual hallucination patterns [37, 95]. However, randomness is not explicitly present in it, and the relationship between the dynamics of individual cells and this macroscopic equation - a central problem in neuroscience [36] - is still elusive.

The present paper pursues the endeavor of addressing rigorously this relationship. Beyond its mathematical interest, this approach would provide a way to understand, from the biological viewpoint, the importance of individual cells or synaptic properties on brain's emerging behaviors. Specifically, this would provide a way to characterize the role of noise and heterogeneity, that were reported to be related to pathologies such as febrile seizures [12]. From the phenomenological viewpoint, it has also been shown that the variance of the weights can notably affect the behavior of the network and lead to phase transitions from trivial to chaotic solutions [221] or synchronization in two-populations networks [134].

The question of characterizing limits of large-scale dynamics of neuronal networks has a long history in neuroscience, and several mathematical and statistical physics methods were introduced. These range from PDE formalisms and kinetic equations [54, 197, 198] with deep applications to the visual system, moment reductions and master equations [35, 162], but also the development of specific Markov chain models reproducing in the thermodynamics limit the dynamics of Wilson-Cowan systems [35, 40, 41, 42, 43]. These techniques were generally developed in order to obtain limits of interconnected neurons through weakly stochastic synapses (typically constant or independent identically distributed synaptic weights with variance $1/N^2$, with N the typical number of incoming connections), and do not hold in the case of *strongly stochastic synapses* whereby synaptic weights have a variance scaled by $1/N$.

4.1.2 Microscopic Neuronal Network Model

The macroscopic activity of cells relies on the collective activity emerging from a large number N of neurons that are distributed over the cortex, seen as a d-dimensional compact set $D \subset \mathbb{R}^d$ (d is generally considered to be equal to 2, sometimes 3). The location of neuron $i \in \{1,\cdots,N\}$ is denoted $r_i \in D$, and we assume that locations are independently drawn according to a probability measure $\pi \in \mathscr{M}_1^+(D)$. The latter represents the density of neurons on the cortex, and is assumed to be absolutely continuous with respect to Lebsegue's measure. The state of neuron i is described by a variable $X^{i,N} \in \mathbb{R}^s$, $s \in \mathbb{N}^*$, and we will assume here for simplicity that $X^{i,N}$ is a scalar variable representing the voltage of each neuron and satisfying the equation:

$$dX_t^{i,N} = \left(f(r_i,t,X_t^{i,N}) + \sum_{j=1}^{N} J_{ij} S(X_{t-\tau_{ij}}^{j,N}) \right) dt + \lambda(r_i) dW_t^i. \tag{4.2}$$

The map $f(r,t,x)$ describes the intrinsic dynamics of a neuron at location r, time t and state x, $\lambda(r)$ the level of noise at location r, and we assume each neuron to be driven by an independent Brownian motions W^i. The interactions between cells are assumed, as in the classical firing-rate formalism [5, 245, 246], to be proportional to a sigmoidal transformation of their membrane potential $S(X_t^{j,N})$. S is a smooth (at least continuously differentiable) increasing map tending to 0 at $-\infty$ and to 1 at ∞. The

synaptic weight J_{ij} represents the amplitude and excitatory or inhibitory nature of the interaction depending on whether $J_{ij} > 0$ or $J_{ij} < 0$. There is no connection form j to i when $J_{ij} = 0$. The parameters τ_{ij} represent the delay of communication between the two neurons, and is assumed to be equal to a deterministic function of the location of neuron i and j: $\tau_{ij} = \tau(r_i, r_j)$ (generally an affine function of the distance $\|r_i - r_j\|_{\mathbb{R}^d}$ between cells when spikes are assumed to be transmitted at constant speed. Here $\|\cdot\|_{\mathbb{R}^d}$ is the Euclidean norm on $\mathbb{R}^d$).

The spatio-temporal activity of the cortex is obtained as a mesoscopic limit of cells activity that resolves distinct locations on the cortex, but where averaging effects related to the large dimension of the network are taken into account. In order to characterize these averaging effects, we will investigate the limit of the network as its size diverges. We thus need to describe how synaptic weights scale with the network size. Consistently with the underlying biological problem (see e.g. [236] and references therein), we assume that the connectivity weights J_{ij} are random variables whose law depends on the location of cells i and j, with mean $J(r_i, r_j)/N$ and variance $\sigma^2(r_i, r_j)/N$. The scaling of the mean ensures that the interaction term does not diverge, while the scaling on the variance, slower than usual cases in $1/N^2$ [231, 236] preserves a non-trivial contribution of the heterogeneous nature of the synaptic weights. Note that biologically, the synaptic weights cannot reach arbitrarily large or small values.

Before we proceed to rigorous developments, let us start by heuristically describing the large N behavior. One can generally get an intuition of the limit of such interacting systems by considering that the $(X^{j,N}, r_j)$ are i.i.d. and independent of the connectivity matrix, and that the network equation converges towards a spatially extended process with law $(\bar{X}(r), r)$. This assumption is known as the Boltzman's "molecular chaos" (*Stoßzahlansatz*) hypothesis. Under these assumptions, one can formally make the conjecture that the network interaction term $\sum_{j=1}^{N} J_{ij} S(X_t^{j,N})$ converges, by virtue of a functional central limit theorem, towards a Gaussian process $U_t^{\bar{X}}(r)$ with mean and covariance that are non-local (i.e. depending on the process X at all other locations), given by:

$$\begin{cases} \int_D J(r_i, r') \mathbb{E}\left[S(\bar{X}_t(r'))\right] d\pi(r') \\ \int_D \sigma(r_i, r')^2 \mathbb{E}\left[S(\bar{X}_t(r')) S(\bar{X}_s(r'))\right] d\pi(r') \end{cases} \tag{4.3}$$

and one thus obtains the implicit equation on $\bar{X}$:

$$d\bar{X}_t(r) = \left(f(r,t,\bar{X}_t(r)) + U_t^{\bar{X}}(r) \right) dt + \lambda(r) dW_t(r). \tag{4.4}$$

Interestingly, we recover an interaction term whose mean is exactly of the Wilson-Cowan type (4.1). Moreover, when $f(r,t,x) = -x$, solutions are Gaussian and their mean satisfies a Wilson-Cowan equation (4.1), in which the sigmoid function depends dynamically on noise and heterogeneity. In particular, we will see that Boltzmann's molecular chaos asymptotically occurs for any finite set of neurons. Note that Boltzmann's Stoßzahlansatz could indicate a certain degree of universality for the limit, as is diplayed by the central limit theorem. In particular, it is possible that this limit remains valid for synaptic weights with bounded second moment. Universality will be partially addressed here, as we will prove the validity of the limit for sub-Gaussian-tailed synaptic weights, as well as for Gaussian-tailed synaptic weights under a short-time hypothesis.

The organization of the paper is as follows. We provide in section 4.2 the notations and main assumptions on the model, as well as a summary of the main results.

Sections 4.3 and 4.4 deal with the case of Gaussian synaptic weights, and are respectively dedicated to the demonstration of a large deviations principle and to the identification of the limit. Section 4.5 is devoted to showing a general convergence result in the case where the synaptic weights are non-Gaussian, including in particular the biologically relevant case of bounded coefficients.

4.2 STATEMENT OF THE RESULTS

We investigate the thermodynamic limit of the neuronal network equations (4.2). These equations are diffusions in random environment, and thus exhibit two distinct sources of randomness:

- *Random environment:* the locations of neurons $\mathbf{r} := (r_i)_{i\leq N}$ and synaptic weights $J := \big(J_{ij}\big)_{i,j\leq N}$ are random variables of a probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \mathscr{P})$. They define the structure of the network, and are independent of the time-fluctuation of the states of the neurons.

- *Stochastic dynamics:* states of neurons are stochastic variables, solutions of a SDE driven by a collection of independent $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$-Brownian motions $(W_t^i)_{i\in\mathbb{N}}$.

The dynamics of the X^i thus depends both on the random environment (i.e., the realization of locations $\mathbf{r}$ and weights J) and noise (the realization of the Brownian motions). We will denote by $\mathscr{E}$ the expectation over the environment (i.e. with respect to the probability distribution $\mathscr{P}$) and introduce the shorthand notation $\mathscr{P}_J$ and $\mathscr{E}_J$ the probability and expectation over the synaptic weights matrix J only (that is, $\mathscr{P}$ and $\mathscr{E}$ conditioned over the positions $\mathbf{r}$). We recall that J depends on $\mathbf{r}$, but that the inverse is not true.

We work under a few regularity assumptions. In particular, we assume that the law of the synaptic weights is continuous in space. In details, although synaptic weights J_{ij} and $J_{i'j'}$ are independent for $i \neq i'$ or $j \neq j'$, we assume that their probability distribution continuously depends on the spatial location of the cells, in the sense that one can find a version $\tilde{J}_{i'j'}$ of $J_{i'j'}$ such that:

$$\mathscr{E}_J\Big(\big|\tilde{J}_{i'j'} - J_{ij}\big|\Big) \leq \frac{C}{N}\Big(\|r_i - r_{i'}\|_{\mathbb{R}^d} + \|r_j - r_{j'}\|_{\mathbb{R}^d}\Big). \tag{4.5}$$

for some $C > 0$ independent of the neurons locations. Moreover, the dynamics of the neurons is assumed to satisfy the following assumptions:

(i) The function f is K_f-Lipschitz continuous in its three variables.

(ii) The mean and variance of the weights J and σ are bounded and, respectively, K_J and K_σ-Lipschitz continuous in their second variable. We denote

$$\|J\|_\infty = \sup_{(r,r')\in D^2} |J(r,r')|, \quad \|\sigma\|_\infty = \sup_{(r,r')\in D^2} \sigma(r,r').$$

(iii) $\tau : D^2 \to \mathbb{R}^+$ is Lipschitz continuous, with constant K_τ. It is in particular bounded, by compactness of D. We denote by τ its supremum.

(iv) The diffusion coefficient $\lambda : D \to \mathbb{R}_+^*$ is K_λ-Lipschitz continuous and uniformly lower-bounded: $\forall r \in D$, $\lambda(r) \geq \lambda_* > 0$.

Let $\mathscr{C}_\tau := \mathscr{C}([-\tau,0],\mathbb{R})$, and $\mu_0 : D \to \mathscr{M}_1^+(\mathscr{C}_\tau)$ be an initial probability distribution mapping continuous in space in the sense that exists a random mapping $\vec{x}^0 : D \to \mathscr{C}_\tau$ on $(\Omega,\mathscr{F},\mathbb{P})$ and $C_0 > 0$ such that:

$$\forall r,r' \in D,\ \mathscr{L}(\vec{x}^0(r)) = \mu_0(r), \quad \mathbb{E}\Big[\sup_{-\tau\le s\le 0}\big|\vec{x}_s^0(r)-\vec{x}_s^0(r')\big|^2\Big] \le C_0\|r-r'\|_{\mathbb{R}^d}^2. \tag{4.6}$$

In this chapter, we will denote by $\|\cdot\|_{\tau,\infty}$ the supremum norm on $\mathscr{C}_\tau$. Throughout the paper, we consider that the network's initial conditions are independent realizations of μ_0:

$$\text{Law of } (x_t)_{t\in[-\tau,0]} = \bigotimes_{i=1}^N \mu_0(r_i). \tag{4.7}$$

It will often be useful to grant the existence of exponential quadratic moments to the solutions, and thus we will make the assumption that initial condition has the following moments condition:

$$\exists \upsilon > 0, \sup_{r\in D}\left\{\int_{\mathscr{C}_\tau}\exp\left\{\upsilon\|x\|_{\tau,\infty}^2\right\}d\big(\mu_0(r)\big)(x)\right\} < \infty. \tag{4.8}$$

We further assume that their trajectories have the same regularity in time as the Brownian motion. The first question that may arise at this point is the well-posedness of the network system. Since the equations driving the network constitute a standard delayed stochastic differential equation in dimension N with Lipschitz-continuous drift and diffusion functions with linear growth property, standard theory on delayed stochastic differential equations [73, 165] ensures existence, uniqueness and square integrability of solutions:

Proposition 4.2.1. For each $\mathbf{r} \in D^N$, and $J \in \mathbb{R}^{N\times N}$ and $T > 0$, there exists a unique weak solution to the system (4.2) defined on $[-\tau,T]$ with initial condition (4.7). Moreover, this solution is square integrable.

Remark 2. Note that if the initial condition was given by $(X_t^{i,N})_{t\in[-\tau,0]} = \zeta^i$ with $\zeta^i \overset{\mathscr{L}}{=} \mu_0(r_i)$, we can of course prove strong existence and uniqueness of solutions.

We now work with an arbitrary fixed time $T > 0$ and denote by $Q_{\mathbf{r}}^N(J)$ the unique law solution of the network equations restricted to the σ-algebra $\sigma(X_s^{i,N}, 1 \le i \le N, -\tau \le s \le T)$. $Q_{\mathbf{r}}^N(J)$ is a probability measure on $\mathscr{C}_\tau^N$ where $\mathscr{C}$ is the space of real valued continuous functions of $[-\tau,T]$. This measure depends on the realizations of both the connectivity matrix J, and the locations of neurons $\mathbf{r}$. For any $t \in [0,T]$, we will denote by $\|\cdot\|_{\infty,t}$ the supremum norm on $\mathscr{C}\big([-\tau,t],\mathbb{R}\big)$. In order to characterize the behavior of the system as the network size diverges, we will show a Large Deviations Principle (LDP) for the empirical measure. This requires delicate estimates, combining different elements.

First, Sanov's theorem states that, for independent copies of the same law μ on a Polish space Σ, the empirical measure satisfies a full LDP with good rate function corresponding to the relative entropy $I(.|\mu)$ defined, for $\nu \in \mathscr{M}_1^+(\Sigma)$, by:

$$I(\nu|\mu) := \begin{cases} \int_\Sigma \log\left(\frac{d\nu}{d\mu}(x)\right)d\nu(x) & \text{if } \nu \ll \mu, \\ \infty & \text{otherwise .} \end{cases}$$

Because of the connections, it is clear that Sanov's theorem does not apply as the states of neurons are not independent. Moreover, symmetry between cells is also

broken by the choice of a realization of the interaction matrix. This motivates us to introduce the system without interaction. When neurons are not coupled (i.e. $J_{ij} = 0$ for all (i, j)), and locations are known, the law of neurons in position $r \in D$ is given by the unique solution P_r of the one-dimensional standard SDE:

$$\begin{cases} dX_t = f(r,t,X_t)dt + \lambda(r)dW_t \\ (X_t)_{t\in[-\tau,0]} \overset{\mathscr{L}}{=} \mu_0(r). \end{cases} \tag{4.9}$$

We denote by P_r the law of this process restricted to the σ-algebra $\mathscr{G}_T = \sigma(X_s, s \leq T)$; it is a probability measure on the space $\mathscr{C}$. Remark that, by a direct application of Girsanov's theorem, $Q^N_{\mathbf{r}}(J)$ is absolutely continuous with respect to $P_{\mathbf{r}} := \bigotimes_{i=1}^N P_{r_i}$, and its density is given by the following equality:

$$\frac{\mathrm{d}Q^N_{\mathbf{r}}(J)}{\mathrm{d}P_{\mathbf{r}}}(\mathbf{x}) = \exp\left(\sum_{i=1}^N \int_0^T \Big(\frac{1}{\lambda(r_i)} \sum_{j=1}^N J_{ij} S(X^{j,N}_{t-\tau(r_i,r_j)})\Big) dW_t(x^i, r_i) \right.$$
$$\left. - \frac{1}{2}\int_0^T \Big(\frac{1}{\lambda(r_i)} \sum_{j=1}^N J_{ij} S(X^{j,N}_{t-\tau(r_i,r_j)})\Big)^2 dt \right), \tag{4.10}$$

where

$$W_t(x,r) := \frac{x_t - x_0}{\lambda(r)} - \int_0^t \frac{f(x_s,r,s)}{\lambda(r)} ds. \tag{4.11}$$

Remark that, by (4.9), $\big(W_t(.,r)\big)_t$ is a P_r-Brownian motion. Moreover, under $P_{\mathbf{r}}$ the Brownian motions $\big(W_t(x^i,r_i), 0 \leq t \leq T\big)_{i\in\{1\cdots N\}}$ are independent.

Under $P_{\mathbf{r}}$ neurons are independent but are not identically distributed as locations are heterogenous. We reduce this difficulty by averaging over locations. Let $P \in \mathscr{M}_1^+(\mathscr{C} \times D)$ be defined by $dP(x,r) := dP_r(x)d\pi(r)$. P properly defines a probability measure on $\mathscr{M}_1^+(\mathscr{C} \times D)$ (see Appendix 4.6.2), and is the law of the pairs $(X^{i,N}, r_i)$ when there is no interaction. We also construct a symmetric law for the coupled network:

Lemma 4.2.2. The map

$$\mathscr{Q} : \begin{cases} D^N \to \mathscr{M}_1^+(\mathscr{C}^N) \\ \mathbf{r} \to Q^N_{\mathbf{r}} \end{cases}$$

where $Q^N_{\mathbf{r}} := \mathscr{E}_J\big(Q^N_{\mathbf{r}}(J)\big)$, is continuous with respect to the weak topology. Moreover,

$$dQ^N(\mathbf{x},\mathbf{r}) := dQ^N_{\mathbf{r}}(\mathbf{x})d\pi^{\otimes N}(\mathbf{r})$$

defines a probability measure on $\mathscr{M}_1^+\big((\mathscr{C} \times D)^N\big)$.

This result is proved in Appendix 4.6.2.

Remark 3. • The probability measure Q^N averages the solutions on the different possible configurations $(J,\mathbf{r})$. Although being a relatively abstract object, it nevertheless provides relevant statistics as we make more explicit now. If $A \subset \big(\mathscr{C} \times D\big)^N$ is an event corresponding to e.g. a pathological behavior, then $Q^N(A)$ corresponds to the proportion of configurations ("brains") presenting this pathology. Conversely, as $Q^N_{\mathbf{r}}(J)$ is the law of one particular individual with a given configuration $(J,\mathbf{r})$, then $Q^N_{\mathbf{r}}(J)(A)$ provides the exact probability for him to suffer from A.

- Results under Q^N are called *averaged*, whereas those under $Q^N_{\mathbf{r}}(J)$ are called *quenched*. Quenched results are much more involved to demonstrate than averaged ones. Several methods have been developed to access these results, particularly based on replica [16, 22, 129]. We do not address these questions in the present chapter.

We are interested in the behavior of the double layer empirical measure:

$$\hat{\mu}_N := \frac{1}{N}\sum_{i=1}^{N} \delta_{(X^{i,N},r_i)}. \tag{4.12}$$

Sanov's theorem ensures that the empirical measure satisfies a full LDP under $P^{\otimes N}$. In the case of Gaussian interactions, we will build upon this result and use an ad hoc version of Varadhan's lemma to derive a weak LDP under Q^N. We will then characterize the possible minima of the associated good rate function, and prove by a fixed point argument that it admits a unique one, denoted Q, characterized as the non-Markovian solution of a MacKean-Vlasov SDE. Large deviations estimates will then ensure that the empirical measure converges toward this minimum. In detail, we show the following:

Theorem 4.2.3. *Suppose interactions are Gaussian. Then, under* $\left(Q^N(\hat{\mu}_N \in \cdot)\right)_{N\in\mathbb{N}^*}$ *satisfies a weak Large Deviations Principle of speed N and converges towards* $\delta_Q \in \mathscr{M}_1^+\left(\mathscr{M}_1^+(\mathscr{C}\times D)\right)$ *as N goes to infinity.*

Remark 4. Note that, for $T < \frac{\lambda_*^2}{2\|\sigma\|_\infty^2}$, a full large deviation principle can be demonstrated, implying exponential speed for the convergence,

$$\mathbb{P}-a.s., \text{ for almost all } J,\mathbf{r}, \qquad \frac{1}{N}\sum_{i=1}^{N}\delta_{(X^{i,N},r_i)} \overset{\mathscr{L}}{\to} Q,$$

that is a quenched convergence. Indeed, under the short-time hypothesis, we can readily prove exponential tightness of the averaged and quenched empirical measures, implying an upper-bound for closed sets, and almost sure convergence by a Borel-Cantelli argument. Exponential tightness crucially relies on the existence of greater moments than 1 of the radon-Nikodym density of $\frac{\mathrm{d}Q^N}{\mathrm{d}P^{\otimes N}}$ under $P^{\otimes N}$ (see [22]). This kind of results will be addressed in chapter 5.

The quantitative estimates leading to this convergence result are summarized in the following two results:

Theorem 4.2.4. *Suppose interactions are Gaussian. There exists a good rate function H on* $\mathscr{M}_1^+(\mathscr{C}\times D)$ *such that for any compact subset K of* $\mathscr{M}_1^+(\mathscr{C}\times D)$*,*

$$\limsup_{N\to\infty}\frac{1}{N}\log Q^N(\hat{\mu}_N \in K) \leq -\inf_K H.$$

The convergence result also relies on the tightness of the sequence of empirical measures:

Theorem 4.2.5. *Suppose interactions are Gaussian. Then, for any real number* $\varepsilon > 0$*, there exists a compact subset* K_ε *such that for any integer N,*

$$Q^N(\hat{\mu}_N \notin K_\varepsilon) \leq \varepsilon.$$

These two results imply convergence of the empirical measure towards the set of minima of the rate function H. Their uniqueness and characterization is subject of the following theorem demonstrated in section 4.4:

Theorem 4.2.6. *The good rate function H achieves its minimal value at the unique probability measure $Q \in \mathscr{M}_1^+(\mathscr{C}\times D)$ satisfying:*

$$Q \simeq P, \qquad \frac{dQ}{dP}(x,r) = \mathscr{E}\left[\exp\left\{\frac{1}{\lambda(r)}\int_0^T G_t^Q(r)dW_t(x,r) - \frac{1}{2\lambda(r)^2}\int_0^T G_t^Q(r)^2 dt\right\}\right]$$

where $(W_t)_{t\in[0,T]}$ is a P-brownian motion, and $G^Q(r)$ is, under $\mathscr{P}$, a Gaussian process with mean and covariance

$$\begin{cases} \mathscr{E}[G_t^Q(r)] = \int_{\mathscr{C}\times D} J(r,r')S(x_{t-\tau(r,r')})dQ(x,r') \\ \mathscr{E}[G_t^Q(r)G_s^Q(r)] = \int_{\mathscr{C}\times D} \sigma(r,r')^2 S(x_{t-\tau(r,r')})S(x_{s-\tau(r,r')})dQ(x,r). \end{cases}$$

For non-Gaussian synaptic weights, the LDP does not hold anymore. Nevertheless, as for central limit theorems, the limit found in the Gaussian case is universal when the weights are sufficiently concentrated. Here, we will handle the case of weights having at most Gaussian tails:

$$\begin{cases} \exists a, D_0 > 0,\ \forall N \geq 1,\ \forall J_1 \in \{J_{ij}(N), i,j \in [\![1,N]\!]\}, \\ \qquad \mathscr{E}_J\Big(\exp\left\{aNJ_1^2\right\}\Big) \leq D_0. \end{cases} \tag{H_J}$$

We will show in section 4.5 that for times $T < \lambda_*^2 a \wedge T^*$, with $T^* = \frac{\lambda_*^2}{2\|\sigma\|_\infty^2}$, the empirical measure converges towards the process described in Theorem 4.2.6. For sub-Gaussian synaptic weights (e.g., with bounded support), this convergence thus holds for any $T < T^*$. This indicates that the limit is universal to some degree. We expect that the universality of the Gaussian case goes beyond this case and may include synaptic weights having bounded polynomial moments (at least the two first moments). These extensions are not in the scope of the present paper, and our exponential convergence result covers all realistic cases arising in neuroscience where synaptic weights are bounded.

By symmetry of the law, the convergence result of Theorem 4.2.3 also implies propagation of chaos, thanks to a result due to A.S. Sznitman [230, Lemma 3.1]:

Theorem 4.2.7. *For any connectivity matrix satisfying hypothesis (H_J), the system enjoys the propagation of chaos property. In other terms, Q^N is Q-chaotic, i.e. for any bounded continuous functions $\varphi_1, \cdots, \varphi_m \in \mathscr{C}_b\big(\mathscr{C}\times D\big)$ of $\mathscr{C}\times D$, and any neuron indexes $(k_1,\cdots,k_m)$, we have:*

$$\lim_{N\to\infty} \int_{(\mathscr{C}\times D)^N} \prod_{j=1}^m \varphi_j(x^{k_j}, r_{k_j}) dQ^N(x,r) = \prod_{j=1}^m \int_{\mathscr{C}\times D} \varphi_j(x,r)dQ(x,r).$$

We now proceed to the proof of our results.

4.3 LARGE DEVIATION PRINCIPLE

The aim of this section is to establish the weak large deviation principle for the network with Gaussian synaptic weights. It relies on three key points. First, we will characterize the good rate function; the intuition for constructing this functional comes from Varadhan's lemma. In our case it does not readily apply and we need to thoroughly demonstrate that the candidate is indeed a good rate function. Second, we will show an upper-bound result for compact sets. The spatially-extended framework will introduce new difficulties, necessitating to introduce an appropriate distance on $\mathscr{C}\times D$. Third, the tightness of our collection of empirical measures will allow to conclude on a weak large-deviations principle.

4.3.1 Construction of the good rate function

Let us consider the interaction term of (4.2):

$$G_t^{i,N}(\mathbf{x},\mathbf{r}) := \frac{1}{\lambda(r_i)}\sum_{j=1}^{N} J_{ij}S\big(x^j_{t-\tau(r_i,r_j)}\big).$$

As stated in section (4.1.2), it shall behave as a Gaussian process in the large N limit, with mean and covariance given by (4.3). With this in mind, we introduce, for $\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)$, the two following functions defined respectively on $[0,T]^2\times D$ and $[0,T]\times D$:

$$\begin{cases} K_\mu(s,t,r) & := \dfrac{1}{\lambda(r)^2}\displaystyle\int_{\mathscr{C}\times D}\sigma(r,r')^2 S(x_{t-\tau(r,r')})S(x_{s-\tau(r,r')})d\mu(x,r'),\\ m_\mu(t,r) & := \dfrac{1}{\lambda(r)}\displaystyle\int_{\mathscr{C}\times D} J(r,r')S(x_{t-\tau(r,r')})d\mu(x,r'). \end{cases}$$

Here, μ can be understood as the putative limit law of the couple (x^j,r_j) if it exists. Covariance and mean functions K_μ and m_μ are well defined as for every fixed $r\in D$ the two maps

$$A_r:(x,r')\to\frac{1}{\lambda(r)}J(r,r')S(x_{.-\tau(r,r')}),\quad \tilde{A}_r(x,r')\to\frac{1}{\lambda(r)^2}\sigma(r,r')^2S(x_{.-\tau(r,r')})S(x_{.-\tau(r,r')})$$

are continuous for the classical product norm $\|(x,r)\|_{\mathscr{C}\times D}:=\|x\|_{\infty,T}+\|r\|_{\mathbb{R}^d}$. Hence, they are Borel-measurable, and integrable with respect to every element of $\mathscr{M}_1^+(\mathscr{C}\times D)$. Remark that, since S takes value in $[0,1]$, both functions are bounded:

$$|K_\mu(s,t,r)|\le\frac{\|\sigma\|_\infty^2}{\lambda_*^2},\quad |m_\mu(t,r)|\le\frac{\|\bar{J}\|_\infty}{\lambda_*}.$$

Moreover, as μ charges continuous functions, K_μ and m_μ are continuous maps by the dominated convergence theorem.

Clearly enough, K_μ has a covariance structure. As a consequence, we can define a probability space $(\hat{\Omega},\hat{\mathscr{F}},\gamma)$ and a family of independent stochastic processes $\big(G^\mu(r)\big)_{\mu\in\mathscr{M}_1^+(\mathscr{C}\times D),r\in D}$ for any measure $\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)$, such that $G^\mu(r)$ is a centered Gaussian process with covariance $K_\mu(.,.,r)$ under γ. This ensures continuity of the map $r\to\mathscr{L}\big(G^\mu(r)\big)$. We will denote by $\mathscr{E}_\gamma$ the expectation under this measure.

Remark 5. (i) As in [22, 51], we could alternatively have defined a family of probability measure $\big(\gamma_\mu\big)_{\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)}$, and a family of Gaussian processes $\big(G(r)\big)_{r\in D}$ with covariance K_μ under γ_μ. This approach is equivalent to ours, but the latter present the advantage of being very adapted to Fubini's theorem.

(ii) The family of processes $\left(G_t^{\mu}(r)\right)_{\mu,r}$ is intended to encompass possible candidates for the effective asymptotic interactions $\lim_N \left(G^{i,N}(\mathbf{x},\mathbf{r})\right)_{i\in\mathbb{N}^*}$. In these interactions, the Gaussian weights are independent for different particles, so that it seems natural to assume independence of $\left(G_t^{\mu}(r)\right)$ for different locations. Notably in our proof, we can swap from a continuous version of $G_t^{\mu}(r)$ to an independent one very easily, as they are never taken jointly under γ. Thus, we can literally choose their covariance structure. For the sake of measurability under any Borel measure of $\mathscr{M}_1^+\left(\mathscr{C}\times D\right)$, we will mainly work with the continuous version of $G^{\mu}(r)$, and will explicitly introduce independent versions when independence is needed.

We recall a few general properties on the relative entropy that are often used throughout the paper. For p and q two probability measures on a Polish space E (see equation (9.10) or e.g. [86, Lemma 3.2.13]), we have the identity:

$$I(q|p) = \sup\left\{\int_E \Phi(x)dq(x) - \log\int_E \exp\left\{\Phi(x)\right\}dp(x)\ ;\ \Phi\in\mathscr{C}_b(E)\right\},$$

which implies in particular that for any bounded measurable function Φ on E,

$$\int_E \Phi(x)dq(x) \le I(q|p) + \log\int_E \exp\left\{\Phi(x)\right\}dp(x). \tag{4.13}$$

If Φ is a lower-bounded (or upper-bounded) measurable function this inequality holds by monotone convergence.

We now state a key result to our analysis

Lemma 4.3.1.

$$\frac{\mathrm{d}Q^N}{\mathrm{d}P^{\otimes N}}(\mathbf{x},\mathbf{r}) = \exp\left\{N\bar{\Gamma}(\hat{\mu}_N)\right\},$$

where

$$\bar{\Gamma}(\hat{\mu}_N) := \frac{1}{N}\sum_{i=1}^N \log\mathscr{E}_{\gamma}\left[\exp\left\{\int_0^T \left(G_t^{\hat{\mu}_N}(r_i) + m_{\hat{\mu}_N}(t,r_i)\right)dW_t(x^i,r_i) - \frac{1}{2}\int_0^T \left(G_t^{\hat{\mu}_N}(r_i) + m_{\hat{\mu}_N}(t,r_i)\right)^2 dt\right\}\right],$$

with W defined as in (4.11).

Proof. Let, for all $(\mathbf{x},\mathbf{r})\in(\mathscr{C}\times D)^N$

$$X_i^N(\mathbf{x},\mathbf{r}) := \int_0^T G_t^{i,N}(\mathbf{x},\mathbf{r})dW_t(x^i,r_i) - \frac{1}{2}\int_0^T G_t^{i,N}(\mathbf{x},\mathbf{r})^2 dt,$$

which is well defined under $P_{\mathbf{r}}$. Going back to equation (4.10), we find:

$$\frac{\mathrm{d}Q_{\mathbf{r}}^N(J)}{\mathrm{d}P_{\mathbf{r}}}(\mathbf{x}) = \exp\Big(\sum_{i=1}^N X_i^N(\mathbf{x},\mathbf{r})\Big).$$

Averaging on J and applying Fubini theorem, we find that $Q_{\mathbf{r}}^N \ll P_{\mathbf{r}}$, with density $\frac{\mathrm{d}Q_{\mathbf{r}}^N}{\mathrm{d}P_{\mathbf{r}}}(\mathbf{x}) = \mathscr{E}_J\Big[\exp\Big(\sum_{i=1}^N X_i^N(\mathbf{x},\mathbf{r})\Big)\Big]$. Moreover, equalities $dQ^N(\mathbf{x},\mathbf{r}) = dQ_{\mathbf{r}}^N(\mathbf{x})d\pi^{\otimes N}(\mathbf{r})$ and $dP^{\otimes N}(\mathbf{x},\mathbf{r}) = dP_{\mathbf{r}}(\mathbf{x})d\pi^{\otimes N}(\mathbf{r})$ give

$$\frac{dQ^N}{dP^{\otimes N}}(\mathbf{x},\mathbf{r}) = \prod_{i=1}^N \mathscr{E}_J\left[\exp\left(X_i^N(\mathbf{x},\mathbf{r})\right)\right] = \exp\left\{\sum_{i=1}^N \log\mathscr{E}_J\left[\exp\left(X_i^N(\mathbf{x},\mathbf{r})\right)\right]\right\},$$

where we have used the independence of the synaptic weights J_{ij}. Note that here $\mathbf{x}$ are coordinates, thus independent of the J_{ij}, and the fact that $\left\{ G_t^{i,N}(\mathbf{x},\mathbf{r}), 0 \le t \le T \right\}$ is, under $\mathscr{P}_J$, a Gaussian process with covariance $K_{\hat{\mu}_N}(t,s,r_i)$, and mean $m_{\hat{\mu}_N}(t,r_i)$. □

Following Varadhan's lemma, this motivates to introduce and precise the domain of definition of the maps:

$$X^\mu(x,r) := \int_0^T \big(G_t^\mu(r) + m_\mu(t,r)\big) dW_t(x,r) - \frac{1}{2}\int_0^T \big(G_t^\mu(r) + m_\mu(t,r)\big)^2 dt, \tag{4.14}$$

$$\Gamma(\mu) := \int_{\mathscr{C}\times D} \log \mathscr{E}_\gamma\Big[\exp\big\{X^\mu(x,r)\big\}\Big] d\mu(x,r),$$

for all $(x,r) \in \mathscr{C}\times D$ and $\mu \in \mathscr{M}_1^+(\mathscr{C}\times D)$.

Proposition 4.3.2. The map

$$\Gamma := \mu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big) \to \begin{cases} \int_{\mathscr{C}\times D} \log \mathscr{E}_\gamma\Big[\exp\big\{X^\mu(x,r)\big\}\Big] d\mu(x,r) & \text{if } I(\mu|P) < \infty, \\ +\infty & \text{otherwise .} \end{cases} \tag{4.15}$$

is well defined in $\mathbb{R}\cup\{+\infty\}$, and satisfies

(i) $\Gamma \le I(\cdot|P)$,

(ii) $\exists \iota \in]0,1[, e \ge 0$, $|\Gamma(\mu)| \le \iota I(\mu|P) + e$.

Proof. If $I(\mu|P) = +\infty$ the result holds. We will thus suppose that $I(\mu|P) < +\infty$, which implies $\mu \ll P$. As $W(\cdot,r)$ is a P_r-Brownian motion, Girsanov's theorem ensures that the stochastic integral $\int_0^T \big(G_t^\mu(r) + m_\mu(t,r)\big) dW_t(x,r)$ is well defined γ-almost surely under μ.

(1):

Let $F_\mu := \log \mathscr{E}_\gamma\Big[\exp\big\{X^\mu(x,r)\big\}\Big]$ denote the integrand in the formulation of Γ (4.15). It is measurable as a continuous function of $(x,r) \to \big(K_\mu(t,s,r), 0 \le t,s \le T\big), \big(m_\mu(t,r), 0 \le t \le T\big), \big(W_t(x,r), 0 \le t \le T\big)$ that are continuous maps. Nevertheless, because of the contribution of the mean term m_μ[1], it is not bounded from below, as was the case in [22]. Let us prove that it is still μ-integrable. In fact, for any $M > 0$

$$-F_\mu^-(x,r) \le F_\mu^+(x,r) - F_\mu^-(x,r) = F_\mu(x,r) \le \log\Big(\mathscr{E}_\gamma\big[\exp\{X^\mu(x,r)\}\big] \vee M^{-1}\Big) =: F_{\mu,M}(x,r),$$

where F_μ^+ and F_μ^- respectively denote the positive and negative part of F_μ. As F_μ^- and $F_{\mu,M}$ are measurable and bounded from below, inequality (4.13) applies. Let $\alpha \ge 1$. On the one hand

$$\begin{aligned} \alpha \int_{\mathscr{C}\times D} F_{\mu,M}(x,r) d\mu(x,r) &\le I(\mu|P) + \log\left\{ \int_{\mathscr{C}\times D} \exp\Big\{\alpha F_{\mu,M}(x,r)\Big\} dP(x,r) \right\} \\ &\overset{\text{Jensen}}{\le} I(\mu|P) + \log\left\{ M^{-\alpha} + \int_{\mathscr{C}\times D} \mathscr{E}_\gamma\Big[\exp\big\{\alpha X^\mu(x,r)\big\}\Big] dP(x,r) \right\} \\ &\overset{\text{Fubini}}{\le} I(\mu|P) + \log\left\{ M^{-\alpha} + \mathscr{E}_\gamma\left[\int_D \int_{\mathscr{C}} \exp\Big\{\alpha X^\mu(x,r)\Big\} dP_r(x) d\pi(r) \right] \right\}, \end{aligned} \tag{4.16}$$

[1] see the expression of Γ_2 in Proposition 4.3.4

with the right-hand side of the two latter inequalities being possibly infinite. On the other hand,

$$\begin{aligned}
&\alpha\int_{\mathscr{C}\times D} F_\mu^-(x,r)d\mu(x,r) = \alpha\int_{\mathscr{C}\times D}\Big(-\log\mathscr{E}_\gamma\big[\exp\big\{X^\mu(x,r)\big\}\big]\Big)^+ d\mu(x,r)\\
&\overset{\text{Jensen}}{\leq} \alpha\int_{\mathscr{C}\times D}\Big(-\mathscr{E}_\gamma\big[X^\mu(x,r)\big]\Big)^+ d\mu(x,r) = \int_{\mathscr{C}\times D}\Bigg(\mathscr{E}_\gamma\Big[-\int_0^T\big(G_t^\mu(r)+m_\mu(t,r)\big)dW_t(x,r)\\
&-\frac{1}{2}\int_0^T\big(G_t^\mu(r)+m_\mu(t,r)\big)^2dt\Big]+\mathscr{E}_\gamma\Big[\int_0^T\big(G_t^\mu(r)+m_\mu(t,r)\big)^2dt\Big]\Bigg)^+ d\mu(x,r)\\
&\overset{(4.13)}{\leq} I(\mu|P)+\log\Bigg\{\int_{\mathscr{C}\times D}\exp\Bigg\{\alpha\Bigg(\mathscr{E}_\gamma\big[X^\mu(x,r)\big]+T\frac{\|J\|_\infty^2+\|\sigma\|_\infty^2}{\lambda_*^2}\Bigg)^+\Bigg\}dP(x,r)\Bigg\}\\
&\overset{\text{Jensen, Fubini}}{\leq} I(\mu|P)+\alpha C_T+\log\Bigg\{\mathscr{E}_\gamma\Big[\int_D\int_{\mathscr{C}}\exp\Big\{\alpha X^\mu(x,r)\Big\}dP_r(x)d\pi(r)\Big]\Bigg\}. \qquad (4.17)
\end{aligned}$$

Remark that in all the manuscript, we shall denote $\overset{(4.13)}{\leq}$ and $\overset{\text{Jensen}}{\leq}$ to respectively indicate that we have relied on equation (4.13) or Jensen's inequality in order to obtain the relation at stake.

Moreover, $W(.,r)$ being a P_r-Brownian motion, the martingale property yields

$$\mathscr{E}_\gamma\Big[\int_D\int_{\mathscr{C}}\exp\Big\{\alpha X^\mu(x,r)\Big\}dP_r(x)d\pi(r)\Big]\leq\int_D\mathscr{E}_\gamma\Big[\exp\Big\{\frac{\alpha^2-\alpha}{2}\int_0^T\big(G_t^\mu(r)+m_\mu(t,r)\big)^2dt\Big\}\Big]d\pi(r). \qquad (4.18)$$

Letting $\alpha=1$, we can see that F_μ is μ-integrable, with

$$\int_{\mathscr{C}\times D}|F_\mu(x,r)|d\mu(x,r)=\int_{\mathscr{C}\times D}F_\mu^-(x,r)+F_{\mu,1}(x,r)d\mu(x,r)\leq 2I(\mu|P)+C_T+\log(2). \qquad (4.19)$$

Moreover,

$$\Gamma(\mu):=\int_{\mathscr{C}\times D}F_\mu(x,r)d\mu(x,r)\overset{(4.16)}{\leq} I(\mu|P)+\log\big\{M^{-1}+1\big\},$$

so that letting $M\to+\infty$ yields the result.

(2):

For $\alpha\geq 1$, inequalities (4.16), (4.17), and (4.18) ensure that

$$\alpha|\Gamma(\mu)|\leq I(\mu|P)+\alpha C_T+\Bigg|\log\Bigg\{\int_D\mathscr{E}_\gamma\Big[\exp\Big\{\frac{\alpha^2-\alpha}{2}\int_0^T\big(G_t^\mu(r)+m_\mu(t,r)\big)^2dt\Big\}\Big]d\pi(r)\Bigg\}\Bigg|.$$

We recall that basic Gaussian calculus gives

$$\mathbb{E}\Big[\exp\Big\{\frac{1}{2}\mathscr{N}(m,v)^2\Big\}\Big]=\frac{1}{\sqrt{1-v}}\exp\Big\{\frac{m^2}{2(1-v)}\Big\}=\exp\Big\{\frac{1}{2}\Big(\frac{m^2}{1-v}-\log(1-v)\Big)\Big\}$$

as soon as $v<1$. Jensen's inequality and Fubini theorem yield

$$\mathscr{E}_\gamma\Big[\exp\Big\{\frac{(\alpha^2-\alpha)T}{2}\int_0^T\big(G_t^\mu(r)+m_\mu(t,r)\big)^2\frac{dt}{T}\Big\}\Big]\leq\int_0^T\mathscr{E}_\gamma\Big[\exp\Big\{\frac{(\alpha^2-\alpha)T}{2}\big(G_t^\mu(r)+m_\mu(t,r)\big)^2\Big\}\Big]\frac{dt}{T}.$$

As $\sqrt{(\alpha^2-\alpha)T}\Big(G_t^\mu(r)+m_\mu(t,r)\Big) \sim \mathscr{N}\Big(\sqrt{(\alpha^2-\alpha)T}m_\mu(t,r),(\alpha^2-\alpha)TK_\mu(t,t,r)\Big)$ under γ then, for $(\alpha-1)$ small enough, exists a constant C_T uniform in space such that

$$\mathscr{E}_\gamma\Big[\exp\Big\{\frac{(\alpha^2-\alpha)}{2}\int_0^T\big(G_t^\mu(r)+m_\mu(t,r)\big)^2dt\Big\}\Big] \leq \exp\Big\{(\alpha-1)C_T+\underbrace{o(\alpha-1)}_{\text{uniform in } r}\Big\}$$
$$\leq \exp\Big\{(\alpha-1)C_T\Big\}.$$

This eventually yields

$$|\Gamma(\mu)| \leq \iota I(\mu|P)+e,$$

with $\iota := \frac{1}{\alpha}$, and $e := (2\alpha-1)C_T$. □

Remark 6. Remark that $\hat{\mu}_N \not\ll P$, and $\Gamma(\hat{\mu}_N)=+\infty$ so that it is not equal to $\Gamma(\hat{\mu}_N) \neq \bar{\Gamma}(\hat{\mu}_N)$. These objects are different in nature, as the latter is random and must be considered under a proper probability measure on $\big(\mathscr{C}\times D\big)^N$, making sense of the stochastic integrals over the $\big(W_t(x^i,r_i),0\leq t\leq T\big)_{i\in\{1\cdots N\}}$ (which are well defined under $P_{\mathbf{r}}$).

As $\mathscr{C}\times D$ and $\mathscr{M}_1^+(\mathscr{C}\times D)$ are Polish spaces, and as the $(X^{i,N},r_i)$ are independent identically distributed random variables under $P^{\otimes N}$, Sanov's theorem ensures that the empirical measure satisfies, under this measure, a LDP with good rate function $I(.|P)$. Furthermore, if Γ was bounded and continuous, Varadhan's lemma would, as a consequence of Lemma (4.3.1), entail a full LDP under Q^N, with good rate function given by

$$H(\mu) := \begin{cases} I(\mu|P)-\Gamma(\mu) & \text{if } I(\mu|P)<\infty, \\ \infty & \text{otherwise .} \end{cases}$$

At this point, it would be easy to conclude would Γ present a few regularity properties. Unfortunately, Varadhan's lemma assumptions fail here, as Γ is not continuous nor bounded from above. Obtaining a weak LDP as well as the convergence of the empirical measure requires to come back to the basics of large deviations theory.

Observe that Γ is a nonlinear function of μ, involving in particular an exponential term depending on the mean and covariance structure of the Gaussian process. In order to handle terms of this type, a key technique proposed by Ben Arous and Guionnet is to linearize this map by considering the terms in the exponential as depending on an additional variable $\nu \in \mathscr{M}_1^+(\mathscr{C}\times D)$ [22, 129]. In our case, this family of linearizations are given by the maps:

$$\Gamma_\nu := \mu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big) \to \begin{cases} \int_{\mathscr{C}\times D}\log\mathscr{E}_\gamma\Big[\exp\big\{X^\nu(x,r)\big\}\Big]d\mu(x,r) & \text{if } I(\mu|P)<\infty, \\ +\infty & \text{otherwise .} \end{cases}$$

where $\mu,\nu\in\mathscr{M}_1^+(\mathscr{C}\times D)$.

Remark 7. Observe that Proposition 4.3.2 also applies to Γ_ν for every $\nu\in\mathscr{M}_1^+(\mathscr{C}\times D)$. Moreover, observe that $\Gamma(\mu)=\Gamma_\mu(\mu)$ for any $\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)$.

Moreover, defining

$$\bar{\Gamma}_\nu(\delta_{(x,r)}) := \log\mathscr{E}_\gamma\Big[\exp\big\{X^\nu(x,r)\big\}\Big], \qquad \bar{\Gamma}_\nu(\hat{\mu}_N) := \frac{1}{N}\sum_{i=1}^N\log\mathscr{E}_\gamma\Big[\exp\big\{X^\nu(x^i,r_i))\big\}\Big],$$

we note that $\bar{\Gamma}_\nu(\hat{\mu}_N) = \frac{1}{N}\sum_{i=1}^N \bar{\Gamma}_\nu(\delta_{(x^i,r_i)})$. Introducing $Q_\nu \in \mathscr{M}_1^+(\mathscr{C}\times D)$ by

$$dQ_\nu(x,r) := \exp\{\bar{\Gamma}_\nu(\delta_{(x,r)})\}dP(x,r) = \mathscr{E}_\gamma\Big[\exp\{X^\nu(x,r)\}\Big]dP(x,r), \tag{4.20}$$

we thus have

$$dQ_\nu^{\otimes N}(\mathbf{x},\mathbf{r}) = \exp\{N\bar{\Gamma}_\nu(\hat{\mu}_N)\}dP^{\otimes N}(\mathbf{x},\mathbf{r}).$$

This equality highlights a connection between two distinct LDP. On the one hand, applying again Sanov's theorem, the empirical measure satisfies a full LDP under $Q_\nu^{\otimes N}$, with good rate function $I(.|Q_\nu)$. On the other hand, Vardhan's lemma suggests that $\hat{\mu}_N$ satisfies, under the same measure, a LDP with rate function

$$H_\nu : \mu \to \begin{cases} I(\mu|P) - \Gamma_\nu(\mu) & \text{if } I(\mu|P) < +\infty, \\ +\infty & \text{otherwise.} \end{cases}$$

This is, for now, only a supposition, as its original counterpart Γ_ν, is not bounded from above nor continuous, and as $\Gamma_\nu(\hat{\mu}_N)$ and $\bar{\Gamma}_\nu(\hat{\mu}_N)$ are not equal. Still, assuming the result is true, uniqueness of the good rate function would imply that H_ν equals $I(.|Q_\nu)$. We shall justify the definition of Q_ν, and proceed to the rigorous demonstration of the latter equality in Theorem 4.3.6. For now, let us introduce some useful objects for the general scope of our demonstration. For any Gaussian process $(G_t)_{t\in[0,T]}$ of $\big(\hat{\Omega},\hat{\mathscr{F}},\gamma\big)$, and any $t\in[0,T]$

$$\Lambda_t(G) := \frac{\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\}}{\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 du\Big\}\Big]}. \tag{4.21}$$

For any $t\in[0,T]$, $r\in D$, and $\nu\in\mathscr{M}_1^+(\mathscr{C}\times D)$ the following defines a probability measure on $\big(\hat{\Omega},\hat{\mathscr{F}}\big)$ (see [183]):

$$d\gamma_{\widetilde{K}^t_{\nu,r}}(\omega) := \Lambda_t(G^\nu(\omega,r))d\gamma(\omega), \quad \forall\omega\in\hat{\Omega}.$$

As proven in [183], $G^\nu(r)$ is still a centered Gaussian process under $\gamma_{\widetilde{K}^t_{\nu,r}}$, with covariance given by:

$$\widetilde{K}^t_{\nu,r}(s,u) := \mathscr{E}_\gamma\Big[G_u^\nu(r)G_s^\nu(r)\Lambda_t\big(G^\nu(r)\big)\Big].$$

We also define for any $\nu\in\mathscr{M}_1^+\big(\mathscr{C}\times D\big)$, $(x,r)\in\mathscr{C}\times D$ and $t\in[0,T]$, the processes

$$L_t^\nu(x,r) := \int_0^t G_s^\nu(r)\Big(dW_s(x,r) - m_\nu(s,r)ds\Big), \quad V_t^\nu(x,r) := W_t(x,r) - \int_0^t m_\nu(s,r)ds. \tag{4.22}$$

Here are a few properties for these objects:

Proposition 4.3.3. Exists a constant $C_T>0$, such that for any $\nu\in\mathscr{M}_1^+\big(\mathscr{C}\times D\big)$, $r\in D$, $t\in[0,T]$,

$$\sup_{0\le s,u\le t}\widetilde{K}^t_{\nu,r}(s,u) \le C_T, \quad \Lambda_t\big(G^\nu(r)\big)\le C_T, \tag{4.23}$$

$$\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^T G_t^\nu(r)^2dt\Big\}\Big] = \exp\Big\{-\frac{1}{2}\int_0^T \widetilde{K}^t_{\nu,r}(t,t)dt\Big\}. \tag{4.24}$$

Moreover, if $(G_t)_{0\le t\le T}$ and $(G'_t)_{0\le t\le T}$ are two centered Gaussian processes of $\big(\hat{\Omega},\hat{\mathscr{F}},\gamma\big)$ with uniformly bounded covariance, then exists $\tilde{C}_T>0$ such that for all $t\in[0,T]$,

$$\big|\Lambda_t(G)-\Lambda_t(G')\big| \le \tilde{C}_T\Big\{\int_0^t \mathscr{E}_\gamma\Big[\big(G_s-G'_s\big)^2\Big]^{\frac{1}{2}}ds + \int_0^t\big|G_s^2-G_s'^2\big|ds\Big\}. \tag{4.25}$$

Proof. Observe that by Jensen inequality:

$$\Lambda_t\big(G^\nu(r)\big) \le \mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^\nu(r)^2 du\Big\}\Big]^{-1} \overset{\text{Jensen}}{\le} \exp\Big\{\frac{1}{2}\int_0^t \mathscr{E}_\gamma\Big[G_s^\nu(r)^2\Big]du\Big\} \le \exp\Big\{\frac{\|\sigma\|_\infty^2 t}{2\lambda_*^2}\Big\}.$$

As a consequence:

$$\begin{aligned}\widetilde{K}_{\nu,r}^t(s,u) = \mathscr{E}_\gamma\Big[G_u^\nu(r)G_s^\nu(r)\Lambda_t\big(G^\nu(r)\big)\Big] &\overset{\text{C.S.}}{\le} \sqrt{K_\nu(s,s,r)K_\nu(t,t,r)}\exp\Big\{\frac{\|\sigma\|_\infty^2 t}{2\lambda_*^2}\Big\}\\ &\le \frac{\|\sigma\|_\infty^2}{\lambda_*^2}\exp\Big\{\frac{\|\sigma\|_\infty^2 t}{2\lambda_*^2}\Big\},\end{aligned}$$

For equality (4.24), let $f(t) := \mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^\nu(r)^2 ds\Big\}\Big]$. As $(t,\omega)\to G_t^\nu(\omega,r)^2\exp\Big\{-\frac{1}{2}\int_0^t G_s^\nu(\omega,r)^2 ds\Big\}$ is a well defined, γ-a.s. continuous, and integrable under γ, we have

$$f'(t) = -\frac{1}{2}\mathscr{E}_\gamma\Big[G_t^\nu(r)^2\exp\Big\{-\frac{1}{2}\int_0^t G_s^\nu(r)^2 ds\Big\}\Big] = -\frac{1}{2}\widetilde{K}_{\nu,r}^t(t,t)f(t),$$

so that integrating $\frac{f'}{f}$ gives the result. Furthermore, letting $(G_t)_{0\le t\le T}$ and $(G'_t)_{0\le t\le T}$ be two centered γ-Gaussian processes with variance bounded by a common constant C_T, we have:

$$\begin{aligned}\big|\Lambda_t(G)-\Lambda_t(G')\big| &= \Big|\frac{\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\}}{\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\}\Big]} - \frac{\exp\Big\{-\frac{1}{2}\int_0^t G_s'^2 ds\Big\}}{\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s'^2 ds\Big\}\Big]}\Big|\\
&\le \exp\Big\{\frac{\|\sigma\|_\infty^2 t}{\lambda_*^2}\Big\}\Big\{\Big|\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\}-\exp\Big\{-\frac{1}{2}\int_0^t G_s'^2 ds\Big\}\Big]\Big|\\
&+\Big|\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\}-\exp\Big\{-\frac{1}{2}\int_0^t G_s'^2 ds\Big\}\Big|\Big\},\\
&\le \frac{1}{2}\exp\Big\{\frac{\|\sigma\|_\infty^2 t}{\lambda_*^2}\Big\}\Big\{\int_0^t \mathscr{E}_\gamma\Big[\big|G_s^2-G_s'^2\big|\Big]ds+\int_0^t\big|G_s^2-G_s'^2\big|ds\Big\},\end{aligned}$$

where we have used the Lipschitz-continuity of exponential on $\mathbb{R}_-$. Consequently, relying on Cauchy-Schwarz inequality, we obtain

$$\big|\Lambda_t(G)-\Lambda_t(G')\big| \overset{\text{C.S.}}{\le} \tilde{C}_T\Big\{\int_0^t \mathscr{E}_\gamma\Big[\big(G_s-G_s'\big)^2\Big]^{\frac{1}{2}}ds+\int_0^t\big|G_s^2-G_s'^2\big|ds\Big\}.$$

□

We now introduce a very useful decomposition of Γ_ν based on Gaussian calculus.

Proposition 4.3.4. For every $\nu\in\mathscr{M}_1^+(\mathscr{C}\times D)$, Γ_ν admits the following decomposition:

$$\Gamma_\nu = \Gamma_{1,\nu}+\Gamma_{2,\nu}, \tag{4.26}$$

where $\forall\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)$

$$\Gamma_{1,\nu}(\mu) := -\frac{1}{2}\int_{\mathscr{C}\times D}\int_0^T\Big\{\widetilde{K}_{\nu,r}^t(t,t)+m_\nu(t,r)^2\Big\}dtd\mu(x,r),$$

and

$$\Gamma_{2,\nu}(\mu) := \begin{cases} \frac{1}{2}\int_{\mathscr{C}\times D}\int_{\hat{\Omega}} L_T^\nu(x,r)^2 d\gamma_{\widetilde{K}^T_{\nu,r}} d\mu(x,r) + \int_{\mathscr{C}\times D}\int_0^T m_\nu(t,r)dW_t(x,r)d\mu(x,r) & \text{if } I(\mu|P)<\infty, \\ +\infty & \text{otherwise .} \end{cases}$$

In particular, we have $\Gamma = \Gamma_1 + \Gamma_2$ where $\Gamma_i(\mu) := \Gamma_{i,\mu}(\mu)$ for $i \in \{1,2\}$ and any $\mu \in \mathscr{M}_1^+(\mathscr{C}\times D)$.

Proof. The equality is clearly satisfied outside $\{I(\cdot|P) < +\infty\}$. Let then $\mu \in \{I(\cdot|P) < +\infty\}$, and let us prove that $\Gamma_{2,\nu}$ is well defined. Both integrands of the sum are measurable, and the first one is positive implying integrability. Moreover, it is easy to see that the second integrand is integrable as

$$\begin{aligned}
&\int_{\mathscr{C}\times D}\Big(\int_0^T m_\nu(t,r)dW_t(x,r)\Big)^+ d\mu(x,r) \overset{(4.13)}{\leq} I(\mu|P) \\
&\quad + \log\bigg(\int_{\mathscr{C}\times D}\exp\Big\{\Big(\int_0^T m_\nu(t,r)dW_t(x,r) - \frac{1}{2}\int_0^T m_\nu(t,r)^2dt\Big)^+ + \frac{\|J\|_\infty^2 T}{\lambda_*^2}\Big\}dP(x,r)\bigg) \\
&\quad \leq I(\mu|P) + \frac{\|J\|_\infty^2 T}{\lambda_*^2} + \log\bigg(1 + \int_{\mathscr{C}\times D}\exp\Big\{\int_0^T m_\nu(t,r)dW_t(x,r) - \frac{1}{2}\int_0^T m_\nu(t,r)^2dt\Big\}dP(x,r)\bigg) \\
&\quad \leq I(\mu|P) + \frac{\|J\|_\infty^2 T}{\lambda_*^2} + \log(2) < +\infty,
\end{aligned}$$

by martingale property, and the negative part can be treated similarly. Now,

$$\begin{aligned}
\Gamma_\nu(\mu) &= \int_{\mathscr{C}\times D}\log\bigg(\int_{\hat{\Omega}}\exp\Big\{\int_0^T \big(G_t^\nu(\omega,r)+m_\nu(t,r)\big)dW_t(x,r) \\
&\quad - \frac{1}{2}\int_0^T \big(G_t^\nu(\omega,r)+m_\nu(t,r)\big)^2 dt\Big\}d\gamma(\omega)\bigg)d\mu(x,r) \\
&= \int_{\mathscr{C}\times D}\log\bigg(\exp\Big\{\int_0^T m_\nu(t,r)dW_t(x,r) - \frac{1}{2}\int_0^T (m_\nu(t,r))^2dt\Big\} \\
&\quad \times \mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_u^\nu(r)^2 du\Big\}\Big]\int_{\hat{\Omega}}\exp\Big\{L_T^\nu(\omega,x,r)\Big\}d\gamma_{\widetilde{K}^T_{\nu,r}}(\omega)\bigg)d\mu(x,r) \\
&\overset{(4.24)}{=} \Gamma_{1,\nu}(\mu) + \int_{\mathscr{C}\times D}\int_0^T m_\nu(t,r)dW_t(x,r)d\mu(x,r) + \int_{\mathscr{C}\times D}\log\Big\{\int_{\hat{\Omega}}\exp L_T^\nu(\omega,x,r)d\gamma_{\widetilde{K}^T_{\nu,r}}(\omega)\Big\}d\mu(x,r)
\end{aligned}$$

But standard Gaussian [183] calculus gives the P-a.s. equality:

$$\int_{\hat{\Omega}}\exp\Big\{L_T^\nu(\omega,x,r)\Big\}d\gamma_{\widetilde{K}^T_{\nu,r}}(\omega) = \exp\Big\{\frac{1}{2}\int_{\hat{\Omega}} L_T^\nu(\omega,x,r)^2 d\gamma_{\widetilde{K}^T_{\nu,r}}\Big\}$$

so that,

$$\begin{aligned}
\Gamma_\nu(\mu) &= \Gamma_{1,\nu}(\mu) + \int_{\mathscr{C}\times D}\int_0^T m_\nu(t,r)dW_t(x,r)d\mu(x,r) + \frac{1}{2}\int_{\mathscr{C}\times D}\int_{\hat{\Omega}} L_T^\nu(\omega,x,r)^2 d\gamma_{\widetilde{K}^T_{\nu,r}}d\mu(x,r) \\
&= \Gamma_{1,\nu}(\mu) + \Gamma_{2,\nu}(\mu).
\end{aligned}$$

□

This decomposition has the interest of splitting the difficulties: while the first term will be relatively easy to handle (see Proposition 4.3.7), the local martingale term will require finer estimates based on Gaussian calculus and a number of tools from stochastic calculus theory. It is also useful to prove the following lemma, central for our analysis:

Lemma 4.3.5.

$$\frac{\mathrm{d}Q_\nu}{\mathrm{d}P}(x,r) = \exp\left\{ \int_0^T O_\nu(t,x,r)dW_t(x,r) - \frac{1}{2}\int_0^T O_\nu^2(t,x,r)dt \right\} \tag{4.27}$$

where

$$O_\nu(t,x,r) = \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)L_t^\nu(x,r)\Big] + m_\nu(t,r).$$

Proof. We reproduce, here, the proof given in [22, Lemma 5.15]. Let us first, for any $r \in D$, introduce the probability measure $Q_{\nu,r}$ on $\mathscr{M}_1^+(\mathscr{C})$ defined by

$$dQ_{\nu,r}(x) := \mathscr{E}_\gamma\Big[\exp\Big\{X^\nu(x,r)\Big\}\Big]\,dP_r(x), \tag{4.28}$$

so that $dQ_\nu(x,r) = dQ_{\nu,r}(x)d\pi(r)$. Using the decomposition of Γ_ν, we find by Gaussian calculus

$$\frac{\mathrm{d}Q_{\nu,r}}{\mathrm{d}P_r}(x) \overset{(4.26)}{=} \exp\left\{ -\frac{1}{2}\int_0^T \Big(\widetilde{K}_{\nu,r}^t(t,t) + m_\nu(t,r)^2\Big)dt \right\} \exp\left\{ \frac{1}{2}\int_{\hat{\Omega}} L_T^\nu(x,r)^2 d\gamma_{\widetilde{K}_{\nu,r}^T} + \int_0^T m_\nu(t,r)dW_t(x,r) \right\}. \tag{4.29}$$

Itô's formula implies

$$L_T^\nu(x,r)^2 = 2\int_0^T G_t^\nu(r)L_t^\nu(x,r)dV_t^\nu(x,r) + \int_0^T G_t^\nu(r)^2 dt, \tag{4.30}$$

$$\Lambda_T(G^\nu(r)) = 1 + \frac{1}{2}\int_0^T \Lambda_t(G^\nu(r))\big(\widetilde{K}_{\nu,r}^t(t,t) - G_t^\nu(r)^2\big)dt,$$

so that

$$\begin{aligned} L_T^\nu(x,r)^2\Lambda_T(G^\nu(r)) = {} & 2\int_0^T \Lambda_t(G^\nu(r))G_t^\nu(r)L_t^\nu(x,r)dV_t^\nu(x,r) + \int_0^T \Lambda_t(G^\nu(r))G_t^\nu(r)^2 dt \\ & + \frac{1}{2}\int_0^T L_t^\nu(x,r)^2\Lambda_t(G^\nu(r))\big(\widetilde{K}_{\nu,r}^t(t,t) - G_t^\nu(r)^2\big)dt. \end{aligned}$$

On the one hand, a stochastic version of Fubini's theorem (8.3.2) gives:

$$\mathscr{E}_\gamma\left[\int_0^T \Lambda_t(G^\nu(r))G_t^\nu(r)L_t^\nu(x,r)dV_t^\nu(x,r)\right] = \int_0^T \mathscr{E}_\gamma\Big[\Lambda_t(G^\nu(r))G_t^\nu(r)L_t^\nu(x,r)\Big]dV_t^\nu(x,r). \tag{4.31}$$

This will be justified at the end of the demonstration. On the other hand, Isserlis' theorem (8.2.4)

$$\mathscr{E}_\gamma\Big[L_t^\nu(x,r)^2\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)^2\Big] = \mathscr{E}_\gamma\Big[L_t^\nu(x,r)^2\Lambda_t(G^\nu(r))\Big]\widetilde{K}_{\nu,r}^t(t,t) + 2\mathscr{E}_\gamma\Big[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)L_t^\nu(x,r)\Big]^2,$$

so that

$$\begin{aligned} \frac{1}{2}\mathscr{E}_\gamma\Big[L_T^\nu(x,r)^2\Lambda_T(G^\nu(r))\Big] - \frac{1}{2}\int_0^T \widetilde{K}_{\nu,r}^t(t,t)dt = {} & \int_0^T \mathscr{E}_\gamma\Big[\Lambda_t(G^\nu(r))G_t^\nu(r)L_t^\nu(x,r)\Big]dV_t^\nu(x,r) \\ & - \frac{1}{2}\int_0^T \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)L_t^\nu(x,r)\Big]^2 dt. \end{aligned} \tag{4.32}$$

Combining (4.29) and (4.32) gives the desired result.
We now justify that Fubini's theorem for stochastic integrals 8.3.2 holds to give (4.31). First observe it is easy to check that $V_t^\nu(x,r)$ is a P_r-semimartingale in $\mathscr{H}^2$, and that the map of $\hat{\Omega}\times[0,T]\times\mathscr{C}$ define as $(\omega,t,x)\to\Lambda_t\big(G^\nu(\omega,r)\big)G_t^\nu(\omega,r)L_t^\nu(\omega,x,r)$ is indeed $\hat{\mathscr{F}}\otimes\mathscr{P}$-measurable[2]. We thus just need to check that $\Lambda_t(G^\nu(r))G_t^\nu(r)L_t^\nu(x,r)$ and $\mathscr{E}_\gamma\Big[\Lambda_t(G^\nu(r))G_t^\nu(r)L_t^\nu(x,r)\Big]^{\frac{1}{2}}$ are in $L\big(V^\nu\big)$, that is:

$$\int_{\mathscr{C}}\int_0^T\Lambda_s(G^\nu(r))^2G_s^\nu(r)^2L_s^\nu(x,r)^2dsdP_r(x)<+\infty,\ \ \gamma-a.s.$$

and

$$\int_{\mathscr{C}}\int_0^T\Big|\mathscr{E}_\gamma\Big[\Lambda_s(G^\nu(r))G_s^\nu(r)L_s^\nu(x,r)\Big]\Big|dsdP_r(x)<+\infty.$$

On the one hand,

$$\begin{aligned}&\int_{\mathscr{C}}\int_0^T\Lambda_s(G^\nu(r))^2G_s^\nu(r)^2L_s^\nu(x,r)^2dsdP_r(x)\overset{\text{C.S.}}{\leq}2T\int_0^T\Lambda_s(G^\nu(r))^2G_s^\nu(r)^2\int_0^sG_u^\nu(r)^2m_\mu(u,r)^2duds\\&+2\int_0^T\Lambda_s(G^\nu(r))^2G_s^\nu(r)^2\int_{\mathscr{C}}\Big(\int_0^sG_u^\nu(x,r)dW_u(x,r)\Big)^2dP_r(x)ds\\&\overset{\text{B.D.G.}}{\leq}C_T\int_0^T\Lambda_s(G^\nu(r))^2G_s^\nu(r)^2\int_0^sG_u^\nu(r)^2duds<+\infty,\ \ \gamma-a.s..\end{aligned}$$

where C.S. stands for Cauchy-Shwarz inequality. On the other hand, by Fubini's theorem

$$\begin{aligned}\int_{\mathscr{C}}\int_0^T\Big|\mathscr{E}_\gamma\Big[\Lambda_s(G^\nu(r))G_s^\nu(r)L_s^\nu(x,r)\Big]\Big|dsdP_r(x)&\overset{\text{C.S.,(4.23)}}{\leq}C_T\Big(\int_0^T\mathscr{E}_\gamma\Big[\int_{\mathscr{C}}\Big(\int_0^sG_u^\nu(r)dW_u(x,r)\Big)^2dP_r(x)\Big]ds\Big)^{\frac{1}{2}}\\&\overset{\text{B.D.G.}}{<}+\infty.\end{aligned}$$

where B.D.G. stands for Burkhölder-Davis-Gundy inequality. □

We are now ready to state one of the main result of the chapter which proves the intuitive equality between the two rate functions $I(\cdot|Q_\nu)$ and H_ν.

Theorem 4.3.6. *Q_ν is a well defined probability measure on $\mathscr{M}_1^+(\mathscr{C}\times D)$, and $H_\nu(\mu)=I(\mu|Q_\nu)$. In particular H_ν is a good rate function.*

Proof. Fix $r\in D$, and define the probability measure $\bar{Q}_{\nu,r}\in\mathscr{M}_1^+(\mathscr{C})$ by:

$$\frac{\mathrm{d}\bar{Q}_{\nu,r}}{\mathrm{d}P_r}(x):=\exp\Big\{\int_0^Tm_\nu(t,r)dW_t(x,r)-\frac{1}{2}\int_0^Tm_\nu(t,r)^2dt\Big\}$$

for which Novikov's criterion holds by boundedness of m_ν. By Girsanov's theorem, $V^\nu(\cdot,r)$ (defined in (4.22)) is a $\bar{Q}_{\nu,r}$-Brownian motion, and we can use Novikov's criterion again to check that:

$$\frac{\mathrm{d}P_r}{\mathrm{d}\bar{Q}_{\nu,r}}(x):=\exp\Big\{-\int_0^Tm_\nu(t,r)dV_t^\nu(x,r)-\frac{1}{2}\int_0^Tm_\nu(t,r)^2dt\Big\},$$

[2]We work on the filtered space $\Big(\mathscr{C},(\mathscr{F}_t^{\mathbf{X}})_{0\leq t\leq T},\mathscr{B}(\mathscr{C}),P_r\Big)$, where $\Big(\mathscr{F}_t^{\mathbf{X}}:=\sigma\big(x_s,0\leq s\leq t\big)\Big)_{0\leq t\leq T}$ denotes the canonical filtration of the coordinate process. In this setting, $\mathscr{P}$ denotes the σ-field generated by adapted continuous processes.

implying $\bar{Q}_{\nu,r} \simeq P_r$. Moreover, let $\bar{Q}_\nu \in \mathscr{M}_1^+(\mathscr{C}\times D)$ be such that $d\bar{Q}_\nu(x,r) = d\bar{Q}_{\nu,r}(x)d\pi(r)$. Then $\bar{Q}_\nu \simeq P$, and by the previous lemma $Q_\nu \ll \bar{Q}_\nu$ with density:

$$\begin{aligned}
\frac{\mathrm{d}Q_\nu}{\mathrm{d}\bar{Q}_\nu}(x,r) &= \mathscr{E}_\gamma\left[\exp\left\{\int_0^T G_t^\nu(r)dV_t^\nu(x,r) - \frac{1}{2}\int_0^T G_t^\nu(r)^2 dt\right\}\right] \\
&\overset{(4.27)}{=} \exp\left\{\int_0^T \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)L_t^\nu(x,r)\Big]dV_t^\nu(x,r) - \frac{1}{2}\int_0^T \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)L_t^\nu(x,r)\Big]^2 dt\right\} \\
&\overset{(4.26)}{=} \exp\left\{\frac{1}{2}\int_{\hat{\Omega}} L_T^\nu(x,r)^2 d\gamma_{\widetilde{K}_{\nu,r}^T} - \frac{1}{2}\int_0^T \widetilde{K}_{\nu,r}^t(t,t)dt\right\} \overset{(4.23)}{\geq} \exp\big\{-C_T\big\} > 0.
\end{aligned}$$

We will first prove that $I(Q_{\nu,r}|\bar{Q}_{\nu,r})$ is finite. This will bring, by applying the exact same reasoning as in [22, Appendix B], the equality:

$$\forall \mu \in \mathscr{M}_1^+(\mathscr{C}\times D), \qquad \bar{H}_\nu(\mu) = I(\mu|Q_\nu),$$

where

$$\bar{H}_\nu : \mu \to \begin{cases} I(\mu|\bar{Q}_\nu) - \int_{\mathscr{C}\times D}\log\left(\frac{\mathrm{d}Q_\nu}{\mathrm{d}\bar{Q}_\nu}(x,r)\right)d\mu(x,r) & \text{if } I(\mu|\bar{Q}_\nu) < +\infty, \\ +\infty & \text{otherwise.} \end{cases}$$

We will then prove that for every $\mu \in \mathscr{M}_1^+(\mathscr{C}\times D)$,

$$\bar{H}_\nu(\mu) = H_\nu(\mu), \tag{4.33}$$

which will conclude the proof.

For the first point, observe that Girsanov's theorem ensures that the process $\Big(B_t^\nu(\cdot,r) := V_t^\nu(\cdot,r) - \int_0^t \mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(r)\big)G_s^\nu(r)L_s^\nu(\cdot,r)\Big]ds\Big)_{0\leq t\leq T}$ is a $Q_{\nu,r}$-Brownian motion, so that

$$\begin{aligned}
&I(Q_{\nu,r}|\bar{Q}_{\nu,r}) \\
&= \int_{\mathscr{C}}\left\{\int_0^T \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)L_t^\nu(x,r)\Big]dV_t^\nu(x,r) - \frac{1}{2}\int_0^T \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)L_t^\nu(x,r)\Big]^2 dt\right\}dQ_{\nu,r}(x) \\
&= \int_{\mathscr{C}}\left\{\int_0^T \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)L_t^\nu(x,r)\Big]dB_t^\nu(x,r) + \frac{1}{2}\int_0^T \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)L_t^\nu(x,r)\Big]^2 dt\right\}dQ_{\nu,r}(x) \\
&= \frac{1}{2}\int_0^T \underbrace{\int_{\mathscr{C}} \mathscr{E}_\gamma\left[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)\left(\int_0^t G_s^\nu(r)dV_s^\nu(x,r)\right)\right]^2 dQ_{\nu,r}(x)}_{\varphi_\nu(t,r)}\,dt.
\end{aligned}$$

We now intend to bound $\varphi_\nu(t,r)$ uniformly in order to obtain the result.

$$\begin{aligned}
\varphi(t,r) &\leq \int_{\mathscr{C}}\left\{\mathscr{E}_\gamma\left[\Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)\left(\int_0^t G_s^\nu(r)dB_s^\nu(x,r)\right)\right]^2\right. \\
&\left. + \left(\int_0^t \widetilde{K}_{\nu,r}^t(t,s)\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(r)\big)G_s^\nu(r)L_s^\nu(x,r)\Big]ds\right)^2\right\}dQ_{\nu,r}(x) \\
&\overset{\text{C.S.},(4.23)}{\leq} C_T\int_{\mathscr{C}}\left\{\widetilde{K}_{\nu,r}^t(t,t)\mathscr{E}_\gamma\left[\left(\int_0^t G_s^\nu(r)dB_s^\nu(x,r)\right)^2\right]\right. \\
&\left. + \int_0^t \widetilde{K}_{\nu,r}^t(t,s)^2\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(r)\big)G_s^\nu(r)L_s^\nu(x,r)\Big]^2 ds\right\}dQ_{\nu,r}(x) \\
&\overset{\text{Fubini},(4.23)}{\leq} C_T\left\{\mathscr{E}_\gamma\left[\int_{\mathscr{C}}\int_0^t G_s^\nu(r)^2 ds dQ_{\nu,r}(x)\right] + \int_0^t \varphi(s,r)ds\right\} \overset{\text{Fubini},(4.23)}{\leq} C_T\left\{1 + \int_0^t \varphi(s,r)ds\right\},
\end{aligned}$$

where we have used Itô isometry, and where C_T is uniform in space. Relying on Gronwall's lemma, we find that $\varphi(t,r)$ is uniformly bounded in space:

$$\sup_{0\le t\le T}\varphi_\nu(t,r)\le C_T\exp C_T.$$

This implies that exists a finite constant $\tilde{C}_T$, uniform in space, such that $I(Q_{\nu,r}|\bar{Q}_{\nu,r})\le \tilde{C}_T$.

Moreover, $I(Q_{\nu,r}|\bar{Q}_{\nu,r})$ is positive and $\pi-$measurable. We can thus integrate on D to find:

$$I(Q_\nu|\bar{Q}_\nu)=\int_D I(Q_{\nu,r}|\bar{Q}_{\nu,r})d\pi(r)\le \tilde{C}_T<\infty.$$

Remark that the proof of Proposition 4.3.2 readily applies to show that exists $0<\iota<1$ and $e>0$ such that

$$\int_{\mathscr{C}\times D}\log\Big(\frac{\mathrm{d}Q_\nu}{\mathrm{d}\bar{Q}_\nu}(x,r)\Big)d\mu(x,r)\le \iota I(\mu|\bar{Q}_\nu)+e.$$

In particular, $\bar{H}_\nu$ is finite whenever $I(\cdot|\bar{Q}_\nu)$ is. Moreover, we can directly apply [22, Appendix B], to obtain:

$$\forall\mu\in\mathscr{M}_1^+(\mathscr{C}\times D),\ \bar{H}_\nu(\mu)=I(\mu|Q_\nu).$$

We now show that equation (4.33) holds. If both $I(\mu|P)$ and $I(\mu|\bar{Q}_\nu)$ are infinite the results is clear. We handle the other cases by remarking that $\{I(\cdot|P)<+\infty\}=\{I(\cdot|\bar{Q}_\nu)<+\infty\}$. In fact, let $\mu\ll P\simeq\bar{Q}_\mu$, suppose that $I(\mu|\bar{Q}_\nu)$ is finite, and observe that

$$I(\mu|P)=\int_{\mathscr{C}\times D}\Big\{\log\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\bar{Q}_\nu}(x,r)\Big)+\log\Big(\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}(x,r)\Big)\Big\}d\mu(x,r).$$

On the one hand, $\log\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\bar{Q}_\nu}\Big)$ is μ-integrable as

$$\int_{\mathscr{C}\times D}\Big|\log\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\bar{Q}_\nu}(x,r)\Big)\Big|d\mu(x,r)=\int_{\mathscr{C}\times D}\Big|\log\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\bar{Q}_\nu}(x,r)\Big)\frac{\mathrm{d}\mu}{\mathrm{d}\bar{Q}_\nu}(x,r)\Big|d\bar{Q}_\nu(x,r),$$

as $\forall x\in\mathbb{R}_+$, $x\log(x)\ge-\frac{1}{\exp(1)}$, and as $I(\mu|\bar{Q}_\nu)<+\infty$. Let us show that $\log\Big(\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}\Big)$ is also μ-integrable. In fact,

$$\begin{aligned}\int_{\mathscr{C}\times D}\log\Big(\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}(x,r)\Big)^-d\mu(x,r)&=\int_{\mathscr{C}\times D}\log\Big(1\vee\frac{\mathrm{d}P}{\mathrm{d}\bar{Q}_\nu}(x,r)\Big)d\mu(x,r)\\&\overset{(4.13)}{\le} I(\mu|\bar{Q}_\nu)+\log\left(1+\int_{\mathscr{C}\times D}\frac{\mathrm{d}P}{\mathrm{d}\bar{Q}_\nu}(x,r)d\bar{Q}_\nu(x,r)\right)\le I(\mu|\bar{Q}_\nu)+\log(2)<+\infty,\end{aligned}$$

whereas

$$\begin{aligned}&\int_{\mathscr{C}\times D}\log\Big(\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}(x,r)\Big)^+d\mu(x,r)\overset{(4.13)}{\le} I(\mu|\bar{Q}_\nu)+\log\left(1+\int_{\mathscr{C}\times D}\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}(x,r)d\bar{Q}_\nu\right)\\&\le I(\mu|\bar{Q}_\nu)+\log\left(1+\int_D\int_{\mathscr{C}}\exp\left\{\int_0^T m_\nu(t,r)dV_t^\nu(x,r)-\frac{1}{2}\int_0^T m_\nu(t,r)^2dt\right\}d\bar{Q}_{\nu,r}(x)e^{\int_0^T m_\nu(t,r)^2dt}d\pi(r)\right)\\&\le I(\mu|\bar{Q}_\nu)+\log\left(1+\exp\Big\{\frac{T\|J\|_\infty^2}{2\lambda_*^2}\Big\}\right)<+\infty.\end{aligned}$$

Hence, $I(\mu|\bar{Q}_\nu) < +\infty$ implies that

$$I(\mu|P) = I(\mu|\bar{Q}_\nu) + \int_{\mathscr{C}\times D} \log\Big(\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}(x,r)\Big) d\mu(x,r) < \infty,$$

and by symmetry, $I(\mu|P) < \infty$ implies finiteness of $I(\mu|\bar{Q}_\nu)$ with same equality. Moreover, we can apply a similar reasoning as in the proof of Proposition 4.3.2 to show that exists constants $0 < \iota < 1$ and $e > 0$ such that

$$\Big| \int_{\mathscr{C}\times D} \log\Big(\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}(x,r)\Big) d\mu(x,r) \Big| \leq \iota I(\mu|\bar{Q}_\nu) + e,$$

and

$$|\Gamma_\nu(\mu)| \leq \iota I(\mu|P) + e.$$

Hence, for $\mu \in \{I(\cdot|P) < +\infty\} = \{I(\cdot|\bar{Q}_\nu) < +\infty\}$ these quantities are finite. Moreover:

$$\Gamma_\nu(\mu) = \int_{\mathscr{C}\times D} \left\{ \log\Big(\frac{\mathrm{d}Q_\nu}{\mathrm{d}\bar{Q}_\nu}(x,r)\Big) + \log\Big(\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}(x,r)\Big) \right\} d\mu(x,r)$$

and we can split this integral as $\frac{\mathrm{d}Q_\nu}{\mathrm{d}\bar{Q}_\nu}$ is bounded away from 0 and $\log\Big(\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}\Big)$ is μ-integrable to obtain:

$$\Gamma_\nu(\mu) = \int_{\mathscr{C}\times D} \log\Big(\frac{\mathrm{d}Q_\nu}{\mathrm{d}\bar{Q}_\nu}(x,r)\Big) d\mu(x,r) + \int_{\mathscr{C}\times D} \log\Big(\frac{\mathrm{d}\bar{Q}_\nu}{\mathrm{d}P}(x,r)\Big) d\mu(x,r).$$

As a consequence, for any $\mu \in \{I(\cdot|P) < +\infty\}$, we have

$$I(\mu|P) - \Gamma_\nu(\mu) = I(\mu|\bar{Q}_\nu) - \int_{\mathscr{C}\times D} \log\Big(\frac{\mathrm{d}Q_\nu}{\mathrm{d}\bar{Q}_\nu}(x,r)\Big) d\mu(x,r),$$

which concludes the proof. □

We have proved that H_ν is a good rate function, and would like to extend this property to H: H_ν is seen in our proof as an intermediate quantity, equal to H when $I(\mu|P) = \infty$, but differing from it of $\Gamma - \Gamma_\nu$ otherwise. We control this difference below in Lemma 4.3.8.

Let us introduce two preliminary objects that will appear in the obtained upper-bound. First, because of spatial extension, it is useful to introduce a proper distance on $\mathscr{C} \times D$:

Definition 4.3.1. The map

$$d_T : \begin{cases} (\mathscr{C}\times D)^2 & \to \mathbb{R}_+ \\ \big((x,r),(y,r')\big) & \to \left\{ \|r-r'\|^2_{\mathbb{R}^d} + \sup\limits_{\substack{a,b\in[-\tau,0],t\in[0,T] \\ |b-a|\leq K\tau\|r-r'\|_{\mathbb{R}^d}}} \big|x_{t+a} - y_{t+b}\big|^2 \right\}^{\frac{1}{2}}, \end{cases} \tag{4.34}$$

is a distance on $\mathscr{C} \times D$. Moreover,

$$d_T\big((x^n,r_n),(x,r)\big) \to 0 \iff \|x - x^n\|_{\infty,T} + \|r - r_n\|_{\mathbb{R}^d} \to 0,$$

and $\big(\mathscr{C}\times D, d_T\big)$ is complete.

Remark 8. In particular, d_T generates the natural Borel σ-field of $\mathscr{C}\times D$, and $\big(\mathscr{C}\times D, d_T\big)$ is a Polish space.

Proof. Symmetry and separation are easy to obtain. The triangular inequality is a consequence of the two following facts. First, for any $(x,r),(y,r'),(z,\tilde{r})$, we have

$$\begin{aligned}\sup_{\substack{a,b\in[-\tau,0],t\in[0,T]\\ |b-a|\le K_\tau\|r-r'\|_{\mathbb{R}^d}}}\big|x_{t+a}-y_{t+b}\big| &\le \sup_{\substack{a,b,c\in[-\tau,0],t\in[0,T]\\ |c-a|\le K_\tau\|r-\tilde{r}\|_{\mathbb{R}^d},|b-c|\le K_\tau\|\tilde{r}-r'\|_{\mathbb{R}^d}}}\big|x_{t+a}-y_{t+b}\big|\\ &\le \sup_{\substack{a,c\in[-\tau,0],t\in[0,T]\\ |c-a|\le K_\tau\|r-\tilde{r}\|_{\mathbb{R}^d}}}\big|x_{t+a}-z_{t+c}\big| + \sup_{\substack{c,b\in[-\tau,0],t\in[0,T]\\ |b-c|\le K_\tau\|\tilde{r}-r'\|_{\mathbb{R}^d}}}\big|z_{t+c}-y_{t+b}\big|,\end{aligned}$$

as the set

$$\Big\{a,b\in[-\tau,0],|b-a|\le K_\tau\|r-r'\|_{\mathbb{R}^d}\Big\}$$

is contained in

$$\Big\{a,b\in[-\tau,0],\exists c\in[-\tau,0],|c-a|\le K_\tau\|r-\tilde{r}\|_{\mathbb{R}^d},|b-c|\le K_\tau\|\tilde{r}-r'\|_{\mathbb{R}^d}\Big\}.$$

Second, the triangular inequality of $\mathbb{R}^2$ for the Euclidean norm gives $\forall a_1,b_1,a_2,b_2\in\mathbb{R}$,

$$\Big\{(a_1+b_1)^2+(a_2+b_2)^2\Big\}^{\frac{1}{2}}\le\big\{a_1^2+a_2^2\big\}^{\frac{1}{2}}+\big\{b_1^2+b_2^2\big\}^{\frac{1}{2}}.$$

Hence, d_T is a distance on $\mathscr{C}\times D$. Let $(x^n,r_n)_{n\in\mathbb{N}}\in\big(\mathscr{C}\times D\big)^{\mathbb{N}}$, and $(x,r)\in\mathscr{C}\times D$. Taking $a=b$ in the supremum, we see that

$$\|r-r'\|_{\mathbb{R}^d}+\|x-y\|_{\infty,T}\le d_T\big((x,r),(y,r')\big), \tag{4.35}$$

we have that $d_T\big((x^n,r_n),(x,r)\big)\to 0$ implies $r_n\to r$, and $x_n\to x$ for the supremum-norm on $[-\tau,T]$. Conversely, suppose that $r_n\to r$, and $x_n\to_{\|\cdot\|_{\infty,T}} x$, and let $\eta>0$ such that

$$s,t\in[-\tau,T],|s-t|\le\eta \quad\Longrightarrow\quad |x_t-x_s|\le\varepsilon.$$

We then see that, we can find a n_0 such that $\forall n\ge n_0$,

$$d_T\big((x^n,r_n),(x,r)\big)^2\le\|r-r_n\|^2_{\mathbb{R}^d}+2\|x-x^n\|^2_{\infty,T}+2\sup_{\substack{a,b\in[-\tau,0],t\in[0,T]\\ |b-a|\le\eta}}\big|x_{t+a}-x_{t+b}\big|^2\le 3\varepsilon^2.$$

The completion of $\big(\mathscr{C}\times D,d_T\big)$ comes from that of $\big(\mathscr{C},\|\cdot\|_{\infty,T}\big)$, and $\big(D,|\cdot|\big)$, and from (4.35). □

We also define the 2-Vaserstein distance on $\mathscr{M}_1^+(\mathscr{C}\times D)$, associated with d_T:

$$d_T^V(\mu,\nu):=\inf_{\xi}\left\{\int_{(\mathscr{C}\times D)^2}d_T\big((x,r),(y,r')\big)^2d\xi\big((x,r),(y,r')\big)\right\}^{\frac{1}{2}}$$

the infimum being taken on the laws $\xi\in\mathscr{C}_{\mu,\nu}$. In the following, we will, for any $t\in[0,T]$, denote by d_t and d_t^V the respective restrictions of d_T and d_T^V on $\big(\mathscr{C}\big([-\tau,t],\mathbb{R}\big)\times D\big)^2$.

Here are a few regularity properties of the covariance and mean of our Gaussian interactions:

Proposition 4.3.7. Exists $C_T > 0$ such that for any $\mu, \nu \in \mathscr{M}_1^+(\mathscr{C} \times D)$, $r \in D$, $t \in [0,T]$ and $u, s \in [0,t]$:

$$\big|m_\mu(t,r) - m_\nu(t,r)\big| + \big|K_\mu(t,s,r) - K_\nu(t,s,r)\big| + \big|\widetilde{K}^t_{\mu,r}(s,u) - \widetilde{K}^t_{\nu,r}(s,u)\big| \leq C_T d_T^V(\mu,\nu). \quad (4.36)$$

Proof. Let $\xi \in \mathscr{C}_{\mu,\nu}$, and let (G,G') be, under γ, a family of independent bi-dimensional centered Gaussian processes with covariance $K_\xi(s,t,r)$ given by:

$$\frac{1}{\lambda(r)^2}\int_{(\mathscr{C}\times D)^2}\begin{pmatrix} \sigma^2_{rr'}S(x_{s-\tau_{rr'}})S(x_{t-\tau_{rr'}}) & \sigma_{rr'}\sigma_{r\tilde{r}'}S(x_{s-\tau_{rr'}})S(y_{t-\tau_{r\tilde{r}'}}) \\ \sigma_{rr'}\sigma_{r\tilde{r}'}S(y_{s-\tau_{r\tilde{r}'}})S(x_{t-\tau_{rr'}}) & \sigma^2_{r\tilde{r}'}S(y_{s-\tau_{r\tilde{r}'}})S(y_{t-\tau_{r\tilde{r}'}}) \end{pmatrix} d\xi\big((x,r'),(y,\tilde{r}')\big). \quad (4.37)$$

with the short-hand notations $\sigma_{rr'} := \sigma(r,r')$, $\tau_{rr'} = \tau(r,r')$. Let us first take care of the mean difference:

$$\begin{aligned}
\big|m_\mu(t,r) - m_\nu(t,r)\big| &= \left|\frac{1}{\lambda(r)}\int_{\mathscr{C}\times D} J(r,r')S(x_{t-\tau(r,r')})d(\mu-\nu)(x,r')\right| \\
&\leq \frac{1}{\lambda_*}\int_{(\mathscr{C}\times D)^2}\Big|J(r,r')S(x_{t-\tau(r,r')}) - J(r,\tilde{r}')S(y_{t-\tau(r,\tilde{r}')})\Big| d\xi\big((x,r'),(y,\tilde{r}')\big) \\
&\leq \frac{1}{\lambda_*}\int_{(\mathscr{C}\times D)^2}\Big\{K_J\|r'-\tilde{r}'\|_{\mathbb{R}^d} + \|J\|_\infty\big|S\big(x_{t-\tau(r,r')}\big) - S\big(y_{t-\tau(r,\tilde{r}')}\big)\big|\Big\} d\xi\big((x,r'),(y,\tilde{r}')\big) \\
&\leq \frac{1}{\lambda_*}\int_{(\mathscr{C}\times D)^2}\Big\{K_J\|r'-\tilde{r}'\|_{\mathbb{R}^d} + \|J\|_\infty K_S \sup_{\substack{a,b\in[-\tau,0],t\in[0,T]\\ |b-a|\leq K_\tau\|r'-\tilde{r}'\|_{\mathbb{R}^d}}}\big|x_{t-a}-y_{t-b}\big|\Big\} d\xi\big((x,r'),(y,\tilde{r}')\big) \\
&\overset{\text{C.S.}}{\leq} C_T\left\{\int_{(\mathscr{C}\times D)^2} d_T\big((x,r'),(y,\tilde{r}')\big)^2 d\xi\big((x,r'),(y,\tilde{r}')\big)\right\}^{\frac{1}{2}}.
\end{aligned}$$

Moreover,

$$\big|K_\mu(t,s,r) - K_\nu(t,s,r)\big| = \Big|\mathscr{E}_\gamma\Big[G_sG_t - G'_sG'_t\Big]\Big| \leq C_T\left\{\mathscr{E}_\gamma\Big[\big(G_t-G'_t\big)^2\Big]^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_s-G'_s\big)^2\Big]^{\frac{1}{2}}\right\}.$$

and

$$\begin{aligned}
\big|\widetilde{K}^t_{\mu,r}(s,u) - \widetilde{K}^t_{\nu,r}(s,u)\big| &\overset{(4.23)}{\leq} C_T\left\{\mathscr{E}_\gamma\Big[\big(\Lambda_t(G)-\Lambda_t(G')\big)^2\Big]^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_s-G'_s\big)^2\Big]^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_u-G'_u\big)^2\Big]^{\frac{1}{2}}\right\} \\
&\overset{(4.25)}{\leq} C_T\left\{\left(\int_0^t \mathscr{E}_\gamma\Big[\big(G_v-G'_v\big)^2\Big]dv\right)^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_s-G'_s\big)^2\Big]^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_u-G'_u\big)^2\Big]^{\frac{1}{2}}\right\}.
\end{aligned}$$

$$\mathscr{E}_\gamma\Big[\big(G_t-G'_t\big)^2\Big] = \frac{1}{\lambda(r)^2}\int_{(\mathscr{C}\times D)^2}\Big(\sigma(r,r')S(x_{t-\tau(r,r')}) - \sigma(r,\tilde{r}')S(y_{t-\tau(r,\tilde{r}')})\Big)^2 d\xi\big((x,r'),(y,\tilde{r}')\big).$$

Splitting the integrand of the right-hand side, we find:

$$\begin{aligned}
\Big(\sigma_{rr'}S(x_{t-\tau_{rr'}}) - \sigma_{r\tilde{r}'}S(y_{t-\tau_{r\tilde{r}'}})\Big)^2 &\leq 2\Big\{\big(\sigma_{rr'}-\sigma_{r\tilde{r}'}\big)^2 S(x_{t-\tau_{rr'}})^2 + \sigma^2_{r\tilde{r}'}\big(S(x_{t-\tau_{rr'}}) - S(y_{t-\tau_{r\tilde{r}'}})\big)^2\Big\} \\
&\leq 2K_\sigma^2\|r'-\tilde{r}'\|^2_{\mathbb{R}^d} + 2\|\sigma\|^2_\infty K_S^2 \sup_{\substack{a,b\in[-\tau,0]\\ |b-a|\leq K_\tau\|r'-\tilde{r}'\|_{\mathbb{R}^d}}}|x_{t+a}-y_{t+b}|^2 \leq C d_t\big((x,r'),(y,\tilde{r}')\big)^2,
\end{aligned}$$

so that

$$\mathscr{E}_\gamma\Big[\big(G_t-G'_t\big)^2\Big] \leq C_T\int_{(\mathscr{C}\times D)^2} d_t\big((x,r'),(y,\tilde{r}')\big)^2 d\xi\big((x,r'),(y,\tilde{r}')\big).$$

Taking the infimum over $\xi \in \mathscr{C}_{\mu,\nu}$ yields (4.36).

□

We have now introduced all the needed elements to state the following technical lemma concluding on the fact that H is a good rate function.

Lemma 4.3.8. Let $\mu,\nu\in\mathscr{M}_1^+\big(\mathscr{C}\times D\big)$, then:

(i) there exists a positive constant C_T such that:

(a) $|\Gamma_{1,\nu}(\mu)-\Gamma_1(\mu)|\leq C_T d_T^V(\mu,\nu)$.

(b) $|\Gamma_{2,\nu}(\mu)-\Gamma_2(\mu)|\leq C_T\big(1+I(\mu|P)\big)d_T^V(\mu,\nu)$.

(ii) H is a good rate function.

Proof. The main techniques were introduced in [22, Lemma 3.3-3.4] and used in a neuroscience setting in [51, Lemma.5], but spatiality induces new issues essentially impacting the proof of point (1.b). For the sake of completeness, we will here reproduce these techniques, and address the specific spatial difficulties.

Proof of Lemma 4.3.8.(1.a).

We define

$$\Gamma_1(\mu,r):=-\frac{1}{2}\int_0^T\Big(\widetilde{K}^t_{\mu,r}(t,t)+m_\mu(t,r)^2\Big)dt,$$

so that

$$\begin{aligned}&\big|\Gamma_{1,\nu}(\mu)-\Gamma_1(\mu)\big|=\left|\int_{\mathscr{C}\times D}\big(\Gamma_1(\nu,r)-\Gamma_1(\mu,r)\big)d\mu(x,r)\right|\\&\leq\frac{1}{2}\int_0^T\Big|\big(m_\mu(t,r)-m_\nu(t,r)\big)\big(m_\mu(t,r)+m_\nu(t,r)\big)\Big|+\Big|\widetilde{K}^t_{\nu,r}(t,t)-\widetilde{K}^t_{\mu,r}(t,t)\Big|dt\overset{(4.36)}{\leq}C_T d_T^V(\mu,\nu).\end{aligned}\tag{4.38}$$

Proof of Lemma 4.3.8.(1.b)

Note that if $I(\mu|P)=\infty$ the inequality is obvious. Let then $\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)$ with $I(\mu|P)<\infty$ implying $\mu\ll P$ and finiteness of $\Gamma_\nu(\mu)$, and $\Gamma(\mu)$. This also implies that μ has a measurable density ρ_μ with respect to $\mathscr{B}\big(\mathscr{C}\times D\big)$:

$$d\mu(x,r)=\rho_\mu(x,r)dP(x,r)=\rho_\mu(x,r)dP_r(x)d\pi(r).$$

Hence, for $r\in D$ such that $c_\mu(r):=\int_{\mathscr{C}}\rho_\mu(x,r)dP_r(x)\notin\{0,+\infty\}$, we can properly define $\mu_r\in\mathscr{M}_1^+(\mathscr{C})$ by $d\mu_r(x):=\frac{\rho_\mu(x,r)}{c_\mu(r)}dP_r(x)$. Of course $\mu_r\ll P_r$, and

$$d\mu(x,r)=d\mu_r(x)c_\mu(r)d\pi(r).\tag{4.39}$$

Remark that c_μ is a measurable function of space such that $\int_D c_\mu(r)d\pi(r)=1$, and that the set $\{r\in D,\ c_\mu(r)\in\{0,+\infty\}\}$ do not impact the value of integrals over μ.

In order to obtain the proper inequality, we split the difference of interest into different terms:

$$\begin{aligned}&\big|\Gamma_{2,\nu}(\mu)-\Gamma_2(\mu)\big|\leq\frac{1}{2}\left|\int_{\mathscr{C}\times D}\left\{\int_{\hat{\Omega}}L_T^\nu(x,r)^2d\gamma_{\widetilde{K}^T_{\nu,r}}-\int_{\hat{\Omega}}L_T^\mu(x,r)^2d\gamma_{\widetilde{K}^T_{\mu,r}}\right\}d\mu(x,r)\right|\\&+\left|\int_{\mathscr{C}\times D}\int_0^T(m_\nu-m_\mu)(t,r)dW_t(x,r)d\mu(x,r)\right|\leq\frac{1}{2}\left|\int_{\mathscr{C}\times D}\int_{\hat{\Omega}}L_T^\nu(x,r)^2d\big(\gamma_{\widetilde{K}^T_{\nu,r}}-\gamma_{\widetilde{K}^T_{\mu,r}}\big)d\mu(x,r)\right|\\&+\frac{1}{2}\left|\int_{\mathscr{C}\times D}\int_{\hat{\Omega}}\Big\{L_T^\mu(x,r)^2-L_T^\nu(x,r)^2\Big\}d\gamma_{\widetilde{K}^T_{\mu,r}}d\mu(x,r)\right|+\left|\int_{\mathscr{C}\times D}\int_0^T(m_\nu-m_\mu)(t,r)dW_t(x,r)d\mu(x,r)\right|.\end{aligned}$$

Let $\xi \in \mathscr{C}_{\mu,\nu}$ be such that

$$\int_{(\mathscr{C}\times D)^2} d_T\big((x,r'),(y,\tilde{r}')\big)^2 d\xi\big((x,r'),(y,\tilde{r}')\big) \leq \big(d_T^V(\mu,\nu)+\varepsilon\big)^2.$$

Moreover, let $\big(G(r),G'(r)\big)_{r\in D}$ be, under γ, a family of independent bi-dimensional centered Gaussian processes with covariance $K_\xi(s,t,r)$ as define in (4.37). Remark that for any map f

$$\mathscr{E}_\gamma\Big[f\big(G^\mu(r)\big)\Big] - \mathscr{E}_\gamma\Big[f\big(G^\nu(r)\big)\Big] = \mathscr{E}_\gamma\Big[f\big(G(r)\big) - f\big(G'(r)\big)\Big]$$

and as proved in Proposition 4.3.7,

$$\mathscr{E}_\gamma\Big[\big(G_t(r)-G'_t(r)\big)^2\Big] \leq C_T\big(d_T^V(\mu,\nu)+\varepsilon\big)^2.$$

Let also

$$L_t(x,r) := \int_0^t G_s(r)dV_s^\mu(x,r),\ \ L'_t(x,r) := \int_0^t G'_s(r)dV_s^\nu(x,r).$$

Using inequality (4.23), we then obtain:

$$\begin{aligned}
|\Gamma_{2,\nu}(\mu)-\Gamma_2(\mu)| \overset{\text{C.S.}}{\leq} C_T\Bigg\{ & \overbrace{\int_{\mathscr{C}\times D} \mathscr{E}_\gamma\Big[\big|\Lambda_T(G(r))-\Lambda_T(G'(r))\big|L'_T(x,r)^2\Big]d\mu(x,r)}^{B_1:=} \\
& + \underbrace{\prod_{\varepsilon=\pm 1}\left(\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\left[\left(\int_0^T (G_t(r)+\varepsilon G'_t)dV_t^\mu(x,r)\right)^2\right]d\mu(x,r)\right)^{\frac{1}{2}}}_{=:B_2} \\
& + \underbrace{\left|\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\left[\left(\int_0^T G_t(r)dV_t^\mu(x,r)\right)^2 - \left(\int_0^T G_t(r)dV_t^\nu(x,r)\right)^2\right]d\mu(x,r)\right|}_{=:B_3} \\
& + \underbrace{\left(\int_{\mathscr{C}\times D}\left|\int_0^T (m_\nu-m_\mu)(t,r)dW_t(x,r)\right|^2 d\mu(x,r)\right)^{\frac{1}{2}}}_{=:B_4}\Bigg\}. \qquad (4.40)
\end{aligned}$$

Before bounding these four terms, we prove a useful inequality. For any $h,m \in L^2([0;T],dt)$, with m bounded, and any $r\in D$ with $c_\mu(r)\notin\{0,+\infty\}$,

$$\int_{\mathscr{C}}\Big(\int_0^T h_t(dW_t(x,r)-m(t)dt)\Big)^2 d\mu_r(x) \leq 2\Big\{\int_{\mathscr{C}}\Big(\int_0^T h_t dW_t(x,r)\Big)^2 + \Big(\int_0^T h_t m_t dt\Big)^2 d\mu_r(x)\Big\}. \qquad (4.41)$$

Moreover, supposing that $h\neq 0_{L^2([0;T],dt)}$, then $\Phi_h(x) = \dfrac{\Big(\int_0^T h_t dW_t(x,r)\Big)^2}{4\Big(\int_0^T h_t^2 dt\Big)}$ is a well-defined, positive and measurable function of the σ-algebra $\mathscr{B}(\mathscr{C})$, so that resorting to (4.13) one obtains

$$\int_{\mathscr{C}}\Phi_h(x)d\mu_r(x) \leq I(\mu_r|P_r) + \log\int_{\mathscr{C}}\exp\Phi_h(x)dP_r(x).$$

As $W(.,r)$ is a Brownian motion under P_r, $\Phi_h \sim \mathscr{N}\big(0,\frac{1}{4}\big)^2$, so that Gaussian calculus gives, for any $C>2$:

$$\int_{\mathscr{C}}\Big(\int_0^T h_t dW_t(x,r)\Big)^2 d\mu_r(x) \leq C\big(I(\mu_r|P_r)+1\big)\Big(\int_0^T h_t^2 dt\Big)$$

Remark that this inequality obviously holds when $h = 0_{L^2([0;T],dt)}$. Applying this result in (4.41) one eventually finds:

$$\int_{\mathscr{C}} \Big(\int_0^T h_t (dW_t(x,r) - m(t)dt) \Big)^2 d\mu_r(x) \overset{\text{C.S.}}{\leq} 2\Big(C\big(1 + I(\mu_r|P_r)\big) + m_\infty^2 T \Big) \Big(\int_0^T h_t^2 dt \Big)$$
$$\leq C_T \big(1 + I(\mu_r|P_r)\big) \Big(\int_0^T h_t^2 dt \Big). \tag{4.42}$$

With this result in mind, we now control the first term. Recall that, by (4.25),

$$\big|\Lambda_T(G(r)) - \Lambda_T(G'(r))\big| \leq C_T \bigg(\int_0^T \big|G_t(r)^2 - {G'}_t(r)^2\big| dt + \int_0^T \mathscr{E}_\gamma \Big[\big(G_t(r) - {G'}_t(r)\big)^2 \Big]^{\frac{1}{2}} dt \bigg).$$

Now, relying on the decomposition of μ, we find

$$\begin{aligned}
B_1 &\overset{\text{Fubini}}{=} \int_D \mathscr{E}_\gamma \bigg[\Big|\Lambda_T(G(r)) - \Lambda_T(G'(r))\Big| \Big\{ \int_{\mathscr{C}} L'_T(x,r)^2 d\mu_r(x) \Big\} \bigg] c_\mu(r) d\pi(r) \\
&\overset{(4.42)}{\leq} \int_D C_T \big(I(\mu_r|P_r) + 1\big) \mathscr{E}_\gamma \bigg[\Big|\Lambda_T(G(r)) - \Lambda_T(G'(r))\Big| \Big\{ \int_0^T G'_t(r)^2 dt \Big\} \bigg] c_\mu(r) d\pi(r) \\
&\leq C_T \int_D \bigg(\int_0^T \int_0^T \mathscr{E}_\gamma \Big[\big|G_s(r)^2 - {G'}_s(r)^2\big| G'_t(r)^2 \Big] \big(I(\mu_r|P_r) + 1\big) ds dt + d_T^V(\mu,\nu) + \varepsilon \bigg) c_\mu(r) d\pi(r) \\
&\leq C_T \bigg(\int_D I(\mu_r|P_r) c_\mu(r) d\pi(r) + 1 \bigg) \big(d_T^V(\mu,\nu) + \varepsilon \big).
\end{aligned}$$

where the last inequality is a consequence of Cauchy-Schwarz's inequality, and Isserlis' theorem 8.2.4. Observe that:

$$\begin{aligned}
\int_D I(\mu_r|P_r) c_\mu(r) d\pi(r) &= \int_D \int_{\mathscr{C}} \log\Big(\frac{\rho_\mu(x,r)}{c_\mu(r)} \Big) d\mu_r(x) c_\mu(r) d\pi(r) \\
&= \int_{\mathscr{C}\times D} \log(\rho_\mu(x,r)) d\mu(x,r) - \int_D \log(c_\mu(r)) c_\mu(r) d\pi(r) \overset{\text{Jensen}}{\leq} I(\mu|P).
\end{aligned}$$

As a consequence,

$$B_1 \leq C_T \big(1 + I(\mu|P)\big) \big(d_T^V(\mu,\nu) + \varepsilon \big).$$

Similarly, there exists a constant c_T such that

$$\begin{aligned}
B_2 &\leq \prod_{\varepsilon=\pm 1} \bigg(\int_D c_T \big(1 + I(\mu_r|P_r)\big) \mathscr{E}_\gamma \Big[\int_0^T \big(G_t(r) + \varepsilon G'_t(r)\big)^2 dt \Big] c_\mu(r) d\pi(r) \bigg)^{\frac{1}{2}} \\
&\leq C_T \big(1 + I(\mu|P)\big)^{\frac{1}{2}} \bigg(\int_D \big(1 + I(\mu_r|P_r)\big) \int_0^T \mathscr{E}_\gamma \Big[\big(G_t(r) - G'_t(r)\big)^2 \Big] dt c_\mu(r) d\pi(r) \bigg)^{\frac{1}{2}} \\
&\leq C_T \big(1 + I(\mu|P)\big) \big(d_T^V(\mu,\nu) + \varepsilon \big).
\end{aligned}$$

To bound B_3, we first use Cauchy-Schwarz inequality:

$$B_3 \leq \prod_{\varepsilon=\pm 1} \bigg\{ \int_{\mathscr{C}\times D} \mathscr{E}_\gamma \bigg[\Big| \int_0^T G_t(r) \Big((1+\varepsilon) dW_t(x,r) - (m_\nu(t,r) + \varepsilon m_\mu(t,r)) dt \Big) \Big|^2 \bigg] d\mu(x,r) \bigg\}^{\frac{1}{2}}. \tag{4.43}$$

Then, again by Cauchy-Schwarz's inequality, one observes that

$$\mathscr{E}_\gamma\Big[\Big|\int_0^T G_t(r)\big(m_\mu(t,r)-m_\nu(t,r)\big)dt\Big|^2\Big] \overset{(4.36)}{\leq} C_T d_T^V(\mu,\nu)^2.$$

Moreover, (4.42) gives:

$$\int_{\mathscr{C}}\Big\{\int_0^T 2G_t(r)\Big(dW_t(x,r)-\frac{m_\mu(t,r)+m_\nu(t,r)}{2}dt\Big)\Big\}^2 d\mu_r(x) \leq c_T\big(1+I(\mu_r|P_r)\big)\int_0^T G_t^2(r)dt.$$

Using Jensen's inequality and injecting the last two inequalities in (4.43) gives:

$$B_3 \leq C_T\big(1+I(\mu|P)\big)^{\frac{1}{2}} d_T^V(\mu,\nu) \leq C_T\big(1+I(\mu|P)\big)d_T^V(\mu,\nu)$$

as $I(.|P)\geq 0$.

As of the last term, we have

$$B_4 \overset{(4.42)}{\leq} \left(\int_D c_T\big(1+I(\mu_r|P_r)\big)\Big\{\int_0^T \Big(m_\mu(t,r)-m_\nu(t,r)\Big)^2 dt\Big\}c_\mu(r)d\pi(r)\right)^{\frac{1}{2}}$$
$$\overset{(4.36)}{\leq} C_T\big(1+I(\mu|P)\big)^{\frac{1}{2}} d_T^V(\mu,\nu) \leq C_T\big(1+I(\mu|P)\big)d_T^V(\mu,\nu).$$

Hence, we conclude that exists a constant C_T satisfying

$$|\Gamma_{2,\nu}(\mu)-\Gamma_2(\mu)| \leq C_T\big(1+I(\mu|P)\big)\big(d_T^V(\mu,\nu)+\varepsilon\big).$$

Sending ε to 0 thus gives the result.

Proof of Lemma4.3.8.(3): We proceed exactly as in Lemma 5.(vi) [51], remarking that $\big(\mathscr{C}\times D, d_T\big)$ is a Polish space with the same topology as the usual Borel one. In order to demonstrate that H is a good rate function, we need to show that it is lower semi-continuous and that it has compact level sets, i.e. $\{H\leq M\}$ is a compact set for any $M>0$.

- lower semi-continuity:
  Let $(\mu_p)_{p\in\mathbb{N}}\in\mathscr{M}_1^+\big(\mathscr{C}\times D\big)^{\mathbb{N}}$ be a sequence of probability measures weakly converging toward $\mu\in\mathscr{M}_1^+\big(\mathscr{C}\times D\big)$, and choose a subsequence $(\mu_{p_m})_m$ such that $\lim_{m\to\infty}H(\mu_{p_m})=\liminf_{p\to\infty}H(\mu_p)$. Suppose first that $I(\mu|P)$ is bounded for large m. we then have:

  $$\liminf_{p\to\infty}H(\mu_p)=\lim_{m\to\infty}(I-\Gamma)(\mu_{p_m})\geq\liminf_{m\to\infty}(I-\Gamma_\mu)(\mu_{p_m})+\liminf_{m\to\infty}(\Gamma_\mu-\Gamma)(\mu_{p_m}).$$

  As $I(\mu_{p_m}|P)$ is finite for big m, Theorem 4.3.6 implies:

  $$\liminf_{m\to\infty}\big(I(\mu_{p_m}|P)-\Gamma_\mu(\mu_{p_m})\big)=\liminf_{m\to\infty}H_\mu(\mu_{p_m})\geq H_\mu(\mu)=H(\mu).$$

  What's more, the first point of this lemma ensures the existence of a finite constant C_T such that,

  $$\big|\Gamma(\mu_{p_m})-\Gamma_\mu(\mu_{p_m})\big|\leq C_T(1+I(\mu_{p_m}|P))d_T^V(\mu,\mu_{p_m})$$

  so that, as $I(\mu_{p_m}|P)$ is bounded for large m, the above difference tends to zero. Hence, $\liminf_{p\to\infty}H(\mu_p)\geq H(\mu)$.
  Suppose now that $I(\mu_{p_m}|P)$ is not bounded for large m, i.e. we can find a subsequence $(p_{m(M)})_{M\in\mathbb{N}}$ such that $\lim_{M\to+\infty}I(\mu_{p_{m(M)}}|P)=+\infty$. Consequently, $H(\mu_{p_{m(M)}})\to_{M\to+\infty}+\infty$, so that

  $$\liminf_{p\to\infty}H(\mu_p)=\lim_{M\to+\infty}H(\mu_{p_{m(M)}})=+\infty\geq H(\mu).$$

  We proved that H is lower semi-continuous.

- compact subsets:
  As H is lower semi-continuous, $\{H \leq M\}$ is a closed set. Moreover, Lemma4.3.2.(ii) ensures that $I(.|P)$ is bounded on this set. As $\mathscr{C} \times D$ is a Polish space, the level sets of the relative entropy are compact for the weak topology (see Proposition 9.5.2 or [84, lemma 6.2.12]) so that $\{H \leq M\}$ is also a compact set.

We have proved that H is a good rate function. □

4.3.2 Upper-bound and Tightness

We have proved that $H = I(\cdot|P) - \Gamma$ is a good rate function, and we now want to show that it is associated with a LDP. We demonstrate here a weak LDP relying on an upper-bound inequality for compact subsets, and tightness of the family $\Big(Q^N(\hat{\mu}_N \in \cdot)\Big)_N$. To prove the first point, we take advantage of the full LDP followed by $\hat{\mu}_N$ under $Q_\nu^{\otimes N}$, and have to control an error term. The second point will rely on the exponential tightness of $P^{\otimes N}$. These proofs are inspired from those of Guionnet in a non-spatial spin-glass model [129].

Theorem 4.3.9. *For any compact subset K of $\mathscr{M}_1^+(\mathscr{C} \times D)$,*

$$\limsup_{N\to\infty} \frac{1}{N} \log Q^N(\hat{\mu}_N \in K) \leq -\inf_K H.$$

Proof. Let $\delta < 0$. as, $\big(\mathscr{M}_1^+(\mathscr{C} \times D), d_T^V\big)$ is a Polish space, we can find an integer M and a family $(\nu_i)_{1\leq i\leq M}$ of $\mathscr{M}_1^+(\mathscr{C} \times D)$ such that

$$K \subset \bigcup_{i=1}^{M} B(\nu_i, \delta),$$

where $B(\nu_i, \delta) := \big\{\mu | d_T^V(\mu, \nu_i) < \delta\big\}$. Lemma (9.2.4), ensures that

$$\limsup_{N\to\infty} \frac{1}{N} \log Q^N\big(\hat{\mu}_N \in K\big) \leq \max_{1\leq i\leq M} \limsup_{N\to\infty} \frac{1}{N} \log Q^N\big(\hat{\mu}_N \in K \cap B(\nu_i, \delta)\big). \tag{4.44}$$

Lemma 4.3.1 yields:

$$\begin{aligned} Q^N\big(\hat{\mu}_N \in K \cap B(\nu, \delta)\big) &= \int_{\hat{\mu}_N \in K\cap B(\nu,\delta)} \exp\Big\{N\bar{\Gamma}(\hat{\mu}_N)\Big\} dP^{\otimes N}(\mathbf{x}, \mathbf{r}) \\ &= \int_{\hat{\mu}_N \in K\cap B(\nu,\delta)} \exp\Big\{N\big(\bar{\Gamma}(\hat{\mu}_N) - \bar{\Gamma}_\nu(\hat{\mu}_N)\big)\Big\} \exp\Big\{N\bar{\Gamma}_\nu(\hat{\mu}_N)\Big\} dP^{\otimes N}(\mathbf{x}, \mathbf{r}). \end{aligned}$$

Recall definition (4.14) and let $(\tilde{X}_i^\mu)_{1\leq i\leq N}$ be a family of independent variables of $(\hat{\Omega}, \hat{\mathscr{F}}, \gamma)$ with same law as $(X_i^\mu)_{1\leq i\leq N}$. We will denote by $\big(\tilde{G}^\mu(r_i)\big)_{1\leq i\leq N}$ the associated independent Gaussian processes. Then, for any conjugate exponents (p, q),

$$\begin{aligned} &Q^N\big(\hat{\mu}_N \in K \cap B(\nu, \delta)\big) = \int_{\hat{\mu}_N \in K\cap B(\nu,\delta)} \exp\Big\{N\big(\bar{\Gamma}(\hat{\mu}_N) - \bar{\Gamma}_\nu(\hat{\mu}_N)\big)\Big\} dQ_\nu^{\otimes N}(\mathbf{x}, \mathbf{r}) \\ &\leq Q_\nu^{\otimes N}\big(\hat{\mu}_N \in K \cap B(\nu, \delta)\big)^{\frac{1}{p}} \left(\int_{\hat{\mu}_N \in K\cap B(\nu,\delta)} \exp\Big\{qN\big(\bar{\Gamma}(\hat{\mu}_N) - \bar{\Gamma}_\nu(\hat{\mu}_N)\big)\Big\} dQ_\nu^{\otimes N}(\mathbf{x}, \mathbf{r}) \right)^{\frac{1}{q}} \\ &\leq Q_\nu^{\otimes N}\big(\hat{\mu}_N \in K \cap B(\nu, \delta)\big)^{\frac{1}{p}} \left(\int_{\hat{\mu}_N \in K\cap B(\nu,\delta)} \prod_{i=1}^{N} \mathscr{E}_\gamma\left(\frac{\exp\big\{\tilde{X}_i^{\hat{\mu}_N}\big\}}{\mathscr{E}_\gamma\big(\exp\big\{\tilde{X}_i^{\nu}\big\}\big)} \right)^q dQ_\nu^{\otimes N}(\mathbf{x}, \mathbf{r}) \right)^{\frac{1}{q}} \end{aligned}$$

$$\overset{\text{Jensen}}{\leq} Q_\nu^{\otimes N}\big(\hat{\mu}_N \in K\cap B(\nu,\delta)\big)^{\frac{1}{p}}\left(\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\mathscr{E}_\gamma\left[\left(\prod_{i=1}^N \exp\left\{q\big(\tilde{X}_i^{\hat{\mu}_N}-\tilde{X}_i^\nu\big)\right\}\right)\frac{\prod_{i=1}^N\exp\tilde{X}_i^\nu}{\mathscr{E}_\gamma\Big[\prod_{i=1}^N\exp\tilde{X}_i^\nu\Big]}\right]dQ_\nu^{\otimes N}(\mathbf{x},\mathbf{r})\right)^{\frac{1}{q}}$$

$$\leq Q_\nu^{\otimes N}\big(\hat{\mu}_N \in K\cap B(\nu,\delta)\big)^{\frac{1}{p}}\Bigg(\underbrace{\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\mathscr{E}_\gamma\left[\prod_{i=1}^N\exp\left\{q\Big(\tilde{X}_i^{\hat{\mu}_N}-\tilde{X}_i^\nu\Big)+\tilde{X}_i^\nu\right\}\right]dP^{\otimes N}(\mathbf{x},\mathbf{r})}_{=:B_N}\Bigg)^{\frac{1}{q}}. \tag{4.45}$$

The first term of the right hand side of (4.45) can be controlled by large deviations estimates. Controlling the second term is the object of the following lemma:

Lemma 4.3.10. For any real number $q>1$, there exists a strictly positive real number $\delta_q>0$ and a function $C_q:]0,\delta_q[\to\mathbb{R}$ such that $\lim_{\delta\to 0}C_q(\delta)=0$ and:

$$B_N\leq\exp\{C_q(\delta)N\}.$$

Proof. Using Hölder inequality with conjugate exponents (σ,η), one finds:

$$B^N\leq\underbrace{\mathscr{E}_\gamma\left[\int_{(\mathscr{C}\times D)^N}\exp\Big\{\sum_{i=1}^N\sigma\tilde{X}_i^\nu\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right]^{\frac{1}{\sigma}}}_{=:(B_1^N)^{\frac{1}{\sigma}}}\underbrace{\mathscr{E}_\gamma\left[\int_{\hat{\mu}_N\in B(\nu,\delta)}\prod_{i=1}^N\exp\Big\{q\eta\big(\tilde{X}_i^{\hat{\mu}_N}-\tilde{X}_i^\nu\big)\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right]^{\frac{1}{\eta}}}_{=:\left(B_2^N\right)^{\frac{1}{\eta}}}. \tag{4.46}$$

The first term is controlled by martingale property:

$$B_1^N=\mathscr{E}_\gamma\Bigg[\int_{D^N}\exp\left\{\sum_{i=1}^N\frac{\sigma^2-\sigma}{2}\int_0^T\big(\tilde{G}_t^\nu(r_i)+m_\nu(t,r_i)\big)^2dt\right\}$$

$$\times\int_{\mathscr{C}^N}\exp\left\{\sum_{i=1}^N\sigma\int_0^T\big(\tilde{G}_t^\nu(r_i)+m_\nu(t,r_i)\big)dW_t(x^i,r_i)-\frac{\sigma^2}{2}\int_0^T\big(\tilde{G}_t^\nu(r_i)+m_\nu(t,r_i)\big)^2dt\right\}dP_{\mathbf{r}}(\mathbf{x})d\pi^{\otimes N}(\mathbf{r})\Bigg]$$

$$\overset{\text{Jensen,Fubini}}{\leq}\int_{D^N}\prod_{i=1}^N\left\{\int_0^T\mathscr{E}_\gamma\left[\exp\left\{\frac{\sigma(\sigma-1)T}{2}\big(\tilde{G}_t^\nu(r_i)+m_\nu(t,r_i)\big)^2\right\}\right]\frac{dt}{T}\right\}d\pi^{\otimes N}(\mathbf{r})\overset{(8.1)}{\leq}\exp\Big\{c_T(\sigma-1)N\Big\},$$

with c_T uniform in space.

Let us control the second term, denoting $\kappa=q\eta$ and supposing that δ is small enough. By Cauchy-Schwarz's inequality and Fubini's theorem:

$$B_2^N\leq\mathscr{E}_\gamma\Bigg[\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N\exp\Big\{2\kappa\int_0^T\Big(\tilde{G}_t^{\hat{\mu}_N}(r_i)-\tilde{G}_t^\nu(r_i)+\big(m_{\hat{\mu}_N}(t,r_i)-m_\nu(t,r_i)\big)\Big)dW_t(x^i,r_i)$$

$$-2\kappa^2\int_0^T\Big(\tilde{G}_t^{\hat{\mu}_N}(r_i)-\tilde{G}_t^\nu(r_i)+\big(m_{\hat{\mu}_N}(t,r_i)-m_\nu(t,r_i)\big)\Big)^2dt\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\Bigg]^{\frac{1}{2}}$$

$$\times\Bigg\{\int_{\hat{\mu}_N\in B(\nu,\delta)}\mathscr{E}_\gamma\Bigg[\prod_{i=1}^N\exp\Big\{2\kappa^2\int_0^T\Big(\tilde{G}_t^{\hat{\mu}_N}(r_i)-\tilde{G}_t^\nu(r_i)+\big(m_{\hat{\mu}_N}(t,r_i)-m_\nu(t,r_i)\big)\Big)^2dt$$

$$-\kappa\int_0^T\Big(\tilde{G}_t^{\hat{\mu}_N}(r_i)+m_{\hat{\mu}_N}(t,r_i)\Big)^2-\Big(\tilde{G}_t^\nu(r_i)+m_\nu(t,r_i)\Big)^2dt\Big\}\Bigg]dP^{\otimes N}(\mathbf{x},\mathbf{r})\Bigg\}^{\frac{1}{2}}$$

The first term is equal to one by martingale property. For the second term, we remark that:

$$-\int_0^T \Big(\tilde{G}_t^{\hat{\mu}_N}(r_i)+m_{\hat{\mu}_N}(t,r_i)\Big)^2-\Big(\tilde{G}_t^{\nu}(r_i)+m_{\nu}(t,r_i)\Big)^2 dt \leq$$
$$\frac{\delta^{\frac{1}{2}}}{2}\Bigg(\frac{1}{\delta}\int_0^T \Big(\tilde{G}_t^{\hat{\mu}_N}(r_i)-\tilde{G}_t^{\nu}(r_i)+\big(m_{\hat{\mu}_N}(t,r_i)-m_{\nu}(t,r_i)\big)\Big)^2 dt$$
$$+\int_0^T \Big(\tilde{G}_t^{\hat{\mu}_N}(r_i)+\tilde{G}_t^{\nu}(r_i)+\big(m_{\hat{\mu}_N}(t,r_i)+m_{\nu}(t,r_i)\big)\Big)^2 dt\Bigg)$$

so that, by Cauchy-Schwarz's inequality:

$$B_2^N \leq \Bigg\{\int_{\hat{\mu}_N\in B(\nu,\delta)} \mathscr{E}_\gamma\Bigg[\prod_{i=1}^N \exp\Big\{\big(4\kappa^2+\kappa\delta^{-\frac{1}{2}}\big)\int_0^T \big(\tilde{G}_t^{\hat{\mu}_N}(r_i)-\tilde{G}_t^{\nu}(r_i)+(m_{\hat{\mu}_N}-m_{\nu})(t,r_i)\big)^2 dt\Big\}\Bigg] dP^{\otimes N}(\mathbf{x},\mathbf{r})\Bigg\}^{\frac{1}{4}}$$
$$\times\Bigg\{\int_{\big(\mathscr{C}\times D\big)^N}\prod_{i=1}^N \underbrace{\mathscr{E}_\gamma\Bigg[\exp\Big\{\kappa\delta^{\frac{1}{2}}\int_0^T \big(\tilde{G}_t^{\hat{\mu}_N}(r_i)+\tilde{G}_t^{\nu}(r_i)+(m_{\hat{\mu}_N}+m_{\nu})(t,r_i)\big)^2 dt\Big\}\Bigg]}_{\overset{(8.1)}{\leq}\exp\left\{c_T\kappa\delta^{\frac{1}{2}}\right\}} dP^{\otimes N}(\mathbf{x},\mathbf{r})\Bigg\}^{\frac{1}{4}}.$$

Let us control the first term of the product, by taking advantage of the fact that $\hat{\mu}_N \in B(\nu,\delta)$. We have, for any $\xi \in \mathscr{M}_1^+\big((\mathscr{C}\times D)^2\big)$ with marginals $\hat{\mu}_N$ and ν:

$$\big|m_{\hat{\mu}_N}-m_{\nu}\big|(t,r_i) \overset{(4.36)}{\leq} C_T d_T^V(\hat{\mu}_N,\nu) \leq C_T\delta,$$

and similarly

$$\mathscr{E}_\gamma\Big[\big(\tilde{G}_t^{\hat{\mu}_N}(r_i)-\tilde{G}_t^{\nu}(r_i)\big)^2\Big] \leq C\delta^2.$$

Moreover, Jensen's inequality gives

$$\Big(\tilde{G}_t^{\hat{\mu}_N}(r_i)-\tilde{G}_t^{\nu}(r_i)+(m_{\hat{\mu}_N}-m_{\nu})(t,r_i)\Big)^2 \leq C\delta^2+2\big(\tilde{G}_t^{\hat{\mu}_N}(r_i)-\tilde{G}_t^{\nu}(r_i)\big)^2,$$

so that by independence of the $\tilde{G}$ for different locations and (8.1)

$$\mathscr{E}_\gamma\Bigg[\prod_{i=1}^N \exp\Big\{\big(4\kappa^2+\kappa\delta^{-\frac{1}{2}}\big)\int_0^T \big(\tilde{G}_t^{\hat{\mu}_N}(r_i)-\tilde{G}_t^{\nu}(r_i)+(m_{\hat{\mu}_N}-m_{\nu})(t,r_i)\big)^2 dt\Big\}\Bigg] \leq \exp\big\{C_T\big(4\kappa^2+\kappa\delta^{-\frac{1}{2}}\big)\delta^2 N\big\}.$$

Hence,

$$B_2^N \leq \exp\big\{C_\kappa(\delta)N\big\}$$

with $C_\kappa(\delta)\to 0$ as $\delta\to 0$. □

Let us now conclude the proof. We recall that $\hat{\mu}_N$ satisfies a full LDP under $Q_\nu^{\otimes N}$, i.e. for any Borel set $A \subset \mathscr{M}_1^+(\mathscr{C}\times D)$:

$$-\inf_{A^o} H_\nu \leq \liminf_{N\to\infty}\frac{1}{N}\log Q_\nu^{\otimes N}(\hat{\mu}_N\in A) \leq \limsup_{N\to\infty}\frac{1}{N}\log Q_\nu^{\otimes N}(\hat{\mu}_N\in A) \leq -\inf_{\bar{A}} H_\nu.$$

Then, taking $\delta < \delta_q$, we find

$$\limsup_{N\to\infty} \frac{1}{N}\log Q^N(\hat{\mu}_N \in K \cap B(\nu,\delta)) \leq \limsup_{N\to\infty} \frac{1}{N}\left(\frac{1}{p}\log Q_\nu^{\otimes N}(\hat{\mu}_N \in K\cap B(\nu,\delta)) + \frac{1}{q}C_q(\delta)N\right)$$
$$\leq -\frac{1}{p}\inf_{K\cap\overline{B(\nu,\delta)}} H_\nu + \frac{1}{q}C_q(\delta)$$

Hence, by (4.44), for $\delta < \delta_q$,

$$\limsup_{N\to\infty} \frac{1}{N}\log Q^N\big(\hat{\mu}_N \in K\big) \leq -\frac{1}{p}\inf_{1\leq i\leq M}\inf_{K\cap\overline{B(\nu_i,\delta)}} H_{\nu_i} + \frac{1}{q}C_q(\delta)$$
$$\leq -\frac{1}{p}\inf_{1\leq i\leq M}\inf_{K\cap\overline{B(\nu_i,\delta)}} \big(I(|P)-\Gamma\big) - \frac{1}{p}\inf_{1\leq i\leq M}\inf_{K\cap\overline{B(\nu_i,\delta)}}(\Gamma-\Gamma_{\nu_i}) + \frac{1}{q}C_q(\delta).$$

Lemma (4.3.8) ensures the existence of a finite constant C_T such that:

$$\big|\Gamma_{\nu_i}(\mu)-\Gamma(\mu)\big| \leq C_T\big(I(\mu|P)+1\big)d_T^V(\nu_i,\mu).$$

In particular,

$$\limsup_{N\to\infty} \frac{1}{N}\log Q^N(\hat{\mu}_N \in K) \leq -\frac{1}{p}\inf_K H - \frac{1}{p}\inf_K\Big(\big(1+I(\mu|P)\big)C_T\delta\Big) + \frac{1}{q}C_q(\delta).$$

Suppose $I(.|P)$ is infinite everywhere on K. Then, proposition 4.3.8.(3) ensures that H is also uniformly infinite on this compact set, so that

$$-\frac{1}{p}\inf_K H - \frac{1}{p}\inf_K\Big(\big(1+I(\mu|P)\big)C_T\delta\Big) + \frac{1}{q}C_q(\delta) = -\inf_K H = -\infty.$$

Now, if exists a probability $\mu \in K$ such that $I(\mu|P) < \infty$. Then, H is also finite by the same argument. In every cases, letting $\delta \searrow 0$ yields

$$\limsup_{N\to\infty} \frac{1}{N}\log Q^N(\hat{\mu}_N \in K) \leq -\frac{1}{p}\inf_K H.$$

One concludes by sending $p \searrow 1$. □

Theorem 4.3.11 (Tightness)**.** *For any real number $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset \mathscr{M}_1^+(\mathscr{C}\times D)$ such that, for any integer N,*

$$Q^N(\hat{\mu}_N \notin K_\varepsilon) \leq \varepsilon.$$

Proof. The proof of this theorem consists in using the relative entropy inequality (4.13) and the exponential tightness of the sequence $\Big(P^{\otimes N}\big(\hat{\mu}_N \in \cdot\big)\Big)_N$. Indeed, defining A an arbitrary set in $\big(\mathscr{C}\times D\big)^N$ and applying (4.13) to the function $\Phi = \log(1+P^{\otimes N}(A)^{-1})\mathbb{1}_A$ yields:

$$\int_A \log(1+P^{\otimes N}(A)^{-1})dQ^N(\mathbf{x},\mathbf{r}) = \log(1+P^{\otimes N}(A)^{-1})Q^N(A) \leq I(Q^N|P^{\otimes N}) + \log(2),$$

so that

$$Q^N(A) \leq \frac{I(Q^N|P^{\otimes N})+\log(2)}{\log(1+P^{\otimes N}(A)^{-1})}.$$

The exponential tightness of $P^{\otimes N}$, which is a consequence of Sanov's theorem and (9.4.2), ensures that for any $\varepsilon > 0$ there exists a compact subset K_ε of $\mathscr{M}_1^+(\mathscr{C}\times D)$ such that

$$P^{\otimes N}\Big(\hat{\mu}_N \in K_\varepsilon^c\Big) \leq \exp\Big\{-\frac{N}{\varepsilon}\Big\}.$$

The theorem is hence proved as soon as we show that there exists a finite constant C, such that for any N, $I(Q^N|P^{\otimes N}) \leq CN$.

$$\frac{\mathrm{d}Q^N}{\mathrm{d}P^{\otimes N}}(\mathbf{x},\mathbf{r}) \overset{(4.3.1)}{=} \prod_{i=1}^N \mathscr{E}_\gamma\Big[\exp\big\{X^{\hat{\mu}_N}(x^i,r_i)\big\}\Big] \overset{(4.3.5)}{=} \prod_{i=1}^N \exp\Big\{\int_0^T O_{\hat{\mu}_N}(t,x^i,r_i)dW_t(x^i,r_i) - \frac{1}{2}\int_0^T O_{\hat{\mu}_N}(t,x^i,r_i)^2 dt\Big\},$$

where

$$O_{\hat{\mu}_N}(t,x,r) := \mathscr{E}_\gamma\Big[\Lambda_t\big(G^{\hat{\mu}_N}(r)\big)G_t^{\hat{\mu}_N}(r)L_t^{\hat{\mu}_N}(x,r)\big)\Big] + m_{\hat{\mu}_N}(t,r).$$

Consequently, there exist N distinct $Q_{\mathbf{r}}^N$-Brownian motion $B^i(.,\mathbf{r}), 1 \leq i \leq N$ such that:

$$W_t(x^i,r_i) = B_t^i(\mathbf{x},\mathbf{r}) + \int_0^t O_{\hat{\mu}_N}(s,x^i,r_i)ds.$$

Using the symmetry of Q^N, and B^1 we have:

$$\begin{aligned}
I(Q^N|P^{\otimes N}) &= N\int_{(\mathscr{C}\times D)^N}\Big\{\int_0^T O_{\hat{\mu}_N}(t,x^1,r_1)dW_t(x^1,r_1) - \frac{1}{2}\int_0^T O_{\hat{\mu}_N}(t,x^1,r_1)^2dt\Big\}dQ^N(\mathbf{x},\mathbf{r})\\
&= N\int_{D^N}\int_{\mathscr{C}^N}\Big\{\int_0^T O_{\hat{\mu}_N}(t,x^1,r_1)dB_t^1(\mathbf{x},\mathbf{r}) + \frac{1}{2}\int_0^T O_{\hat{\mu}_N}(t,x^1,r_1)^2dt\Big\}dQ_{\mathbf{r}}^N(\mathbf{x})d\pi^{\otimes N}(\mathbf{r})\\
&= \frac{N}{2}\int_{D^N}\int_{\mathscr{C}^N}\int_0^T\Big(\mathscr{E}_\gamma\Big[\Lambda_t\big(G^{\hat{\mu}_N}(r_1)\big)G_t^{\hat{\mu}_N}(r_1)L_t^{\hat{\mu}_N}(x^1,r_1)\big)\Big] + m_{\hat{\mu}_N}(t,r_1)\Big)^2 dtdQ_{\mathbf{r}}^N(\mathbf{x})d\pi^{\otimes N}(\mathbf{r})\\
&\overset{\text{Fubini}}{\leq} N\Big\{\int_{D^N}\int_0^T\underbrace{\int_{\mathscr{C}^N}\mathscr{E}_\gamma\Big[\Lambda_t\big(G^{\hat{\mu}_N}(r_1)\big)G_t^{\hat{\mu}_N}(r_1)L_t^{\hat{\mu}_N}(x^1,r_1)\big)\Big]^2 dQ_{\mathbf{r}}^N(\mathbf{x})}_{\varphi(t,\mathbf{r})}dtd\pi^{\otimes N}(\mathbf{r}) + \frac{\|J\|_\infty^2 T}{\lambda_*^2}\Big\}.
\end{aligned} \tag{4.47}$$

We now intend to bound $\varphi(t,\mathbf{r})$ uniformly in order to obtain the result.

$$\begin{aligned}
\varphi(t,\mathbf{r}) &= 2\int_{\mathscr{C}^N}\Big\{\mathscr{E}_\gamma\Big[\Lambda_t\big(G^{\hat{\mu}_N}(r_1)\big)G_t^{\hat{\mu}_N}(r_1)\Big(\int_0^t G_s^{\hat{\mu}_N}(r_1)dB_s^1(\mathbf{x},\mathbf{r})\Big)\Big]^2\\
&+\Big(\int_0^t \widetilde{K}_{\hat{\mu}_N,r_1}^t(t,s)\mathscr{E}_\gamma\Big[\Lambda_s\big(G^{\hat{\mu}_N}(r_1)\big)G_s^{\hat{\mu}_N}(r_1)L_s^{\hat{\mu}_N}(x^1,r_1)\Big]ds\Big)^2\Big\}dQ_{\mathbf{r}}^N(\mathbf{x})\\
&\overset{\text{C.S.,(4.23)}}{\leq} C_T\int_{\mathscr{C}^N}\Big\{\mathscr{E}_\gamma\Big[\Big(\int_0^t G_s^{\hat{\mu}_N}(r_1)dB_s^1(\mathbf{x},\mathbf{r})\Big)^2\Big] + \int_{\mathscr{C}^N}\int_0^t\mathscr{E}_\gamma\Big[\Lambda_s\big(G^{\hat{\mu}_N}(r_1)\big)G_s^{\hat{\mu}_N}(r_1)L_s^{\hat{\mu}_N}(x^1,r_1)\Big]^2 ds\Big\}dQ_{\mathbf{r}}^N(\mathbf{x})\\
&\overset{\text{Fubini}}{\leq} C_T\Big\{\mathscr{E}_\gamma\Big[\int_{\mathscr{C}^N}\Big(\int_0^t G_s^{\hat{\mu}_N}(r_1)^2 ds\Big)dQ_{\mathbf{r}}^N(\mathbf{x})\Big] + \int_0^t\varphi(s,\mathbf{r})ds\Big\} \overset{\text{Fubini}}{\leq} C_T\Big\{1 + \int_0^t\varphi(s,\mathbf{r})ds\Big\}.
\end{aligned}$$

So that Gronwall's lemma ensures that exists a constant $C_T > 0$ independent of $\mathbf{r}$ such that:

$$\sup_{t\leq T}\varphi(t,\mathbf{r}) \leq C_T\exp\big\{C_T\big\}.$$

□

4.4 IDENTIFICATION OF THE MEAN-FIELD EQUATIONS

In the Gaussian interaction case, we have seen that the series of empirical measures $\left(\hat{\mu}_N\right)_N$ satisfies a large deviations principle of speed N, and with good rate function H. In order to identify the limit of the system, we study in this section the minima of the functions H, and characterize them through an implicit equation. In the spin-glass model investigated in [22], existence and uniqueness of solutions was made difficult by the fact that the drift was not considered Lipschitz continuous. Moreover, the characterization of the possible minima of the good rate function H was achieved through an intricate variational study. Here, we propose another approach that substantially simplifies this characterization. Moreover, because of the regularity of our dynamics, we propose an original contraction argument to show that the good rate function H admits a unique minimum, proof that was yet to be developed in the context of the neuronal equations[3].

Lemma 4.4.1. Let Q be a probability measure on $\mathscr{C}\times D$ which minimizes H. Then

$$Q\simeq P,\qquad \frac{\mathrm{d}Q}{\mathrm{d}P}(x,r)=\mathscr{E}_\gamma\Big[\exp\left\{X^Q(x,r)\right\}\Big]. \tag{4.48}$$

Proof. Let $Q\in\mathscr{M}_1^+(\mathscr{C}\times D)$ that minimizes H, and define the probability measure $Q_Q\in\mathscr{M}_1^+(\mathscr{C}\times D)$ as in (4.20):

$$\forall(x,r)\in\mathscr{C}\times D,\quad \frac{\mathrm{d}Q_Q}{\mathrm{d}P}(x,r):=\mathscr{E}_\gamma\Big[\exp\left\{X^Q(x,r)\right\}\Big],$$

which is equivalent to P by Theorem 4.3.6. As H is a good rate function, its minimal value must be 0, so that $H(Q)=I(Q|P)-\Gamma(Q)=0$. This imply by Proposition 4.3.2 that $I(Q|P)=\Gamma(Q)<+\infty$, which in turn implies $Q\ll P$. Theorem 4.3.6 ensures that $\forall\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)$, $H_Q(\mu)=I(\mu|Q_Q)=I(\mu|P)-\Gamma_Q(\mu)$. In particular,

$$I(Q|Q_Q)=H_Q(Q)=I(Q|P)-\Gamma_Q(Q)=I(Q|P)-\Gamma(Q)=H(Q)=0,$$

so that $Q=Q_Q$ by Proposition 9.5.2. Furthermore $Q_Q\simeq P$ is a consequence of $\bar{Q}_Q\simeq P$, $I(Q_Q|\bar{Q}_Q)<+\infty$ and $\frac{\mathrm{d}Q_Q}{\mathrm{d}\bar{Q}_Q}>0$. □

We now prove that there exists a unique probability measure satisfying (4.48). Let the map

$$L:=\left\{\begin{array}{ll}\mathscr{M}_1^+\big(\mathscr{C}\times D\big) & \to\mathscr{M}_1^+\big(\mathscr{C}\times D\big)\\ \mu & \to Q_\mu.\end{array}\right.$$

where Q_μ is defined as in (4.20) by

$$\frac{\mathrm{d}L(\mu)}{\mathrm{d}P}(x,r):=\frac{\mathrm{d}Q_\mu}{\mathrm{d}P}(x,r)=\mathscr{E}_\gamma\Big[\exp\left\{X^\mu(x,r)\right\}\Big].$$

The main result of the section is the following:

Theorem 4.4.2. *L admits a unique fixed point on $\mathscr{M}_1^+(\mathscr{C}\times D)$.*

[3]For instance, in the non-spatialized case treated in [51] was used a strong assumption of linearity of the intrinsic dynamics (our function f) which implied that solutions were Gaussian, special case for which moment methods were used (see [103]).

Proof. Lemma (4.3.5) gives

$$\frac{\mathrm{d}L(\mu)}{\mathrm{d}P}(x,r) = \exp\left\{ \int_0^T O_\mu(t,x,r)dW_t(x,r) - \frac{1}{2}\int_0^T O_\mu(t,x,r)^2 dt \right\},$$

where

$$O_\mu(t,x,r) := \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\mu(r)\big)G_t^\mu(r)L_t^\mu(x,r)\Big] + m_\mu(t,r).$$

Let $\mu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big), r \in D$, and remark that $x \to \frac{\mathrm{d}L(\mu)}{\mathrm{d}P}(x,r)$ is a P_r-martingale. Hence one can properly define $dL(\mu)_r(x) := \frac{\mathrm{d}L(\mu)}{\mathrm{d}P}(x,r)dP_r(x) \in \mathscr{M}_1^+(\mathscr{C})$. By Girsanov's theorem $L(\mu)_r$ is the law of $\big(x_t^\mu(r)\big)_{t\in[0,T]}$, the unique strong solution of the SDE (see lemma 4.4.3)

$$\begin{cases} dx_t^\mu(r) = f(r,t,x_t^\mu(r))dt + O_\mu^{\tilde{W}}(t,r)dt + \lambda(r)d\tilde{W}_t \\ \big(x_s^\mu(r)\big)_{-\tau\le s\le 0} = \big(\bar{x}_s^0(r)\big)_{-\tau\le s\le 0}. \end{cases} \tag{4.49}$$

where $\tilde{W}$ is a $\mathbb{P}$-Brownian motion,

$$O_\mu^{\tilde{W}}(t,r) := \lambda(r)\mathscr{E}_\gamma\Big[\Lambda_t\big(G^\mu(r)\big)G_t^\mu(r)\tilde{L}_t^\mu(r)\Big] + \lambda(r)m_\mu(t,r),$$

$$\tilde{L}_t^\mu(r) := \int_0^t G_s^\mu(r)\Big(d\tilde{W}_s - m_\mu(s,r)ds\Big),$$

and $\bar{x}^0(r) \in \mathscr{C}_\tau$ is the version of $\mu_0(r)$ of hypothesis (4.6). Let also $\nu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big)$, and define similarly $x_t^\nu(r)$ the process defined by:

$$\begin{cases} dx_t^\nu(r) = f(r,t,x_t^\nu(r))dt + O_\nu^{\tilde{W}}(t,r)dt + \lambda(r)d\tilde{W}_t \\ \big(x_s^\nu(r)\big)_{-\tau\le s\le 0} = \big(\bar{x}_s^0(r)\big)_{-\tau\le s\le 0}, \end{cases}$$

where both the initial condition $\bar{x}^0(r)$ and the driving Brownian motion $(\tilde{W}_t)$ are the same as for the definition of (4.49). We have

$$\begin{aligned} \big(x_t^\mu(r) - x_t^\nu(r)\big) &= \int_0^t \big(f(r,s,x_s^\mu(r)) - f(r,s,x_s^\nu(r)) + \lambda(r)(m_\mu(s,r) - m_\nu(s,r)\big)ds \\ &+ \lambda(r)\int_0^t \mathscr{E}_\gamma\Big[\Lambda_s\big(G^\mu(r)\big)G_s^\mu(r)\tilde{L}_s^\mu(r) - \Lambda_s\big(G^\nu(r)\big)G_s^\nu(r)\tilde{L}_s^\nu(r)\Big]ds. \end{aligned} \tag{4.50}$$

Let another $\xi \in \mathscr{M}_1^+\big((\mathscr{C}\times D)^2\big)$ with marginals μ and ν, and let $\big(G,G'\big)$ be a bi-dimensional centered Gaussian process on the probability space $\big(\hat{\Omega},\hat{\mathscr{F}},\gamma\big)$ with covariance $K_\xi\big(\cdot,\cdot,r\big)$ given in (4.37).

As in proposition 4.3.7 we can obtain:

$$\lambda(r)(m_\mu(s,r) - m_\nu(s,r)) \le C_T d_s^V(\mu,\nu).$$

Moreover, observe that

$$\begin{aligned} &\mathscr{E}_\gamma\Big[\Lambda_t\big(G^\mu(r)\big)G_t^\mu(r)\tilde{L}_t^\mu(r) - \Lambda_t\big(G^\nu(r)\big)G_t^\nu(r)\tilde{L}_t^\nu(r)\Big] = \mathscr{E}_\gamma\Big[\Lambda_s(G)G_sL_s - \Lambda_s(G')G'_sL'_s\Big] \\ &= \mathscr{E}_\gamma\Big[\big(\Lambda_t(G) - \Lambda_t(G')\big)G_tL_t\Big] + \mathscr{E}_\gamma\Big[\Lambda_t(G')\big(G_t - G'_t\big)L_t\Big] + \mathscr{E}_\gamma\Big[\Lambda_t(G')G'_t\big(L_t - L'_t\big)\Big] \\ &\overset{\text{C.S.}}{\le} \mathscr{E}_\gamma\big[L_t^2\big]^{\frac{1}{2}}\left(\mathscr{E}_\gamma\Big[\big(\Lambda_t(G) - \Lambda_t(G')\big)^2G_t^2\Big]^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\Lambda_t(G')^2\big(G_t - G'_t\big)^2\Big]^{\frac{1}{2}}\right) + \mathscr{E}_\gamma\Big[\Lambda_t(G')^2G_t'^2\Big]^{\frac{1}{2}}\mathscr{E}_\gamma\Big[\big(L_t - L'_t\big)^2\Big]^{\frac{1}{2}} \end{aligned}$$

where $L_t := \int_0^t G_s d\big(d\tilde{W}_s - m_\mu(s,r)ds\big)$, and $L'_t := \int_0^t G'_s d\big(d\tilde{W}_s - m_\mu(s,r)ds\big)$.

On the one hand, relying on (4.25), (4.23) and Isserlis' theorem, we can show as in Proposition 4.3.7 that there exists $C_T > 0$ such that:

$$\mathscr{E}_\gamma\left[\Big(\Lambda_t(G) - \Lambda_t(G')\Big)^2 G_t^2\right] + \mathscr{E}_\gamma\Big[\Lambda_t(G')^2\big(G_t - G'_t\big)^2\Big] \leq C_T\left(\int_{(\mathscr{C}\times D)^2} d_t\big((y,r'),(z,\tilde{r})\big)^2 d\xi\big((y,r'),(z,\tilde{r})\big)\right).$$

and

$$\lambda(r)\mathscr{E}_\gamma\Big[\Lambda_t(G')^2 G_t'^2\Big]^{\frac{1}{2}} \leq C_T.$$

On the other hand, remark that

$$\begin{aligned}
\mathscr{E}_\gamma\Big[\big(L_t - L'_t\big)^2\Big] &\leq 2\mathscr{E}_\gamma\left[\left(\int_0^t G_s - G'_s d\tilde{W}_s\right)^2\right] + 2\mathscr{E}_\gamma\left[\left(\int_0^t G_s m_\mu(s,r) - G'_s m_\nu(s,r)ds\right)^2\right] \\
&\overset{\text{Jens., Fub.}}{\leq} 2\mathscr{E}_\gamma\left[\left(\int_0^t G_s - G'_s d\tilde{W}_s\right)^2\right] + 4t\int_0^t \left\{\mathscr{E}_\gamma\Big[\big(G_s - G'_s\big)^2 m_\mu(s,r)^2\Big] + \mathscr{E}_\gamma\Big[G_s'^2\big(m_\mu(s,r) - m_\nu(s,r)\big)^2\Big]\right\} ds \\
&\overset{(4.36)}{\leq} C_T\left\{\mathscr{E}_\gamma\left[\left(\int_0^t G_s - G'_s d\tilde{W}_s\right)^2\right] + \left(\int_{(\mathscr{C}\times D)^2} d_t\big((y,r'),(z,\tilde{r})\big)^2 d\xi\big((y,r'),(z,\tilde{r})\big)\right)\right\},
\end{aligned}$$

and also that

$$\mathscr{E}_\gamma[L_t^2] \overset{\text{C.S.}}{\leq} 2\mathscr{E}_\gamma\left[\left(\int_0^t G_s d\tilde{W}_s\right)^2\right] + \underbrace{2t\int_0^t \mathscr{E}_\gamma\Big[G_t^2 m_\mu(t,r)^2\Big]}_{\leq C_T}.$$

Injecting these result in (4.50), we obtain:

$$\begin{aligned}
\big\|x^\mu(r) - x^\nu(r)\big\|_{\infty,t}^2 \leq C_T\int_0^t \Bigg\{ &\big\|x^\mu(r) - x^\nu(r)\big\|_{\infty,s}^2 + \mathscr{E}_\gamma\left[\sup_{v\leq s}\Big(\int_0^v G_u - G'_u d\tilde{W}_u\Big)^2\right] \\
&+ \left(1 + \mathscr{E}_\gamma\left[\sup_{v\leq s}\Big(\int_0^v G_u d\tilde{W}_u\Big)^2\right]\right)\left(\int_{(\mathscr{C}\times D)^2} d_t\big((y,r'),(z,\tilde{r})\big)^2 d\xi\big((y,r'),(z,\tilde{r})\big)\right)\Bigg\} ds,
\end{aligned}$$

so that by Gronwall's lemma

$$\begin{aligned}
d_t\Big((x^\mu(r),r),(x^\nu(r),r)\Big)^2 \leq C_T\int_0^t \Bigg\{ &\mathscr{E}_\gamma\left[\sup_{v\leq s}\Big(\int_0^v G_u - G'_u d\tilde{W}_u\Big)^2\right] \\
&+ \left(1 + \mathscr{E}_\gamma\left[\sup_{v\leq s}\Big(\int_0^v G_u d\tilde{W}_u\Big)^2\right]\right)\left(\int_{(\mathscr{C}\times D)^2} d_t\big((y,r'),(z,\tilde{r})\big)^2 d\xi\big((y,r'),(z,\tilde{r})\big)\right)\Bigg\} ds.
\end{aligned}$$

Taking the expectation over the Brownian path and initial condition, and using Fubini's theorem and Burkholder-Davis-Gundy's inequality, we obtain

$$\mathbb{E}\left[d_t\Big((x^\mu(r),r),(x^\nu(r),r)\Big)^2\right] \leq C_T\int_0^t d_s^V(\mu,\nu)^2 ds. \tag{4.51}$$

Let us now show that we can integrate the term of the left-hand side over π. To this purpose, fix $r' \neq r \in D$, and let $x^\mu(r')$ be the strong solution of (4.49) with same $\tilde{W}$ but

initial condition given by $\vec{x}^0(r')$ and intrinsic dynamics $f(r',\cdot,x^\mu_\cdot(r'))$. Then

$$\begin{aligned}
&\left|x^\mu_t(r)-x^\mu_t(r')\right| \le \left|\vec{x}^0_0(r)-\vec{x}^0_0(r')\right| + \left(K_f+K_\lambda\left|\tilde{W}_t\right|\right)\|r-r'\|_{\mathbb{R}^d} + \int_0^t \Big(K_f\left\|x^\mu(r)-x^\mu(r')\right\|_{\infty,s} \\
&+C_T\left|\lambda(r)m_\mu(s,r)-\lambda(r')m_\mu(s,r')\right|\Big)ds \\
&+\int_0^t \left|\mathscr{E}_\gamma\left[\lambda(r)\Lambda_t\big(G^\mu(r)\big)G^\mu_t(r)\tilde{L}^\mu_t(r)-\lambda(r')\Lambda_t\big(G^\mu(r')\big)G^\mu_t(r')\tilde{L}^\mu_t(r')\right]\right|ds.
\end{aligned}$$

First, observe that

$$\begin{aligned}
&\left|\lambda(r)m_\mu(s,r)-\lambda(r')m_\mu(s,r')\right| = \left|\int_{\mathscr{C}\times D}\Big(J(r,\tilde{r})S(y_{s-\tau(r,\tilde{r})})-J(r',\tilde{r})S(y_{s-\tau(r',\tilde{r})})\Big)d\mu(y,\tilde{r})\right| \\
&\le K_J\|r-r'\|_{\mathbb{R}^d}+\|J\|_\infty K_S\int_{\mathscr{C}\times D}\left|y_{s-\tau(r,\tilde{r})}-y_{s-\tau(r',\tilde{r})}\right|d\mu(y,\tilde{r}) \le C\int_{\mathscr{C}\times D} d_s\big((y,r),(y,r')\big)d\mu(y,\tilde{r}),
\end{aligned}$$

and we obtain similarly by choosing the proper covariance structure between $G^\mu(r)$ and $G^\mu(r')$:

$$\begin{aligned}
&\mathscr{E}_\gamma\left[\Big(\lambda(r)\Lambda_t\big(G^\mu(r)\big)-\lambda(r')\Lambda_t\big(G^\mu(r')\big)\Big)^2 G^\mu_t(r)^2\right]+\mathscr{E}_\gamma\left[\Lambda_t\big(G^\mu(r')\big)^2\big(\lambda(r)G^\mu_t(r)-\lambda(r)G^\mu_t(r')\big)^2\right] \\
&\qquad\qquad \le C_T\int_{\mathscr{C}\times D} d_t\big((y,r),(y,r')\big)^2 d\mu(y,\tilde{r}).
\end{aligned}$$

As a consequence, by a similar analysis as previously, we obtain

$$\begin{aligned}
\left\|x^\mu(r)-x^\mu(r')\right\|^2_{\infty,t} \le C_T\Bigg\{&\left\|\vec{x}^0(r)-\vec{x}^0(r')\right\|^2_{\tau,\infty}+\left(1+\tilde{W}^{*2}_t\right)\|r-r'\|^2_{\mathbb{R}^d}+\int_0^t\left\|x^\mu(r)-x^\nu(r)\right\|^2_{\infty,s}ds \\
&+\int_0^t\left(\int_{\mathscr{C}\times D} d_s\big((y,r),(y,r')\big)^2 d\mu(y,\tilde{r})\right)\left(1+\mathscr{E}_\gamma\left[\sup_{v\le s}\Big(\int_0^v G^\mu_u(r)d\tilde{W}_u\Big)^2\right]\right)ds \\
&+\int_0^t\mathscr{E}_\gamma\left[\sup_{v\le s}\Big(\int_0^v\lambda(r)G^\mu_u(r)-\lambda(r')G^\mu_u(r')d\tilde{W}_u\Big)^2\right]ds\Bigg\},
\end{aligned}$$

where $\tilde{W}^*_t=\sup_{0\le s\le t}|\tilde{W}_s|$. Hence, using Gronwall's lemma, taking the expectation and relying again on Fubini's theorem and Burkholder-Davis-Gundy's inequality, we obtain:

$$\begin{aligned}
\mathbb{E}\left[\left\|x^\mu(r)-x^\mu(r')\right\|^2_{\infty,t}\right] \le C_T\Bigg\{&\mathbb{E}\left[\left\|\vec{x}^0(r)-\vec{x}^0(r')\right\|^2_{\tau,\infty}\right]+\left(1+\mathbb{E}\left[\tilde{W}^{*2}_t\right]\right)\|r-r'\|^2_{\mathbb{R}^d} \\
&+\int_0^t\int_{\mathscr{C}\times D} d_s\big((y,r),(y,r')\big)^2 d\mu(y,\tilde{r})ds\Bigg\}.
\end{aligned}$$

Hence $\mathbb{E}\left[\left\|x^\mu(r)-x^\mu(r')\right\|^2_{\infty,t}\right]\to 0$ as $\|r'-r\|_{\mathbb{R}^d}\searrow 0$, by using (4.6), and the Monotone Convergence Theorem. Now, observe that,

$$\underbrace{\mathbb{E}\left[d_t\Big((x^\mu(r),r),(x^\nu(r),r)\Big)^2\right]}_{=:\phi^{\mu,\nu}_t(r)} = \mathbb{E}\left[\left\|x^\mu(r)-x^\nu(r)\right\|^2_{\infty,t}\right],$$

so that

$$\begin{aligned}
&\left|\phi_t^{\mu,\nu}(r)-\phi_t^{\mu,\nu}(r')\right| = \left|\mathbb{E}\Big[\big\|x^\mu(r)-x^\nu(r)\big\|_{\infty,t}^2-\big\|x^\mu(r')-x^\nu(r')\big\|_{\infty,t}^2\Big]\right|,\\
&\overset{\text{C.S.}}{\leq} \prod_{\varepsilon=\pm1}\mathbb{E}\left[\Big(\big\|x^\mu(r)-x^\nu(r)\big\|_{\infty,t}+\varepsilon\big\|x^\mu(r')-x^\nu(r')\big\|_{\infty,t}\Big)^2\right]^{\frac{1}{2}}\\
&\leq \sqrt{2\big(\phi_t^{\mu,\nu}(r')+\phi_t^{\mu,\nu}(r)\big)}\,\mathbb{E}\left[\Big(\big\|x^\mu(r)-x^\mu(r')\big\|_{\infty,t}+\big\|x^\nu(r)-x^\nu(r')\big\|_{\infty,t}\Big)^2\right]^{\frac{1}{2}}\\
&\leq 2\sqrt{C_T\int_0^t d_s^V(\mu,\nu)^2ds}\sqrt{\mathbb{E}\Big[\big\|x^\mu(r)-x^\mu(r')\big\|_{\infty,t}^2\Big]+\mathbb{E}\Big[\big\|x^\nu(r)-x^\nu(r')\big\|_{\infty,t}^2\Big]}.
\end{aligned}$$

so that $r\to\phi_t^{\mu,\nu}(r)$ is continuous, and we can integrate inequality (4.51) over space yielding:

$$d_t^V\big(L(\mu),L(\nu)\big)^2\leq\mathbb{E}\left[\int_D d_t\Big((x^\mu(r),r),(x^\nu(r),r)\Big)^2d\pi(r)\right]\leq C_T\int_0^t d_s^V(\mu,\nu)^2ds.$$

This inequality allows to develop the classical Picard's iterations method to conclude on the existence and uniqueness of solution to the fixed point equation in $\mathscr{M}_1^+(\mathscr{C}\times D)$.

□

Lemma 4.4.3. For any $r\in D$ and $\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)$, there exists a unique strong solution to the SDE:

$$\begin{cases} dx_t^\mu(r)=f(r,t,x_t^\mu(r))dt+\lambda(r)O_\mu^{\tilde W}(t,r)dt+\lambda(r)d\tilde W_t\\ \big(x_s^\mu(r)\big)_{-\tau\leq s\leq0}=\bar x^0(r).\end{cases}$$

where $\tilde W$ is a $\mathbb{P}$-Brownian motion, $\bar x^0(r)\in\mathscr{C}_\tau$ is the continuous realization of the initial law $\mu_0(r)$ of (4.6), and

$$O_\mu^{\tilde W}(t,r):=\mathscr{E}_\gamma\left[\Lambda_t\big(G^\mu(r)\big)G_t^\mu(r)\left(\int_0^t G_s^\mu(r)\big(d\tilde W_s-m_\mu(s,r)ds\big)\right)\right]+m_\mu(t,r).$$

Proof. The proof relies on Picard's iterations. Let $x^0\in\mathscr{C}_\tau$ with $x^0=\bar x^0(r)$, and define recursively the sequence $\big(x_t^n,0\leq t\leq T\big)_{n\in\mathbb{N}^*}$ by $(x_s^n)_{-\tau\leq s\leq0}=\bar x^0(r)$, and

$$x_t^{n+1}=\bar x_0^0(r)+\int_0^t f(r,s,x_s^n)ds+\int_0^t\lambda(r)O_\mu^{\tilde W}(s,x_s^n)ds+\lambda(r)\tilde W_t,\ \ \forall t\in[0,T].$$

Then, using the same analysis as in the precedent theorem, we find (see inequality (4.51))

$$\mathbb{E}\Big[\sup_{s\leq t}\big|x_s^{n+1}-x_s^n\big|^2\Big]\leq C_T\int_0^t\mathbb{E}\Big[\sup_{u\leq s}\big|x_u^n-x_u^{n-1}\big|^2\Big]ds.$$

The conclusion relies on classical arguments. □

4.4.1 Convergence of the process

We are now in a position to prove theorem 4.2.3.

Proof. Let $\delta > 0$ and $B(Q,\delta)$ the open ball of radius δ centered in Q for the Vaserstein distance d_T^V. We prove that $Q^N(\hat{\mu}_N \notin B(Q,\delta))$ tends to zero as N goes to infinity. Indeed, for K_ε a compact defined in theorem 4.3.11, we have for any $\varepsilon > 0$:

$$Q^N\big(\hat{\mu}_N \notin B(Q,\delta)\big) \leq \varepsilon + Q^N\big(\hat{\mu}_N \in B(Q,\delta)^c \cap K_\varepsilon\big). \tag{4.52}$$

The set $B(Q,\delta)^c \cap K_\varepsilon$ is a compact, and theorem 4.3.9 now ensures that

$$\limsup_{N\to\infty} \frac{1}{N} \log Q^N\big(\hat{\mu}_N \in B(Q,\delta)^c \cap K_\varepsilon\big) \leq - \inf_{B(Q,\delta)^c \cap K_\varepsilon} H$$

and eventually, theorem 4.2.6 ensures that the right-hand side of the inequality is strictly negative, which implies that

$$\lim_{N\to\infty} Q^N\big(\hat{\mu}_N \notin B(Q,\delta)\big) \leq \varepsilon,$$

that is:

$$\lim_{N\to\infty} Q^N\big(\hat{\mu}_N \notin B(Q,\delta)\big) = 0.$$

□

4.5 Non Gaussian connectivity weights

In this section, we relax the hypothesis that the synaptic weights are Gaussian. We only consider that the J_{ij} are i.i.d. random variables with sub-Gaussian tails (condition (H_J)), mean $\frac{J(r_i,r_j)}{N}$ and variance $\frac{\sigma(r_i,r_j)^2}{N}$, whereas the $\tilde{J}_{ij}$ will denote independent Gaussian variables with the same mean and variance. For technical reasons, we also assume here that the map σ is bounded away from zero: $\exists \sigma_* > 0, \sigma(r,r') \geq \sigma_*$. In this new setting, the LDP upper-bound of Theorem 4.2.4 no longer holds, its proof made important use of Gaussian properties. Nonetheless, we will show that the empirical measure for the network with non-Gaussian interactions still converges towards the same process as in the Gaussian case: the unique minimum of the good rate function H.

We revisit technical tools developed by Moynot and Samuelides in [178] where they demonstrate similar results in a discrete time, non-spatialized setting. Their central idea is to show that the non-Gaussian and Gaussian density are exponentially close to one another, so that their quotient can be controlled by the exponential convergence of the Gaussian empirical measure toward Q. In that purpose, we will work with a finite discretization of the time interval, as their approach cannot be readily applied to a continuous-time settings. Technically, the estimates in [178] contain a sum of squares of standard centered Gaussian variables over all the partition's times, which would diverge with the discretization step going to zero. An additional error term - comparing continuous and discrete Gaussian densities - arises from the discretization which we need to control. Nonetheless, we will show that, under a short-time hypothesis and when the partition is fine enough, the error becomes uniformly controllable. In all the demonstration, it is of crucial importance to track the effect of the size of the partition in every constant obtained for our upper-bounds.

The synaptic weights are assumed independent and with a law satisfying the Lindenberg-type hypothesis (H_J) introduced above and that we repeat here in an

equivalent manner (see Appendix of [178]):

$$\begin{cases} \exists a, D_0 > 0,\ \forall N \geq 1,\ \forall m \leq N,\ \forall (J_1, \cdots, J_m) \in \{\mathscr{L}(J_{ij}(N)), i, j \in [\![1,N]\!]\}^m \text{ independent}, \\ \forall (\lambda_1, \cdots, \lambda_m) \in [-1,1]^m, \\ \qquad\qquad \mathscr{E}_J\left(\exp \frac{aN}{m} (\lambda_1 J_1 + \cdots + \lambda_m J_m)^2 \right) \leq D_0. \end{cases} \tag{4.53}$$

In the whole section, we will denote by $Q_{\mathbf{r}}^N(J)$ the solution of the original network's equation (6.1) with non-Gaussian synaptic weights $\left(J_{ij}\right)_{1\leq i,j\leq N}$ and by $Q_{\mathbf{r}}^N(\tilde{J})$ its Gaussian counterpart, involving Gaussian weights $\left(\tilde{J}_{ij}\right)_{1\leq i,j\leq N}$. Moreover, there averaged versions will be denoted Q^N and Q_0^N respectively. In the whole section we will work on a finite time interval $[0,T]$, with the following time condition:

$$T < \frac{\lambda_*^2}{2} (\|\sigma\|_\infty^2 \wedge a). \tag{4.54}$$

Moreover, we introduce our regular discretization of time: we choose $\frac{1}{\delta} \in \mathbb{N}^*$, and define $\Delta_\delta := \{t_l = l\delta T, l \in [\![0, \frac{1}{\delta}]\!]\}$, a regular partition of $[0,T]$. We denote by $Q_{\mathbf{r}}^{N,\delta}(J)$ and $Q^{N,\delta}$ the quenched and averaged solutions of the following discretized SDE with non-Gaussian independent connectivity coefficients J_{ij}

$$\begin{cases} dX_t^{i,N} = \left(f(r_i, t, X_t^{i,N}) + \sum_{j=1}^N J_{ij} S(X_{t^{(l)}-\tau(r_i,r_j)}^{j,N}) \right) dt + \lambda(r_i) dW_t^i, \\ t^{(l)} := \sup\left\{ t_l \in \Delta_\delta | t_l \leq t \right\}, \\ \text{Law of } (x_t)_{t\in[-\tau,0]} = \bigotimes_{i=1}^N \mu_0(r_i), \end{cases}$$

and by $Q_{\mathbf{r}}^{N,\delta}(\tilde{J})$, $Q_0^{N,\delta}$ their Gaussian counterparts.

For simplicity of notations, we introduce

$$Y_i(\mathbf{x},\mathbf{r}) := \int_0^T \underbrace{\left(\sum_{j=1}^N \frac{1}{\lambda(r_i)} J_{ij} S\left(x_{t-\tau(r_i,r_j)}^j\right) \right)}_{=:\hat{G}_t(r_i)} dW_t(x^i, r_i) - \frac{1}{2} \int_0^T \left(\sum_{j=1}^N \frac{1}{\lambda(r_i)} J_{ij} S\left(x_{t-\tau(r_i,r_j)}^j\right) \right)^2 dt,$$

$$Y_i^\delta(\mathbf{x},\mathbf{r}) := \int_0^T \underbrace{\left(\sum_{j=1}^N \frac{1}{\lambda(r_i)} J_{ij} S\left(x_{t^{(l)}-\tau(r_i,r_j)}^j\right) \right)}_{=:\hat{G}_t^\delta(r_i)} dW_t(x^i, r_i) - \frac{1}{2} \int_0^T \left(\sum_{j=1}^N \frac{1}{\lambda(r_i)} J_{ij} S\left(x_{t^{(l)}-\tau(r_i,r_j)}^j\right) \right)^2 dt,$$

and

$$\tilde{Y}_i^\delta(\mathbf{x},\mathbf{r}) := \int_0^T \underbrace{\left(\sum_{j=1}^N \frac{1}{\lambda(r_i)} \tilde{J}_{ij} S\left(x_{t^{(l)}-\tau(r_i,r_j)}^j\right) \right)}_{=:\tilde{G}_t^\delta(r_i)} dW_t(x^i, r_i) - \frac{1}{2} \int_0^T \left(\sum_{j=1}^N \frac{1}{\lambda(r_i)} \tilde{J}_{ij} S\left(x_{t^{(l)}-\tau(r_i,r_j)}^j\right) \right)^2 dt.$$

For $\varepsilon > 0$, we are interested in the probability $Q^N\left(\hat{\mu}_N \in B(Q,\varepsilon)^c\right)$. First remark that, for any $A \in \mathscr{B}(\mathscr{C} \times D)$

$$Q^N\left(\hat{\mu}_N \in B(Q,\varepsilon)^c\right) \leq Q^N\left(A \cap \left\{\hat{\mu}_N \in B(Q,\varepsilon)^c\right\}\right) + Q^N\left(A^c\right).$$

In the same spirit as done in (4.45), let some conjugate exponents (ω_1, ω_2) and remark that:

$$
\begin{aligned}
Q^N\big(A \cap \{\hat{\mu}_N \in B(Q,\varepsilon)^c\}\big) &= \int_{\{\hat{\mu}_N \notin B(Q,\varepsilon)\}\cap A} \frac{\mathrm{d}Q^N}{\mathrm{d}Q^{N,\delta}}(\mathbf{x},\mathbf{r}) dQ^{N,\delta}(\mathbf{x},\mathbf{r}) = \int_{\{\hat{\mu}_N \notin B(Q,\varepsilon)\}\cap A} \prod_{i=1}^N \frac{\mathscr{E}_J\left[e^{Y_i(\mathbf{x},\mathbf{r})}\right]}{\mathscr{E}_J\left[e^{Y_i^\delta(\mathbf{x},\mathbf{r})}\right]} dQ^{N,\delta}(\mathbf{x},\mathbf{r}) \\
&\leq \Bigg(\underbrace{\int_{(\mathscr{C}\times D)^N} \mathscr{E}_J\left(\mathbb{1}_A \prod_{i=1}^N \exp\{\omega_1(Y_i - Y_i^\delta)\} \exp\{Y_i^\delta\} \right) dP^{\otimes N}(\mathbf{x},\mathbf{r})}_{=:B_A^\delta} \Bigg)^{\frac{1}{\omega_1}} \\
&\times \left(\int_{\{\hat{\mu}_N \notin B(Q,\varepsilon)\}\cap A} \frac{\mathrm{d}Q^{N,\delta}}{\mathrm{d}Q_0^{N,\delta}}(\mathbf{x},\mathbf{r}) dQ_0^{N,\delta}(\mathbf{x},\mathbf{r}) \right)^{\frac{1}{\omega_2}} \\
&\leq B_A^{\delta\,\frac{1}{\omega_1}} \Bigg(\underbrace{\int_A \left(\frac{\mathrm{d}Q^{N,\delta}}{\mathrm{d}Q_0^{N,\delta}}\right)^{q-1} dQ^{N,\delta}(\mathbf{x},\mathbf{r})}_{=:Z_N} \Bigg)^{\frac{1}{q\omega_2}} \left(\int_{\{\hat{\mu}_N \notin B(Q,\varepsilon)\}\cap A} \frac{\mathrm{d}Q_0^{N,\delta}}{\mathrm{d}Q_0^N}(\mathbf{x},\mathbf{r}) dQ_0^N(\mathbf{x},\mathbf{r}) \right)^{\frac{1}{p\omega_2}} \\
&\leq B_A^{\delta\,\frac{1}{\omega_1}} Z_N^{\frac{1}{q\omega_2}} \Bigg(\underbrace{\int_{(\mathscr{C}\times D)^N} \mathscr{E}_J\left(\mathbb{1}_A \prod_{i=1}^N \exp\{\omega_1(\tilde{Y}_i^\delta - \tilde{Y}_i)\} \exp\{\tilde{Y}_i\} \right) dP^{\otimes N}(\mathbf{x},\mathbf{r})}_{=:\tilde{B}_A^\delta} \Bigg)^{\frac{1}{p\omega_1\omega_2}} Q_0^N\Big(\hat{\mu}_N \notin B(Q,\varepsilon)\Big)^{\frac{1}{p\omega_2^2}}.
\end{aligned}
\tag{4.55}
$$

Building on the LDP satisfied by the averaged network with Gaussian synaptic weights, we hope that $Q_0^N\Big(\hat{\mu}_N \in B(Q,\varepsilon)^c\Big)$ decays exponentially fast. The idea would then be to show that the terms B_A^δ, $\tilde{B}_A^\delta$, and Z_N grow sufficiently slowly to be controlled by that exponential. Yet, as appear in (4.52), we have no information on the decay of $Q_0^N\Big(\hat{\mu}_N \in B(Q,\varepsilon)^c\Big)$. To cope with this issue, a solution is to work with T satisfying (4.54). In fact, as proved in (5.3.8), this short time condition implies an exponential tightness result for the sequence $Q_0^N\Big(\hat{\mu}_N \in \cdot\Big)$. This, along with the upper-bound for compact sets (4.45), and goodness of H implies an upper-bound for any closed set following the result of [84, Lemma 1.2.18], and thus, the exponential decay of $Q_0^N\Big(\hat{\mu}_N \in B(Q,\varepsilon)^c\Big)$.

Hence, proving that the quotients $\frac{\mathrm{d}Q^{N,\delta}}{\mathrm{d}Q_0^{N,\delta}}$, $\frac{\mathrm{d}Q^N}{\mathrm{d}Q^{N,\delta}}$, and $\frac{\mathrm{d}Q_0^{N,\delta}}{\mathrm{d}Q_0^N}$ are sufficiently close to 1 on a suitable set A, the terms in Z_N, B_A^δ and $\tilde{B}_A^\delta$ would be overridden by the exponential decay. The proof will then be completed by showing that the extra term vanishes.

This is proven in the three following lemmas. The first one controls the term Z_N, the second one copes with the terms B_A^δ and $\tilde{B}_A^\delta$, whereas the third proves that the extra term vanishes. All this is done for a suitable choice of the set A.

Let

$$
a_i^\delta(\mathbf{x},\mathbf{r}) := \mathscr{E}_J\Big[\exp\big(Y_i^\delta(\mathbf{x},\mathbf{r})\big)\Big], \quad b_i^\delta(\mathbf{x},\mathbf{r}) := \mathscr{E}_J\Big[\exp\big(\tilde{Y}_i^\delta(\mathbf{x},\mathbf{r})\big)\Big],
$$

so that

$$
\frac{\mathrm{d}Q^{N,\delta}}{\mathrm{d}Q_0^{N,\delta}}(\mathbf{x},\mathbf{r}) = \prod_{i=1}^N \frac{a_i^\delta(\mathbf{x},\mathbf{r})}{b_i^\delta(\mathbf{x},\mathbf{r})}.
$$

Lemma 4.5.1. There exists a set $A^2_{N,\delta} \in \mathscr{B}\Big(\big(\mathscr{C}\times D\big)^N\Big)$ with $P^{\otimes N}(A^2_{N,\delta})=1$, on which the (a_i^δ) and (b_i^δ) satisfy the following properties:

$$\exists A,B>0,\forall N,\frac{1}{\delta}\in\mathbb{N},i\le N,\quad a_i^\delta(\mathbf{x},\mathbf{r})\ge A\exp\Big(-\sqrt{\delta}B\sum_{l=1}^{\frac{1}{\delta}}|B_{t_l}(x_i,r_i)|\Big) \tag{H1}$$

$$\exists\lambda<1,C>0,\forall N,\frac{1}{\delta}\in\mathbb{N},i\le N,\quad a_i^\delta(\mathbf{x},\mathbf{r})\le C\exp\Big(\frac{\lambda}{2}\sum_{l=1}^{\frac{1}{\delta}}B^2_{t_l}(x_i,r_i)\Big) \tag{H2}$$

$$\forall\eta>0,\exists\alpha\in[0,1],\forall N\ge 1,\forall\frac{1}{\delta}\in\mathbb{N},k\le N,\text{ if }\frac{k}{N}\le\alpha\text{ then, }\forall s,\text{ injection from} \tag{H3}$$

$\{1,\dots,k\}$ into $\{1,\dots N\},\forall i\notin Im(s)=\{s(1),\dots,s(k)\},\exists\tilde{a}_i^\delta$ depending only on $j\notin Im(s),$ such that,

$$\sup\Big(\frac{a_i^\delta(\mathbf{x},\mathbf{r})}{\tilde{a}_i^\delta(\mathbf{x},\mathbf{r})},\frac{\tilde{a}_i^\delta(\mathbf{x},\mathbf{r})}{a_i^\delta(\mathbf{x},\mathbf{r})}\Big)\le(1+\eta)\exp\Big(\frac{\eta}{2}\sum_{l=1}^{\frac{1}{\delta}}B^2_{t_l}(x_i,r_i)\Big)$$

$$\exists D>0,\forall E>0,\exists N_0,\forall\frac{1}{\delta}\in\mathbb{N},\forall N\ge N_0,\forall i\in\{1,\dots,N\}, \tag{H4}$$

$$\frac{a_i^\delta(\mathbf{x},\mathbf{r})}{b_i^\delta(\mathbf{x},\mathbf{r})}\le 1+E\exp\Big\{\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}}\Big(B^2_{t_l}(x_i,r_i)+D\sqrt{\delta}|B_{t_l}(x_i,r_i)|\Big)\Big\},$$

where $B_{t_l}(x,r)=\Big(\frac{W_{t_{l+1}}(x,r)-W_{t_l}(x,r)}{\sqrt{\delta T}}\Big)$.
Moreover, on $A^2_{N,\delta}$, choosing $\eta_0>0$ such that $\alpha+2\eta_0<1$, $\eta\le\eta_0$, $\alpha\le\eta$, and conjugate exponents (p,q) satisfying $q\in]1,\frac{3}{2}[$ and $\lambda+(q-1)<1$, then $\exists C_1,C_2>0,\xi<1$ such that

$$Z_N\le\big(1+E^{q-1}C_{\alpha,\delta}\big)^N(1+\eta)^{2N}\left(\exp\Big\{N\big(\frac{\sqrt{\eta}}{\delta}+B\frac{\eta^{\frac{1}{4}}}{\delta^{\frac{1}{2}}}\big)\Big\}+C_2^{\frac{N}{\delta}}\exp\Big\{N\frac{\xi-1}{4\delta\sqrt{\eta}}\Big\}\right),$$

where $C_{\alpha,\delta}:=\max\big(C_1^{\frac{1}{\alpha\delta}},\frac{C}{A}C_1^{\frac{1}{\delta}},1\big)$.

Remark 9. Sums of squares of centered standard Gaussian appear in an exponential for the upperbound of (H2), (H3) and (H4). There are exactly as many Gaussians as points in the partition Δ_δ, so that in the continuous limit, these terms will diverge.

We will demonstrate that all the hypotheses of lemma 4.5.1 are valid in the case of the randomly connected network. Their proof are postponed in Appendix 4.6.3.

Lemma 4.5.2. Suppose that $\frac{T}{\lambda_*^2}<a$ and let $\frac{1}{4}<\beta<\frac{1}{2}$. Then for δ small enough, exists a constant C_T independent of N and δ, and a set $A^1_{N,\delta}\in\mathscr{B}\big(\big(\mathscr{C}\times D\big)^N\big)$ such that:

$$\exists\delta_0>0,\forall\delta<\delta_0,\forall N,\quad B^\delta_{A^1_{N,\delta}}+\tilde{B}^\delta_{A^1_{N,\delta}}\le\exp\Big\{C_TN\delta^\beta\Big\}.$$

Restricting the integral on the set $A_{N,\delta}:=A^1_{N,\delta}\cap A^2_{N,\delta}$ will allow obtaining a proper control on both the a_i^δ, the b_i^δ and the term $B^\delta_{A^1_{N,\delta}}$. It also makes $Q^N\big(A^c_{N,\delta}\big)$ appear in (4.55). We must justify that this quantity goes to zero as N grows to infinity. This is the purpose of the following lemma, whose proof is postponed to the end of the section:

Lemma 4.5.3. For $T < \lambda_*^2 a$, exists a constant $\delta_0 > 0$ such that $\forall \delta < \delta_0$, $Q^N\big(A_{N,\delta}^c\big)$ decreases exponentially fast to zero as N goes to infinity, where $A_{N,\delta}$ is the intersection of $A_{N,\delta}^1$ and $A_{N,\delta}^2$ from lemmas (4.5.1) and (4.5.2).

We now state and prove the main result of the section:

Theorem 4.5.4. *Let (J_{ij}) satisfy condition (H_J) and δ be small enough. Let λ be the constant of condition* (H2)*, and suppose that the conjugate exponents (p,q) satisfy $q \in]1, \frac{3}{2}[$ and $\lambda + (q-1) < 1$. Then, under a short time hypothesis, the empirical measure undergoes, under $Q^N\big(\hat{\mu}_N \in \cdot\big) \overset{\mathscr{L}}{\to} \delta_Q$. In details, if $\frac{T}{\lambda_*^2} < a$, then,*

$$\exists b > 0, \exists N_0 \in \mathbb{N}, \forall N \geq N_0, \quad Q^N\big(\hat{\mu}_N \in B(Q,\varepsilon)^c\big) \leq \exp(-bN).$$

Proof of Theorem 4.5.4. We start by choosing in a specific order the parameters appearing in the previous estimates. Let $\frac{1}{4} < \beta < \frac{1}{2}$, and the associated set $A_{N,\delta}^1$ as in lemma 4.5.2. Let also $b_0 > 0$ such that

$$Q_0^N\big(\hat{\mu}_N \in B(Q,\varepsilon)^c\big) \leq \exp\big\{-b_0 N\big\}. \tag{4.56}$$

Let $\eta = \delta^3$, and $\alpha_\delta \leq \eta$ as in hypothesis (H3). Remark that it is valid for any pair $(\frac{1}{\delta}, N)$ with $N \in \mathbb{N}$. Moreover, let $0 < \gamma < \frac{1}{4} < \beta$ and fix $\omega_1 = 1 + \delta^{\frac{\gamma}{2}}$. Then, for δ small enough and under the short time condition $\frac{T}{\lambda_*^2} < a$, the conditions of Lemmas 4.5.2 and 4.5.3, as well as $\eta < \eta_0$ will be satisfied, with

$$B_A^\delta + \tilde{B}_A^\delta \leq \exp\Big\{C_T N \delta^\beta\Big\}.$$

Furthermore, inequality (4.55) gives

$$Q^N\Big(\hat{\mu}_N \in B(Q,\varepsilon)^c\Big) \leq \exp\Big(-\frac{b_0 N(\omega_1 - 1)^2}{p\omega_1^2}\Big) \exp\Big\{C_T N \delta^{\frac{1}{4}}\Big\} Z_N^{\frac{\omega_1 - 1}{q\omega_1}} + Q^N(A_{N,\delta}^c).$$

Lemma 4.5.1 ensures that for δ small enough, one has $\frac{1-\xi}{4\sqrt{\eta}} > \log(C_2)$, so that

$$Z_N \leq 2\exp\Big\{N\log(1 + E^{q-1}C_{\alpha,\delta})\Big\} \exp\Big\{2N\big((1+B)\delta^{\frac{1}{4}} + \log(1+\delta^3)\big)\Big\}.$$

Moreover, for δ small enough, $C_{\alpha,\delta} = C_1^{\frac{1}{\alpha\delta}}$, so that choosing $E = C_{\alpha,\delta}^{\frac{-2}{q-1}}$, Z_N is at most of order $\exp\big\{C_T \delta^{\frac{1}{4}} N\big\}$. As $0 < \gamma < \frac{1}{4} < \beta$, one can see that for δ small enough, exists N_0 such that $\forall N \geq N_0$

$$\begin{aligned} Q^N\big(\hat{\mu}_N \in B(Q,\varepsilon)^c\big) &\leq \exp\big(-\frac{CN\delta^\gamma}{p}\big) \exp\big(CN\delta^\beta\big) \exp\big(CN\frac{\delta^{\frac{2\gamma+1}{4}}}{q}\big) + Q^N(A_{N,\delta}^c) \\ &\leq \exp\big\{-\tilde{C}N\delta^\gamma\big\} + Q^N(A_{N,\delta}^c). \end{aligned}$$

□

We now prove Lemmas 4.5.2, 4.5.3 and then 4.5.1.

Proof of Lemma 4.5.2. By symmetry, we only prove the inequality for the term B_A^δ. We use Hölder inequality with conjugate exponents (κ_1, κ_2), to split B_A^δ in two terms:

$$B_A^\delta \leq \Bigg\{ \underbrace{\int_{(\mathscr{C}\times D)^N} \prod_{i=1}^N \exp\big\{\kappa_1 Y_i\big\} dP^{\otimes N}(\mathbf{x},\mathbf{r})}_{B_1^\delta} \Bigg\}^{\frac{1}{\kappa_1}} \Bigg\{ \underbrace{\int_A \prod_{i=1}^N \mathscr{E}_J\Big(\exp\big\{(\omega_1 - 1)\kappa_2\big(Y_i - Y_i^\delta\big)\big\}\Big) dP^{\otimes N}(\mathbf{x},\mathbf{r})}_{B_2^\delta} \Bigg\}^{\frac{1}{\kappa_2}}.$$

To control each of these terms, we will mainly rely on martingale property, as well as on the hypothesis (H_J). For the control of the terms B_1^δ, the idea is to chose κ_1 sufficiently close to 1 so that the expectation of the power of the martingale $\prod_{i=1}^N \exp\{Y_i\}$ will be almost equal to 1. The smallness of the second term will be a consequence of the Hölder continuity of solutions under P. In detail, we have:

$$\begin{aligned}
B_1^\delta =& \mathscr{E}_J\left(\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N \exp\Big\{\kappa_1\int_0^T \hat{G}_t^\delta(r_i)dW_t(x^i,r_i) - \frac{\kappa_1}{2}\int_0^T \hat{G}_t^\delta(r_i)^2dt\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right)\\
\overset{\text{Hölder}}{\leq}& \mathscr{E}_J\left(\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N \exp\Big\{\kappa_1^2\int_0^T \hat{G}_t^\delta(r_i)dW_t(x^i,r_i) - \frac{\kappa_1^4}{2}\int_0^T \hat{G}_t^\delta(r_i)^2dt\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right)^{\frac{1}{\kappa_1}}\\
&\times \mathscr{E}_J\left(\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N \exp\Big\{\frac{\kappa_1}{\kappa_1-1}\frac{\kappa_1^3-\kappa_1}{2\lambda_*^2}\int_0^T\Big(\sum_{j=1}^N J_{ij}S\big(x^j_{t^{(l)}-\tau(r_i,r_j)}\big)\Big)^2dt\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right)^{\frac{\kappa_1-1}{\kappa_1}}.
\end{aligned}$$

The first term of the right-hand side is equal to one by martingale property, so that using Fubini theorem and Jensen's inequality yields

$$B_1^\delta \leq \left\{\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N\int_0^T \mathscr{E}_J\left(\exp\Big\{\frac{\kappa_1^2(\kappa_1+1)T}{2\lambda_*^2}\Big(\sum_{j=1}^N J_{ij}S\big(x^j_{t^{(l)}-\tau(r_i,r_j)}\big)\Big)^2\Big\}\right)\frac{dt}{T}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right\}^{\frac{\kappa_1-1}{\kappa_1}}.$$

Moreover, using hypothesis (H_J) and the inequality $\frac{T}{\lambda_*^2} < a$, one can choose $\kappa_1 - 1$ small enough so to obtain

$$\big(B_1^\delta\big)^{\frac{1}{\kappa_1}} \leq \exp\big\{C_T(\kappa_1-1)N\big\}. \tag{4.57}$$

We now deal with the second term:

$$\begin{aligned}
B_2^\delta =& \mathscr{E}_J\left(\int_A\prod_{i=1}^N \exp\Bigg\{(\omega_1-1)\kappa_2\int_0^T\Big(\hat{G}_t(r_i)-\hat{G}_t^\delta(r_i)\Big)dW_t(x^i,r_i)\right.\\
&\left.-\frac{(\omega_1-1)\kappa_2}{2}\int_0^T \hat{G}_t(r_i)^2-\hat{G}_t^\delta(r_i)^2dt\Bigg\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right)\\
\overset{\text{C.S.}}{\leq}& \mathscr{E}_J\left(\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N \exp\Bigg\{2(\omega_1-1)\kappa_2\int_0^T\Big(\hat{G}_t(r_i)-\hat{G}_t^\delta(r_i)\Big)dW_t(x^i,r_i)\right.\\
&\left.-4(\omega_1-1)^2\kappa_2^2\int_0^T\Big(\hat{G}_t(r_i)-\hat{G}_t^\delta(r_i)\Big)^2dt\Bigg\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right)^{\frac{1}{2}}\\
&\times\mathscr{E}_J\left(\int_A\prod_{i=1}^N \exp\Bigg\{4(\omega_1-1)^2\kappa_2^2\int_0^T\Big(\hat{G}_t(r_i)-\hat{G}_t^\delta(r_i)\Big)^2dt\right.\\
&\left.-(\omega_1-1)\kappa_2\int_0^T \hat{G}_t(r_i)^2-\hat{G}_t^\delta(r_i)^2dt\Bigg\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right)^{\frac{1}{2}}.
\end{aligned}$$

As previously, the first term of the right-hand side is equal to 1 by the martingale property. Moreover,

$$-\big(\hat{G}^2-(\hat{G}^\delta)^2\big) = (\hat{G}^\delta-\hat{G})(\hat{G}+\hat{G}^\delta) \leq \frac{\kappa_2}{2}(\hat{G}^\delta-\hat{G})^2+\frac{1}{2\kappa_2}(\hat{G}^\delta+\hat{G})^2.$$

Hypothesis (H_J) allow to control the first of these terms. Let $\frac{1}{4} < \beta < \frac{1}{2}$. Then:

$$(\hat{G}^\delta - \hat{G})^2 \leq \frac{K_S^2 \delta^{2\beta}}{\lambda_*^2} \left(\sum_{j=1}^N J_{ij} \underbrace{\frac{S\big(x^j_{t-\tau(r_i,r_j)}\big) - S\big(x^j_{t^{(l)}-\tau(r_i,r_j)}\big)}{K_S \delta^\beta}}_{\lambda_{ij}(t)} \right)^2.$$

$$\begin{aligned} B_2^\delta &\overset{\text{Fubini}}{\leq} \left\{ \int_A \prod_{i=1}^N \mathscr{E}_J \left(\exp\left\{ \frac{K_S^2(\omega_1-1)\kappa_2^2\delta^{2\beta}\big(8(\omega_1-1)+1\big)}{2\lambda_*^2} \int_0^T \Big(\sum_{j=1}^N J_{ij}\lambda_{ij}(t) \Big)^2 dt \right.\right. \\ &\qquad \left.\left. + \frac{\omega_1-1}{2} \int_0^T \big(\hat{G}_t(r_i) + \hat{G}_t^\delta(r_i)\big)^2 dt \right\} \right) dP^{\otimes N}(\mathbf{x},\mathbf{r}) \right\}^{\frac{1}{2}} \\ &\overset{\text{C.S.}}{\leq} \left\{ \int_A \prod_{i=1}^N \mathscr{E}_J \left(\int_0^T \exp\left\{ C_T(\omega_1-1)\kappa_2^2\delta^{2\beta} \Big(\sum_{j=1}^N J_{ij}\lambda_{ij}(t) \Big)^2 \right\} \frac{dt}{T} \right) dP^{\otimes N}(\mathbf{x},\mathbf{r}) \right\}^{\frac{1}{4}} \\ &\times \left\{ \int_{(\mathscr{C}\times D)^N} \prod_{i=1}^N \mathscr{E}_J \left(\int_0^T \exp\left\{ C_T(\omega_1-1) \Big(\sum_{j=1}^N J_{ij} \frac{S\big(x^j_{t-\tau(r_i,r_j)}\big) + S\big(x^j_{t^{(l)}-\tau(r_i,r_j)}\big)}{2} \Big)^2 \right\} \frac{dt}{T} \right) dP^{\otimes N}(\mathbf{x},\mathbf{r}) \right\}^{\frac{1}{4}}. \end{aligned}$$

One sees that, for $\omega_1 - 1$ small enough, the second term in the right-hand side can be handled using again Fubini theorem and hypothesis (H_J). We split the other term into two parts: one in which we keep only the λ_{ij} that behave nicely, so that we can rely on hypothesis (H_J), and the other one in which only pathological λ_{ij} appear: these may bring large contributions, but they appear infrequently. Moreover, even for such λ_{ij}, $K_S\delta^\beta|\lambda_{ij}| \leq 2$. Let then $E^j_{N,\delta} := \Big\{ \sup_{t,s\in[-\tau,T],|t-s|\leq\delta} \big|x_t^j - x_s^j\big| \leq \delta^\beta \Big\}$. Let also $c_{N,\delta} := \sum_{j=1}^N \mathbf{1}_{(E^j_{N,\delta})^c}$. It is the number of indices j for which problems appear. Let $A^1_{N,\delta} = \{c_{N,\delta} \leq \delta^{2\beta} N\}$. In particular, on the event $E^j_{N,\delta}$, every quantity $|\lambda_{ij}(t)|$ is smaller than 1 for δ small enough. Then

$$\begin{aligned} &\left\{ \int_{A^1_{N,\delta}} \prod_{i=1}^N \int_0^T \mathscr{E}_J \left(\exp\left\{ C_T(\omega_1-1)\kappa_2^2\delta^{2\beta} \Big(\sum_{j=1}^N J_{ij}\lambda_{ij}(t) \Big)^2 \right\} \right) \frac{dt}{T} dP^{\otimes N}(\mathbf{x},\mathbf{r}) \right\}^{\frac{1}{4}} \\ &\overset{\text{C.S.}}{\leq} \left\{ \int_{(\mathscr{C}\times D)^N} \prod_{i=1}^N \int_0^T \mathscr{E}_J \left(\exp\left\{ C_T(\omega_1-1)\kappa_2^2\delta^{2\beta} \Big(\sum_{j=1}^N \mathbf{1}_{E^j_{N,\delta}} J_{ij}\lambda_{ij}(t) \Big)^2 \right\} \right) \frac{dt}{T} dP^{\otimes N}(\mathbf{x},\mathbf{r}) \right\}^{\frac{1}{8}} \\ &\times \left\{ \int_{A^1_{N,\delta}} \prod_{i=1}^N \int_0^T \mathscr{E}_J \left(\int_0^T \exp\left\{ C_T(\omega_1-1)\kappa_2^2 \Big(\sum_{j=1}^N \mathbf{1}_{(E^j_{N,\delta})^c} J_{ij} \frac{K_S\delta^\beta\lambda_{ij}(t)}{2} \Big)^2 \right\} \frac{dt}{T} \right) dP^{\otimes N}(\mathbf{x},\mathbf{r}) \right\}^{\frac{1}{8}}. \end{aligned}$$

Remark that, considered under $P^{\otimes N}$, the λ_{ij} are independent of the matrix J. Then, for $\kappa_2 = O(\delta^{-\beta})$, hypothesis (H_J) yields

$$\begin{aligned} &\left\{ \int_{(\mathscr{C}\times D)^N} \prod_{i=1}^N \int_0^T \mathscr{E}_J \left(\exp\left\{ C_T(\omega_1-1)\kappa_2^2\delta^{2\beta} \Big(\sum_{j=1}^N \mathbf{1}_{E^j_{N,\delta}} J_{ij}\lambda_{ij}(t) \Big)^2 \right\} \right) \frac{dt}{T} dP^{\otimes N} \right\}^{\frac{1}{8}} \\ &\qquad \leq \exp\Big\{ C_T(\omega_1-1)\kappa_2^2\delta^{2\beta}N \Big\}. \end{aligned}$$

Now for the other term, as $c_{N,\delta} \leq \delta^{2\beta} N$ on $A^1_{N,\delta}$,

$$\left\{ \int_{A^1_{N,\delta}} \prod_{i=1}^{N} \mathscr{E}_J \left(\int_0^T \exp\left\{ C_T(\omega_1 - 1)\kappa_2^2 \left(\sum_{j=1}^{N} \mathbf{1}_{(E^j_{N,\delta})^c} J_{ij} \frac{K_S \delta^\beta \lambda_{ij}(t)}{2} \right)^2 \right\} \frac{dt}{T} \right) dP^{\otimes N}(\mathbf{x},\mathbf{r}) \right\} \leq$$
$$\left\{ \int_{A^1_{N,\delta}} \prod_{i=1}^{N} \int_0^T \mathscr{E}_J \left(\exp\left\{ C_T(\omega_1 - 1)\kappa_2^2 \delta^{2\beta} \frac{N}{c_{N,\delta}} \left(\sum_{j=1}^{N} \mathbf{1}_{(E^j_{N,\delta})^c} J_{ij} \frac{K_S \delta^\beta \lambda_{ij}(t)}{2} \right)^2 \right\} \right) \frac{dt}{T} dP^{\otimes N}(\mathbf{x},\mathbf{r}) \right\}.$$

Moreover, $\frac{K_S \delta^\beta |\lambda_{ij}(t)|}{2} \leq 1$, so that for $\kappa_2 = O(\delta^{-\beta})$, we are also in the scope of hypothesis (H_J).

To summarize, as soon as $\omega_1 - 1$ is small enough and for $\kappa_2 = O(\delta^{-\beta})$ one can use Fubini Theorem and hypothesis (H_J) to obtain

$$\left(B^\delta_2\right)^{\frac{1}{\kappa_2}} \leq \exp\left\{ C_T \left((\omega_1 - 1)\kappa_2 \delta^{2\beta} + \frac{(\omega_1 - 1)}{\kappa_2} \right) N \right\} \leq \exp\left\{ \frac{C_T(\omega_1 - 1)N}{\kappa_2} \right\}. \tag{4.58}$$

Hence, using inequalities (4.57) and (4.58) with $\kappa_1 = 1 + \delta^\beta$, $\omega_1 - 1$ small enough, then $\kappa_2 = O(\delta^{-\beta})$, and under a short time hypothesis $\frac{T}{\lambda_*^2} < a$, there exists a constant C_T independent of N and δ such that

$$B^\delta_{A^1_{N,\delta}} \leq \exp\left\{ C_T \left((\kappa_1 - 1) + \frac{\omega_1 - 1}{\kappa_2} \right) N \right\} \leq \exp\left\{ C_T(\kappa_1 - 1)N \right\} = \exp\left\{ C_T \delta^\beta N \right\}.$$

□

Let us now take care of Z_N appearing in the right-hand side of (4.55).

Proof of Lemma 4.5.1. We will demonstrate assumptions (H1)- (H4), and define the set $A^2_{N,\delta} \in \mathscr{B}\left(\left(\mathscr{C} \times D \right)^N \right)$ in Appendix 4.6.3.
We prove here the inequality involving Z_N. Let λ be the constant of condition (H2). Then, we chose in (4.55) conjugate exponents (p,q) satisfying $q \in]1, \frac{3}{2}[$ and $\lambda + (q-1) < 1$. Then,

$$Z_N \leq \int_{A^2_{N,\delta}} \prod_{i=1}^{N} \left(1 + \frac{a^\delta_i(\mathbf{x},\mathbf{r}) - b^\delta_i(\mathbf{x},\mathbf{r})}{b^\delta_i(\mathbf{x},\mathbf{r})} \right)^{q-1} dQ^{N,\delta}(\mathbf{x},\mathbf{r}).$$

Property (H4) implies that $\left| \frac{a^\delta_i(\mathbf{x},\mathbf{r}) - b^\delta_i(\mathbf{x},\mathbf{r})}{b^\delta_i(\mathbf{x},\mathbf{r})} \right| \leq E \exp\left\{ \frac{1}{2} \sum_{l=0}^{\frac{1}{\delta}} \left(B^2_{t_l}(x_i,r_i) + D\sqrt{\delta} \left| B_{t_l}(x_i,r_i) \right| \right) \right\}$.
Moreover, as $(x+y)^{q-1} \leq x^{q-1} + y^{q-1}$ for any $x, y > 0$, one has

$$Z_N \leq 1 + \sum_{k=1}^{N} \frac{E^{k(q-1)}}{k!} \sum_{s \in \mathscr{I}^k_N} O_{s,k}, \tag{4.59}$$

where $\mathscr{I}^k_N$ is the set of injective application from $[\![1,k]\!]$ to $[\![1,N]\!]$, and

$$O_{s,k} = \int_{A^2_{N,\delta}} \prod_{i=1}^{k} \exp\left\{ \frac{q-1}{2} \sum_{l=1}^{\frac{1}{\delta}} \left(B^2_{t_l}\left(x_{s(i)}, r_{s(i)}\right) + \sqrt{\delta} D \left| B_{t_l}\left(x_{s(i)}, r_{s(i)}\right) \right| \right) \right\} dQ^{N,\delta}(\mathbf{x},\mathbf{r}).$$

Let $\eta > 0$, and $\alpha \leq \eta$ be as in $(H3)$. Then, if $\frac{k}{N} > \alpha$ we can apply (H2) to obtain:

$$O_{s,k} \le O_{s,N} \le C^N \int_{A^2_{N,\delta}} \prod_{i=1}^{N} \left(\exp\Big\{ \frac{\lambda+(q-1)}{2} \sum_{l=1}^{\frac{1}{\delta}} B^2_{t_l}(x_i,r_i) + \sqrt{\delta} D \frac{q-1}{2} \sum_{l=1}^{\frac{1}{\delta}} \big| B_{t_l}(x_i,r_i) \big| \Big\} \right) dP^{\otimes N}(\mathbf{x},\mathbf{r})$$

$$\le \left(C \prod_{l=1}^{\frac{1}{\delta}} \int_{\mathscr{C}\times D} \exp\Big\{ \frac{\lambda+(q-1)}{2} B^2_{t_l}(x,r)) + \sqrt{\delta} D \frac{q-1}{2} \big| B_{t_l}(x,r) \big| \Big\} dP(x,r) \right)^N \le C_1^{\frac{N}{\delta}}$$

so that

$$O_{s,k} \le \big(C_1^{\frac{1}{\alpha\delta}} \big)^k.$$

Suppose now that $\frac{k}{N} \le \alpha$. We then use property (H3) for a_i^δ, with $i \notin \{s(1),\dots,s(k)\}$, and property (H2) for the other i, and obtain by independence

$$O_{s,k} \le (1+\eta)^N F_N G_N$$

where

$$F_N = C^k \int_{A^2_{N,\delta}} \prod_{i=1}^{k} \left(\exp\Big\{ \frac{\lambda+(q-1)}{2} \sum_{l=1}^{\frac{1}{\delta}} B^2_{t_l}(x_{s(i)},r_{s(i)}) + \sqrt{\delta} D \frac{q-1}{2} \sum_{l=1}^{\frac{1}{\delta}} \big| B_{t_l}(x_{s(i)},r_{s(i)}) \big| \Big\} \right) dP^{\otimes N}(\mathbf{x},\mathbf{r}),$$

and

$$G_N = \int_{A^2_{N,\delta}} \prod_{i \notin Im(s)} \tilde{a}_i^\delta(\mathbf{x},\mathbf{r}) \exp\Big\{ \frac{\eta}{2} \sum_{l=1}^{\frac{1}{\delta}} B^2_{t_l}(x_i,r_i) \Big\} dP^{\otimes N}(\mathbf{x},\mathbf{r}).$$

As previously, $F_N \le C_1^{\frac{k}{\delta}}$. Moreover, using (H3) and (H1) to recover every a_i^δ, we obtain

$$G_N \le \frac{(1+\eta)^N}{A^k} \int_{A^2_{N,\delta}} \left(\prod_{i=1}^{N} \exp\Big\{ \eta \sum_{l=1}^{\frac{1}{\delta}} B^2_{t_l}(x_i,r_i) \Big\} \right) \left(\prod_{i=1}^{k} \exp\Big\{ B\sqrt{\delta} \sum_{l=1}^{\frac{1}{\delta}} \big| B_{t_l}(x_i,r_i) \big| \Big\} \right) dQ^{N,\delta}(\mathbf{x},\mathbf{r}).$$

Let now

$$I_N = \int_{A^2_{N,\delta}} \exp\Big\{ \eta \sum_{\substack{1\le l\le \frac{1}{\delta} \\ 1\le i\le N}} B^2_{t_l}(x_i,r_i) + \sqrt{\delta} B \sum_{\substack{1\le l\le \frac{1}{\delta} \\ 1\le i\le k}} \big| B_{t_l}(x_{s(i)},r_{s(i)}) \big| \Big\} \mathbf{1}_{\left\{ \sum_{\substack{1\le l\le \frac{1}{\delta} \\ 1\le i\le N}} B^2_{t_l}(x_i,r_i) \le \frac{N}{\sqrt{\eta}\delta} \right\}} dQ^{N,\delta}(\mathbf{x},\mathbf{r}),$$

and

$$J_N = \int_{A^2_{N,\delta}} \exp\Big\{ \eta \sum_{\substack{1\le l\le \frac{1}{\delta} \\ 1\le i\le N}} B^2_{t_l}(x_i,r_i) + \sqrt{\delta} B \sum_{\substack{1\le l\le \frac{1}{\delta} \\ 1\le i\le k}} \big| B_{t_l}(x_{s(i)},r_{s(i)}) \big| \Big\} \mathbf{1}_{\left\{ \sum_{\substack{1\le l\le \frac{1}{\delta} \\ 1\le i\le N}} B^2_{t_l}(x_i,r_i) > \frac{N}{\sqrt{\eta}\delta} \right\}} dQ^{N,\delta}(\mathbf{x},\mathbf{r}),$$

so that

$$G_N \le \frac{(1+\eta)^N}{A^k} \big(I_N + J_N \big).$$

As $\alpha \le \eta$, and

$$\sum_{\substack{1\le l\le \frac{1}{\delta} \\ 1\le i\le k}} \big| B_{t_l}(x_{s(i)},r_{s(i)}) \big| \overset{\text{C.S.}}{\le} \left(\sum_{\substack{1\le l\le \frac{1}{\delta} \\ 1\le i\le N}} B^2_{t_l}(x_i,r_i) \right)^{\frac{1}{2}} \cdot \sqrt{\frac{k}{\delta}},$$

we finds

$$I_N \le \exp\Big\{ N \big(\frac{\sqrt{\eta}}{\delta} + B \frac{\eta^{\frac{1}{4}}}{\delta^{\frac{1}{2}}} \big) \Big\}.$$

Moreover, for $\eta \leq \eta_0$, with $\lambda + 2\eta_0 < 1$, using (H2) we have

$$J_N \leq \int_{A^2_{N,\delta}} \exp\Big\{ \frac{\lambda+2\eta}{2} \sum_{\substack{1\leq l\leq \frac{1}{\delta}\\ 1\leq i\leq N}} B^2_{t_l}(x_i,r_i) + B\sqrt{\delta} \sum_{\substack{1\leq l\leq \frac{1}{\delta}\\ 1\leq i\leq N}} \big|B_{t_l}(x_i,r_i)\big| \Big\} \mathbf{1}_{\left\{\sum_{\substack{1\leq l\leq \frac{1}{\delta}\\ 1\leq i\leq N}} B^2_{t_l}(x_i,r_i) > \frac{N}{\sqrt{\eta\delta}}\right\}} dP^{\otimes N}(\mathbf{x},\mathbf{r}).$$

Under $P^{\otimes N}$ the $\Big(B^2_{t_l}(x_i,r_i)\Big)_{i,l}$ are independent centered standard Gaussian variables. Hence, writting down their density, we see that exists $\xi < 1$ such that

$$J_N \leq \int_{A^2_{N,\delta}} \exp\Big\{ \frac{\xi-1}{2} \sum_{\substack{1\leq l\leq \frac{1}{\delta}\\ 1\leq i\leq N}} u^2_{i,l} + B\sqrt{\delta} \sum_{\substack{1\leq l\leq \frac{1}{\delta}\\ 1\leq i\leq N}} \big|u_{i,l}\big| \Big\} \mathbf{1}_{\left\{\sum_{\substack{1\leq l\leq \frac{1}{\delta}\\ 1\leq i\leq N}} u^2_{i,l} > \frac{N}{\sqrt{\eta\delta}}\right\}} \frac{du_{1,1}\dots du_{N,\frac{1}{\delta}}}{\sqrt{(2\pi)^{\frac{N}{\delta}}}},$$

$$J_N \leq \exp\Big\{ N\frac{\xi-1}{4\delta\sqrt{\eta}} \Big\} \int_{A^2_{N,\delta}} \exp\Big\{ \frac{\xi-1}{4} \sum_{\substack{1\leq l\leq \frac{1}{\delta}\\ 1\leq i\leq N}} u^2_{i,l} + B\sqrt{\delta} \sum_{\substack{1\leq l\leq \frac{1}{\delta}\\ 1\leq i\leq N}} \big|u_{i,l}\big| \Big\} \frac{du_{1,1}\dots du_{N,\frac{1}{\delta}}}{\sqrt{(2\pi)^{\frac{N}{\delta}}}},$$

$$J_N \leq C_2^{\frac{N}{\delta}} \exp\Big\{ N\frac{\xi-1}{4\delta\sqrt{\eta}} \Big\}.$$

Therefore, letting $C_{\alpha,\delta} = \max\big(C_1^{\frac{1}{\alpha\delta}}, \frac{C}{A}C_1^{\frac{1}{\delta}}, 1\big)$, we obtain $\forall k \in \{1,\dots,N\}$

$$O_{s,k} \leq (1+\eta)^{2N} C^k_{\alpha,\delta} \left(\exp\Big\{ N\big(\frac{\sqrt{\eta}}{\delta} + B\frac{\eta^{\frac{1}{4}}}{\delta^{\frac{1}{2}}}\big) \Big\} + C_2^{\frac{N}{\delta}} \exp\Big\{ N\frac{\xi-1}{4\delta\sqrt{\eta}} \Big\} \right)$$

and injecting this upperbound in (4.59) we get

$$Z_N \leq \big(1 + E^{q-1} C_{\alpha,\delta}\big)^N (1+\eta)^{2N} \left(\exp\Big\{ N\big(\frac{\sqrt{\eta}}{\delta} + B\frac{\eta^{\frac{1}{4}}}{\delta^{\frac{1}{2}}}\big) \Big\} + C_2^{\frac{N}{\delta}} \exp\Big\{ N\frac{\xi-1}{4\delta\sqrt{\eta}} \Big\} \right).$$

□

Proof of Lemma 4.5.3. As $P^{\otimes N}(A^2_{N,\delta}) = 1$ as soon as $N \geq \frac{1}{\delta}$ (see (4.62)), it suffices to show the result on $A^1_{N,\delta}$. Hölder inequality yields

$$\int_{(C\times D)^N} \mathbf{1}_{(A^1_{N,\delta})^c} \frac{\mathrm{d}Q^N}{\mathrm{d}P^{\otimes N}}(\mathbf{x},\mathbf{r}) dP^{\otimes N}(\mathbf{x},\mathbf{r}) \overset{\text{C.S.}}{\leq} \left(\int_{(C\times D)^N} \exp\{\kappa_1 N \bar{\Gamma}(\hat{\mu}_N)\} dP^{\otimes N}(\mathbf{x},\mathbf{r}) \right)^{\frac{1}{\kappa_1}} P^{\otimes N}\big((A^1_{N,\delta})^c\big)^{\frac{1}{\kappa_2}}.$$

As done in the proof of Lemma 4.5.2, we obtain under a short-time hypothesis that

$$\left(\int_{(C\times D)^N} \exp\{\kappa_1 N \bar{\Gamma}(\hat{\mu}_N) dP^{\otimes N}(\mathbf{x},\mathbf{r})\} \right)^{\frac{1}{\kappa_1}} \leq \exp\big\{C_T(\kappa_1 - 1)N\big\},$$

where $\tilde{C}_T$ is independent of N and δ. Furthermore,

$$P^{\otimes N}\Big((A^1_{N,\delta})^c\Big) = P^{\otimes N}\big(c_{N,\delta} > \delta^{2\beta} N\big) = P^{\otimes N}\left(\frac{\sum_{i=1}^N \mathbf{1}_{(E^j_{N,\delta})^c} - P\big((E^1_{N,\delta})^c\big)}{N} > \delta^{2\beta} - P\big((E^1_{N,\delta})^c\big) \right).$$

Let us show that $\forall m \in \mathbb{N}^*, \exists C_{T,m} > 0$ such that $P\Big((E^1_{N,\delta})^c\Big) < C_{T,m}\delta^{m(\frac{1}{2}-\beta)-1}$, so that for m big enough and δ small enough, $P\Big((E^1_{N,\delta})^c\Big) < \delta^{2\beta}$.

Remember that the semi-martingale decomposition of x under P_r

$$x_t - x_s = \int_s^t f(r,u,x_u)du + \lambda(r)\big(W_t(x,r) - W_s(x,r)\big)$$

so that using the Lipschitz continuity of S, one has

$$P\big((E^1_{N,\delta})^c\big) \leq P\left(\sup_{s\in[-\tau,T-\delta],t\in[s,s+\delta]} \Big|\int_s^t f(r_1,u,x^1_u)du\Big| > \frac{\delta^\beta}{2}\right) + P\left(\sup_{s\in[-\tau,T-\delta],t\in[s,s+\delta]} \big|W_t(x^1,r_1) - W_s(x^1,r_1)\big| > \frac{\delta^\beta}{2\lambda^*}\right).$$

In fact suppose we are on $(E^1_{N,\delta})^c$, that is there exist $s\in[-\tau,T-\delta]$ and $t\in[s,s+\delta]$ such that $|x^1_t - x^1_s| > \delta^\beta$. Then, one scenario is that $s^{(l)} = t^{(l)}$ or $t = s^{(l)} + \delta$, so that s and t are in the same interval $[s^{(l)}, s^{(l)}+\delta]$, while the other possibility is that they belong to two different consecutive such intervals. Then, by triangular inequality

$$\begin{aligned}
P\Big(\sup_{s\in[-\tau,T-\delta],t\in[s,s+\delta]} \big|W^1_t - W^1_s\big| > \frac{\delta^\beta}{2\lambda^*}\Big) &\leq P\left(\sup_{t\in[-\tau,T]} \max\Big(\big|W^1_t - W^1_{t^{(l)}}\big|, \big|W^1_t - W^1_{t^{(l)}+\delta}\big|\Big) > \frac{\delta^\beta}{4\lambda^*}\right) \\
&\overset{Markov}{\leq} (4\lambda^*)^m \frac{\mathbb{E}\Big[\sup_{t\in[-\tau,T]} \max\Big(\big|W^1_t - W^1_{t^{(l)}}\big|^m, \big|W^1_t - W^1_{t^{(l)}+\delta}\big|^m\Big)\Big]}{\delta^{m\beta}} \\
&\leq (4\lambda^*)^m \frac{\sum_{l=\frac{-\tau}{\delta}}^{T/\delta} \mathbb{E}\Big[\sup_{t\in[t^{(l)},t^{(l)}+\delta]} \big|W^1_t - W^1_{t^{(l)}}\big|^m\Big] + \mathbb{E}\Big[\sup_{t\in[t^{(l)}-\delta,t^{(l)}]} \big|W^1_t - W^1_{t^{(l)}}\big|^m\Big]}{\delta^{m\beta}} \\
&\leq \frac{2(T+\tau)}{\delta^{m\beta+1}} (4\lambda^*)^m \mathbb{E}\Big[\sup_{t\in[0,\delta]} \big|W^1_t\big|^m\Big] \overset{\text{B.D.G.}}{\leq} C_{T,m}\delta^{m(\frac{1}{2}-\beta)-1},
\end{aligned}$$

where $C_{T,m}$ is a constant independent of δ and N. Moreover, Markov inequality gives:

$$P\Big(\sup_{s\in[-\tau,T-\delta],t\in[s,s+\delta]} \Big|\int_s^t f(r_1,u,x^1_u)du\Big| > \frac{\delta^\beta}{2}\Big) \leq 2^m\delta^{m(1-\beta)}\mathbb{E}\Big[\sup_{t\in[-\tau,T]} \big|f\big(r_1,t,X^{1,N}_t\big)\big|^m\Big].$$

Furthermore, it is proven in the Appendix 4.6.1 that $\sup_{t\in[-\tau,T]} \big|f\big(r_1,t,X^{1,N}_t\big)\big|$ admits exponential moments under $\mathbb{P}$. Let $\sigma^2_{N,\delta} := Var(\mathbf{1}_{(E^1_{N,\delta})^c}) = P^{\otimes N}\big((E^1_{N,\delta})^c\big)\big(1 - P^{\otimes N}\big((E^1_{N,\delta})^c\big)\big) \leq C_{T,m}\delta^{m(\frac{1}{2}-\beta)-1}$. Then, for δ small enough,

$$\begin{aligned}
P^{\otimes N}\big((A^1_{N,\delta})^c\big) &\leq P^{\otimes N}\left(\frac{1}{2\sigma_{N,\delta}\sqrt{N}}\Big|\sum_{i=1}^N \mathbf{1}_{(E^j_{N,\delta})^c} - P\big((E^1_{N,\delta})^c\big)\Big| > \frac{1}{4\sigma_{N,\delta}}\sqrt{N}\delta^{2\beta}\right) \\
&\leq \exp\{-\tilde{C}_{T,m}N\delta^{2+4\beta-2m(\frac{1}{2}-\beta)}\}\mathbb{E}\Big[\exp\Big\{\frac{1}{4}\mathcal{N}(0,1)^2\Big\}\Big].
\end{aligned}$$

Hence

$$\int_{(\mathscr{C}\times D)^N} \mathbf{1}_{(A^1_{N,\delta})^c}\frac{\mathrm{d}Q^N}{\mathrm{d}P^{\otimes N}}dP^{\otimes N}(\mathbf{x},\mathbf{r}) \leq \exp\big\{-\tilde{C}_{T,m}\Big(\delta^{2+4\beta-2m(\frac{1}{2}-\beta)}+1\Big)(\kappa_1-1)N\big\}.$$

For δ small enough, and m big enough, this term goes to zero exponentially fast with $N\to\infty$. □

4.6 APPENDIX

4.6.1 A priori estimates for single neurons

Lemma 4.6.1. $\forall r \in D$, $\|f(r,\cdot,x_\cdot)\|^2_{\infty,T}$ admits exponential moments under P_r.

Proof. Fix $r \in D$, and suppose first that $t \in [0,T]$. Using the Lipschitz continuity of $f(r,.,.)$, we have

$$\begin{aligned}\big|f(r,t,x_t)\big| &\leq \|f(r,\cdot,x_\cdot)\|_{\tau,\infty} + \big|f(r,t,x_t) - f(r,0,\vec{x}_0^0)\big| \\ &\leq \big|f(r,0,\vec{x}_0^0)\big| + K_f T + K_f \int_0^t \big|f(r,u,x_u)\big| du + K_f \lambda^* \big|W_t(x,r)\big| \\ &\leq \big|f(0,0,0)\big| + K_f(|\vec{x}_0^0| + T + d_D) + K_f \int_0^t \big|f(r,u,x_u)\big| du + K_f \lambda^* \big|W_t(x,r)\big|\end{aligned}$$

so that by Gronwall's lemma:

$$\sup_{t\in[0,T]} |f(r,t,x_t)|^2 \leq C_T\big(|\vec{x}^0|^2 + 1 + W_T^*(r)^2\big), \tag{4.60}$$

with $W_T^*(r) := \sup_{t\in[0,T]} |W_u(x,r)|$. Moreover,

$$\|f(r,\cdot,x_\cdot)\|^2_{\tau,\infty} \leq 3K_f^2\big(\|r\|^2_{\mathbb{R}^d} + \tau^2 + \|\vec{x}^0(r)\|^2_{\tau,\infty}\big),$$

so that

$$\|f(r,\cdot,x_\cdot)\|^2_{\infty,T} \leq C_T\big(\|\vec{x}^0(r)\|^2_{\tau,\infty} + 1 + W_T^*(r)^2\big).$$

As $W_T^*(r)^2 \leq 2\Big(\sup_{t\in[0,T]} W_t(r)\Big)^2 + 2\Big(\sup_{t\in[0,T]} -W_t(r)\Big)^2$, where each of the two terms of the left-hand side of the last sum has the law of $2|W_T(r)|^2$ under P_r, we have for $\alpha > 0$

$$\int_{\mathscr{C}} \exp\big\{\alpha\|f(r,\cdot,x_\cdot)\|^2_{\infty,T}\big\} dP_r(x) \overset{\text{Hölder}}{\leq} e^{\tilde{\alpha}} \left(\int_{\mathscr{C}_\tau} \exp\big\{3\tilde{\alpha}\|x\|^2_{\tau,\infty}\big\} d\mu_0(r)(x)\right)^{\frac{1}{3}} \left(\int_{\mathscr{C}} \exp\big\{6\tilde{\alpha}|W_T(r)|^2\big\} dP_r(x)\right)^{\frac{2}{3}},$$

where $\tilde{\alpha} = \alpha C_T$. For α small enough, $\int_{\mathscr{C}} \exp\big\{3\tilde{\alpha}|x_0|^2\big\} dP_r(x)$ is finite by hypothesis (4.8), and $\int_{\mathscr{C}} \exp\big\{6\tilde{\alpha}|W_T(r)|^2\big\} dP_r(x)$ is by (8.1), so that $\|f(r,\cdot,x_\cdot)\|^2_{\infty,T}$ admits exponential moments. □

4.6.2 Proof of lemma 4.2.2: regularity of the solutions for the limit equation

In this appendix we demonstrate the regularity in space of the solutions that is expressed in lemma 4.2.2. We start by showing a technical lemma on the uncoupled system before proceeding to the proof of that result.

Lemma 4.6.2. (i) The map:

$$\mathscr{P} : \begin{cases} D \to \mathscr{M}_1^+(\mathscr{C}) \\ r \to P_r \end{cases}$$

is continuous with respect to the borel topology on D, and the weak topology on $\mathscr{M}_1^+(\mathscr{C})$, e.g. $r_n \to r \implies P_{r_n} \overset{\mathscr{L}}{\to} P_r$.

(ii) Let $\mathscr{W}$ be the Wiener measure on $\mathscr{C}$. Then, $\forall A \in \mathscr{B}(\mathscr{C}), \mathscr{W}(A) = 0 \implies P_r(A) = 0$.

(iii) P is a well defined probability measure on $\mathscr{C} \times D$.

Proof. The first point is the consequence of a coupling argument. Let W be a P-Brownian motion, $\vec{x}^0 : D \to \mathscr{C}_\tau$ be as in (4.6), and $(r_n)_{n\in\mathbb{N}} \in D^{\mathbb{N}}$ a sequence of positions that converges toward $r \in D$. We consider X^n and X, the respective strong solutions of the SDEs:

$$\begin{cases} dX_t^n = f(r_n,t,X_t^n)dt + \lambda(r_n)dW_t \\ (X_t^n)_{t\in[-\tau,0]} = \vec{x}^0(r_n) \end{cases} \qquad \begin{cases} dX_t = f(r,t,X_t)dt + \lambda(r)dW_t \\ (X_t)_{t\in[-\tau,0]} = \vec{x}^0(r) \end{cases}$$

driven by the same Brownian motion W.
Then, by Gronwall lemma, letting $W_T^* = \sup_{t\in[-\tau,T]} |W_t|$,

$$\|X^n - X\|_{\infty,T} \leq \Big(\|\vec{x}^0(r_n) - \vec{x}^0(r)\|_{\tau,\infty} + \|r - r_n\|_{\mathbb{R}^d} K_f T + K_\lambda \|r - r_n\|_{\mathbb{R}^d} W_T^* \Big) e^{\{K_f T\}},$$

Hence, by (4.6):

$$\mathbb{E}\Big[\|X^n - X\|_{\infty,T}^2 \Big] \to 0, \text{ as } r_n \to r,$$

so that $P_{r_n} = \mathscr{L}(X_n) \implies \mathscr{L}(X) = P_r$ as $r_n \to r$.
In order to prove the second point, let W_r be the unique strong solution of

$$\begin{cases} dX_t = \lambda(r)dW_t \\ (X_t)_{t\in[-\tau,0]} = \vec{x}^0(r). \end{cases}$$

Following Exercise (2.10) of [200], we remark, by Lipschitz continuity of f, that explosion of P_r almost surely never occurs in finite time, so that Girsanov's theorem applies:

$$P_r \ll W_r, \qquad \frac{\mathrm{d}P_r}{\mathrm{d}W_r} = \exp\Big\{ \int_0^T \frac{f(r,t,X_t)}{\lambda(r)} dX_t - \frac{1}{2}\int_0^T \Big(\frac{f(r,t,X_t)}{\lambda(r)}\Big)^2 dt \Big\}.$$

Consequently, $\forall A \in \mathscr{B}(\mathscr{C})$,

$$P_r(A) = \mathbb{E}_{W_r}\Big(\frac{\mathrm{d}P_r}{\mathrm{d}W_r}\mathbf{1}_{\mathbf{A}}\Big)$$

so that $P_r(A) = 0$ as soon as $W_r(A) = 0$. As $\lambda(r) > \lambda_*$, $W_r(A) = 0 \iff \mathscr{W}(A) = 0$.
The third point is now easy to settle. In fact, for any $y \in \mathscr{C}$ and $\varepsilon > 0$, $\mathscr{W}(\partial\mathscr{B}(y,\varepsilon)) = \mathscr{W}(\{x \in \mathscr{C}, \|x - y\|_{\infty,T} = \varepsilon\}) = 0$. Hence, Portmanteau implies that $r \to P_r(\mathscr{B}(y,\varepsilon))$ is a continuous map, so that we can define $\int_D P_r(\mathscr{B}(y,\varepsilon))d\pi(r)$ univocally. As $\{\mathscr{B}(y,\varepsilon) \times B, y \in \mathscr{C}, \varepsilon > 0, B \in \mathscr{B}(D)\}$ form a Π-system that generates $\mathscr{B}(\mathscr{C} \times D)$, P is a well defined probability measure on $\mathscr{C} \times D$. □

We now proceed to prove lemma 4.2.2 that we repeat below:

Lemma 4.6.3. The map

$$\mathscr{Q} : \begin{cases} D^N \to \mathscr{M}_1^+(\mathscr{C}^N) \\ \mathbf{r} \to Q_{\mathbf{r}}^N \end{cases}$$

where $Q_{\mathbf{r}}^N := \mathscr{E}_J\big(Q_{\mathbf{r}}^N(J)\big)$, is continuous with respect to the weak topology. Moreover,

$$dQ^N(\mathbf{x},\mathbf{r}) := dQ_{\mathbf{r}}^N(\mathbf{x})d\pi^{\otimes N}(\mathbf{r})$$

defines a probability measure on $\mathscr{M}_1^+\big((\mathscr{C} \times D)^N\big)$.

Remark 10. $\mathscr{Q}$ maps the positions $\mathbf{r}$ to the Gaussian averaged of the solutions $Q^N_{\mathbf{r}}(J)$, so that its continuity seems to be a consequence of Cauchy-Lipschitz theorem with parameter $\mathbf{r}$. Yet, the equation depends on $\mathbf{r}$ through the synaptic weights J_{ij} which only satisfy a continuity in law. Meanwhile the proof is not difficult, it must rely on another argument. The one developed here is a coupling method.

Proof. We insist on the fact that N remains constant in this proof. Fix a deterministic sequence $\left(\mathbf{r}^n = (r_i^n)_{1\le i\le N}\right)_{n\in\mathbb{N}^*} \to_n \mathbf{r} = (r_i)_{1\le i\le N} \in D^N$, let $(W_t^i, 0\le t\le T)_{i\in[\![1,N]\!]}$ be a family of independent $\mathbb{P}$-Brownian motions, and $\bar{x}^{0,i} : D \to \mathscr{C}_\tau, 1\le i\le N$, be N independent initial condition as in (4.6). Let now $X^N_{\mathbf{r}^n} = \left(X^{i,N}_{\mathbf{r}^n}\right)_{i\in[\![1,N]\!]}$ and $X^N_{\mathbf{r}} = \left(X^{i,N}_{\mathbf{r}}\right)_{i\in[\![1,N]\!]}$ be the respective strong solutions of the two following stochastic differential equations:

$$\begin{cases} dX^{i,N}_{\mathbf{r}^n}(t) = \left(f(r_i^n,t,X^{i,N}_{\mathbf{r}^n}(t)) + \sum_{j=1}^N \tilde{J}^{\mathbf{r}^n}_{ij} S\big(X^{j,N}_{\mathbf{r}^n}(t-\tau_{r_i^n r_j^n})\big)\right)dt + \lambda(r_i^n)dW_t^i \\ (X^N_{\mathbf{r}^n}(t))_{t\in[-\tau,0]} = \left(\bar{x}^{0,i}(r_i^n)\right)_{1\le i\le N}, \end{cases}$$

$$\begin{cases} dX^{i,N}_{\mathbf{r}}(t) = \left(f(r_i,t,X^{i,N}_{\mathbf{r}}(t)) + \sum_{j=1}^N J^{\mathbf{r}}_{ij} S\big(X^{j,N}_{\mathbf{r}}(t-\tau_{r_i r_j})\big)\right)dt + \lambda(r_i)dW_t^i \\ (X^N_{\mathbf{r}}(t))_{t\in[-\tau,0]} = \left(\bar{x}^{0,i}(r_i)\right)_{1\le i\le N}. \end{cases}$$

where $J^{\mathbf{r}}_{ij} \sim \mathscr{N}\left(\frac{J(r_i,r_j)}{N}, \frac{\sigma(r_i,r_j)^2}{N}\right)$, $\tilde{J}^{\mathbf{r}^n}_{ij} \sim J^{\mathbf{r}^n}_{ij}$ satisfy (4.5), and where we used the short-hand notation $\tau_{rr'} := \tau(r,r')$. In particular, $X^{i,N}_{\mathbf{r}^n}$ has law $Q^N_{\mathbf{r}^n}(J^{\mathbf{r}^n})$, and $X^{i,N}_{\mathbf{r}}$ has law $Q^N_{\mathbf{r}}(J^{\mathbf{r}})$.

Then, we have for every $t\in[0,T]$,

$$\begin{aligned} \big(X^{i,N}_{\mathbf{r}^n}(t) - X^{i,N}_{\mathbf{r}}(t)\big) &= \big(\bar{x}_0^{0,i}(r_i^n) - \bar{x}_0^{0,i}(r_i)\big) + \Bigg(\int_0^t \Big(f(r_i^n,s,X^{i,N}_{\mathbf{r}^n}(s)) - f(r_i,s,X^{i,N}_{\mathbf{r}}(s))\Big)ds \\ &\quad + \sum_{j=1}^N \Bigg\{ \big(\tilde{J}^{\mathbf{r}^n}_{ij} - J^{\mathbf{r}}_{ij}\big)\int_0^t S\big(X^{j,N}_{\mathbf{r}^n}(s-\tau_{r_i^n r_j^n})\big)ds \\ &\quad + J^{\mathbf{r}}_{ij}\int_0^t \Big(S\big(X^{j,N}_{\mathbf{r}^n}(s-\tau_{r_i^n r_j^n})\big) - S\big(X^{j,N}_{\mathbf{r}}(s-\tau_{r_i r_j})\big)\Big)ds\Bigg\}\Bigg) + \big(\lambda(r_i^n)-\lambda(r_i)\big)W_t^i. \end{aligned}$$

Let $W_T^{*,i} = \sup_{t\in[0,T]} |W_t^i|$. Then using Lipschitz continuity of f, λ, S, the fact that $|S|\le 1$, and taking the supremum in time one obtains

$$\begin{aligned} &\left\|X^{i,N}_{\mathbf{r}^n} - X^{i,N}_{\mathbf{r}}\right\|_{\infty,t} \le \|\bar{x}^{0,i}(r_i^n) - \bar{x}^{0,i}(r_i)\|_{\tau,\infty} + \left(K_f t + K_\lambda W_T^{*,i}\right)\|r_i^n - r_i\|_{\mathbb{R}^d} \\ &+ \int_0^t \left\{K_f\left\|X^{i,N}_{\mathbf{r}^n} - X^{i,N}_{\mathbf{r}}\right\|_{\infty,s} + K_S\sum_{j=1}^N |J^{\mathbf{r}}_{ij}|\left\|X^{j,N}_{\mathbf{r}^n} - X^{j,N}_{\mathbf{r}}\right\|_{\infty,s}\right\}ds + \sum_{j=1}^N \left\{t\left|\tilde{J}^{\mathbf{r}^n}_{ij} - J^{\mathbf{r}}_{ij}\right|\right. \\ &\left. + K_S\int_0^t |J^{\mathbf{r}}_{ij}| \sup_{\substack{a,b\in[-\tau,0]\\ |a-b|\le 2K_\tau\|\mathbf{r}^n-\mathbf{r}\|_\infty}} \left|X^{j,N}_{\mathbf{r}}(s+a) - X^{j,N}_{\mathbf{r}}(s+b)\right|ds\right\}. \end{aligned}$$

where $\|\mathbf{r}\|_\infty = \sup_{1\le i\le N}\|r_i\|_{\mathbb{R}^d}$. Let us denote, for any $X = \left(X^i\right)_{1\le i\le N} \in \mathscr{C}^N$, $t\in[-\tau,T]$, $\|X\|^1_{\infty,t} = \sum_{1\le i\le N}\|X^i\|_{\infty,t}$, and for any $\mathbf{r}\in D^N$, $\|\mathbf{r}\|_1 = \sum_{1\le i\le N}\|r_i\|_{\mathbb{R}^d}$. Summing over $i\in[\![1,N]\!]$ and using Gronwall's inequality now yields

$$\big\|X^N_{\mathbf{r}^n}-X^N_{\mathbf{r}}\big\|^1_{\infty,t}\le C_T\exp\Big\{C_T\sum_{i,j=1}^N|J^{\mathbf{r}}_{ij}|\Big\}\Bigg(\|\bar{\mathbf{x}}^0(\mathbf{r}^n)-\bar{\mathbf{x}}^0(\mathbf{r})\|^1_{\tau,\infty}+\big(\sup_{1\le i\le N}W_T^{*,i}\big)\|\mathbf{r}^n-\mathbf{r}\|_1$$
$$+\sum_{i,j=1}^N t\big|\tilde{J}^{\mathbf{r}^n}_{ij}-J^{\mathbf{r}}_{ij}\big|+\Big(\sum_{i,j=1}^N|J^{\mathbf{r}}_{ij}|\Big)\int_0^t\sup_{1\le j\le N}\Big\{\sup_{\substack{a,b\in[-\tau,0]\\|a-b|\le 2K_\tau\|\mathbf{r}^n-\mathbf{r}\|_\infty}}\Big|X^{j,N}_{\mathbf{r}}(s+a)-X^{j,N}_{\mathbf{r}}(s+b)\Big|\Big\}ds\Bigg). \tag{4.61}$$

Observe that exists a map $\chi:D^N\times D^N\to\mathbb{R}$ such that $\chi(\mathbf{r}^n,\mathbf{r})\to 0$, when $n\to\infty$, and a constant $C_{T,N}$ such that:

$$\mathscr{E}_J\Big[\exp\Big\{2C_T\sum_{i,j=1}^N|J^{\mathbf{r}}_{ij}|\Big\}\Big]^{\frac12}+\mathbb{E}\Big[\big(\sup_{1\le i\le N}W_T^{*,i}\big)^2\Big]+\mathscr{E}_J\Big[\Big(\sum_{i,j=1}^N|J^{\mathbf{r}}_{ij}|\Big)^2\Big]\overset{(H_J)}{\le}C_{T,N}$$

$$\mathbb{E}\Big[\Big(\|\bar{\mathbf{x}}^0(\mathbf{r}^n)-\bar{\mathbf{x}}^0(\mathbf{r})\|^1_{\tau,\infty}\Big)^2\Big]+\big(\|\mathbf{r}^n-\mathbf{r}\|_1\big)^2+\sum_{i,j=1}^N t^2\mathscr{E}_J\Big[\big(\tilde{J}^{\mathbf{r}^n}_{ij}-J^{\mathbf{r}}_{ij}\big)^2\Big]\overset{(4.6),(4.5)}{\le}C_{T,N}\chi(\mathbf{r}^n,\mathbf{r}).$$

Denoting $\mathbb{E}_J\big[\,\cdot\,\big]:=\mathbb{E}\big[\mathscr{E}_J\big[\,\cdot\,\big]\big]$, we find taking the expectation in (4.61) and relying on Cauchy-Schwarz's inequality:

$$\mathbb{E}_J\Big[\Big[\big\|X^N_{\mathbf{r}^n}-X^N_{\mathbf{r}}\big\|^1_{\infty,t}\Big]\le\tilde{C}_{T,N}\Bigg(\chi(\mathbf{r}^n,\mathbf{r})+\mathbb{E}_J\Bigg[\sup_{1\le j\le N}\sup_{\substack{a,b\in[-\tau,0],s\in[0,t]\\|a-b|\le 2K_\tau\|\mathbf{r}^n-\mathbf{r}\|_\infty}}\Big|X^{j,N}_{\mathbf{r}}(s+a)-X^{j,N}_{\mathbf{r}}(s+b)\Big|^2\Bigg]\Bigg)^{\frac12}.$$

As solution are $\mathbb{P}$-almost surely continuous, and N remains (here) finite, the Monotone Convergence Theorem ensures that the right-hand side tends toward 0 when n goes to infinity. It implies in particular that $Q^N_{\mathbf{r}^n}$ converges in law toward $Q^N_{\mathbf{r}}$ when $n\to\infty$, so that the map $\mathbf{r}\to\mathscr{E}_J\Big[\int_{\mathscr{C}^N}\phi(\mathbf{x})dQ^N_{\mathbf{r}}(\mathbf{x})\Big]$ is continuous and integrable with respect to $\pi^{\otimes N}$. In particular, $dQ^N(\mathbf{x},\mathbf{r}):=dQ^N_{\mathbf{r}}(\mathbf{x})d\pi^{\otimes N}(\mathbf{r})$ defines a probability measure on $(\mathscr{C}\times D)^N$.

□

4.6.3 Non-Gaussian estimates

We prove that the different assumptions (H1)-(H4) are valid.
(H1): By a direct application of Jensen's inequality

$$a^\delta_i(\mathbf{x},\mathbf{r})\ge\exp\Big\{\int_0^t m_{\hat{\mu}_N}(t^{(l)},r_i)dW_t(x^i,r_i)-\frac12\int_0^T\big(m_{\hat{\mu}_N}(t^{(l)},r_i)^2+K_{\hat{\mu}_N(t^{(l)},t^{(l)},r_i)}\big)dt\Big\}$$
$$\ge\exp\Big\{-\frac12\Big(\frac{\|J\|^2_\infty+\|\sigma\|^2_\infty}{\lambda_*^2}\Big)T\Big\}\exp\Big\{-\frac{\|J\|_\infty\sqrt{\delta T}}{\lambda_*}\sum_{l=1}^{\frac{1}{\delta}}|B_{t_l}(x_i,r_i)|\Big\}.$$

(H2): Remark that

$$a_i^\delta(\mathbf{x},\mathbf{r}) = \exp\Big\{\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}} B_{t_l}(x_i,r_i)^2\Big\}\mathscr{E}_J\left(\exp\Big\{-\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}}\Big(\sqrt{\delta T}\hat{G}^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r}) - B_{t_l}(x^i,r_i)\Big)^2\Big\}\right),$$
$$\overset{\text{Hölder}}{\leq} \exp\Big\{\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}} B_{t_l}(x_i,r_i)^2\Big\}\prod_{l=1}^{\frac{1}{\delta}}\mathscr{E}_J\left(\exp\Big\{-\frac{1}{2\delta}\Big(\sqrt{\delta T}\hat{G}^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r}) - B_{t_l}(x^i,r_i)\Big)^2\Big\}\right)^\delta.$$

Suppose first that $B_{t_l}(x^i,r_i)\geq 0$. Then

$$\Big(B_{t_l}(x_i,r_i) - \sqrt{\delta T}\hat{G}^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r})\Big)^2 \geq \Big(B_{t_l}(x_i,r_i) - \sqrt{\delta T}\hat{G}^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r})\Big)^2\mathbf{1}_{\left\{2\sqrt{\delta T}\hat{G}^{\delta,i}_{t^{(l)}}\leq B_{t_l}\right\}}$$
$$\geq \frac{B_{t_l}(x_i,r_i)^2}{4}\mathbf{1}_{\left\{2\sqrt{\delta T}\hat{G}^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r})\leq B_{t_l}(x_i,r_i)\right\}},$$

so that

$$\mathscr{E}_J\left(\exp\Big\{-\frac{1}{2\delta}\Big(\sqrt{\delta T}\hat{G}^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r}) - B_{t_l}(x_i,r_i)\Big)^2\Big\}\right) \overset{(H_J)}{\leq} \exp\Big\{-\frac{B^2_{t_l}(x_i,r_i)}{8\delta}\Big\} + D_0\exp\Big\{-\frac{\lambda_*^2 a}{4\delta T}B^2_{t_l}(x_i,r_i)\Big\}.$$

We obtain the same inequality under the hypothesis $B_{t_l}\leq 0$, so that, making use of hypothesis (H_J)

$$a_i^\delta(\mathbf{x},\mathbf{r}) \leq \max\big(1,D_0\big)\exp\Big\{\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}} B_{t_l}(x_i,r_i)^2\Big\}\prod_{l=1}^{\frac{1}{\delta}}\exp\Big\{-\frac{1}{4}\min\big(\frac{1}{2},\frac{\lambda_*^2 a}{T}\big)B^2_{t_l}(x_i,r_i)\Big\}.$$

(H3): Following the exact proof of Moynot and Samuelides [178], with constants $q = \frac{a\lambda_*^2}{\sqrt{2}T}\sqrt{\frac{N}{k}}$, and $\varepsilon = \sqrt{2\frac{k}{N}}$, we obtain

$$\tilde{a}_i^\delta(\mathbf{x},\mathbf{r}) \leq \exp\Big\{\big(\frac{1}{2q} + \frac{a\lambda_*^2}{2Tq}\big)\sum_{l=1}^{\frac{1}{\delta}} B^2_{t_l}(x_i,r_i)\Big\}\big(\frac{D_0}{A}\big)^{\frac{1}{q}}\exp\Big\{\frac{B\sqrt{\delta}}{q}\sum_{l=1}^{\frac{1}{\delta}}|B_{t_l}(x_i,r_i)|\Big\}.$$

Remark that $\sum_{l=1}^{\frac{1}{\delta}}\sqrt{\delta}|B_{t_l}(x_i,r_i)| \leq \frac{1}{2}\big(1+\sum_{l=1}^{\frac{1}{\delta}} B^2_{t_l}(x_i,r_i)\big)$, so that taking q large enough, i.e. $\frac{k}{N}$ small enough, yields the result.
(H4):

$$b_i^\delta(\mathbf{x},\mathbf{r}) = \exp\Big\{\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}} B^2_{t_l}(x_i,r_i)\Big\}\mathscr{E}_J\left(\exp\Big\{-\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}}\Big(\sqrt{\delta T}\tilde{G}^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r}) - B_{t_l}(x_i,r_i)\Big)^2\Big\}\right).$$

Under $\mathscr{E}_J$, $U_i^\delta(\mathbf{x},\mathbf{r}) := \Big(\sqrt{\delta T}\tilde{G}^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r})\Big)_{l\in[\![1,\frac{1}{\delta}]\!]}$ is a Gaussian vector of mean $\bar{U}_i^\delta(\mathbf{x},\mathbf{r}) := \big(\sqrt{\delta T}m_{\hat{\mu}_N}(t^{(l)},r_i)\big)_{l\in[\![1,\frac{1}{\delta}]\!]}$ and variance-covariance matrix $\Sigma_i^\delta(\mathbf{x},\mathbf{r}) := \Big(\delta T K_{\hat{\mu}_N}(t_l,t_m,r_i)\Big)_{(l,m)\in[\![1,\frac{1}{\delta}]\!]^2}$.
Let

$$A^2_{N,\delta} := \left\{(\mathbf{x},\mathbf{r})\in\big(\mathscr{C}\times D\big)^N, \forall Y\in\mathbb{R}^{\frac{1}{\delta}}\setminus\{0\}, \forall i\in[\![1,N]\!], \exists j\in[\![1,N]\!], \Big(\sum_{l=1}^{\frac{1}{\delta}} y_l S(x^j_{t_l-\tau(r_i,r_j)})\Big)^2 > 0\right\}. \tag{4.62}$$

As

$$ {}^tY\Sigma_i^\delta(\mathbf{x},\mathbf{r})Y = \frac{\delta T}{\lambda(r_i)^2 N}\sum_{j=1}^N \sigma(r_i,r_j)^2\Big(\sum_{l=1}^{\frac{1}{\delta}} y_l S(x^j_{t_l-\tau(r_i,r_j)})\Big)^2, $$

the matrix Σ_i^δ is positive definite on $A^2_{N,\delta}$. Being on this set basically mean that the N random vectors $\big(S(x^j_{t_l-\tau(r_i,r_j)})\big)_{l\in[\![1,\frac{1}{\delta}]\!]}, 1\le j\le N$ are not contained in an hyperplane of $\mathbb{R}^{\frac{1}{\delta}}$. For $N\ge\frac{1}{\delta}$, $P^{\otimes N}\big(A^2_{N,\delta}\big)=1$ as, under $P_{\mathbf{r}}$, the x^j are independent semi-martingales. Let $B:=\big(B_{t_l}\big)_{l\in[\![1,\frac{1}{\delta}]\!]}$. Hence, denoting $\big\|X\big\|_{2,\delta}=\sqrt{\sum_{l=1}^{\frac{1}{\delta}}x_l^2}$ for any $X\in\mathbb{R}^{\frac{1}{\delta}}$, we have

$$ b_i^\delta(\mathbf{x},\mathbf{r}) = \exp\Big\{\frac{1}{2}\big\|B\big\|^2_{2,\delta}(x_i,r_i)\Big\}\mathscr{E}_J\bigg(\exp\Big\{-\frac{1}{2}\big\|U_i^\delta(\mathbf{x},\mathbf{r})-\bar U_i^\delta(\mathbf{x},\mathbf{r})+\bar U_i^\delta(\mathbf{x},\mathbf{r})-B(x_i,r_i)\big\|^2_{2,\delta}\Big\}\bigg), $$

$$ \begin{aligned} b_i^\delta(\mathbf{x},\mathbf{r}) =& \exp\Big\{\frac{1}{2}\big\|B\big\|^2_{2,\delta}(x_i,r_i)\Big\}\exp\Big\{-\frac{1}{2}\big\|\bar U_i^\delta(\mathbf{x},\mathbf{r})-B(x_i,r_i)\big\|^2_{2,\delta}\Big\}\\ &\times\int_{\mathbb{R}^{\frac{1}{\delta}}}\exp\Big\{-X.\Big(\bar U_i^\delta(\mathbf{x},\mathbf{r})-B(x_i,r_i)\Big)\Big\}\frac{\exp\Big\{-\frac{1}{2}X.\Big(I_{\frac{1}{\delta}}+(\Sigma_i^\delta(\mathbf{x},\mathbf{r}))^{-1}\Big).X\Big\}}{\sqrt{(2\pi)^N\det\big(\Sigma_i^\delta(\mathbf{x},\mathbf{r})\big)}}dX,\\ =&\sqrt{\det\Big(I_{\frac{1}{\delta}}+\Sigma_i^\delta(\mathbf{x},\mathbf{r})\Big)^{-1}}\exp\Big\{\frac{1}{2}\big\|B\big\|^2_{2,\delta}(x_i,r_i)\Big\}\\ &\times\exp\Big\{\frac{1}{2}\big(\bar U_i^\delta(\mathbf{x},\mathbf{r})-B(x_i,r_i)\big).\Big(A_i^\delta(\mathbf{x},\mathbf{r})-I_{\frac{1}{\delta}}\Big).\big(\bar U_i^\delta(\mathbf{x},\mathbf{r})-B(x_i,r_i)\big)\Big\}, \end{aligned} $$

where $A_i^\delta(\mathbf{x},\mathbf{r})=\Sigma_i^\delta(\mathbf{x},\mathbf{r})\Big(I_{\frac{1}{\delta}}+\Sigma_i^\delta(\mathbf{x},\mathbf{r})\Big)^{-1}$ is a definite postive matrix with eigenvalues strictly smaller than 1. Consequently,

$$ (b_i^\delta(\mathbf{x},\mathbf{r}))^{-1}\le\underbrace{\sqrt{\det\Big(I_{\frac{1}{\delta}}+\Sigma_i^\delta(\mathbf{x},\mathbf{r})\Big)}\exp\Big\{\frac{1}{2}\big\|\bar U_i^\delta(\mathbf{x},\mathbf{r})-B(x^i,r_i)\big\|^2_{2,\delta}\Big\}}_{=:N_i(\mathbf{x},\mathbf{r})}\exp\Big\{-\frac{1}{2}\big\|B\big\|^2_{2,\delta}(x_i,r_i)\Big\}. $$

As

$$ |a_i^\delta(\mathbf{x},\mathbf{r})-b_i^\delta(\mathbf{x},\mathbf{r})| = e^{\Big\{\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}}B^2_{t_l}(x_i,r_i)\Big\}}O_i(\mathbf{x},\mathbf{r}), $$

with

$$ O_i(\mathbf{x},\mathbf{r}) := \bigg|\mathscr{E}_J\bigg(e^{\Big\{-\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}}\Big(\sqrt{\delta T}\hat G^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r})-B_{t_l}(x_i,r_i)\Big)^2\Big\}}-e^{\Big\{-\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}}\Big(\sqrt{\delta T}\tilde G^{\delta,i}_{t^{(l)}}(\mathbf{x},\mathbf{r})-B_{t_l}(x_i,r_i)\Big)^2\Big\}}\bigg)\bigg| $$

then

$$ \Big|\frac{a_i^\delta(\mathbf{x},\mathbf{r})-b_i^\delta(\mathbf{x},\mathbf{r})}{b_i^\delta(\mathbf{x},\mathbf{r})}\Big|\le O_i(\mathbf{x},\mathbf{r})N_i(\mathbf{x},\mathbf{r}). $$

Remark that

$$ \sqrt{\det\Big(I_{\frac{1}{\delta}}+\Sigma_i^\delta(\mathbf{x},\mathbf{r})\Big)}=\exp\Big\{\frac{1}{2}\sum_{\lambda\in sp\left(\Sigma_i^\delta(\mathbf{x},\mathbf{r})\right)}\log(1+\lambda)\Big\}\le\exp\Big\{\frac{1}{2}tr\big(\Sigma_i^\delta(\mathbf{x},\mathbf{r})\big)\Big\}\le\exp\Big\{\frac{\|\sigma\|^2_\infty T}{2\lambda_*^2}\Big\}, $$

so that

$$ N_i(\mathbf{x},\mathbf{r})\le\exp\Big\{T\frac{\|\sigma\|^2_\infty+\|J\|^2_\infty}{2\lambda_*^2}\Big\}\exp\Big\{\frac{1}{2}\sum_{l=1}^{\frac{1}{\delta}}\Big(B_{t_l}(x^i,r_i)^2+2\sqrt{\delta T}\frac{\|J\|_\infty}{\lambda_*}\big|B_{t_l}(x^i,r_i)\big|\Big)\Big\}. $$

To obtain an upperbound for O_i we rely on [178, Lemma 4.2] where the fixed and finite $\frac{1}{\delta}$ corresponds to their T. Following its proof, we define the function:

$$\Phi_{\mathbf{x},\mathbf{r}}(y_1,\dots,y_{\frac{1}{\delta}}) = \prod_{l=1}^{\frac{1}{\delta}} \phi\big(y_l + a_l(\mathbf{x},\mathbf{r})\big),$$

where $\phi(z) := \exp -\frac{z^2}{2}$ and $a_l(\mathbf{x},\mathbf{r}) := \frac{\sqrt{\delta T}}{\lambda(r_i)} \sum_{j=1}^N \frac{J(r_i,r_j)}{N} S\big(x^j_{t^{(l)}-\tau(r_i,r_j)}\big) - B_{t_l}(x^i, r_i)$. One easily sees that the three first derivatives of $\Phi_{\mathbf{x},\mathbf{r}}$ are bounded by some constant C_3 independent of δ and $(\mathbf{x},\mathbf{r})$. Let

$$V_j(\mathbf{x},\mathbf{r}) = \left(\frac{\sqrt{\delta T}}{\lambda(r_i)}\Big(J_{ij} - \frac{J(r_i,r_j)}{N}\Big) S\big(x^j_{t^{(l)}-\tau(r_i,r_j)}\big)\right)_{1\le l\le \frac{1}{\delta}},$$

$$\tilde V_j(\mathbf{x},\mathbf{r}) = \left(\frac{\sqrt{\delta T}}{\lambda(r_i)}\Big(\tilde J_{ij} - \frac{J(r_i,r_j)}{N}\Big) S\big(x^j_{t^{(l)}-\tau(r_i,r_j)}\big)\right)_{1\le l\le \frac{1}{\delta}},$$

so that

$$O_i(\mathbf{x},\mathbf{r}) = \left|\mathscr{E}_J\Big(\Phi_{\mathbf{x},\mathbf{r}}\Big(\sum_{j=1}^N V_j(\mathbf{x},\mathbf{r})\Big) - \Phi_{\mathbf{x},\mathbf{r}}\Big(\sum_{j=1}^N \tilde V_j(\mathbf{x},\mathbf{r})\Big)\Big)\right|.$$

Let $\varepsilon > 0$. Then

$$O_i(\mathbf{x},\mathbf{r}) \le C_3\left(\frac{\varepsilon}{6}\sum_{j=1}^N \mathscr{E}_J\Big(\big\|V_j(\mathbf{x},\mathbf{r})\big\|^2_{2,\delta}\Big) + \sum_{j=1}^N \mathscr{E}_J\big(\big\|\tilde V_j(\mathbf{x},\mathbf{r})\big\|^3_{2,\delta}\Big) + \sum_{j=1}^N \mathscr{E}_J\Big(\big\|V_j(\mathbf{x},\mathbf{r})\big\|^2_{2,\delta}\mathbf{1}_{\left\{\|V_j(\mathbf{x},\mathbf{r})\|_{2,\delta}\ge\varepsilon\right\}}\Big)\right).$$

But

$$\mathscr{E}_J\Big(\big\|V_j(\mathbf{x},\mathbf{r})\big\|^2_{2,\delta}\Big) \le \frac{\delta T}{\lambda_*^2}\mathscr{E}_J\left[\Big(J_{ij} - \frac{J(r_i,r_j)}{N}\Big)^2\right]\sum_{l=1}^{\frac{1}{\delta}} S\big(x^j_{t^{(l)}-\tau(r_i,r_j)}\big)^2 \le \frac{\|\sigma\|^2_\infty T}{N\lambda_*^2},$$

$$\mathscr{E}_J\Big(\big\|\tilde V_j(\mathbf{x},\mathbf{r})\big\|^3_{2,\delta}\Big) = \left(\frac{\sqrt{\delta T}}{\lambda(r_i)}\Big(\sum_{l=1}^{\frac{1}{\delta}} S\big(x^j_{t^{(l)}-\tau(r_i,r_j)}\big)^2\Big)^{\frac12}\right)^3 \mathscr{E}_J\Big(\Big|\tilde J_{ij} - \frac{J(r_i,r_j)}{N}\Big|^3\Big) \le \frac{\|\sigma\|^3_\infty T^{\frac32}}{N^{\frac32}\lambda_*^3}\mathscr{E}_J\Big(\big|\mathscr{N}(0,1)\big|^3\Big),$$

$$\mathscr{E}_J\Big(\big\|V_j(\mathbf{x},\mathbf{r})\big\|^2_{2,\delta}\mathbf{1}_{\{\|V_j(\mathbf{x},\mathbf{r})\|_{2,\delta}\ge\varepsilon\}}\Big) \le \frac{1}{N}\mathscr{E}_J\left(\frac{\delta TN}{\lambda_*^2}\Big(J_{ij} - \frac{J(r_i,r_j)}{N}\Big)^2 \mathbf{1}_{\left\{\frac{\delta T}{\lambda_*^2}\left(J_{ij} - \frac{J(r_i,r_j)}{N}\right)^2 \ge \varepsilon^2\right\}}\right).$$

Let $C_a = \sup_{x\ge0} x^2 \exp\big\{-a\frac{\lambda_*^2}{2T}x\big\}$. Then

$$\begin{aligned}\mathscr{E}_J\Big(\big\|V_j(\mathbf{x},\mathbf{r})\big\|^2_{2,\delta}\mathbf{1}_{\left\{\|V_j(\mathbf{x},\mathbf{r})\|_{2,\delta}\ge\varepsilon\right\}}\Big) &\le \frac{\delta\lambda_*^2 C_a}{TN^2}\mathscr{E}_J\left(\Big(J_{ij} - \frac{J(r_i,r_j)}{N}\Big)^{-2} e^{\frac{aN}{2}\left(J_{ij} - \frac{J(r_i,r_j)}{N}\right)^2}\mathbf{1}_{\left\{\frac{\delta T}{\lambda_*^2}(J_{ij} - \frac{J(r_i,r_j)}{N})^2 \ge \varepsilon^2\right\}}\right)\\ &\le \frac{C_a\delta^2}{N^2\varepsilon^2}\exp\Big\{a\frac{\|J\|^2_\infty}{N}\Big\}\mathscr{E}_J\Big(\exp\Big\{aNJ_{ij}^2\Big\}\Big).\end{aligned}$$

Choosing $\varepsilon = N^{-\frac14}$, and using hypothesis (H_J) yields the result.

CHAPTER 5

NETWORK WITH STATE-DEPENDENT SYNAPSES

We investigate the role of the interaction amplitude disorder on the thermodynamic limit of a class of particle systems arising in the modeling of neuronal networks and interacting oscillators. The originality of the model is that the directed impact of one particle onto another (i) depends on the state of both particles, and (ii) have an heterogeneous amplitude varying between the different cells considered. We focus here on the case where the scaling of the fluctuations of interaction amplitude disorder provides this parameter a non-trivial role, i.e. when the variance of the synaptic weights decays as the inverse of the network size. Under sufficient regularity assumptions, we show that the empirical measure satisfies a large-deviation principle with good rate function achieving its minimum at a unique probability measure. This technical results implies in particular the convergence of the empirical measure, of the law of a given particle, as well as a propagation of chaos property. The limit is characterized through a complex non Markovian implicit equation in which the network interaction term is replaced by a Gaussian field depending on the state of the particle.

5.1 INTRODUCTION

Interacting particle systems in random environments are ubiquitous in the theory of complex systems. They are useful to model a broad range of phenomena, from neural networks of the brain [221] to communication networks [30], internet traffic [122], disordered physical systems [22], economics and social science [188]. A particularly important model in life science is the Kuramoto model of coupled oscillators [152] which is one of the seldom models that are completely solvable in a mean-field scaling (see the outstanding review of Strogatz [227] and, more recently, [160]). All these models have in common to be described by a variable $(X_t^{i,N})_{i=1\cdots N} \in \mathbb{R}^N$ characterizing each particle's state, and that satisfies a stochastic differential equation of type:

$$dX_t^{i,N} = \left(f(r_i,t,X_t^{i,N}) + \sum_{j=1}^{N} J_{ij} b(X_t^{i,N}, X_t^{j,N}) \right) dt + \lambda dW_t^i, \tag{5.1}$$

where:

- the map f denotes each particle's intrinsic dynamics, which depends on time, on the state of the particle, as well as on its intrinsic properties taken into account through a disorder parameter $r_i \in D \subseteq \mathbb{R}^d$ that vary across the different particles;
- the terms $(J_{ij} b(x^i, x^j))$ represent the impact of particle j with state x^j onto the

particle i with state x^i. The amplitude of this interaction is modulated by the coefficient $J_{ij} \in \mathbb{R}$ that incorporates both the network topology and the variability of interactions intensity;

- the stochastic fluctuations are driven by independent standard real-valued Brownian motions $(W_t^i)_{t\geq 0}$, and $\lambda > 0$.

The main contribution of the present manuscript is to characterize the thermodynamic limit of these systems when the interaction coefficients $(J_{ij})_{i,j\in\{1,\cdots,N\}^2}$ have a slowly decaying variance, equivalent to σ^2/N when $N \to \infty$.

In contrast to the well-studied mean-field regime where the variance typically scales as $\mathcal{O}(1/N^2)$ [85, 161, 171, 231] and where the fluctuations of the interaction coefficients disappear in the limit, the present setting preserves a nontrivial contribution of the microscopic disorder of the J_{ij}. This regime has been studied for several decades: it is particularly rich, somewhat nonstandard, and still not fully understood. Notable advances in the characterization of such systems revealed that level of disorder σ has an important impact on the macroscopic behavior, governing the glassy transition in spin glass systems [216] or a transition to chaos in randomly connected neural networks [221]. The Kuramoto model also displays a transition as a function of the disorder level, the nature of which remains debated [76, 77, 224, 225].

Mathematically, important advances in the understanding of large systems in this scaling regime were achieved in the context of Langevin spin glass systems by Ben Arous, Dembo and Guionnet [14, 22, 23, 129]. These works are fundamental in the field in that they introduce a general methodology to characterize systems of bounded spins with linear interactions only depending on the state of the other particles (i.e., $b(x,y) = y$). They prove a Large Deviations Principle (LDP) for the empirical measure, averaged propagation of chaos and convergence towards a non-Markov implicit equation. The same technique was used in neuroscience in order to understand the dynamics of rate models with interactions of type $b(x,y) = S(y)$ with S a sigmoidal function. This was first performed in discrete-time systems [104, 178], and then extended to continuous time settings with spatial extension and delays [51, 52].

Here, we undertake the analysis of systems with general interactions $b(x,y)$. The main motivation for this extension is to address limit properties of a few biologically relevant models such as coupled oscillators (e.g., the canonical Kuramoto model [152]), collective animal behavior models [56] or neuronal networks [100, 117]. We will show that the empirical measure of system (5.1), averaged over the disorder parameters, satisfies a LDP, with an explicit good rate function that has a unique minimum. This approach readily proves convergence of the network equations towards a non-Markovian complex mean-field equation as in [14, 22, 23, 129]. Taking into account general interactions introduces a number of specific difficulties. In particular, it induces complex dependences between processes that prevents from isolating exponential martingales terms, which leads us to work on a finite time interval. Besides, the regularity assumptions of biological models allows to propose an alternative and somewhat more standard approach for showing existence and uniqueness of solutions to the limit system.

The paper is organized as follows. We start by introducing the mathematical setting and main results in Section 5.2. The proofs are found in the following sections. Section 5.3 establishes a partial LDP for the averaged empirical measure, which relies on the identification of the good rate function as well as on exponential tightness

and upper-bounds on closed sets for the sequence of empirical measures. In Section 5.4, we demonstrate that the good rate function admits a unique minimum Q and prove the averaged and quenched convergence of the empirical measure towards Q. We discuss these perspectives as well as a few open research directions in the conclusion.

These results will also be applied to the randomly connected Kuramoto model in Chapter 7. The study of the resulting equations exhibit the competition taking place between the mean and the variance of the interactions intensity for reaching synchrony. This new phenomenon may have applications in biology where interactions are generally highly heterogeneous.

5.2 Mathematical setting and main results

Throughout the paper, we will denote by $\mathscr{M}_1^+(\Sigma)$ for a Polish space Σ the set of Borel probability measures on Σ. The interacting particle system (5.1) is a diffusion in random environment, and as such involves two probability spaces:

- The disorder parameter of the particles as well as their interaction amplitudes are random variables on a complete probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathscr{P}})$. These heterogeneities are taken into account through:
  - independent identically distributed random variables $(r_i)_{i\in\mathbb{N}^*} \in D^{\mathbb{N}^*}$ with distribution $\pi \in \mathscr{M}_1^+(D)$ absolutely continuous with respect to Lebesgue's measure.
  - the interaction coefficients $J := (J_{ij})_{i,j\in\{1\cdots N\}}$ that are independent and identically distributed real Gaussian random variables with mean $\frac{\bar{J}}{N}$ and variance $\frac{\sigma^2}{N}$ for all $N \in \mathbb{N}^*$. These are assumed independent of the r_i parameters[1].

  These parameters constitute the random environment of the dynamic and are frozen in time. Their realization do not depend on the evolution of the system. We will denote $\tilde{\mathscr{E}}$ the expectation under $\tilde{\mathscr{P}}$, and by $\mathscr{E}_J$ and $\mathscr{P}_J$ the expectation and probability conditionally on the variables $(r_i)_{i\in\mathbb{N}^*}$ (i.e., over the variables J_{ij} only).

- The particles are driven by independent standard Brownian motions (W_t^i) on the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, \mathbb{P})$ satisfying the usual conditions.

The particles state, given by equation (5.1), thus depends both on the random environment and on the Brownian motions. We assume that the parameters of the equation driving the dynamics of the network enjoy the following regularity assumptions:

(i) The map $f : D \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ is K_f-Lipschitz continuous in all variables;

(ii) The map $b : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ is bounded ($\|b\|_\infty := \sup_{x,y\in\mathbb{R}\times\mathbb{R}} |b(x,y)| < \infty$) and K_b-Lipschitz-continuous in all variables.

[1] We emphasize that the law of the J_{ij} depends on N. This does not necessary require to draw distinct sequences of random variables for each N. For instance, one may define $J_{ij} = \frac{\bar{J}}{N} + \frac{\sigma}{\sqrt{N}}\xi_{ij}$ with ξ_{ij} a sequence of iid standard Gaussian random variables.

The initial conditions of the network will be considered independent for the different neurons. The law of the initial condition of a given neuron at position $r \in D$ may however depend on this heterogeneity parameter: we denote it by $\mu_0(r) \in \mathscr{M}_1^+(\mathbb{R})$. Moreover, we will assume that the collection of measures $\big(\mu_0(r)\big)_{r\in D}$ is continuous with respect to $r \in D$, in the sense that there exists a collection of random variable $(\bar{x}_0(r))_{r\in D}$ such that for any $r \in D$, $\bar{x}_0(r)$ has law $\mu_0(r)$ and, $r \mapsto \bar{x}_0(r)$ is continuous:

$$\forall r, r' \in D, \qquad \mathscr{L}(\bar{x}_0(r)) = \mu_0(r), \quad \lim_{\|r-r'\|_{\mathbb{R}^d}\to 0} |\bar{x}_0(r) - \bar{x}_0(r')| \to 0. \tag{5.2}$$

where $\|\cdot\|_{\mathbb{R}^d}$ denotes the Euclidean norm on $\mathbb{R}^d$. The network equations (5.1) are thus completed by the initial condition:

$$(X_0^i)_{i=1\cdots N} \overset{\mathscr{L}}{=} \bigotimes_{i=1}^{N} \mu_0(r_i), \tag{5.3}$$

The regularity assumptions on f and b classically ensure well-posedness of the network equation:

Proposition 5.2.1. For any given $J \in \mathbb{R}^{N\times N}$, $\mathbf{r} := (r_i)_{i\in\{1\cdots N\}} \in D^N$, and $T > 0$, there exists a unique weak solution to the system (5.1) defined on $[0,T]$ with initial condition (5.3). Moreover, this solution is square integrable.

Let $T > 0$, and $Q_{\mathbf{r}}^N(J)$ be this unique weak solution up to time T. $Q_{\mathbf{r}}^N(J)$ is a probability measure on $\mathscr{C}^N$, where $\mathscr{C} := \mathscr{C}\big([0,T],\mathbb{R}\big)$ is the set of continuous function from $[0,T]$ to $\mathbb{R}$. This solution depends on both realizations of J and $\mathbf{r}$. We are interested in proving a LDP for the double-layer empirical measure $\hat{\mu}_N \in \mathscr{M}_1^+(\mathscr{C}\times D)$:

$$\hat{\mu}_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{(X^{i,N},r_i)}, \tag{5.4}$$

where $\delta_{(x,r)} \in \mathscr{M}_1^+(\mathscr{C}\times D)$ denotes the degenerate probability measure at $(x,r) \in \mathscr{C}\times D$. In the proof of the LDP, it will be useful to consider the system of N particles with no interaction, i.e. system (5.1) with $J_{ij} = 0$ for all $(i,j) \in \{1\cdots N\}^2$. In this case, the law of a node with heterogeneity parameter $r \in D$ is given by the unique solution of the one-dimensional SDE:

$$\begin{cases} dX_t = f(r,t,X_t)dt + \lambda dW_t \\ X_0 \overset{\mathscr{L}}{=} \mu_0(r). \end{cases} \tag{5.5}$$

We denote by $P_r \in \mathscr{M}_1^+(\mathscr{C})$ the law of the solution up to time T. This family of probabilities $(P_r)_{r\in D}$ is regular in r (see Appendix B. of chapter 4), and we can thus introduce the probability measure $P \in \mathscr{M}_1^+\big(C\times D\big)$ defined, for any couple (A,B) of Borel sets of $\mathscr{C}$ and D respectively, by the equality:

$$P(A,B) := \int_B P_r(A)d\pi(r).$$

We will denote $dP(x,r) = dP_r(x)d\pi(r)$. Under $P^{\otimes N}$, state and heterogeneity parameter of the particles are i.i.d. so that Sanov's theorem ensures the existence of a full LDP for the double-layer empirical measure of unconnected particles, with good

rate function given by the relative entropy[2] $I(.|P)$. Following the approach proposed in [22, 129], we will rely on this LDP to derive an analogous result for the interacting system.

A direct application of Girsanov's theorem yields that $Q_{\mathbf{r}}^N(J)$ is absolutely continuous with respect to $P_{\mathbf{r}} := \bigotimes_{i=1}^N P_{r_i}$, with density:

$$\frac{\mathrm{d}Q_{\mathbf{r}}^N(J)}{\mathrm{d}P_{\mathbf{r}}}(\mathbf{x}) = \exp\left(\sum_{i=1}^N \left\{ \int_0^T \Big(\frac{1}{\lambda}\sum_{j=1}^N J_{ij} b(x_t^i, x_t^j)\Big) dW_t(x^i, r_i) - \frac{1}{2}\int_0^T \Big(\frac{1}{\lambda}\sum_{j=1}^N J_{ij} b(x_t^i, x_t^j)\Big)^2 dt \right\}\right), \tag{5.6}$$

where $W_t(x,r) := \frac{x_t - x_0}{\lambda} - \int_0^t \frac{f(r,s,x_s)}{\lambda} ds$, $\forall (x,r) \in \mathscr{C} \times D$ and $t \in [0,T]$, so that $W(.,r)$ is a P_r-Brownian motion. Moreover, as done in Appendix B. of chapter 4, we can properly introduce the averaged probability measure $Q^N \in \mathscr{M}_1^+\big((\mathscr{C} \times D)^N\big)$ define for any Borel sets $A \subset \mathscr{C}^N$, and $B \subset D^N$ by

$$Q^N(A \times B) := \int_B \mathscr{E}_J\big(Q_{\mathbf{r}}^N(J)(A)\big) d\pi^{\otimes N}(\mathbf{r}) = \int_{D^N} \mathscr{E}_J\left[\int_{\mathscr{C}^N} \mathbb{1}_{A\times B}(\mathbf{x},\mathbf{r}) dQ_{\mathbf{r}}^N(J)(\mathbf{x})\right] d\pi^{\otimes N}(\mathbf{r}).$$

Remark 11. We have, for any Borel set $A \subset \mathscr{C}^N$,

$$Q^N(A \times D^N) = \mathscr{E}\big[Q_{\mathbf{r}}^N(J)(A)\big].$$

Our results will hold under the condition that the horizon time T is such that

$$\frac{2\sigma^2 \|b\|_\infty^2 T}{\lambda^2} < 1. \tag{5.7}$$

These may be summarized as follows:

Theorem 5.2.2. *For T small enough for inequality* (5.7) *to hold, there exists a double-layer probability distribution $Q \in \mathscr{M}_1^+(\mathscr{C} \times D)$ such that:*

$$Q^N(\hat{\mu}_N \in \cdot) \xrightarrow{\mathscr{L}} \delta_Q(\cdot) \in \mathscr{M}_1^+\big(\mathscr{M}_1^+(\mathscr{C} \times D)\big),$$

exponentially fast.

The existence of Q and the exponential convergence results follow from three points: (i) the exponential tightness of the sequence $Q^N\big(\hat{\mu}_N \in \cdot\big)$, (ii) a partial LDP for the empirical measure relying on an upper-bound for closed sets, and (iii) a characterization of the set of minima of the good rate function.

Theorem 5.2.3 (Partial Large Deviation Principle)**.**
For T small enough for inequality (5.7) *to hold,*

(i) *for any real number $M \in \mathbb{R}$, there exists a compact subset K_M such that for any integer N,*

$$\frac{1}{N}\log Q^N(\hat{\mu}_N \notin K_M) \leq -M.$$

[2] We recall that if Σ is a Polish space, the relative entropy of $\nu \in \mathscr{M}_1^+(\Sigma)$ with respect to $\mu \in \mathscr{M}_1^+(\Sigma)$ is defined by:

$$I(\nu|\mu) := \begin{cases} \int_\Sigma \log\big(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(x)\big) d\nu(x) & \text{if } \nu \ll \mu, \\ \infty & \text{otherwise .} \end{cases}$$

(ii) *there exists a good rate function* $H : \mathscr{M}_1^+(\mathscr{C}\times D)$ *such that for any closed subset* F *of* $\mathscr{M}_1^+(\mathscr{C}\times D)$*:*

$$\limsup_{N\to\infty}\frac{1}{N}\log Q^N(\hat{\mu}_N\in F)\leq -\inf_F H.$$

This theorem is proved in section 5.3.

Theorem 5.2.4 (Minima of the rate function)**.** *The good rate function* H *achieves its minimal value at a unique probability measure* $Q\in\mathscr{M}_1^+(\mathscr{C}\times D)$ *satisfying:*

$$Q\simeq P,\qquad \frac{dQ}{dP}(x,r)=\mathscr{E}\left[\exp\left\{\frac{1}{\lambda}\int_0^T G_t^Q(x)dW_t(x,r)-\frac{1}{2\lambda^2}\int_0^T (G_t^Q(x))^2dt\right\}\right]$$

where $(W_t(.,r))_{t\in[0,T]}$ *is a* P_r*-Brownian motion, and* $G^Q(x)$ *is a* $(\tilde{\Omega},\tilde{\mathscr{F}},\mathscr{P})$*, a Gaussian process with mean:*

$$\mathscr{E}[G_t^Q(x)]=\int_{\mathscr{C}\times D}\bar{J}b(x_t,y_t)dQ(y,r')$$

and covariance:

$$\mathscr{E}[G_t^Q(x)G_s^Q(x)]=\int_{\mathscr{C}\times D}\sigma^2 b(x_t,y_t)b(x_s,y_s)dQ(y,r').$$

This theorem will be demonstrated in section 5.4. Combining both results, the general result of Sznitman [230, Lemma 3.1] implies that:

Theorem 5.2.5 (Propagation of chaos)**.** *For* T *small enough for inequality* (5.7) *to hold,* Q^N *is* Q*-chaotic in the sense that for any* $m\in\mathbb{N}^*$*, any collection of bounded continuous functions* $\varphi_1,\dots,\varphi_m:\mathscr{C}\times D\to\mathbb{R}$ *and any set of nonzero distinct integers* $k_1,\dots,k_m$*, we have:*

$$\lim_{N\to\infty}\int_{(\mathscr{C}\times D)^N}\prod_{j=1}^m\varphi_j(x^{k_j},r_{k_j})dQ^N(\mathbf{x},\mathbf{r})=\prod_{j=1}^m\int_{\mathscr{C}\times D}\varphi_j(x,r)dQ(x,r).$$

Our results partially extends to the quenched case as stated in the following theorem:

Theorem 5.2.6 (Quenched results)**.** *For* T *small enough for inequality* (5.7) *to hold, we have the following quenched upper-bound:*

$$\mathscr{P}-a.s.,\ \forall\ \text{closed } F\subset\mathscr{M}_1^+(\mathscr{C}\times D),\ \limsup_{N\to\infty}\frac{1}{N}\log Q_{\mathbf{r}}^N(J)(\hat{\mu}_N\in F)\leq -\inf_F H,$$

where H *is the good rate function introduced in theorem 5.2.3. In particular, for almost every realization of* $\mathbf{r}$ *and* J*,* $Q_{\mathbf{r}}^N(J)(\hat{\mu}_N\in\cdot)$ *is exponentially tight and converges in law toward* δ_Q *exponentially fast. Eventually, this implies joint* $\mathbb{P}$ *and* $\mathscr{P}$*-almost sure convergence of the empirical measure to* Q*.*

These theoretical results are applied to the disordered Kuramoto model in chapter 7 to provide an alternative representation to the ones proposed in the physics literature [76, 77, 224, 225], and allow investigating the interplay between the mean and variance parameters of the interactions coefficients on the synchronization of oscillators.

5.3 LARGE DEVIATION PRINCIPLE

This section is devoted to proving the existence of a partial large deviations principle for the averaged empirical measure. We start by constructing the appropriate good rate function before obtaining an upper-bound and an exponential tightness result. Many points of the proof proceed as in precedent works [22, 51, 52, 129]. To avoid reproducing fastidious demonstrations we will often refer to precedent contributions and focus our attention on the new difficulties arising in our setting.

5.3.1 Construction of the good rate function

For $\mu \in \mathscr{M}_1^+(\mathscr{C}\times D)$, we define the two following functions respectively on $[0,T]^2\times\mathscr{C}$ and $[0,T]\times\mathscr{C}$:

$$\begin{cases} K_\mu(s,t,x) & := \dfrac{\sigma^2}{\lambda^2}\displaystyle\int_{\mathscr{C}\times D} b(x_t,y_t)b(x_s,y_s)d\mu(y,r') \\ m_\mu(t,x) & := \dfrac{\bar{J}}{\lambda}\displaystyle\int_{\mathscr{C}\times D} b(x_t,y_t)d\mu(y,r'). \end{cases}$$

Both functions are well defined as $(y,r)\to b(x_t,y_t)b(x_s,y_s)$ and $(y,r)\to b(x_t,y_t)$ are continuous for $\|\cdot\|_\infty$ the uniform norm on $\mathscr{C}\times D$, and μ is a Borel measure. They are bounded: $|K_\mu(s,t,x)|\leq \frac{\sigma^2\|b\|_\infty^2}{\lambda^2}$ and $|m_\mu(t,x)|\leq \frac{\bar{J}\|b\|_\infty}{\lambda}$. Moreover, as μ charges continuous functions, K_μ and m_μ are continuous maps by the dominated convergence theorem.

Since K_μ has a covariance structure, we can define a probability space $(\hat{\Omega},\hat{\mathscr{F}},\gamma)$ and a family of stochastic processes $\left(G^\mu(x)\right)_{x\in\mathscr{C},\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)}$ continuous in x, and such that $G^\mu(x)$ is a centered Gaussian process with covariance $K_\mu(.,.,x)$ under measure γ. We denote $\mathscr{E}_\gamma$ the expectation under γ.

Remark 12.

- Let $\mu\in\mathscr{M}_1^+\left(\mathscr{C}\times D\right)$, and let $(e_i^\mu)_{i\in\mathbb{N}^*}$ be an orthonormal basis of $L^2_\mu\left(\mathscr{C}\times D\right)$. Let also for any $x\in\mathscr{C}, t\in[0,T]$, $\rho_{t,x}\in L^2_\mu\left(\mathscr{C}\times D\right)$ such that $\rho_{t,x}(y,r):=b(x_t,y_t)$. As stated in [129, Remark 2.14], a possible explicit construction for the $G^\mu(x)$ is given by

$$G_t^\mu(x):=\sum_{i\in\mathbb{N}} J_i\langle\rho_{t,x},e_i^\mu\rangle_{L^2_\mu(\mathscr{C}\times D)}=\sum_{i\in\mathbb{N}}J_i\int_{\mathscr{C}\times D}b(x_t,y_t)e_i^\mu(y,r)d\mu(y,r),$$

  where the $\left(J_i\right)_{i\in\mathbb{N}^*}$ are independent centered Gaussian variables of the probability space $(\hat{\Omega},\hat{\mathscr{F}},\gamma)$ and with variance σ^2.

- The family of processes $\left(G_t^\mu(x)\right)_{\mu,x}$ is intended to encompass possible candidates for the effective asymptotic interactions $\lim_N\left(\sum_{j=1}^N J_{ij}b(x_t^i,x_t^j)\right)_{i\in\mathbb{N}^*}$. In these interactions, the Gaussian weights are independent for different particles, so that it would seem more natural to consider independent families of Gaussian weights labeled by the particle $\left(J_i(x)\right)_{i\in\mathbb{N}^*,x\in\mathscr{C}}$ in the previous construction. Notably, we can swap from the continuous version to the independent one very easily in our proof. Indeed, the $G_t^\mu(x)$ are never taken jointly under γ, so that we can literally choose their covariance structure. For the sake of measurability under any Borel measure of $\mathscr{M}_1^+\left(\mathscr{C}\times D\right)$, we will exclusively work with the continuous version of $G^\mu(x)$ which proves most convenient.

We define for any $N \in \mathbb{N}^*$, $(\mathbf{x},\mathbf{r}) \in (\mathscr{C}\times D)^N$

$$X_i^N(\mathbf{x},\mathbf{r}) := \int_0^T G_t^{i,N}(\mathbf{x})dW_t(x^i,r_i) - \frac{1}{2}\int_0^T G_t^{i,N}(\mathbf{x})^2 dt$$

where $G_t^{i,N}(\mathbf{x}) := \frac{1}{\lambda}\sum_{j=1}^N J_{ij}b(x_t^i,x_t^j)$. As for the spatial case, we have the following expression for the density of the averaged network:

Lemma 5.3.1.

$$\frac{dQ^N}{dP^{\otimes N}}(\mathbf{x},\mathbf{r}) = \exp\Big\{N\bar{\Gamma}(\hat{\mu}_N)\Big\}.$$

where,

$$\bar{\Gamma}(\hat{\mu}_N) := \frac{1}{N}\sum_{i=1}^N \log\mathscr{E}_\gamma\Big[\exp\Big\{\int_0^T \big(G_t^{\hat{\mu}_N}(x^i)+m_{\hat{\mu}_N}(t,x^i)\big)dW_t(x^i,r_i) - \frac{1}{2}\int_0^T \big(G_t^{\hat{\mu}_N}(x^i)+m_{\hat{\mu}_N}(t,x^i)\big)^2 dt\Big\}\Big]. \tag{5.8}$$

Proof. Averaging the expression (5.6) on J and applying Fubini's theorem, we find that $Q^N \ll P^{\otimes N}$ with

$$\frac{dQ^N}{dP^{\otimes N}}(\mathbf{x},\mathbf{r}) = \mathscr{E}_J\Big[\frac{\mathrm{d}Q_{\mathbf{r}}^N(J)}{\mathrm{d}P_{\mathbf{r}}}(\mathbf{x})\Big] = \mathscr{E}_J\Big[\exp\Big\{\sum_{i=1}^N X_i^N(\mathbf{x},\mathbf{r})\Big\}\Big] = \prod_{i=1}^N \mathscr{E}_J\Big[\exp\big\{X_i^N(\mathbf{x},\mathbf{r})\big\}\Big],$$

by independence of the J_{ij}. Here, $\mathbf{x}$ is the coordinate process taken under $P^{\otimes N}$. It displays no dependence with the J_{ij}. To conclude, we remark that $\Big\{G_t^{i,N}(\mathbf{x}), 0\le t\le T\Big\}$ is, under $\mathscr{P}_J$, a Gaussian process with covariance $K_{\hat{\mu}_N}(t,s,x^i)$, and mean $m_{\hat{\mu}_N}(t,x^i)$. □

Following Varadhan's lemma, this motivates to introduce and precise the domain of definition of the maps:

$$X^\mu(x,r) := \int_0^T \big(G_t^\mu(x)+m_\mu(t,x)\big)dW_t(x,r) - \frac{1}{2}\int_0^T \big(G_t^\mu(x)+m_\mu(t,x)\big)^2 dt,$$

and

$$\Gamma(\mu) := \int_{\mathscr{C}\times D} \log\mathscr{E}_\gamma\Big[\exp\big\{X^\mu(x,r)\big\}\Big]d\mu(x,r),$$

for all $(x,r)\in\mathscr{C}\times D$ and $\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)$.

Proposition 5.3.2. The map

$$\Gamma := \mu\in\mathscr{M}_1^+\big(\mathscr{C}\times D\big) \to \begin{cases} \int_{\mathscr{C}\times D}\log\mathscr{E}_\gamma\Big[\exp\big\{X^\mu(x,r)\big\}\Big]d\mu(x,r) & \text{if } I(\mu|P)<\infty,\\ +\infty & \text{otherwise .}\end{cases} \tag{5.9}$$

is well defined in $\mathbb{R}\cup\{+\infty\}$, and satisfies

(i) $\Gamma \le I(\cdot|P)$,

(ii) If $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$, $\exists \iota\in]0,1[, e\ge 0$, $|\Gamma(\mu)| \le \iota I(\mu|P)+e$.

Proof. If $I(\mu|P) = +\infty$ the result holds. We will thus suppose that $I(\mu|P) < +\infty$, which implies $\mu \ll P$. As $W(\cdot, r)$ is a P_r-Brownian motion, Girsanov's theorem ensures that the stochastic integral $\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big) dW_t(x,r)$ is well defined γ-almost surely under μ.

(1):

Let $F_\mu := \log \mathscr{E}_\gamma\Big[\exp\big\{X^\mu(x,r)\big\}\Big]$ denote the integrand in the formulation of Γ (5.9). It is measurable as a continuous function of $(x,r) \to \big(K_\mu(t,s,x), 0 \le t,s \le T\big), \big(m_\mu(t,x), 0 \le t \le T\big), \big(W_t(x,r), 0 \le t \le T\big)$ that are continuous maps. Nevertheless, because of the mean term m_μ it is not bounded from below, as was the case in [22]. Let us prove that it is still μ-integrable. In fact,

$$-F_\mu^-(x,r) \le F_\mu^+(x,r) - F_\mu^-(x,r) = F_\mu(x,r) \le \log\Big(\mathscr{E}_\gamma\big[\exp\{X^\mu(x,r)\}\big] \vee M^{-1}\Big) =: F_{\mu,M}(x,r),$$

where F_μ^+ and F_μ^- respectively denote the positive and negative part of F_μ. As F_μ^- and $F_{\mu,M}$ are measurable and bounded from below, inequality (4.13) applies. Let $\alpha \ge 1$. On the one hand

$$\begin{aligned}
\alpha \int_{\mathscr{C}\times D} F_{\mu,M}(x,r) d\mu(x,r) &\le I(\mu|P) + \log\left\{ \int_{\mathscr{C}\times D} \exp\Big\{\alpha F_{\mu,M}(x,r)\Big\} dP(x,r) \right\} \\
&\overset{\text{Jensen}}{\le} I(\mu|P) + \log\left\{ M^{-\alpha} + \int_{\mathscr{C}\times D} \mathscr{E}_\gamma\Big[\exp\big\{\alpha X^\mu(x,r)\big\}\Big] dP(x,r) \right\} \\
&\overset{\text{Fubini}}{\le} I(\mu|P) + \log\left\{ M^{-\alpha} + \mathscr{E}_\gamma\left[\int_D \int_{\mathscr{C}} \exp\Big\{\alpha X^\mu(x,r)\Big\} dP_r(x) d\pi(r) \right] \right\},
\end{aligned} \tag{5.10}$$

with the right-hand side of the two latter inequalities being possibly infinite. On the other hand,

$$\begin{aligned}
\alpha \int_{\mathscr{C}\times D} F_\mu^-(x,r) d\mu(x,r) &= \alpha \int_{\mathscr{C}\times D} \Big(-\log \mathscr{E}_\gamma\big[\exp\big\{X^\mu(x,r)\big\}\big] \Big)^+ d\mu(x,r) \\
&\overset{\text{Jensen}}{\le} \alpha \int_{\mathscr{C}\times D} \Big(-\mathscr{E}_\gamma\big[X^\mu(x,r)\big] \Big)^+ d\mu(x,r) = \int_{\mathscr{C}\times D} \left(\mathscr{E}_\gamma\left[-\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big) dW_t(x,r) \right.\right. \\
&\left.\left. - \frac{1}{2} \int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big)^2 dt \right] + \mathscr{E}_\gamma\left[\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big)^2 dt \right] \right)^+ d\mu(x,r) \\
&\overset{(4.13)}{\le} I(\mu|P) + \log\left\{ \int_{\mathscr{C}\times D} \exp\left\{ \alpha \left(\mathscr{E}_\gamma\big[X^\mu(x,r)\big] + T \frac{(J^2+\sigma^2)\|b\|_\infty^2}{\lambda^2} \right)^+ \right\} dP(x,r) \right\} \\
&\overset{\text{Jensen, Fubini}}{\le} I(\mu|P) + \alpha C_T + \log\left\{ \mathscr{E}_\gamma\left[\int_D \int_{\mathscr{C}} \exp\Big\{\alpha X^\mu(x,r)\Big\} dP_r(x) d\pi(r) \right] \right\}.
\end{aligned} \tag{5.11}$$

Moreover, $W(.,r)$ being a P_r-Brownian motion, the martingale property yields for $\alpha = 1$

$$\mathscr{E}_\gamma\left[\int_D \int_{\mathscr{C}} \exp\Big\{\alpha X^\mu(x,r)\Big\} dP_r(x) d\pi(r) \right] = 1,$$

so that we can see that F_μ is μ-integrable, with

$$\int_{\mathscr{C}\times D} |F_\mu(x,r)| d\mu(x,r) = \int_{\mathscr{C}\times D} F_\mu^-(x,r) + F_{\mu,1}(x,r) d\mu(x,r) \le 2I(\mu|P) + C_T + \log(2). \tag{5.12}$$

Moreover,

$$\Gamma(\mu) := \int_{\mathscr{C}\times D} F_\mu(x,r) d\mu(x,r) \overset{(5.10)}{\le} I(\mu|P) + \log\big\{M^{-1} + 1\big\},$$

so that letting $M \to +\infty$ yields the result.

(2):

Let $\alpha > 1$. As $G^\nu(x)$ cannot be extracted from the integral on $dP_r(x)$, we rely on Hölder inequality with conjugate exponents $(\alpha, \frac{\alpha}{\alpha-1})$ to make use of a martingale property:

$$\int_{\mathscr{C}\times D} \mathscr{E}_\gamma\Big[\exp\big\{\alpha X^\nu(x,r)\big\}\Big] dP(x,r) \le \bigg\{ \int_{\mathscr{C}\times D} \mathscr{E}_\gamma\bigg[\exp\Big\{\frac{\alpha^2(\alpha+1)}{2}\int_0^T \big(G_t^\nu(x)+m_\nu(t,x)\big)^2 dt\Big\}\bigg] dP(x,r)\bigg\}^{\frac{\alpha-1}{\alpha}}$$
$$\times \bigg\{ \int_{\mathscr{C}\times D} \mathscr{E}_\gamma\bigg[\exp\Big\{\alpha^2\int_0^T \big(G_t^\nu(x)+m_\nu(t,x)\big) dW_t(x,r) - \frac{\alpha^4}{2}\int_0^T \big(G_t^\nu(x)+m_\nu(t,x)\big)^2 dt\Big\}\bigg] dP(x,r)\bigg\}^{\frac{1}{\alpha}}.$$

The second term in the product of the right-hand side is equal to 1 by Fubini's theorem and the martingale property. The short-time hypothesis $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$ ensures finiteness of the second term for $\alpha - 1$ small enough. Indeed, by Jensen and Fubini's inequalities, we have

$$\mathscr{E}_\gamma\bigg[\exp\Big\{\frac{\alpha^2(\alpha+1)T}{2}\int_0^T \big(G_t^\nu(x)+m_\nu(t,x)\big)^2 \frac{dt}{T}\Big\}\bigg] \le \int_0^T \mathscr{E}_\gamma\bigg[\exp\Big\{\frac{\alpha^2(\alpha+1)T}{2}\big(G_t^\nu(x)+m_\nu(t,x)\big)^2\Big\}\bigg]\frac{dt}{T}.$$

Moreover, since $\sqrt{\alpha^2(\alpha+1)T}\Big(G_t^\nu(x)+m_\nu(t,x)\Big) \sim \mathscr{N}\Big(\sqrt{\alpha^2(\alpha+1)T}m_\nu(t,x), \alpha^2(\alpha+1)TK_\nu(t,t,x)\Big)$ under γ, we are able, for $\alpha-1$ small enough and under the short time hypothesis, to use (8.1). We thus conclude that as soon as $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$, there exists a constant c_T, uniform in $x \in \mathscr{C}$, such that:

$$\int_{\mathscr{C}\times D} \mathscr{E}_\gamma\Big[\exp\big\{\alpha X^\nu(x,r)\big\}\Big] dP(x,r) \le \exp\big\{(\alpha-1)c_T\big\}. \tag{5.13}$$

Inequalities (5.10), (5.11), and (5.13) ensure that, for $\alpha > 1$ and under the condition $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$

$$|\Gamma(\mu)| \le \iota I(\mu|P) + e,$$

with $\iota := \frac{1}{\alpha}$, and $e := (2\alpha-1)C_T$. □

Let

$$H(\mu) := \begin{cases} I(\mu|P) - \Gamma(\mu) & \text{if } I(\mu|P) < \infty, \\ \infty & \text{otherwise}, \end{cases}$$

for any $\nu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big)$:

$$\Gamma_\nu := \mu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big) \to \begin{cases} \int_{\mathscr{C}\times D} \log \mathscr{E}_\gamma\Big[\exp\big\{X^\nu(x,r)\big\}\Big] d\mu(x,r) & \text{if } I(\mu|P) < \infty, \\ +\infty & \text{otherwise}, \end{cases}$$

as well as the following probability measure on $\mathscr{C}\times D$

$$dQ_\nu(x,r) := \exp\big\{\bar\Gamma_\nu(\delta_{(x,r)})\big\} dP(x,r) := \mathscr{E}_\gamma\Big[\exp\Big\{X^\nu(x,r)\Big\}\Big] dP(x,r). \tag{5.14}$$

A key observation is that

$$dQ_\nu^{\otimes N}(\mathbf{x},\mathbf{r}) = \Big(\exp\big\{\bar\Gamma_\nu(\delta_{(x,r)})\big\} dP(x,r)\Big)^{\otimes N} = \exp\big\{N\bar\Gamma_\nu(\hat\mu_N)\big\} dP^{\otimes N}(\mathbf{x}),$$

where

$$\bar\Gamma_\nu(\hat\mu_N) := \frac{1}{N}\sum_{i=1}^N \log \mathscr{E}_\gamma\Big[\exp\big\{X^\nu(x^i,r_i)\big\}\Big].$$

Sanov's theorem ensures that $Q_\nu^{\otimes N}\big(\hat{\mu}_N \in \cdot\big)$ satisfies a full LDP with good rate function $I(.|Q_\nu)$. Shall Varadhan's lemma apply, and $\sup_{\mathscr{M}_1^+(\mathscr{C}\times D)}\big\{\Gamma_\nu - I(\cdot|P)\big\} = 0$, the good rate function would also be given by the map H_ν, defined for any $\mu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big)$ by

$$H_\nu : \mu \to \begin{cases} I(\mu|P) - \Gamma_\nu(\mu) & \text{if } I(\mu|P) < +\infty, \\ +\infty & \text{otherwise.} \end{cases}$$

This will be proven in theorem 5.3.4. For any Gaussian process $(G_t)_{t\in[0,T]}$ of $\big(\hat{\Omega},\hat{\mathscr{F}},\gamma\big)$, and any $t \in [0,T]$ we define

$$\Lambda_t(G) := \frac{\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\}}{\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 du\Big\}\Big]}. \tag{5.15}$$

For any $t \in [0,T]$, $x \in \mathscr{C}$, and $\nu \in \mathscr{M}_1^+(\mathscr{C}\times D)$ the following defines a probability measure on $\big(\hat{\Omega},\hat{\mathscr{F}}\big)$ (see [183]):

$$d\gamma_{\widetilde{K}^t_{\nu,x}}(\omega) := \Lambda_t(G^\nu(\omega,x))d\gamma(\omega), \quad \forall \omega \in \hat{\Omega}.$$

As proven in [183], $G^\nu(x)$ is still a centered Gaussian process under $\gamma_{\widetilde{K}^t_{\nu,x}}$, with covariance given by:

$$\widetilde{K}^t_{\nu,x}(s,u) := \mathscr{E}_\gamma\Big[G_u^\nu(x)G_s^\nu(x)\Lambda_t\big(G^\nu(x)\big)\Big].$$

We also define for any $\nu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big)$, $(x,r) \in \mathscr{C}\times D$ and $t \in [0,T]$, the processes

$$L_t^\nu(x,r) := \int_0^t G_s^\nu(x)\Big(dW_s(x,r) - m_\nu(s,x)ds\Big), \quad V_t^\nu(x,r) := W_t(x,r) - \int_0^t m_\nu(s,x)ds. \tag{5.16}$$

Here are a few properties for these objects:

Proposition 5.3.3. Exists a constant $C_T > 0$, such that for any $\nu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big)$, $x \in \mathscr{C}$, $t \in [0,T]$,

$$\sup_{0\le s,u\le t} \widetilde{K}^t_{\nu,x}(s,u) \le C_T, \quad \Lambda_t\big(G^\nu(x)\big) \le C_T, \tag{5.17}$$

$$\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^T G_t^\nu(x)^2 dt\Big\}\Big] = \exp\Big\{-\frac{1}{2}\int_0^T \widetilde{K}^t_{\nu,x}(t,t)dt\Big\}. \tag{5.18}$$

Moreover, if $(G_t)_{0\le t\le T}$ and $(G'_t)_{0\le t\le T}$ are two centered Gaussian processes of $\big(\hat{\Omega},\hat{\mathscr{F}},\gamma\big)$ with uniformly bounded covariance, then exists $\tilde{C}_T > 0$ such that for all $t \in [0,T]$,

$$\big|\Lambda_t(G) - \Lambda_t(G')\big| \le \tilde{C}_T\Big\{\int_0^t \mathscr{E}_\gamma\Big[\big(G_s - G'_s\big)^2\Big]^{\frac{1}{2}} ds + \int_0^t \big|G_s^2 - {G'_s}^2\big| ds\Big\}. \tag{5.19}$$

Proof. Observe that by Jensen inequality:

$$\Lambda_t\big(G^\nu(x)\big) \le \mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^\nu(x)^2 du\Big\}\Big]^{-1} \overset{\text{Jensen}}{\le} \exp\Big\{\frac{1}{2}\int_0^t \mathscr{E}_\gamma\Big[G_s^\nu(x)^2\Big]du\Big\} \le \exp\Big\{\frac{\sigma^2\|b\|_\infty^2 t}{2\lambda^2}\Big\}.$$

As a consequence:

$$\widetilde{K}^t_{\nu,x}(s,u) = \mathscr{E}_\gamma\Big[G_u^\nu(x)G_s^\nu(x)\Lambda_t\big(G^\nu(x)\big)\Big] \overset{\text{C.S.}}{\le} \sqrt{K_\nu(s,s,x)K_\nu(t,t,x)}\exp\Big\{\frac{\sigma^2\|b\|_\infty^2 t}{2\lambda^2}\Big\} \le \frac{\sigma^2\|b\|_\infty^2}{\lambda^2}\exp\Big\{\frac{\sigma^2\|b\|_\infty^2 t}{2\lambda^2}\Big\},$$

For equality (5.18), let $f(t) := \mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^\nu(x)^2 ds\Big\}\Big]$. As $(t,\omega) \to G_t^\nu(\omega,x)^2 \exp\Big\{-\frac{1}{2}\int_0^t G_s^\nu(\omega,x)^2 ds\Big\}$ is a well defined, γ-a.s. continuous, and integrable under γ, we have

$$f'(t) = -\frac{1}{2}\mathscr{E}_\gamma\Big[G_t^\nu(x)^2 \exp\Big\{-\frac{1}{2}\int_0^t G_s^\nu(x)^2 ds\Big\}\Big] = -\frac{1}{2}\widetilde{K}_{\nu,x}^t(t,t) f(t),$$

so that integrating $\frac{f'}{f}$ gives the result. Furthermore, letting $(G_t)_{0\le t\le T}$ and $(G'_t)_{0\le t\le T}$ be two centered γ-Gaussian processes with variance bounded by a common constant C_T, we have:

$$\begin{aligned}
\big|\Lambda_t(G) - \Lambda_t(G')\big| &= \Big|\frac{\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\}}{\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\}\Big]} - \frac{\exp\Big\{-\frac{1}{2}\int_0^t {G'_s}^2 ds\Big\}}{\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t {G'_s}^2 ds\Big\}\Big]}\Big| \\
&\le \exp\Big\{\frac{\sigma^2\|b\|_\infty^2 t}{\lambda^2}\Big\}\Big\{\Big|\mathscr{E}_\gamma\Big[\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\} - \exp\Big\{-\frac{1}{2}\int_0^t {G'_s}^2 ds\Big\}\Big]\Big| \\
&+ \Big|\exp\Big\{-\frac{1}{2}\int_0^t G_s^2 ds\Big\} - \exp\Big\{-\frac{1}{2}\int_0^t {G'_s}^2 ds\Big\}\Big|\Big\}, \\
&\le \frac{1}{2}\exp\Big\{\frac{\sigma^2\|b\|_\infty^2 t}{\lambda^2}\Big\}\Big\{\int_0^t \mathscr{E}_\gamma\Big[\big|G_s^2 - {G'_s}^2\big|\Big] ds + \int_0^t \big|G_s^2 - {G'_s}^2\big| ds\Big\},
\end{aligned}$$

where we have used the Lipschitz-continuity of exponential on $\mathbb{R}_-$. Consequently, relying on Cauchy-Schwarz inequality, we obtain

$$\big|\Lambda_t(G) - \Lambda_t(G')\big| \overset{\text{C.S.}}{\le} \tilde{C}_T\Big\{\int_0^t \mathscr{E}_\gamma\Big[\big(G_s - G'_s\big)^2\Big]^{\frac{1}{2}} ds + \int_0^t \big|G_s^2 - {G'_s}^2\big| ds\Big\}.$$

□

We can show as in Theorem 4.3.6 this intuitive result:

Theorem 5.3.4. *Q_ν is a well defined probability measure on $\mathscr{M}_1^+(\mathscr{C}\times D)$, and the two maps H_ν and $I(.|Q_\nu)$ are equal on $\mathscr{M}_1^+(\mathscr{C}\times D)$. In particular H_ν is a good rate function reaching its unique minimum at Q_ν.*

We introduce the Vaserstein distance on $\mathscr{M}_1^+(\mathscr{C}\times D)$, compatible with the weak topology:

$$d_T^V(\mu,\nu) := \inf_\xi\Big\{\int_{(\mathscr{C}\times D)^2}\Big(\|x-y\|_{\infty,T}^2 + \|r-r'\|_{\mathbb{R}^d}^2\Big) d\xi\big((x,r),(y,r')\big)\Big\}^{\frac{1}{2}}$$

the infimum being taken on the laws $\xi \in \mathscr{C}_{\mu,\nu}$. Moreover we will denote for any $t \in [0,T]$, and any $(x,r),(y,r') \in \mathscr{C}\times D$,

$$d_t\big((x,r),(y,r')\big) := \Big(\|x-y\|_{\infty,t}^2 + \|r-r'\|_{\mathbb{R}^d}^2\Big)^{\frac{1}{2}},$$

and also

$$d_t^V(\mu,\nu) := \inf_\xi\Big\{\int_{(\mathscr{C}\times D)^2} d_t\big((x,r),(y,r')\big)^2 d\xi\big((x,r),(y,r')\big)\Big\}^{\frac{1}{2}}.$$

As in the delayed spatial case, the metric d_T^V will control the regularity of the mean and variance structure of the Gaussian interactions and, in the long run (see Theorem 5.3.6), of the error between H and its approximation H_ν:

Proposition 5.3.5. Exists $C_T > 0$ such that for any $\mu, \nu \in \mathscr{M}_1^+(\mathscr{C} \times D)$, $x \in C$, $t \in [0,T]$ and $u, s \in [0,t]$:

$$\big|m_\mu(t,x) - m_\nu(t,x)\big| + \big|K_\mu(t,s,x) - K_\nu(t,s,x)\big| + \big|\widetilde{K}^t_{\mu,x}(s,u) - \widetilde{K}^t_{\nu,x}(s,u)\big| \leq C_T d_t^V(\mu,\nu). \quad (5.20)$$

Proof. Let $\xi \in \mathscr{C}_{\mu,\nu}$, and let $\big(G, G'\big)$ be, under γ, a family of independent bi-dimensional centered Gaussian processes with covariance $K_\xi(s,t,x)$ given by:

$$\frac{\sigma^2}{\lambda^2} \int_{(\mathscr{C}\times D)^2} \begin{pmatrix} b(x_s,y_s)b(x_t,y_t) & b(x_s,y_s)b(x_t,z_t) \\ b(x_s,z_s)b(x_t,y_t) & b(x_s,z_s)b(x_t,z_t) \end{pmatrix} d\xi\big((y,r'),(z,\tilde{r}')\big). \quad (5.21)$$

Let us first take care of the mean difference:

$$\big|m_\mu(t,x) - m_\nu(t,x)\big| = \left|\frac{\bar{J}}{\lambda}\int_{\mathscr{C}\times D} b(x_t,y_t)d(\mu-\nu)(y,r')\right| \leq \frac{\bar{J}}{\lambda}\int_{(\mathscr{C}\times D)^2}\Big|b(x_t,y_t) - b(x_t,z_t)\Big| d\xi\big((y,r'),(z,\tilde{r}')\big)$$
$$\overset{\text{C.S.}}{\leq} \frac{\bar{J}K_b}{\lambda}\left\{\int_{(\mathscr{C}\times D)^2} \|y-z\|^2_{\infty,t} d\xi\big((y,r'),(z,\tilde{r}')\big)\right\}^{\frac{1}{2}}.$$

Moreover,

$$\big|K_\mu(t,s,x) - K_\nu(t,s,x)\big| = \Big|\mathscr{E}_\gamma\Big[G_sG_t - G'_sG'_t\Big]\Big| \overset{\text{C.S.}}{\leq} C_T\left\{\mathscr{E}_\gamma\Big[\big(G_t - G'_t\big)^2\Big]^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_s - G'_s\big)^2\Big]^{\frac{1}{2}}\right\}.$$

and

$$\big|\widetilde{K}^t_{\mu,x}(s,u) - \widetilde{K}^t_{\nu,x}(s,u)\big| \overset{(5.17)}{\leq} C_T\left\{\mathscr{E}_\gamma\Big[\big(\Lambda_t(G) - \Lambda_t(G')\big)^2\Big]^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_s - G'_s\big)^2\Big]^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_u - G'_u\big)^2\Big]^{\frac{1}{2}}\right\}$$
$$\overset{(5.19)}{\leq} C_T\left\{\left(\int_0^t \mathscr{E}_\gamma\Big[\big(G_v - G'_v\big)^2\Big]dv\right)^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_s - G'_s\big)^2\Big]^{\frac{1}{2}} + \mathscr{E}_\gamma\Big[\big(G_u - G'_u\big)^2\Big]^{\frac{1}{2}}\right\}.$$

$$\mathscr{E}_\gamma\Big[\big(G_t - G'_t\big)^2\Big] = \frac{\sigma^2}{\lambda^2}\int_{(\mathscr{C}\times D)^2}\Big(b(x_t,y_t) - b(x_t,z_t)\Big)^2 d\xi\big((y,r'),(z,\tilde{r}')\big)$$
$$\leq \frac{\sigma^2K_b^2}{\lambda^2}\int_{(\mathscr{C}\times D)^2} d_t\big((y,r'),(z,\tilde{r}')\big)^2 d\xi\big((y,r'),(z,\tilde{r}')\big).$$

Taking the infimum over $\xi \in \mathscr{C}_{\mu,\nu}$ yields (5.20). □

Theorem 5.3.6.

(i) *$\exists C_T > 0$, such that for every $\mu, \nu \in \mathscr{M}_1^+\big(\mathscr{C} \times D\big)$,*

$$|\Gamma_\nu(\mu) - \Gamma(\mu)| \leq C_T\big(1 + I(\mu|P)\big)d_T^V(\mu,\nu).$$

(ii) *If $\frac{2\sigma^2\|b\|^2_\infty T}{\lambda^2} < 1$, H is a good rate function.*

Proof. The basic mechanism for the proof is similar as in Proposition 4.3.8 or [15, Lemma 3.3-3.4]. However, the dependence in x of the Gaussian $G^\mu(x)$ is problematic, as we cannot take it out of integrals on x. To cope with this difficulty, we will rely on tools from probability theory, such as Fubini's theorem for stochastic integrals, or Dambis-Dubins-Schwarz theorem. We focus our attention on point 1., whereas point 2. previously shown without restriction on time in cases where $b(x,y) = S(y)$, is now only valid under the short-time hypothesis of Proposition 5.3.2 point 2.

As proven in Proposition 4.3.4, Γ_ν writes $\Gamma_\nu(\mu) = \Gamma_{1,\nu}(\mu) + \Gamma_{2,\nu}(\mu)$ with

$$\Gamma_{1,\nu}(\mu) := -\frac{1}{2}\int_{\mathscr{C}\times D}\int_0^T \Big(\widetilde{K}^t_{\nu,x}(t,t) + m_\nu(t,x)^2\Big)dtd\mu(x,r),$$

and

$$\Gamma_{2,\nu}(\mu) := \begin{cases} \frac{1}{2}\int_{\mathscr{C}\times D}\int_{\hat{\Omega}} L^\nu_T(x,r)^2 d\gamma_{\widetilde{K}^T_{\nu,x}} d\mu(x,r) + \int_{\mathscr{C}\times D}\int_0^T m_\nu(t,x)dW_t(x,r)d\mu(x,r) & \text{if } I(\mu|P) < \infty, \\ +\infty & \text{otherwise .} \end{cases}$$

The previous decomposition has the interest of splitting the difficulties: $|\Gamma_\nu(\mu) - \Gamma(\mu)| \le |\Gamma_{1,\nu}(\mu) - \Gamma_1(\mu)| + |\Gamma_{2,\nu}(\mu) - \Gamma_2(\mu)|$. The first term is easily controlled by $C_T d_T^V(\mu,\nu)$ using Proposition 5.3.5. Let us prove that

$$|\Gamma_{2,\nu}(\mu) - \Gamma_2(\mu)| \le C_T(1 + I(\mu|P))d_T^V(\mu,\nu).$$

The inequality is trivial when $I(\mu|P) = \infty$. We now assume that $I(\mu|P) < \infty$ implying $\mu \ll P$ and finiteness of $\Gamma(\mu)$ and $\Gamma_\nu(\mu)$. In particular, μ has a Borel-measurable density ρ_μ with respect to P:

$$d\mu(x,r) = \rho_\mu(x,r)dP(x,r).$$

Let $\varepsilon > 0$, and let $\xi \in \mathscr{C}_{\mu,\nu}$ be such that

$$\left\{\int_{(\mathscr{C}\times D)^2} d_T\big((y,r'),(z,\tilde{r})\big)^2 d\xi\big((y,r'),(z,\tilde{r})\big)\right\}^{\frac{1}{2}} \le d_T^V(\mu,\nu) + \varepsilon.$$

Let also $\big(G(x), G'(x)\big)_{x\in\mathscr{C}}$ a family of bi-dimensional centered Gaussian process from the probability space $\big(\hat{\Omega}, \hat{\mathscr{F}}, \gamma\big)$ with covariance K_ξ defined by (5.21). In the expression of $\Gamma_{2,\nu}(\mu)$ and $\Gamma_2(\mu)$ we can then replace the triplet (G^μ, G^ν, γ) by (G, G', γ), so that we choose their covariance to be given by K_ξ (see remark 12). As proved in Proposition 5.3.5, we can show that exist a constant $C_T > 0$ such that for any $t \in [0,T]$, $x \in \mathscr{C}$,

$$\mathscr{E}_\gamma\Big[\big(G_t(x) - G'_t(x)\big)^2\Big] \le \big(d_T^V(\mu,\nu) + \varepsilon\big)^2.$$

Let also for any $t \in [0,T]$

$$L_t(x,r) := \int_0^t G_s(x)dV_s^\mu(x,r), \quad L'_t(x,r) := \int_0^t G'_s(x)dV_s^\nu(x,r)$$

Then,

$$|\Gamma_{2,\nu}(\mu) - \Gamma_2(\mu)| \le \frac{1}{2}\left|\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\Big[L'_T(x,r)^2\big(\Lambda_T(G'(x)) - \Lambda_T(G(x))\big)\Big]d\mu(x,r)\right|$$
$$+\frac{1}{2}\left|\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\Big[\big(L_T(x,r)^2 - L'_T(x,r)^2\big)\Lambda_T(G(x))\Big]d\mu(x,r)\right| + \left|\int_{\mathscr{C}\times D}\int_0^T (m_\nu - m_\mu)(t,x)dW_t(x,r)d\mu(x,r)\right|$$

Observe that by inequality (5.19) we have

$$\left|\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\Big[L'_T(x,r)^2\big(\Lambda_T(G'(x))-\Lambda_T(G(x))\big)\Big]d\mu(x,r)\right| \le C_T\Bigg\{\big(d_T^V(\mu,\nu)+\varepsilon\big)\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\Big[L'_T(x,r)^2\Big]d\mu(x,r)$$
$$+\int_{\mathscr{C}\times D}\int_0^T\mathscr{E}_\gamma\Big[\big|G_t(x)^2-G'_t(x)^2\big|L'_T(x,r)^2\Big]dtd\mu(x,r)\Bigg\},$$

as Isserlis' theorem (Theorem 8.2.4) ensures that,

$$\mathscr{E}_\gamma\Big[\big(G'_t(x)-G_t(x)\big)^2L'_T(x,r)^2\Big] = \mathscr{E}_\gamma\Big[\big(G'_t(x)-G_t(x)\big)^2\Big]\mathscr{E}_\gamma\Big[L'_T(x,r)^2\Big]+2\mathscr{E}_\gamma\Big[\big(G'_t(x)-G_t(x)\big)L'_T(x,r)\Big]^2$$
$$\overset{\text{C.S.}}{\le} 3\mathscr{E}_\gamma\Big[\big(G'_t(x)-G_t(x)\big)^2\Big]\mathscr{E}_\gamma\Big[L'_T(x,r)^2\Big] \le 3\big(d_T^V(\mu,\nu)+\varepsilon\big)^2\mathscr{E}_\gamma\Big[L'_T(x,r)^2\Big],$$

and similarly

$$\mathscr{E}_\gamma\Big[\big(G'_t(x)+G_t(x)\big)^2L'_T(x,r)^2\Big] \le C_T\mathscr{E}_\gamma\Big[L'_T(x,r)^2\Big].$$

As a consequence,

$$|\Gamma_{2,\nu}(\mu)-\Gamma_2(\mu)| \overset{\text{C.S.}}{\le} C_T\Bigg\{\big(d_T^V(\mu,\nu)+\varepsilon\big)\overbrace{\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\Big[L'_T(x,r)^2\Big]d\mu(x,r)}^{B_1}$$
$$+\underbrace{\prod_{\varepsilon=\pm1}\left(\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\Big[\Big(\int_0^T\big(G_t(x)+\varepsilon G'_t(x)\big)dV_t^\nu(x,r)\Big)^2\Big]d\mu(x,r)\right)^{\frac12}}_{B_2}$$
$$+\underbrace{\left(\int_{\mathscr{C}\times D}\Big|\int_0^T(m_\nu-m_\mu)(t,x)dW_t(x,r)\Big|^2d\mu(x,r)\right)^{\frac12}}_{B_3}$$
$$+\underbrace{\prod_{\varepsilon=\pm1}\left(\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\Big[\Big\{\int_0^T G_t(x)\Big((1+\varepsilon)dW_t(x,r)-\big(m_\mu(t,x)+\varepsilon m_\nu(t,x)\big)dt\Big)\Big\}^2\Big]d\mu(x,r)\right)^{\frac12}}_{B_4}\Bigg\}. \tag{5.22}$$

Remark that these four terms can be cast in the form

$$\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\Big[\Big(\int_0^T H_t(G,G',\mu,\nu)(x)\big(\alpha dW_t(x,r)-M_t(\mu,\nu)(x)dt\big)\Big)^2\Big]d\mu(x,r)$$

with α equals 0 or 1. Controlling such terms is the aim of the following technical lemma.

Lemma 5.3.7. Let $\mu\in\mathscr{M}_1^+(\mathscr{C}\times D)$, with $\mu\ll P$ and let the filtration $\big(\mathscr{F}_t^{\mathbf{x}}\big)_{t\in[0,T]}$ on $\mathscr{C}$, where $\mathscr{F}_t^{\mathbf{x}}:=\sigma\big(x_s,0\le s\le t\big)$ is the σ-algebra on $\mathscr{C}$ generated by the coordinate process up to time t. Let also

- $x\in\mathscr{C}\to\big(M_t(x)\big)_{t\in[0,T]}$ a bounded time-continuous process progressively measurable for the filtration $(\mathscr{F}_t^{\mathbf{x}})_{t\in[0,T]}$ and continuous in x,

- $(x,\omega) \in \mathscr{C} \times \hat{\Omega} \to \Big(H_t(x,\omega)\Big)_{t\in[0,T]}$ a progressively measurable process for the filtration $\big(\mathscr{F}_t^{\mathbf{x}} \otimes \hat{\mathscr{F}}\big)_{t\in[0,T]}$, such that $\Big(H_t(x,\cdot), t\in[0,T]\Big)_{x\in\mathscr{C}}$ is a continuous family of γ-Gaussian processes (possibly deterministic) with uniformly bounded covariance,

and define

$$A(\mu) := \int_{\mathscr{C}\times D}\int_{\hat{\Omega}} \Big(\int_0^T H_t(x,\omega)\big(\alpha dW_t(x,r) - M_t(x)dt\big)\Big)^2 d\gamma(\omega)d\mu(x,r)$$

with $\alpha \in \{0,1\}$. Then, there exists a constant $C_T > 0$ independent of μ such that

$$A(\mu) \le C_T\Big\{\alpha\big(I(\mu|P)+1\big) + \sup_{x\in\mathscr{C},t\in[0,T]} M_t^2(x)\Big\} \sup_{x\in\mathscr{C},t\in[0,T]} \mathscr{E}_\gamma\big[H_t^2(x)\big], \tag{5.23}$$

with the right-hand side being possibly infinite.

Proof. As $(a+b)^2 \le 2a^2+2b^2$, $\forall a,b\in\mathbb{R}$,

$$A(\mu,\nu) \le 2\int_{\mathscr{C}\times D}\int_{\hat{\Omega}} \Big\{\alpha\Big(\underbrace{\int_0^T H_t(x,\omega)dW_t(x,r)}_{=:N_T(x,\omega,r)}\Big)^2 + \Big(\int_0^T H_t(x,\omega)M_t(x)dt\Big)^2\Big\}d\gamma(\omega)d\mu(x,r)$$

$$\overset{\text{Fubini,C.S.}}{\le} 2\alpha\int_{\hat{\Omega}}\int_{\mathscr{C}\times D} N_T^2(x,\omega,r)d\mu(x,r)d\gamma(\omega) + 2\int_{\mathscr{C}\times D}\int_0^T M_t^2(x)\mathscr{E}_\gamma\big[H_t^2(x)\big]dtd\mu(x,r).$$

Define the Radon-Nikodym density $\rho_\mu(x,r) := \frac{d\mu}{dP}(x,r)$ and remark that for every $r\in D$, $\big(N_t(,\cdot,\cdot,r)\big)$ is, γ-a.s., a well-defined P_r-martingale. Itô calculus gives, γ-a.s., the indistinguishable equality

$$N_T^2(x,\omega,r) = 2\int_0^T H_t(x,\omega)N_t(x,\omega,r)dW_t(x,r) + \int_0^T H_t^2(x,\omega)dt, \tag{5.24}$$

under P_r so that, γ-a.s.,

$$\int_{\mathscr{C}\times D} N_T^2(x,\omega,r)\rho_\mu(x,r)dP(x,r) = 2\int_{\mathscr{C}\times D}\int_0^T H_t(x,\omega)N_t(x,\omega,r)dW_t(x,r)\rho_\mu(x,r)dP(x,r)$$
$$+ \int_{\mathscr{C}\times D}\int_0^T H_t^2(x,\omega)dt\rho_\mu(x,r)dP(x,r).$$

Relying again on Fubini's Theorem,

$$A(\mu,\nu) \le 4\alpha\int_{\mathscr{C}\times D}\mathscr{E}_\gamma\Big[\int_0^T H_t(x)N_t(x,r)dW_t(x,r)\Big]\rho_\mu(x,r)dP(x,r)$$
$$+ 2\int_{\mathscr{C}\times D}\int_0^T \mathscr{E}_\gamma\Big[\alpha H_t^2(x)\Big]dtd\mu(x,r) + 2T\int_{\mathscr{C}\times D}\int_0^T M_t^2(x)\mathscr{E}_\gamma\big[H_t^2(x)\big]dtd\mu(x,r). \tag{5.25}$$

Under the favorable assumptions of the lemma, the last two terms of the right-hand side of (5.25) are easily controlled taking the supremum of their integrand on $\mathscr{C}\times[0,T]$. In order to control the first term, we rely on the stochastic Fubini's theorem [195, Theorem IV.65], to show that the equality

$$\tilde{N}_T(x,r) := \int_0^T \mathscr{E}_\gamma\Big[H_t(x)N_t(x,r)\Big]dW_t(x,r) = \mathscr{E}_\gamma\Big[\int_0^T H_t(x)N_t(x,r)dW_t(x,r)\Big],$$

is well-defined, and holds P-almost surely. To do so, we need to ensure that:

(i) $\forall r \in D$, $(x,\omega) \to \Big(\tilde{H}_t(x,\omega,r) := H_t(x,\omega)N_t(x,\omega,r)\Big)_{t\in[0,T]}$ is $\hat{\mathscr{F}} \otimes \mathscr{P}$ measurable, where $\mathscr{P}$ is the σ-algebra generated by continuous $\big(\mathscr{F}_t^{\mathbf{x}}\big)_{t\in[0,T]}$-adapted processes,

(ii) the following integrability condition holds $\forall r \in D$:

$$\int_{\mathscr{C}} \int_0^T \int_{\hat{\Omega}} \tilde{H}_t(x,\omega,r)^2 d\gamma(\omega) dt dP_r(x) < \infty.$$

The first hypothesis is a direct consequence of the regularity and measurability hypotheses of the lemma. We now demonstrate that the second hypothesis is valid. Indeed, for any $t \in [0,T]$,

$$\begin{aligned}
&\int_{\mathscr{C}} \int_{\hat{\Omega}} \tilde{H}_t(x,\omega,r)^2 d\gamma(\omega) dP_r(x) = \int_{\mathscr{C}} \mathscr{E}_\gamma\Big[H_t(x,r)^2 N_t(x,r)^2\Big] dP_r(x) \\
&\overset{\text{C.S.,Fub.}}{\leq} \left\{ \int_{\mathscr{C}} \mathscr{E}_\gamma\big[H_t^4(x)\big] dP_r(x) \right\}^{\frac{1}{2}} \mathscr{E}_\gamma\left[\int_{\mathscr{C}} N_t^4(x,r) dP_r(x) \right]^{\frac{1}{2}} \overset{\text{B.D.G.}}{\leq} C_T \mathscr{E}_\gamma\left[\int_{\mathscr{C}} \langle N \rangle_t^2(x) dP_r(x) \right]^{\frac{1}{2}} \\
&\overset{\text{C.S.,Fubini}}{\leq} C_T \left\{ \int_{\mathscr{C}} \int_0^t \mathscr{E}_\gamma\Big[H_s^4(x)\Big] ds dP_r(x) \right\}^{\frac{1}{2}} < +\infty.
\end{aligned}$$

Hence, the theorem applies so that

$$\int_{\mathscr{C}\times D} \mathscr{E}_\gamma\left[\int_0^T H_t(x) N_t(x,r) dW_t(x,r) \right] d\rho_\mu(x,r) dP(x,r) = \int_{\mathscr{C}\times D} \tilde{N}_T(x,r) d\mu(x,r).$$

Observe that inequality (4.42) brings

$$\int_{\mathscr{C}\times D} \tilde{N}_T(x,r) d\mu(x,r) \overset{\text{C.S.}}{\leq} 2\Big(\int_{\mathscr{C}\times D} \langle \tilde{N} \rangle_T(x,r) d\mu(x,r) \Big)^{\frac{1}{2}} \left(I(\mu|P) + \log\left\{ \int_{\mathscr{C}\times D} \exp\left\{ \frac{\tilde{N}_T^2(x,r)}{4\langle \tilde{N} \rangle_T(x,r)} \right\} dP(x,r) \right\} \right)^{\frac{1}{2}}.$$

As $\tilde{N}(\cdot,r)$ is a P_r-local martingale for every $r \in D$, Dambis-Dubins-Schwarz (D.D.S.) theorem ensures that $\frac{\tilde{N}_T(\cdot,r)^2}{4\langle \tilde{N}(\cdot,r) \rangle_T}$ has the same law as $\frac{B^2_{\langle \tilde{N} \rangle_T}}{4\langle \tilde{N} \rangle_T}$, where B is some P_r-Brownian motion, so that exists a constant $C > 0$ satisfying

$$\log\left\{ \int_{\mathscr{C}\times D} \exp\left\{ \frac{\tilde{N}_T^2(x,r)}{4\langle \tilde{N} \rangle_T(x,r)} \right\} dP(x,r) \right\} \leq C.$$

We can therefore conclude that there exists two constants: $\tilde{C} > 0$ independent of time, and $C_T > 0$ increasing with T such that:

$$\begin{aligned}
&\int_{\mathscr{C}\times D} \tilde{N}_t(x,r) d\mu(x,r) \leq \tilde{C} \left(\int_{\mathscr{C}\times D} \int_0^T \mathscr{E}_\gamma\Big[H_t(x) N_t(x,r) \Big]^2 dt d\mu(x,r) \right)^{\frac{1}{2}} \big(I(\mu|P) + 1 \big)^{\frac{1}{2}} \\
&\overset{\text{C.S.,Fubini}}{\leq} 2\tilde{C} \sup_{(x,t)\in\mathscr{C}\times[0,T]} \left\{ \mathscr{E}_\gamma\Big[\langle N \rangle_t(x) H_t^2(x) \Big] \right\}^{\frac{1}{2}} \left(\int_0^T \mathscr{E}_\gamma\left[\int_{\mathscr{C}\times D} \frac{N_t^2(x,r)}{4\langle N \rangle_t(x)} d\mu(x,r) \right] dt \right)^{\frac{1}{2}} \big(I(\mu|P) + 1 \big)^{\frac{1}{2}} \\
&\overset{\text{D.D.S,(4.42)}}{\leq} C_T \sup_{\mathscr{C}\times[0,T]} \left\{ \mathscr{E}_\gamma\Big[H_s^2(x) H_t^2(x) \Big] \right\}^{\frac{1}{2}} \big(I(\mu|P) + 1 \big).
\end{aligned}$$

We conclude by relying on Isserlis' theorem 8.2.4. □

It is easy to check that $B_1,\dots,B_4$ are of the form of the terms handled in lemma (5.3.7), satisfying in particular the adaptability conditions (keep in mind that the law of $G_t^\nu(x)$ depends on the trajectory of x up to time t). To conclude, we then have to underline that the quantities $\sup_{x\in\mathscr{C},t\in[0,T]} \mathscr{E}_\gamma\Big[\big(G_t(x)-G'_t(x)\big)^2\Big]$, and $\sup_{x\in\mathscr{C},t\in[0,T]} \big(m_\mu(t,x)-m_\nu(t,x)\big)^2$, are bounded by $\big(d_T^V(\mu,\nu)+\varepsilon\big)^2$ (see equation (5.20) for the term involving means). □

5.3.2 Upper-bound and Tightness

We are now in a position to demonstrate a partial LDP relying on an upper-bound inequality for closed subsets, and exponential tightness of the family $\Big(Q^N\big(\hat{\mu}_N\in\cdot\big)\Big)_N$. To prove the first point, we take advantage of the full LDP satisfied by $\hat{\mu}_N$ under $Q_\nu^{\otimes N}$, and control an error. The second point will rely on the exponential tightness of $P^{\otimes N}$ induced by the short time hypothesis (5.7). These developments follow the approach proposed by Ben Arous and Guionnet in [22, 129].

Theorem 5.3.8. *Under the condition* $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2}<1$*, we have:*

(i) *For any real number* $M\in\mathbb{R}$*, there exists a compact set* K_M *of* $\mathscr{M}_1^+(\mathscr{C}\times D)$ *such that, for any integer* N*,*

$$\frac{1}{N}\log Q^N(\hat{\mu}_N\notin K_M)\le -M.$$

(ii) *For any closed subset* F *of* $\mathscr{M}_1^+(\mathscr{C}\times D)$*,*

$$\limsup_{N\to\infty}\frac{1}{N}\log Q^N(\hat{\mu}_N\in F)\le -\inf_F H.$$

Proof. (1):

The proof of this theorem consists in using the exponential tightness of the sequence $(P^{\otimes N})_N$ provided by Sanov's Theorem and 9.4.2. Let K_M be a compact of $\mathscr{M}_1^+(\mathscr{C}\times D)$ such that

$$\frac{1}{N}\log P^{\otimes N}(\hat{\mu}_N\notin K_M)\le -M,$$

and remark that Hölder inequality yields for any conjugate exponents (p,q) with $\frac{(p+1)p^2\sigma^2\|b\|_\infty^2 T}{\lambda^2}<1$:

$$\begin{aligned}Q^N(\hat{\mu}_N\notin K_M)&\le\left(\int_{(\mathscr{C}\times D)^N}\exp\Big\{pN\bar{\Gamma}(\hat{\mu}_N)\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right)^{\frac{1}{p}}P^{\otimes N}(\hat{\mu}_N\notin K_M)^{\frac{1}{q}}\\&\overset{\text{Jensen}}{\le}\left(\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N\mathscr{E}_\gamma\left(\exp\Big\{pX^{\hat{\mu}_N}(x^i,r_i)\Big\}\right)dP^{\otimes N}(\mathbf{x},\mathbf{r})\right)^{\frac{1}{p}}P^{\otimes N}(\hat{\mu}_N\notin K_M)^{\frac{1}{q}}\end{aligned}$$

Let $(\tilde{X}^{\hat{\mu}_N,i})_{1\le i\le N}$ be independent copies of $X^{\hat{\mu}_N}$ under the measure γ. Then, by independence, Hölder inequality and martingale property, we have

$$\begin{aligned}\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N\mathscr{E}_\gamma\left(\exp\Big\{pX^{\hat{\mu}_N}(x^i,r_i)\Big\}\right)dP^{\otimes N}(\mathbf{x},\mathbf{r})&=\mathscr{E}_\gamma\left[\int_{(\mathscr{C}\times D)^N}\exp\Big\{p\sum_{i=1}^N\tilde{X}^{\hat{\mu}_N,i}(x^i,r_i)\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\right]\\&\le\left(\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N\mathscr{E}_\gamma\left[\exp\left\{\frac{p^2(p+1)}{2}\int_0^T\big(G_t^{\hat{\mu}_N}(x^i)+m_{\hat{\mu}_N}(t,x^i)\big)^2dt\right\}\right]dP^{\otimes N}(\mathbf{x},\mathbf{r})\right)^{\frac{p-1}{p}}.\end{aligned}\tag{5.26}$$

We can now proceed as in the proof of Proposition 5.3.2. point 2. to find that exists a constant c_T such that

$$\int_{(\mathscr{C}\times D)^N}\prod_{i=1}^N \mathscr{E}_\gamma\Big(\exp\Big\{pX^{\hat{\mu}_N}(x^i,r_i)\Big\}\Big)dP^{\otimes N}(\mathbf{x},\mathbf{r})\leq e^{(p-1)c_T N}.$$

As a consequence,

$$\limsup_{N\to+\infty}\frac{1}{N}\log Q^N(\hat{\mu}_N\notin K_M)\leq (p-1)c_T-\frac{M}{q}.$$

(2): As $\big(Q^N(\hat{\mu}_N\in\cdot)\big)_N$ is exponentially tight and H is good, it is sufficient to prove the upper-bound for compact sets (see [84, Lemma 1.2.18 (a)]). Let then K be a compact subset of $\mathscr{M}_1^+(\mathscr{C}\times D)$ and $\delta<0$. We can find an integer M and a family $(\nu_i)_{1\leq i\leq M}$ of $\mathscr{M}_1^+(\mathscr{C}\times D)$ such that

$$K\subset\bigcup_{i=1}^M B(\nu_i,\delta),$$

where $B(\nu_i,\delta)=\big\{\mu|d_T^V(\mu,\nu_i)<\delta\big\}$. A very classical result (see 9.2.4 or e.g. [84, lemma 1.2.15]), ensures that

$$\limsup_{N\to\infty}\frac{1}{N}\log Q^N(\hat{\mu}_N\in K)\leq\max_{1\leq i\leq p}\limsup_{N\to\infty}\frac{1}{N}\log Q^N(\hat{\mu}_N\in K\cap B(\nu_i,\delta)).$$

Lemma 5.3.1 yields:

$$\begin{aligned}Q^N(\hat{\mu}_N\in K\cap B(\nu,\delta))&=\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\exp\Big\{N\bar{\Gamma}(\hat{\mu}_N)\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r})\\&=\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\exp\Big\{N\big(\bar{\Gamma}(\hat{\mu}_N)-\bar{\Gamma}_\nu(\hat{\mu}_N)\big)\Big\}\exp\Big\{N\bar{\Gamma}_\nu(\hat{\mu}_N)\Big\}dP^{\otimes N}(\mathbf{x},\mathbf{r}).\end{aligned}$$

Hence, for any conjugate exponents (p,q),

$$\begin{aligned}Q^N(\hat{\mu}_N\in K\cap B(\nu,\delta))&=\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\exp\Big\{N\big(\bar{\Gamma}(\hat{\mu}_N)-\bar{\Gamma}_\nu(\hat{\mu}_N)\big)\Big\}dQ_\nu^{\otimes N}(\mathbf{x},\mathbf{r})\\&\leq Q_\nu^{\otimes N}\big(\hat{\mu}_N\in K\cap B(\nu,\delta)\big)^{\frac{1}{p}}\bigg(\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\exp\Big\{qN\big(\bar{\Gamma}(\hat{\mu}_N)-\bar{\Gamma}_\nu(\hat{\mu}_N)\big)\Big\}dQ_\nu^{\otimes N}(\mathbf{x},\mathbf{r})\bigg)^{\frac{1}{q}}.\end{aligned}\tag{5.27}$$

Then, by definition of Γ and Γ_ν:

$$\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\exp\Big\{qN\big(\bar{\Gamma}(\hat{\mu}_N)-\bar{\Gamma}_\nu(\hat{\mu}_N)\big)\Big\}dQ_\nu^{\otimes N}(\mathbf{x},\mathbf{r})=\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\left(\prod_{i=1}^N\frac{\mathscr{E}_\gamma\Big[\exp\big\{X^{\hat{\mu}_N}(x^i,r_i)\big\}\Big]}{\mathscr{E}_\gamma\Big[\exp\big\{X^{\nu}(x^i,r_i)\big\}\Big]}\right)^q dQ_\nu^{\otimes N}(\mathbf{x},\mathbf{r}).$$

Let $(\tilde{X}^{\nu,i})_{1\leq i\leq N}$ (resp. $(\tilde{X}^{\hat{\mu}_N,i})_{1\leq i\leq N}$) be independent copies of X^ν (resp. $X^{\hat{\mu}_N}$) under the measure γ. Then,

$$\begin{aligned}&\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\exp\Big\{qN\big(\bar{\Gamma}(\hat{\mu}_N)-\bar{\Gamma}_\nu(\hat{\mu}_N)\big)\Big\}dQ_\nu^{\otimes N}(\mathbf{x},\mathbf{r})\qquad(5.28)\\&=\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\mathscr{E}_\gamma\left[\prod_{i=1}^N\left(\exp\Big\{\tilde{X}^{\hat{\mu}_N,i}(x^i,r_i)-\tilde{X}^{\nu,i}(x^i,r_i)\Big\}\right)\frac{\prod_{i=1}^N\exp\big\{\tilde{X}^{\nu,i}(x^i,r_i)\big\}}{\mathscr{E}_\gamma\Big[\prod_{i=1}^N\exp\big\{\tilde{X}^{\nu,i}(x^i,r_i)\big\}\Big]}\right]^q dQ_\nu^{\otimes N}(\mathbf{x},\mathbf{r})\\&\overset{\text{Jensen}}{\leq}\Bigg\{\underbrace{\int_{\hat{\mu}_N\in K\cap B(\nu,\delta)}\mathscr{E}_\gamma\left[\prod_{i=1}^N\exp\Big\{q\Big(\tilde{X}^{\hat{\mu}_N,i}(x^i,r_i)-\tilde{X}^{\nu,i}(x^i,r_i)\Big)\Big\}\exp\big\{\tilde{X}^{\nu,i}(x^i,r_i)\big\}\right]dP^{\otimes N}(\mathbf{x},\mathbf{r})}_{B_N}\Bigg\}^{\frac{1}{q}},\end{aligned}$$

so that

$$Q^N(\hat{\mu}_N \in K \cap B(\nu,\delta)) \leq Q_\nu^{\otimes N}\big(\hat{\mu}_N \in K \cap B(\nu,\delta)\big)^{\frac{1}{p}} B_N^{\frac{1}{q}}.$$

The first term of the right-hand side can be controlled by large deviations estimates. The boundedness of the second term ensues from the following lemma. Based on these two results, we can conclude exactly as in 4.3.9. □

Lemma 5.3.9. For any real number $q > 1$, if $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$, then exist a real number $\delta_q > 0$ and a function $C_q :]0,\delta_q[\to \mathbb{R}^+$ such that $\lim_{\delta\to 0} C_q(\delta) = 0$:

$$B_N \leq \exp\{C_q(\delta)N\}.$$

Proof. Using Holder inequality with conjugate exponents (ρ,η), one finds:

$$B_N \leq \overbrace{\left\{\int_{(\mathscr{C}\times D)^N} \prod_{i=1}^N \mathscr{E}_\gamma\Big[\exp\rho X^\nu(x^i,r_i)\Big] dP^{\otimes N}(\mathbf{x},\mathbf{r})\right\}^{\frac{1}{\rho}}}^{B_1^N}$$
$$\times \underbrace{\left\{\int_{\hat{\mu}_N\in B(\nu,\delta)} \mathscr{E}_\gamma\Big[\prod_{i=1}^N \exp q\eta\left(\tilde{X}^{\hat{\mu}_N,i}(x^i,r_i) - \tilde{X}^{\nu,i}(x^i,r_i)\right)\Big] dP^{\otimes N}(\mathbf{x},\mathbf{r})\right\}^{\frac{1}{\eta}}}_{B_2^N}. \quad (5.29)$$

On the one hand, we can proceed exactly as in calculus (5.26) to obtain, under our short time hypothesis (4.54), the existence of a constant c_T uniform in ρ and N such that

$$B_1^N \leq e^{N(\rho-1)c_T},$$

so that one has to choose the proper relation between $\rho - 1$ and δ. On the other hand, no new difficulty arises from the second term, and we can show, as in [51], that exists a function $C_2(\delta) \underset{\delta\to 0}{\to} 0$ such that

$$B_2^N \leq \exp\{C_2(\delta)N\}.$$

Properly tuning the parameter (e.g. choosing $\rho = 1+\delta^{\frac{1}{4}}$, so that $q\eta\delta^{\frac{1}{2}} \leq 2q\delta^{\frac{1}{4}}$) yields the result. □

5.4 EXISTENCE, UNIQUENESS, AND CHARACTERIZATION OF THE LIMIT

In this section follows the steps of 4.4. Whereas the characterization of the minima of Q can be obtained exactly as in Lemma 4.4.1, to obtain the following equation

$$Q \simeq P, \qquad \frac{\mathrm{d}Q}{\mathrm{d}P}(x,r) = \mathscr{E}_\gamma\Big[\exp\big\{X^Q(x,r)\big\}\Big], \quad \forall (x,r) \in \mathscr{C}\times D, \quad (5.30)$$

the fixed point argument will be harder to develop because of the strong dependence of $O_\mu(t,x,r)$ in x.

The map $(x,r) \to \mathscr{E}_\gamma\Big[\exp\Big\{X^\mu(x,r)\Big\}\Big]$ is non-negative and measurable for every $\mu \in \mathscr{M}_1^+(\mathscr{C}\times D)$. Hence, we can properly define

$$L := \left\{ \begin{array}{l} \mathscr{M}_1^+\big(\mathscr{C}\times D\big) \to \mathscr{M}_1^+\big(\mathscr{C}\times D\big) \\ \mu \to dL(\mu)(x,r) := \mathscr{E}_\gamma\Big[\exp\Big\{X^\mu(x,r)\Big\}\Big] dP(x,r). \end{array} \right.$$

In fact, as $\exp\Big\{X^\mu(x,r)\Big\}$ is γ-almost surely finite, one can use Novikov criterion to show that $L(\mu)$ defines a probability measure on $\mathscr{C}\times D$. Equation (5.30) can be reformulated as follow: any minimum of H must satisfy:

$$Q \simeq P, \qquad Q = L(Q) = Q_Q.$$

with Q_Q defined as in Theorem 5.3.4.

Remark 13. It will appear in the proof of the below theorem that $\forall r \in D$, $\exists Q_r \in \mathscr{M}_1^+(\mathscr{C})$ such that $Q_r \ll P_r$, $\frac{\mathrm{d}Q_r}{\mathrm{d}P_r}(x) = \frac{\mathrm{d}Q}{\mathrm{d}P}(x,r) = \mathscr{E}_\gamma\Big[\exp\big\{X^Q(x,r)\big\}\Big]$, and $dQ(x,r) = dQ_r(x)d\pi(r)$.

Theorem 5.4.1. *The map L admits a unique fixed point.*

Proof. The proof of Lemma (4.3.5), can be readily apply to show that

$$\frac{\mathrm{d}L(\mu)}{\mathrm{d}P}(x,r) = \exp\left\{ \int_0^T O_\mu(t,x,r)dW_t(x,r) - \frac{1}{2}\int_0^T O_\mu^2(t,x,r)dt \right\},$$

where

$$O_\mu(t,x,r) = \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\mu(x)\big)G_t^\mu(x)L_t^\mu(x,r)\Big] + m_\mu(t,x).$$

Let $\mu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big), r \in D$, and remark that $x \to \frac{\mathrm{d}L(\mu)}{\mathrm{d}P}(x,r)$ is a P_r-martingale. Hence one can properly define $dL(\mu)_r(x) := \frac{\mathrm{d}L(\mu)}{\mathrm{d}P}(x,r)dP_r(x) \in \mathscr{M}_1^+(\mathscr{C})$. Because of the form of the density, Girsanov's theorem naturally leads to introduce the following SDE whose putative solution have a law equal to $L(\mu)_r$:

$$\left\{ \begin{array}{l} dx_t^\mu(r) = f(r,t,x_t^\mu(r))dt + \lambda O_\mu^{\tilde{W}}(t,x_t^\mu(r))dt + \lambda d\tilde{W}_t \\ x_0^\mu(r) = \bar{x}_0(r). \end{array} \right. \tag{5.31}$$

where $\tilde{W}$ is a $\mathbb{P}$-Brownian motion,

$$O_\mu^{\tilde{W}}(t,x) := \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\mu(x)\big)G_t^\mu(x)\tilde{L}_t^\mu(x)\Big] + m_\mu(t,x),$$

$$\tilde{L}_t^\mu(x) := \int_0^t G_s^\mu(x)\big(d\tilde{W}_s - m_\mu(s,x)ds\big),$$

and $\bar{x}_0(r) \in \mathbb{R}$ is the realization of the continuous version for the family of initial laws $\big(\mu_0(r)\big)_{r\in D}$ evaluated at r (see (5.2)). We show in Lemma 5.4.2 that for any $(r,\mu) \in D\times\mathscr{M}_1^+(\mathscr{C}\times D)$, there exists a unique strong solution $(x_t^\mu(r))_{t\in[0,T]}$ to equation (5.31). Let also $\nu \in \mathscr{M}_1^+\big(\mathscr{C}\times D\big)$, and define similarly $x_t^\nu(r)$ the process defined by:

$$\left\{ \begin{array}{l} dx_t^\nu(r) = f(r,t,x_t^\nu(r))dt + \lambda O_\nu^{\tilde{W}}(t,x_t^\nu(r))dt + \lambda d\tilde{W}_t \\ x_0^\nu(r) = \bar{x}_0(r), \end{array} \right.$$

where both the initial condition $\bar{x}_0(r)$ and the driving Brownian motion $(\tilde{W}_t)$ are the same as for the definition of $x_t^\mu(r)$. We have

$$\begin{aligned}
&\big(x_t^\mu(r)-x_t^\nu(r)\big)=\int_0^t\Big(f(r,s,x_s^\mu(r))+\lambda m_\mu(s,x_\cdot^\mu(r))-f(r,s,x_s^\nu(r))-\lambda m_\nu(s,x_\cdot^\nu(r)\Big)ds\\
&+\lambda\int_0^t\Big\{\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\mu(x_\cdot^\mu(r))\big)G_s^\mu(x_\cdot^\mu(r))\tilde{L}_s^\mu(x_\cdot^\mu(r))\Big]-\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(x_\cdot^\mu(r))\big)G_s^\nu(x_\cdot^\mu(r))\tilde{L}_s^\nu(x_\cdot^\mu(r))\Big]\Big\}ds\\
&+\lambda\int_0^t\int_0^s\Big(\tilde{K}^s_{\nu,x_\cdot^\mu(r)}(s,v)m_\nu\big(v,x_\cdot^\mu(r)\big)-\tilde{K}^s_{\nu,x_\cdot^\nu(r)}(s,v)m_\nu\big(v,x_\cdot^\nu(r)\big)\Big)dvds\\
&+\lambda\int_0^t\Big\{\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(x_\cdot^\mu(r))\big)G_s^\nu(x_\cdot^\mu(r))\Big(\int_0^s G_v^\nu(x_\cdot^\mu(r))d\tilde{W}_v\Big)\Big]\\
&-\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(x_\cdot^\nu(r))\big)G_s^\nu(x_\cdot^\nu(r))\Big(\int_0^s G_v^\nu(x_\cdot^\nu(r))d\tilde{W}_v\Big)\Big]\Big\}ds. \qquad (5.32)
\end{aligned}$$

Let $\xi\in\mathscr{C}_{\mu,\nu}$. We have:

$$\begin{aligned}
&\lambda\big(m_\mu(t,x_\cdot^\mu(r))-m_\nu(t,x_\cdot^\nu(r))\big)=\bar{J}\int_{(\mathscr{C}\times D)^2}\Big(b(x_t^\mu(r),y_t)-b(x_t^\nu(r),y_t)\Big)\\
&+\Big(b(x_t^\nu(r),y_t)-b(x_t^\nu(r),z_t)\Big)d\xi\big((y,r'),(z,\tilde{r})\big)\\
&\le K_b\bar{J}\Big(\big|x_t^\mu(r)-x_t^\nu(r)\big|+\int_{(\mathscr{C}\times D)^2}\|y-z\|_{\infty,t}d\xi\big((y,r'),(z,\tilde{r})\big)\Big)\le C\Big(\big|x_t^\mu(r)-x_t^\nu(r)\big|+d_t^V(\mu,\nu)\Big)
\end{aligned}$$

where we took the infimum on ξ.

Furthermore, let $(\tilde{G},\tilde{G}')$ be a bidimensional centered Gaussian process on the probability space $(\hat{\Omega},\hat{\mathscr{F}},\gamma)$ with covariance given by:

$$\frac{\sigma^2}{\lambda^2}\int_{(\mathscr{C}\times D)^2}\begin{pmatrix} b(x_s^\mu(r),y_s)b(x_t^\mu(r),y_t) & b(x_s^\mu(r),y_s)b(x_t^\nu(r),y_t)\\ b(x_s^\nu(r),y_s)b(x_t^\mu(r),y_t) & b(x_s^\nu(r),y_s)b(x_t^\nu(r),y_t)\end{pmatrix}d\nu(y,r). \qquad (5.33)$$

Then

$$\Big|\tilde{K}^t_{\nu,x_\cdot^\mu(r)}(t,s)-\tilde{K}^t_{\nu,x_\cdot^\nu(r)}(t,s)\Big|=\Big|\mathscr{E}_\gamma\Big[\big(\Lambda_t(\tilde{G})-\Lambda_t(\tilde{G}')\big)\tilde{G}_t\tilde{G}_s+\Lambda_t(\tilde{G}')\big(\tilde{G}_t-\tilde{G}_t'\big)\tilde{G}_s+\Lambda_t(\tilde{G}')\tilde{G}_t'\big(\tilde{G}_s-\tilde{G}_s'\big)\Big]\Big|$$

Moreover,

$$\begin{aligned}
\mathscr{E}_\gamma\Big[\Lambda_t(\tilde{G})\big(\tilde{G}_t-\tilde{G}_t'\big)\tilde{G}_s\Big]&\overset{\text{C.S.,(5.17)}}{\le}C_T\mathscr{E}_\gamma\Big[\big(\tilde{G}_t-\tilde{G}_t'\big)^2\Big]^{\frac{1}{2}}\le C_T\left(\int_{\mathscr{C}\times D}\Big(b\big(x_t^\mu(r),y_t\big)-b\big(x_t^\nu(r),y_t\big)\Big)^2d\nu(y,r')\right)^{\frac{1}{2}}\\
&\le C_T\big|x_t^\mu(r)-x_t^\nu(r)\big| \qquad (5.34)
\end{aligned}$$

and

$$\mathscr{E}_\gamma\Big[\big(\Lambda_t(\tilde{G})-\Lambda_t(\tilde{G}')\big)\tilde{G}_t\tilde{G}_s\Big]\overset{(5.19),\text{C.S.}}{\le}C_T\big|x_t^\mu(r)-x_t^\nu(r)\big|,$$

so that

$$\Big|\tilde{K}^t_{\nu,x_\cdot^\mu(r)}(t,s)-\tilde{K}^t_{\nu,x_\cdot^\nu(r)}(t,s)\Big|\le C_T\big|x_t^\mu(r)-x_t^\nu(r)\big|.$$

We now focus on controlling the second term of (5.32). Let another $\xi\in\mathscr{C}_{\mu,\nu}$, and let (G,G') be a bidimensional centered Gaussian process on the probability space

$(\hat{\Omega},\hat{\mathscr{F}},\gamma)$ with covariance $K_\xi\big(\cdot,\cdot,x_\cdot^\mu(r)\big)$ given in (5.21). Observe that

$$\begin{aligned}&\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\mu(x_\cdot^\mu(r))\big)G_s^\mu(x_\cdot^\mu(r))\tilde{L}_s^\mu(x_\cdot^\mu(r))\Big]-\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(x_\cdot^\mu(r))\big)G_s^\nu(x_\cdot^\mu(r))\tilde{L}_s^\nu(x_\cdot^\mu(r))\Big]\\&=\mathscr{E}_\gamma\Big[\Lambda_s(G)G_sL_s-\Lambda_s(G')G'_sL'_s\Big]\end{aligned}$$

where $L_t:=\int_0^t G_s\big(d\tilde{W}_s-m_\mu(s,x_\cdot^\mu(r))ds\big)$, and $L'_t:=\int_0^t G'_s\big(d\tilde{W}_s-m_\nu(s,x_\cdot^\mu(r))ds\big)$.

As in the proof of Theorem 4.4.2, we have

$$\begin{aligned}\mathscr{E}_\gamma\Big[\Lambda_t(G)G_tL_t-\Lambda_t(G')G'_tL'_t\Big]&=\mathscr{E}_\gamma\Big[\big(\Lambda_t(G)-\Lambda_t(G')\big)G_tL_t\Big]+\mathscr{E}_\gamma\Big[\Lambda_t(G')\big(G_t-G'_t\big)L_t\Big]+\mathscr{E}_\gamma\Big[\Lambda_t(G')G'_t\big(L_t-L'_t\big)\Big]\\&\overset{\text{C.S.}}{\leq}\mathscr{E}_\gamma\big[L_t^2\big]^{\frac{1}{2}}\left(\mathscr{E}_\gamma\Big[\big(\Lambda_t(G)-\Lambda_t(G')\big)^2G_t^2\Big]^{\frac{1}{2}}+\mathscr{E}_\gamma\Big[\Lambda_t(G')^2\big(G_t-G'_t\big)^2\Big]^{\frac{1}{2}}\right)+\mathscr{E}_\gamma\Big[\Lambda_t(G')^2G_t'^2\Big]^{\frac{1}{2}}\mathscr{E}_\gamma\Big[\big(L_t-L'_t\big)^2\Big]^{\frac{1}{2}}\end{aligned}$$

We can show, as in Proposition 5.3.5

$$\mathscr{E}_\gamma\Big[\Lambda_t(G)^2\big(G_t-G'_t\big)^2\Big]\leq C_T\left(\int_{(\mathscr{C}\times D)^2}\|y-z\|_{\infty,t}^2d\xi\big((y,r'),(z,\tilde{r})\big)\right),$$

and

$$\mathscr{E}_\gamma\Big[\big(\Lambda_t(G)-\Lambda_t(G')\big)^2G_t^2\Big]\overset{(5.19),(8.2.4)}{\leq}C_T\left(\int_{(\mathscr{C}\times D)^2}\|y-z\|_{\infty,t}^2d\xi\big((y,r'),(z,\tilde{r})\big)\right).$$

On the other hand,

$$\begin{aligned}&\mathscr{E}_\gamma\Big[\big(L_t-L'_t\big)^2\Big]\leq 2\mathscr{E}_\gamma\left[\left(\int_0^t\big(G_s-G'_s\big)d\tilde{W}_s\right)^2\right]+4t\int_0^t\left\{\mathscr{E}_\gamma\Big[\big(G_s-G'_s\big)^2m_\mu(s,x_\cdot^\mu(r))^2\Big]\right.\\&\left.+\mathscr{E}_\gamma\Big[G_s'^2\Big(m_\mu\big(s,x_\cdot^\mu(r)\big)-m_\nu\big(s,x_\cdot^\mu(r)\big)\Big)^2\Big]\right\}ds\\&\overset{(5.20)}{\leq}C_T\left\{\mathscr{E}_\gamma\left[\left(\int_0^t\big(G_s-G'_s\big)d\tilde{W}_s\right)^2\right]+\int_{(\mathscr{C}\times D)^2}\|y-z\|_{\infty,t}^2d\xi\big((y,r'),(z,\tilde{r})\big)\right\},\end{aligned}$$

We thus obtain:

$$\begin{aligned}&\left|\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\mu(x_\cdot^\mu(r))\big)G_s^\mu(x_\cdot^\mu(r))\tilde{L}_s^\mu(x_\cdot^\mu(r))\Big]-\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(x_\cdot^\mu(r))\big)G_s^\nu(x_\cdot^\mu(r))\tilde{L}_s^\nu(x_\cdot^\mu(r))\Big]\right|^2\\&\leq C_T\left(1+\mathscr{E}_\gamma\left[\left(\int_0^sG_vd\tilde{W}_v\right)^2\right]\right)\left(\int_{(\mathscr{C}\times D)^2}\|y-z\|_{\infty,s}^2d\xi\big((y,r'),(z,\tilde{r})\big)\right)\\&+C_T\mathscr{E}_\gamma\left[\left(\int_0^s\big(G_v-G'_v\big)d\tilde{W}_v\right)^2\right].\end{aligned}$$

Injecting these result in (5.32), we obtain:

$$\begin{aligned}\|x^\mu(r)-x^\nu(r)\|_{\infty,t}^2\leq C_T\int_0^t\Bigg\{&\|x^\mu(r)-x^\nu(r)\|_{\infty,s}^2+\mathscr{E}_\gamma\left[\left(\int_0^s\big(G_v-G'_v\big)d\tilde{W}_v\right)^2\right]\\&+\left(1+\mathscr{E}_\gamma\left[\left(\int_0^sG_vd\tilde{W}_v\right)^2\right]\right)\left(\int_{(\mathscr{C}\times D)^2}\|y-z\|_{\infty,s}^2d\xi\big((y,r'),(z,\tilde{r})\big)\right)\\&+\lambda^2\left|\mathscr{E}_\gamma\left[\Lambda_s\big(G^\nu(x_\cdot^\mu(r))\big)G_s^\nu(x_\cdot^\mu(r))\Big(\int_0^sG_v^\nu(x_\cdot^\mu(r))d\tilde{W}_v\Big)\right]\right.\\&\left.-\mathscr{E}_\gamma\left[\Lambda_s\big(G^\nu(x_\cdot^\nu(r))\big)G_s^\nu(x_\cdot^\nu(r))\Big(\int_0^sG_v^\nu(x_\cdot^\nu(r))d\tilde{W}_v\Big)\right]\right|^2\Bigg\}ds,\end{aligned}$$

so that relying on Gronwall's lemma, taking the expectation over initial conditions and the Brownian path, and making use of Fubini's theorem, Itô isometry, and eventually taking the infimum in ξ yields:

$$\mathbb{E}\Big[\|x^\mu(r)-x^\nu(r)\|_{\infty,t}^2\Big] \le C_T\int_0^t\Big\{d_s^V(\mu,\nu)^2+\lambda^2\mathbb{E}\Big[\Big|\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(x_\cdot^\mu(r))\big)G_s^\nu(x_\cdot^\mu(r))\Big(\int_0^s G_v^\nu(x_\cdot^\mu(r))d\tilde W_v\Big)\Big] - \mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(x_\cdot^\nu(r))\big)G_s^\nu(x_\cdot^\nu(r))\Big(\int_0^s G_v^\nu(x_\cdot^\nu(r))d\tilde W_v\Big)\Big]\Big|^2\Big]\Big\}ds,$$

To cope with the last term of the right-hand side, let again $(\tilde G,\tilde G')$ be a bidimensional centered Gaussian process on the probability space $(\hat\Omega,\hat{\mathscr{F}},\gamma)$ with covariance given by (5.33). Let also $\mathbb{E}_\gamma[\,\cdot\,]:=\mathbb{E}\Big[\mathscr{E}_\gamma[\,\cdot\,]\Big]$. Then

$$\mathbb{E}\Big[\Big|\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(x_\cdot^\mu(r))\big)G_s^\nu(x_\cdot^\mu(r))\Big(\int_0^s G_v^\nu(x_\cdot^\mu(r))d\tilde W_v\Big)\Big]-\mathscr{E}_\gamma\Big[\Lambda_s\big(G^\nu(x_\cdot^\nu(r))\big)G_s^\nu(x_\cdot^\nu(r))\Big(\int_0^s G_v^\nu(x_\cdot^\nu(r))d\tilde W_v\Big)\Big]\Big|^2\Big] =$$

$$\mathbb{E}\Big[\mathscr{E}_\gamma\Big[\Lambda_s(\tilde G)\tilde G_s\Big(\int_0^s\tilde G_v d\tilde W_v\Big)-\Lambda_s(\tilde G')\tilde G'_s\Big(\int_0^s\tilde G'_v d\tilde W_v\Big)\Big]^2\Big] \tag{5.35}$$

$$\overset{\text{C.S.}}{\le}\mathbb{E}_\gamma\Big[\Big\{\Lambda_s(\tilde G)\tilde G_s\Big(\int_0^s\tilde G_v d\tilde W_v\Big)-\Lambda_s(\tilde G')\tilde G'_s\Big(\int_0^s\tilde G'_v d\tilde W_v\Big)\Big\}^2\Big]$$

$$\overset{\text{C.S.}}{\le} 3\mathbb{E}_\gamma\Big[\Big(\int_0^s\tilde G_v d\tilde W_v\Big)^4\Big]^{\frac12}\Big(\mathbb{E}_\gamma\Big[\big(\Lambda_t(\tilde G)-\Lambda_t(\tilde G')\big)^4\tilde G_t^4\Big]^{\frac12}+\mathbb{E}_\gamma\Big[\Lambda_t(\tilde G')^4\big(\tilde G_t-\tilde G'_t\big)^4\Big]^{\frac12}\Big)$$

$$+3\mathbb{E}_\gamma\Big[\Lambda_t(\tilde G')^4\tilde G_t'^4\Big]^{\frac12}\mathbb{E}_\gamma\Big[\Big(\int_0^s\big(\tilde G_v-\tilde G'_v\big)d\tilde W_v\Big)^4\Big]^{\frac12}. \tag{5.36}$$

Gaussian calculus and (5.34) gives

$$\mathscr{E}_\gamma\Big[\big(\tilde G_t-\tilde G'_t\big)^4\Big]=C\mathscr{E}_\gamma\Big[\big(\tilde G_t-\tilde G'_t\big)^2\Big]^2\le C_T\big|x_t^\mu(r)-x_t^\nu(r)\big|^2.$$

Then relying on (5.17), (5.19) and Burkhölder Davis Gundi inequality, we obtain:

$$\mathbb{E}\Big[\|x^\mu(r)-x^\nu(r)\|_{\infty,t}^2\Big]\le C_T\int_0^t\Big\{d_s^V(\mu,\nu)^2+\mathbb{E}\Big[\|x^\mu(r)-x^\nu(r)\|_{\infty,s}^2\Big]\Big\}ds.$$

Another use of Gronwall's lemma then gives:

$$\mathbb{E}\Big[\|x^\mu(r)-x^\nu(r)\|_{\infty,t}^2\Big]\le C_T\int_0^t d_s^V(\mu,\nu)^2ds.$$

Let us now show the regularity in space of left-hand side in the above inequality. To this purpose, fix $r'\neq r\in D$, and consider $x_\cdot^\mu(r')$ be the strong solution of (5.31) with same $\tilde W$ but initial condition given by $\bar x_0(r')$ and intrinsic dynamics $f(r',\cdot,x_\cdot^\mu(r'))$. Using the same analysis, one observes that

$$\|x^\mu(r)-x^\mu(r')\|_{\infty,t}^2\le C_T\Big\{\big(\bar x_0(r)-\bar x_0(r')\big)^2+\|r-r'\|_{\mathbb{R}^d}^2+\int_0^t\|x^\mu(r)-x^\mu(r')\|_{\infty,s}^2ds\Big\},$$

and thus conclude on continuity of $r\mapsto(x_t^\mu(r))_{t\in[0,T]}$ by Gronwall's lemma and continuity of $r\mapsto\bar x_0(r)$. Hence the maps $r\to x_\cdot^\mu(r)$ and $r\to x_\cdot^\mu(r)-x_\cdot^\nu(r)$ are continuous for

the uniform norm in $\mathscr{C}$. As a consequence, $r \to \|x^\mu(r) - x^\nu(r)\|_{\infty,t}$ is also continuous, so that Fatou's lemma ensures that $r \to \mathbb{E}\Big[\|x^\mu(r) - x^\nu(r)\|^2_{\infty,t}\Big]$ is at least lower semi-continuous. Integrating on $r \in D$, and remarking that (x^μ_r, x^ν_r) is a particular coupling of $\big(L(\mu)_r, L(\nu)_r\big)$ yields:

$$d_t^V\big(L(\mu), L(\nu)\big)^2 \leq C_T \int_0^t d_s^V(\mu,\nu)^2 ds.$$

This inequality allows to develop the classical Picard's iterations method to conclude on the existence and uniqueness of solution to the fixed point equation in $\mathscr{M}_1^+(\mathscr{C} \times D)$. □

Lemma 5.4.2. For any $r \in D$ and $\mu \in \mathscr{M}_1^+(\mathscr{C} \times D)$, there exists a unique strong solution to the SDE:

$$\begin{cases} dx_t^\mu(r) = f(r,t,x_t^\mu(r))dt + \lambda O_\mu^{\tilde{W}}(t, x_t^\mu(r))dt + \lambda d\tilde{W}_t \\ x_0^\mu(r) = \bar{x}_0(r). \end{cases}$$

where $\tilde{W}$ is a $\mathbb{P}$-Brownian motion, $\bar{x}_0(r) \in \mathbb{R}$ is the realization of the continuous version for the family of initial laws $\big(\mu_0(r)\big)_{r\in D}$, and

$$O_\mu^{\tilde{W}}(t,x) := \mathscr{E}_\gamma\Big[\Lambda_t\big(G^\mu(x)\big)G_t^\mu(x)\int_0^t G_s^\mu(x)\big(d\tilde{W}_s - m_\mu(s,x)ds\big)\Big] + m_\mu(t,x).$$

Proof. The proof relies on Picard's iterations. Let $x^0 \in \mathscr{C}$ with $x_0^0 = \bar{x}_0(r)$, and define recursively the sequence $\big(x_t^n, 0 \leq t \leq T\big)_{n\in\mathbb{N}^*}$ by

$$x_t^{n+1} = \bar{x}_0(r) + \int_0^t f(r,s,x_s^n)ds + \int_0^t \lambda O_\mu^{\tilde{W}}(s, x_s^n)ds + \lambda\tilde{W}_t, \ \ \forall t \in [0,T].$$

Then, using the estimate (5.36) obtained in the proof of theorem 5.4.1 and taking the expectation, we find

$$\mathbb{E}\Big[\|\big|x^{n+1} - x^n\|^2_{\infty,t}\Big] \leq C_T \int_0^t \mathbb{E}\Big[\sup_{u\leq s}\big|x_u^n - x_u^{n-1}\big|^2\Big]ds.$$

□

5.4.1 Convergence of the process and Quenched results.

We are now in a position to prove theorem 5.2.2.

Proof of Theorem (5.2.2). Let $\delta > 0$ and $B(Q,\delta)$ the open ball of radius δ centered in Q for the Vaserstein distance. We prove that $Q^N(\hat{\mu}_N \notin B(Q,\delta))$ tends to zero exponentially fast as N goes to infinity. In fact, the upper-bound of the LDP for the closed set $B(Q,\delta)^c$ yields

$$\limsup_{N\to\infty} \frac{1}{N}\log Q^N(\hat{\mu}_N \notin B(Q,\delta)) \leq -\inf_{B(Q,\delta)^c} H < 0$$

where the last inequality comes from the fact that H attains its unique minimum at Q. This implies that $Q^N(\hat{\mu}_N \notin B(Q,\delta)) \to 0$ at least exponentially fast, so that the result is proved. □

Proof of Theorem 5.2.6. For a given closed set F, we can obtain a quenched upper-bound as a consequence of Theorem.3 and Borel-Cantelli, by proceeding exactly as in [22, Theorem 2.7 of Appendix C.]. As $\mathscr{M}_1^+(\mathscr{C}\times D)$ is Polish, we are able to define a sequence of closed sets $(F_i)_{i\in\mathbb{N}}$ of $\mathscr{M}_1^+(\mathscr{C}\times D)$ such that for all closed set $F\subset\mathscr{M}_1^+(\mathscr{C}\times D)$ there exists $A_F\subset\mathbb{N}$, and

$$F=\bigcap_{i\in A_F}F_i.$$

Moreover, as $\left\{F\subset\mathscr{M}_1^+(\mathscr{C}\times D),\exists A_F\subset\mathbb{N},F=\bigcap_{i\in A_F}F_i\right\}$ is countable and contains every closed set, we obtain an $\mathscr{P}$-almost sure upper-bound for every closed set:

$$\mathscr{P}-a.s,\ \forall \text{ closed set } F\subset\mathscr{M}_1^+(\mathscr{C}\times D),\ \limsup_{N\to\infty}\frac{1}{N}\log Q_{\mathbf{r}}^N(J)(\hat{\mu}_N\in F)\leq-\inf_F H.$$

H being a good rate function, the $\mathscr{P}$-almost sure exponential tightness is a consequence of [84, Exercice 4.1.10 (c)] (citing the results of [163, Lemma 2.6] and [196, Theorem P]), whereas the $\mathscr{P}$-almost sure convergence of the empirical measure stems from Borel-Cantelli lemma, noting that for any $\varepsilon>0$,

$$Q_{\mathbf{r}}^N(J)(\hat{\mu}_N\notin B(Q,\varepsilon))=Q_{\mathbf{r}}^N(J)(d_T(\hat{\mu}_N,Q)\geq\varepsilon)$$

is summable.

□

Part III

Phenomenology of Random Networks

CHAPTER 6

NUMERICAL STUDY OF A NEURAL NETWORK

6.1 INTRODUCTION

In [51], we analyzed randomly coupled neuronal networks and derived the limit of the empirical averages as the number of neurons tends to infinity. To this purpose, we showed that the system satisfies a large deviation principle and exhibited the related good rate function. This approach generalized the work of Gerard Ben Arous and Alice Guionnet [22, 129] developed for spin glasses in three main directions: (i) the synaptic weights are not centered, introducing additional, deterministic terms in the coupling, (ii) interactions are delayed, which projects the problem into infinite dimensions, and (iii) the system is composed of several populations, which was handled showing that empirical measures on each populations simultaneously satisfy a large deviation principle. The proof is made on a particular model very popular in physics and neurosciences, the Wilson and Cowan system, which is close of the famous Hopfield model, but as mentioned, can be easily generalized to nonlinear neuron models. Indeed, most of the proofs deal with a quantity which is related to the density of the coupled network with respect to the uncoupled dynamics, and this quantity is independent of the dynamics of individual cells. This approach can be also readily extended to networks with distributed delays. Eventually, let us note that this result provides large-deviations estimates on the convergence of deterministically coupled networks as studied in [239].

For the neuroscience viewpoint, this approach had the interest to justify an approach widely used in the analysis of large-scale networks, and introduce in the seminal paper of Sompolinsky and collaborators [221].

In this chapter, we will study the possible solution of the found limit equation. Under the choice of Gaussian initial conditions and linear intrinsic dynamics, we can show existence and uniqueness for the solutions, as these solutions are also Gaussian. These Gaussian solutions are unequivocally characterized by their mean and covariance functions, and these variables satisfy a closed set of deterministic equations. In the stochastic synaptic noise case, we will show in the second part of this study that the mean and standard deviations satisfy a closed set of delayed differential equations where the heterogeneity and the noise coefficients appear as a param-

eter. This will allow to uncover, using a bifurcation analysis, the qualitative effects of noise on the solutions. In the static random heterogeneity case, the mean satisfies a delayed differential equation coupled to the variance of the solution. Unfortunately, in contrast with the case of stochastic noise, the variance does not satisfies an ordinary differential equation but can be written as the solution of a fixed point equation. This distinction is fundamentally related to the non-Markovian nature of the asymptotic equation. We will then develop a few analytical approaches to characterize the solutions and a heuristic argument will allow demonstrating that noise is directly related to the emergence of synchronized oscillations, a highly relevant macroscopic state related to fundamental cortical functions such as memory, attention, sleep and consciousness, and its impairments relate to serious pathologies such as epilepsy or Parkinson's disease [50, 240], and that may account for the results of Aradi and colleagues showing that increased heterogeneity was related with epilepsy.

6.2 MATHEMATICAL SETTING AND GAUSSIAN CHARACTERIZATION OF THE LIMIT

In all this chapter, we are working in a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ endowed with a filtration $\left(\mathscr{F}_t\right)_t$ satisfying the usual conditions. We consider a network composed of N neurons falling into P populations. We define by $p : \mathbb{N} \mapsto \{1, \cdots, P\}$ the *population function* associating to a neuron index the population label it belongs to. The state of each neuron i in population $p(i) = \alpha \in \{1, \cdots, P\}$ is described by its membrane potential $x^i \in \mathbb{R}$. Its evolution is driven by the inputs it receive from the other neurons, through a sigmoidal transformation of their membrane potential $S_{\alpha\gamma}(x_t^j)$, as well as external inputs $I_{p(i)}(t)$. This sigmoidal function is specific to both post- and pre-synaptic populations $p(i)$ and $p(j)$, and is assume to be smooth increasing functions that tends to 0 at $-\infty$ and to 1 at ∞. Moreover, this input from pre-synaptic neuron j to post-synaptic neuron i is modulated by the synaptic efficiency J_{ij}. Without inputs, the membrane potential exponentially decreases with characteristic time θ_α. We can now introduce the dynamic of x_t^i of population $p(i) = \alpha$:

$$dx_t^i = \left(-\frac{1}{\theta_\alpha} x_t^i + \sum_{\gamma=1}^{P} \sum_{j:p(j)=\gamma} J_{ij} S_{\alpha\gamma}(x_{t-\tau_{\alpha\gamma}}^j) \right) dt + \lambda_\alpha dW_t^i \tag{6.1}$$

The Brownian motions W_t^i are independent and account for the noisy input received by all neurons. Moreover, λ_α denote the noise intensity specific of population α, $\tau_{\alpha\gamma}$ is the propagation delay between neurons of population γ and neurons of population α, and where the synaptic weights J_{ij} are assumed to be independent Gaussian random variables, with law $\mathscr{N}\left(\frac{\bar{J}_{p(i)p(j)}}{N_{p(j)}}, \frac{\sigma^2_{p(i)p(j)}}{N_{p(j)}}\right)$.

In [51], it was shown that the empirical measure of population α was converging (in an averaged sense) toward a solution of:

$$\begin{cases} d\bar{X}_t^\alpha = \left(-\frac{1}{\theta_\alpha} \bar{X}_t^\alpha + U_t^{\alpha,\bar{X}} \right) dt + \lambda_\alpha dW_t^\alpha \\ U_t^{\alpha,\bar{X}} \sim \mathscr{N}\left(\sum_{\gamma=1}^P \bar{J}_{\alpha\gamma} \mathbb{E}\left[S_{\alpha\gamma}(\bar{X}_{t-\tau_{\alpha\gamma}}^\gamma) \right], \sum_{\gamma=1}^P \sigma^2_{\alpha\gamma} \mathbb{E}\left[S_{\alpha\gamma}(\bar{X}_{t-\tau_{\alpha\gamma}}^\gamma)^2 \right] \right) \textit{ pairwise independent}. \end{cases} \tag{6.2}$$

Nevertheless, existence and uniqueness of solution was not addressed in the general case. An interesting property of these P implicit SDE is that its solutions are Gaussian processes, provided that the initial conditions are Gaussian. In fact, if we write down $\bar{X}_t^\alpha$ we find a sum of integrated independent Gaussian processes. Hence, characterizing the solutions of (6.2) only requires to determine its mean and covariance. The above theorem hence characterizes unequivocally the limits of the network equations considered. The proof of this proposition was done in [106] starting from equations (6.2) which were introduced using a heuristic argument.

Theorem 6.2.1. *Consider that the initial conditions of* (6.2) *are Gaussian, and denote by* $\mu^\alpha(t)$, $C^{\alpha\beta}(t,s) = \mathbb{E}[X_t^\alpha X_s^\beta]$ *the mean and variance of its solution. Then:*

$$\dot{\mu}^\alpha(t) = -\frac{1}{\theta_\alpha}\mu^\alpha(t) + \sum_{\gamma=1}^{P} \bar{J}_{\alpha\gamma} f_{\alpha\gamma}(\mu^\gamma(t-\tau_{\alpha\gamma}), C^{\alpha\alpha}(t-\tau_{\alpha\gamma}, t-\tau_{\alpha\gamma})) \tag{6.3}$$

where $f_{\alpha\gamma}(\mu, v) = \int_{\mathbb{R}} S_{\alpha\gamma}(x) \frac{e^{-(x-\mu)^2/2v}}{\sqrt{2\pi v}}$. *The covariance is equal to zero when* $\beta \neq \alpha$ *and:*

$$C^\alpha(t,s) = e^{-(t+s)/\theta_\alpha}\left[C^\alpha(0,0) + \frac{\theta_\alpha \lambda_\alpha^2}{2}(\exp 2(t\wedge s)/\theta_\alpha - 1) + \sum_{\gamma=1}^{P} \sigma_{\alpha\gamma}^2 \int_0^t \int_0^s e^{(u+v)/\theta_\alpha} \Delta_{\mu,C}^{\alpha\gamma}(u-\tau_{\alpha\gamma}, v-\tau_{\alpha\gamma}) du dv\right] \tag{6.4}$$

where $\Delta_{\mu,C}^{\alpha\gamma}(u,v) = \mathbb{E}\left[S_{\alpha\gamma}(X_u^\gamma) S_{\alpha\gamma}(V_v^\gamma)\right]$ *is a nonlinear function of* $\mu^\gamma(u)$, $\mu^\gamma(v)$, $C^{\gamma\gamma}(u,v)$, $C^{\gamma\gamma}(u,u)$ *and* $C^{\gamma\gamma}(v,v)$.

Moreover, there exists a unique solutions to these self consistent equations (6.3) *and* (6.4).

Note also that if the initial condition is not Gaussian, the solutions are not Gaussian. However, as time goes by, solutions get exponentially fast attracted to the Gaussian solutions described in theorem 6.2.1. That description hence provides a handy procedure to analyze the solutions of the mean-field equations and their dynamics as a function of the parameters. In particular, we observe that the levels of heterogeneity, $(\sigma_{\alpha\gamma})$, appear as parameters of the equations. The moment equations provided above hence allow analyzing the qualitative effects of heterogeneity on the behavior of the network.

All these results can be readily confirmed by numerical simulations of the network equations. Considering for instance a two-populations network with parameters given in section 6.3.3, we simulated a network of 12000 neurons (6000 in each population) and considered the distribution of the values of the membrane potentials as a statistical sample. The empirical distribution, superimposed with the theoretical Gaussian distribution, is plotted in figure Fig. 6.1 and shows a very clear fit, which we confirmed using the Kolmogorov-Smirnov test. For each population, the Kolmogorov-Smirnov test comparing the sample obtained by numerical simulations with the predicted Gaussian distribution ensures that the sample has indeed the Gaussian distribution, with a p-value equal to 1. Moreover, we used a chi-square test of independence which validates the independence between the two populations and this independence test was validated with a p-value of 0.87.

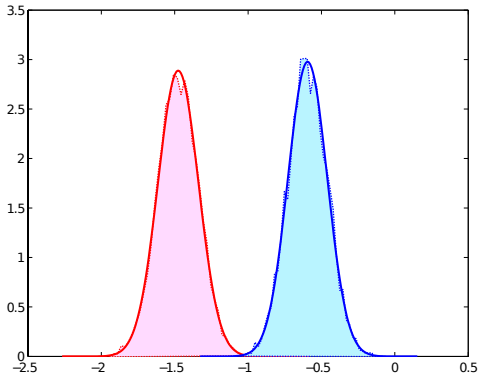


Figure 6.1: Empirical distribution (colored histogram with dotted lines) and theoretical Gaussian distributions for a 2 populations neuronal network (parameters given in section 6.3.3 with common heterogeneity parameter $\sigma = 1$ and noise $\lambda = 0.5$).

6.2.1 Qualitative effects of the heterogeneity parameters

In this section, we use the above derived limit equations to analyze the behavior of networks, with a particular focus on the effect of the disorder parameters, namely the standard deviations of the synaptic weights. This study was done in a one population model with centered coefficients, centered sigmoidal functions S and no delays, in the outstanding work of Sompolinsky, Crisanti and Sommers [221]. We start by generalizing their equations in our context and discussing a particularly interesting phase transition identified in their article, namely the transition from a stationary to a chaotic behavior. We will then discuss the persistence of that phase transition when the assumptions on the parameters (non-centered synaptic weights or sigmoids, delays, multiple populations...) are relaxed. We will then turn to the analysis of multi-populations networks with no delay and we will show that the disorder parameter can trigger or destroy oscillations.

In the present section, in order to revisit the results of SCS, we will sometimes deal with deterministic equations with random coefficients and consider $\lambda_\alpha = 0$. Rigorously, in that case, the proof of the large deviation principle and of the existence and uniqueness of solutions to the mean-field equations no more hold. These non-noisy regimes correspond to limits of the mean-field equations where $\lambda_\alpha \to 0$, and to a sort of viscosity solution of the system: all the properties of convergence, existence and uniqueness of solution hold for arbitrarily small λ_α and provide solutions that have a limit when $\lambda_\alpha \to 0$.

6.2.2 The generalized Somplinsky-Crisanti-Sommers Equations

In their article, Sompolinsky, Crisanti and Sommers (SCS) introduce a set of equations governing the dynamics of covariance of possible stationary solutions to the mean-field equations. These equations are used to analyze the dynamics of the limit process and in particular to show a striking transition between stationary and chaotic solutions. We derive here a generalized equation of the type of the SCS equations in our framework with multiple populations and delays, and use these equations to ex-

plore the boundaries of the SCS phase transition when considering different models.

Proposition 6.2.2. Possible stationary solutions are Gaussian with mean $\bar{\mu}^\alpha$ and covariance $\bar{C}^\alpha(\tau) = C^{\alpha\alpha}(t+\tau,t)$ for any $t \geq 0$. These two variables satisfy the system of equations:

$$\begin{cases} 0 = -\dfrac{1}{\theta_\alpha}\bar{\mu}^\alpha + \displaystyle\sum_{\gamma=1}^{P} \bar{J}_{\alpha\gamma} f_{\alpha\gamma}(\bar{\mu}^\gamma, \bar{C}^\alpha(0)) \\ \ddot{\bar{C}}^\alpha(\zeta) = \dfrac{\bar{C}^\alpha(\zeta)}{\theta_\alpha^2} + \displaystyle\sum_{\gamma=1}^{P} \bar{\Delta}^{\alpha\gamma}_{\bar{\mu},\bar{C}}(\zeta) \end{cases} \tag{6.5}$$

Remark 14. Note that the above equations do not constitute a dynamical system, but rather correspond to implicit equations. In particular, an important difficulty is the choice of the initial condition $C^\alpha(0)$ which corresponds to the variance of the stationary solution, which is obviously unknown. This quantity parametrizes both the equation on the first moment and the form of the term $\Delta^{\alpha\gamma}$ on the second moment equations.

Proof. The equation on the mean $\bar{\mu}^\alpha$ is a simple rewriting of equation (6.3) under stationarity condition. The equation on the stationary covariance requires more care. For arbitrary time t, denoting X_t^α the solution of the mean-field equation with for all α, $\lambda_\alpha = 0$, we have, using equation (6.2):

$$\begin{aligned} \dot{C}^\alpha(\zeta) &= \frac{\mathrm{d}}{\mathrm{d}\zeta}\mathbb{E}[(\bar{X}^\alpha(t+\zeta) - \mu^\alpha(t+\zeta))(\bar{X}^\alpha(t) - \mu^\alpha(t))] \\ &= -\frac{C^\alpha(\zeta)}{\theta_\alpha} + \mathbb{E}[\bar{X}^\alpha(t) U^{\alpha,\bar{X}}(t+\zeta)]. \end{aligned} \tag{6.6}$$

The second term is not easy to characterize. The method used by Sompolinsky and collaborators to deal with this term is to derive a second time with respect to ζ. However, the differential of U^α is unknown. Fortunately, we can express this term as a function of $\delta^\alpha(\zeta) = \mathbb{E}[\bar{X}^\alpha(t+\zeta)U^\alpha(t)]$. This function is way easier to handle since using the differential equation (6.2) and differentiating this expression with respect to ξ, one obtains:

$$\dot{\delta}^\alpha(\zeta) = -\frac{\delta^\alpha}{\theta_\alpha} + \sum_{\beta=1}^{P} \sigma^2_{\alpha\beta} \Delta^{\alpha\beta}_{\bar{\mu},\bar{C}}(\zeta)$$

where we denoted with a slight abuse of notations $\Delta^{\alpha\beta}_{\bar{\mu},\bar{C}}(\zeta)$ the common value of $\Delta^{\alpha\beta}_{\bar{\mu},\bar{C}}(t+\zeta,t)$ for any $t > 0$ using the assumed stationarity of the solution.

In order to relate the second term of the righthand side of (6.6) with δ^α, we compute $\dot{C}(\zeta+\xi)$ expressing it the differential with respect to ξ of $\mathbb{E}[\bar{X}^\alpha(t+\zeta+\xi/2)\bar{X}^\alpha(t-\xi/2)]$. In this computation, most of the terms cancel out and we obtain the simple expression at $\xi = 0$:

$$2\dot{C}^\alpha(\zeta) = \mathbb{E}[\bar{X}^\alpha(t) U^{\alpha,\bar{X}}(t+\zeta)] - \delta^\alpha(\zeta).$$

Plugging this expression into (6.6) we obtain:

$$\dot{C}^\alpha(\zeta) = \frac{C^\alpha(\zeta)}{\theta_\alpha} - \delta^\alpha(\zeta).$$

Differentiating this expression with respect to ζ and reinjecting the latter equation in the obtained expression, we get:

$$\begin{aligned}\ddot{C}^{\alpha}(\zeta) &= \frac{\dot{C}^{\alpha}(\zeta)}{\theta_{\alpha}} - \left(-\frac{\delta^{\alpha}}{\theta_{\alpha}} + \sum_{\beta=1}^{P} \sigma_{\alpha\beta}^2 \Delta^{\alpha\beta}(\zeta)\right) \\ &= \frac{C^{\alpha}(\zeta)}{\theta_{\alpha}^2} - \sum_{\beta=1}^{P} \sigma_{\alpha\beta}^2 \Delta^{\alpha\beta}(\zeta). \end{aligned} \tag{6.7}$$

□

This equation is very similar to the original SCS equation. As they remarked, this equation does not characterize the process. Indeed, we know that $\dot{C}^{\alpha}(0) = 0$ using the fact that the covariance is even, but the initial condition $C^{\alpha}(0)$ is not fixed: it is the asymptotic stationary variance of the process, when it exists, and this initial condition is a parameter of both the stationary mean equation and stationary covariance equation. However, Sompolinsky and collaborators showed very elegantly an important phase transition taking place in this system, analyzing the shape of the potential together with a stability analysis of the solutions. We revisit their results in our more general framework, first in one population systems, and then in higher dimensional systems, and particularly focus on the effects of delays, non-zero mean connectivity and non-centered sigmoids.

6.3 NUMERICAL RESULTS ON RANDOM NEURAL NETWORKS

6.3.1 One population networks

The heterogeneity level appears as a parameter in (6.7). In SCS one-population setting, equation (6.7) can be written as the equation of the position of a particle submitted to a force deriving from a potential Φ_1 (the label 1 denotes the number of populations) which is equal to $-\frac{1}{2}C^2 + \sigma^2\psi$ where ψ is a primitive of Δ considered as a function of C. The shape of the potential for $\sigma < 1/(\theta S'(0))$ ensures that the only possible stationary solution correspond to the case where all neurons are equal to zero. When $\sigma > 1/(\theta S'(0))$, a number of solutions can appear depending on the value of $C(0)$, but all are shown unstable except one corresponding to chaotic regime where the covariance is non-zero. In what follows, we will consider that $S(x) = \mathrm{erf}(gx) = \int_0^{gx} e^{-x^2/2}$. Then it is easy to show using a change of variables (see [239]) that

$$f(x,\sigma) = \mathrm{erf}\left(\frac{gx}{\sqrt{1+g^2\sigma}}\right).$$

Non-delayed networks with non-centered synapses

Let us start by a one-population network with no delays, $\theta = 1$ and $\bar{J} \neq 0$. In that case, it is easy to see that fixed point with mean $\mu = 0$, $C(\zeta) = 0$ is a stationary solution whatever the parameters. It is stable if and only if $\bar{J}\frac{\partial f}{\partial \mu}(0,0) = \bar{J}S'(0) < 1$. Therefore, for fixed σ, the system undergoes a pitchfork bifurcation for $\bar{J} = \bar{J}^* = \frac{\partial f}{\partial \mu}(0,0)^{-1} = 1/g$,

the system undergoes a pitchfork bifurcation, and two new equilibria $\mu^+ > 0$ and $\mu^- < 0$ appear, which are stable. For these equilibria, the null covariance is no more a solution to the equations, and we observe a stationary behavior of neurons with a non-zero standard deviation, i.e. a dispersion of the individual trajectories displaying constant time course.

For fixed $\bar{J}$, the system undergoes SCS-like phase transitions from stationary to chaotic activity when the heterogeneity coefficient σ is increased. This bifurcation occurs, for $\bar{J} < \bar{J}^*$, when σ exceeds $1/(S'(0))$, and for $\bar{J} > \bar{J}^*$ when $\sigma > (\frac{\partial f}{\partial \mu}(\mu^{\pm}, C(0)))^{-1}$, implicitly defining a critical value for σ since $C(0)$ depends on σ.

Since $C(0)$ is an increasing function of function and the differential of S takes its maximum at 0 and decreases to zero at $\pm\infty$, the value of σ corresponding to the secondary phase transition to chaos is an increasing function of σ. Moreover, in that case, the chaotic activity will be no more centered around zero but around the new fixed point $\mu^{\pm}$. A hand-drawn bifurcation diagram reflecting this behavior, together with simulations of the trajectories, is plotted in figure Fig. 6.2.

Delay-induced oscillations

We now consider a one-population network with delays. Without loss of generality, we consider that the time constant is equal to 1. The solutions of the mean-field equations with no heterogeneity are Gaussian processes whose moments reduce to a dynamical system:

$$\begin{cases} \dot{\mu} &= -\mu + \bar{J} f(\mu(t-\tau), v(t-\tau)) \\ \dot{v} &= -2v + \lambda^2 \end{cases}$$

and hence the variance converges towards $\lambda^2/2$. To fix ideas, we consider $S(x) = \mathrm{erf}(gx)$, so that $f(x,v) = \mathrm{erf}(\frac{gx}{\sqrt{1+g^2 v}})$. Since for any v, $f(0,v) = 0$, the null mean is a stationary solution of the equation. Its stability depends on the roots of the characteristic equation (or dispersion relationship):

$$\xi = -1 + \bar{J} \left.\frac{\partial f}{\partial \mu}\right|_{0, \frac{\lambda^2}{2}} e^{-\xi\tau} = -1 + \bar{J} \frac{g}{\sqrt{1+g^2\frac{\lambda^2}{2}}} e^{-\xi\tau}.$$

If all characteristic roots have negative real part, the fixed point $\mu = 0$ is stable. As a function of the parameters of the system, characteristic roots can cross the imaginary axis and yield a destabilization of the fixed point. Turing-Hopf instabilities arise when there exists purely imaginary characteristic roots $\xi = \mathbf{i}\omega$. In that case, we obtain the following equivalent system:

$$\begin{cases} -1 + \bar{J}\frac{g}{\sqrt{1+g^2\frac{\lambda^2}{2}}} \cos(\omega\tau) = 0 \\ \omega = -\bar{J}\frac{g}{\sqrt{1+g^2\frac{\lambda^2}{2}}} \sin(\omega\tau) \end{cases}$$

which has real solutions only for $\bar{J}\frac{g}{\sqrt{1+g^2\frac{\lambda^2}{2}}} > 1$. It is then easy to show that Turing-

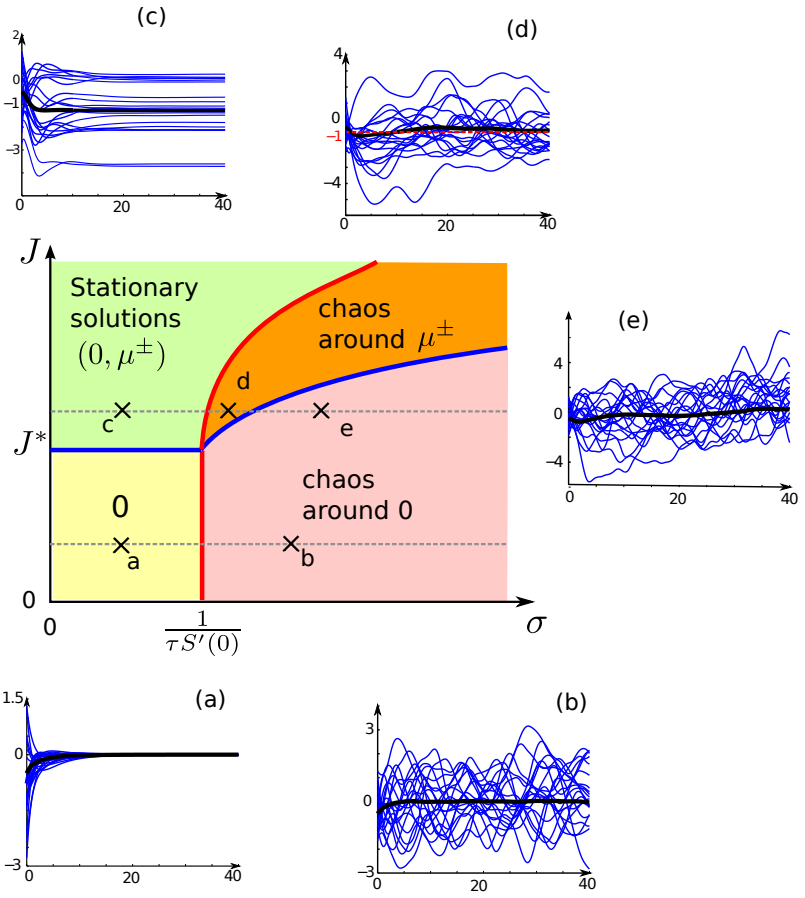
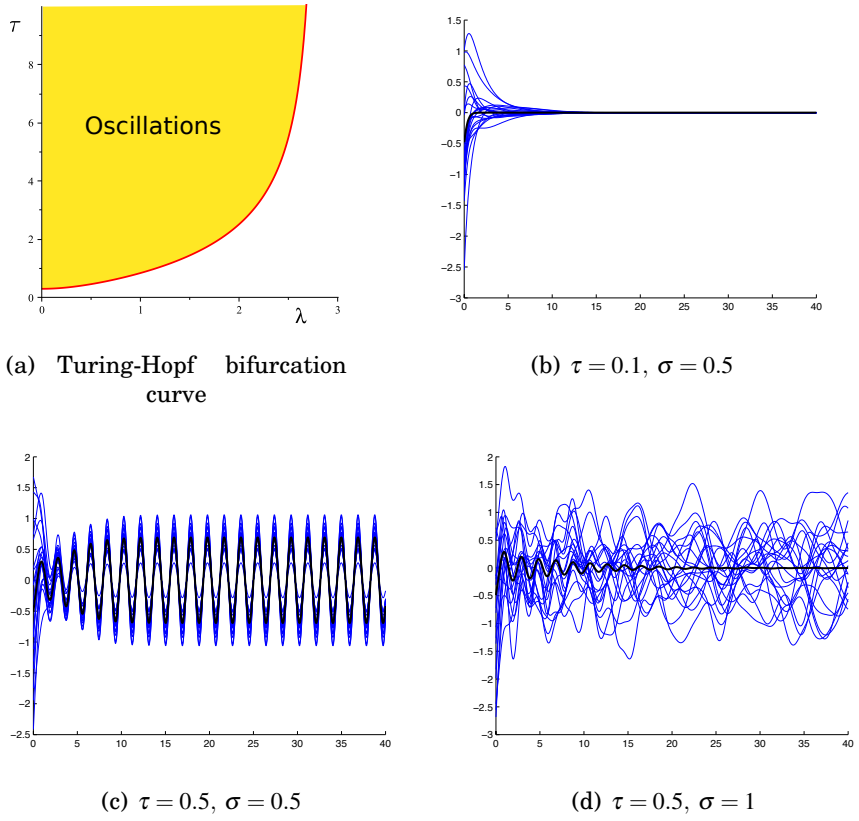


Figure 6.2: Behavior of a one-population system with non-centered synaptic coefficients. Center: Bifurcation diagram (hand-drawn) segmented into four regions: two regions of stationary behavior (yellow: centered at zero and green: centered on $\mu^{\pm}$) and two chaotic regions (pink: centered around zero and orange: centered around $\mu^{\pm}$). The boundaries of these regions are: a pitchfork bifurcation (blue curve) separating the stationary or chaotic regions centered on 0 to the ones centered on $\mu^{\pm}$, and a generalized SCS phase transition (red curve) separating the stationary and chaotic regimes. The subfigures (a)-(e) show the time course of 30 arbitrarily chosen neurons among 10 000 neurons in the network corresponding to the points a-e of the diagram: $\theta = 1$, $S'(0) = 1$, (a): $J = 0.5$, $\sigma = 0.5$, (b): $J = 0.5$, $\sigma = 1.5$, (c): $J = 1.5$, $\sigma = 1.5$, (d): $J = 1.5$, $\sigma = 1.7$, (e): $J = 1.5$, $\sigma = 2$.

Hopf bifurcations arise when the parameters satisfy the relationship:

$$\tau = \frac{\arccos\left(\frac{\sqrt{1+g^2\frac{\lambda_2^2}{2}}}{\bar{J}g}\right)}{\sqrt{\frac{\bar{J}^2g^2}{1+g^2\frac{\lambda_2^2}{2}} - 1}}$$

and these correspond to characteristic roots $\omega = \sqrt{\frac{\bar{J}^2g^2}{1+g^2\frac{\lambda_2^2}{2}} - 1}$. These result into oscillations of the solutions at a pulsation equal by ω.



(a) Turing-Hopf bifurcation curve

(b) $\tau = 0.1,\ \sigma = 0.5$

(c) $\tau = 0.5,\ \sigma = 0.5$

(d) $\tau = 0.5,\ \sigma = 1$

Figure 6.3: One population delayed system, $\theta = 1$, $S(x) = \mathrm{erf}(gx)$, $g = 3$ and $\bar{J} = -2$. (a) represents the curve of Turing-Hopf bifurcations in the plane (τ, λ) obtained analytically. (b): τ small: no oscillations. As the delays are increased, a Turing Hopf bifurcation occurs and oscillations appear (b), which disappear when the heterogeneity σ is increased beyond a critical value in favor a chaotic activity (c).

Let us now return to the case of random coefficients with variance σ and no additive noise $\lambda = 0$. The mean of the Gaussian solution satisfy the same equation as the one studied above with $\lambda = C(0)$, and as noted in the previous section, the stationary covariance is an increasing function of σ. For $\bar{J} > \frac{\partial f}{\partial \mu}(0,0)$, the fixed point 0 is unstable, and the covariance is non-zero. This implies that for sufficiently large values of the delay, the network displays oscillations. Thanks to the propagation of chaos property, all neurons have the same distribution, which is a Gaussian with oscillatory mean, and hence the network displays phase-locked oscillations. Eventually, as noise is increased beyond a critical value, a SCS phase transition occurs and the system no more displays phase locked oscillations but asynchronous chaotic activity. This is illustrated in figure Fig. 6.3.

We conclude that in one population networks, the presence of non-centered synaptic coefficients or of interaction delays qualitatively shape the dynamics of the network. We now turn to study related questions for multi-populations networks.

6.3.2 Multi-population networks

In this section, we analyze the dynamics of randomly coupled neuronal networks in the case of the deterministic coupling of several original SCS networks, before turning the analysis of the dynamics of a more biologically plausible neuronal network composed of an excitatory and an inhibitory population. As demonstrated by Sompolinsky and coworkers in [221], the study of the stationary states using equations (6.5) is very useful to analyze the dynamics of their networks. Unfortunately, this method does not persists in higher dimensions, since the equation does not necessarily derive from a potential. Indeed, in order for the equation to derive from a potential $\Phi_P : \mathbb{R}^P \mapsto \mathbb{R}$ in dimension P greater than 1, the following relationship shall be satisfied for any $\alpha \in \{1,\dots,P\}$:

$$\frac{\partial \Phi_P}{\partial C^\alpha} = -C^\alpha + \sum_{\beta=1}^{P} \sigma_{\alpha\beta}^2 \Delta^\beta$$

The only case where this is possible is the case where $\sigma_{\alpha\beta}^2 = 0$ for any $\alpha \neq \beta$. Indeed, shall the above relationship be true, the equality $\frac{\partial^2 \Phi_P}{\partial C^\alpha \partial C^\beta} = \frac{\partial^2 \Phi_P}{\partial C^\beta \partial C^\alpha}$ directly yields $\sigma_{\alpha\beta}^2 \frac{\partial \Delta^\beta}{\partial C^\beta} = \sigma_{\beta\alpha}^2 \frac{\partial \Delta^\alpha}{\partial C^\alpha}$. The lefthand side is a function of C^α only and the righthand side a function of C^β only, they for $\alpha \neq \beta$ these functions are necessarily constant. For regular functions S, this necessitates to have $\sigma_{\alpha\beta} = \sigma_{\beta\alpha} = 0$. This is precisely the case of deterministic lateral connections between randomly coupled networks, which will now study.

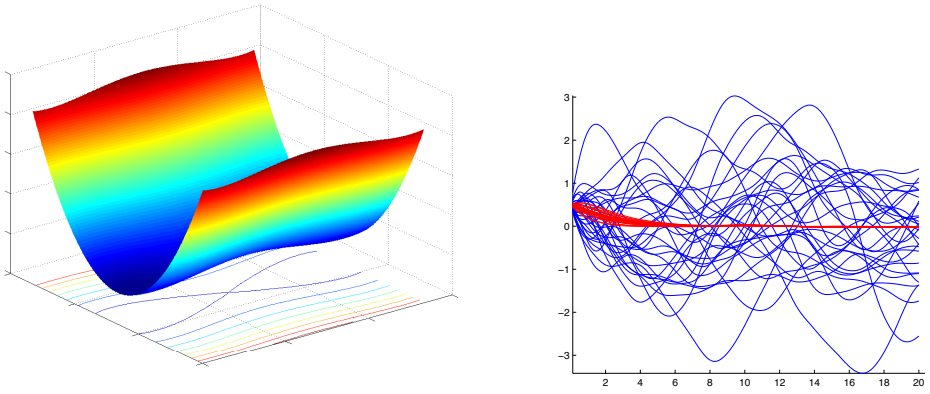
Deterministic lateral coupling of SCS networks

In this section we analyze the coupling of different SCS networks, called lateral coupling, with deterministic coefficients. The only randomness in the models is included in the random synaptic coefficients between neurons belonging to the same population. In that particular case, equation (6.7) derives from the potential $\Phi_P(C^1 \cdots C^P) = \sum_{\alpha=1}^{P} \Phi_1(C^\alpha)$, and in that case the analysis driven by Sompolinsky and collaborators can be adapted to the multi-dimensional case. Since the potential is now the sum of the individual potentials at each population, we observe a strange phenomenon of localization of chaos in the populations that display a large heterogeneity (namely, in our notations, when the SCS condition $\sigma_{\alpha\alpha}^2 S'_{\alpha\alpha}(0)\tau_\alpha > 1$ is satisfied). Only the populations that individually would be in a chaotic state are in a chaotic state, and the other populations converge to zero with a Dirac delta covariance at zero, and the input received by such populations from chaotic populations do not perturb this state. Let us for instance illustrate this phenomenon on a two-populations network with parameters:

$$\bar{J} = \begin{pmatrix} 0 & J_{21} \\ J_{12} & 0 \end{pmatrix} \qquad \text{and} \qquad \sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$$

Each population receives input from the neurons of the other population, with a constant synaptic weight equal to $J_{\alpha\beta}$, and the intra-population synaptic weights are noisy.

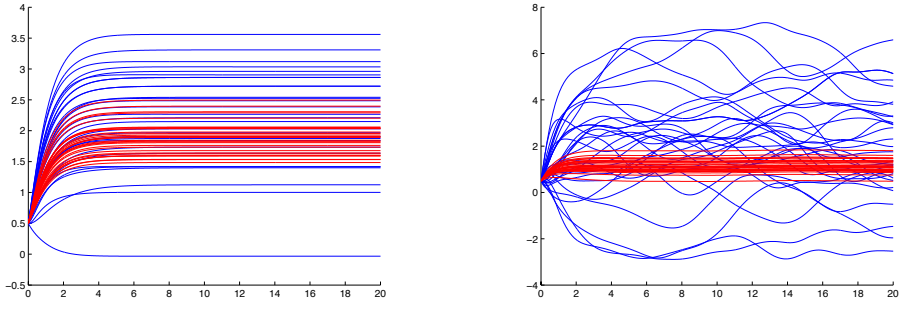
Further analysis of this networks as a function of the coupling reveals a similar phenomenon as the one described in the one-population network of section 6.3.1.



(a) Shape of the potential (b) Trajectories

Figure 6.4: Dynamics for a 2-populations network with Hamiltonian dynamics (no variance on the cross-population synaptic weights): $\theta_1 = \theta_2 = 1$, $\sigma_1 = 3 > 1$, $\sigma_2 = 0.5 < 1$, $J_{12} = J_{21} = 3$. The potential shows a double-well shape, corresponding to a chaotic state on population 1 and a stationary state on population 2. Simulation of a 4000 neurons network illustrate this phenomenon (right): blue (resp. red): 30 arbitrarily chosen trajectories population from 1 (resp. 2).

Indeed, as the strength of the lateral coupling J_{12} and J_{21} are increased, additional stationary solutions with non-zero covariance appear. Let us for instance denote by μ^* the mean of one of these stationary solutions. Following SCS analysis, we are ensured that the behavior of the trajectories of neurons in population α around μ^*_α is stationary as long as $\sigma_\alpha < \frac{1}{\tau_\alpha S'(\mu^*_\alpha)}$ and chaotic otherwise, and this independently of the behavior of the other population. This phenomenon is illustrated in figure Fig. 6.5.

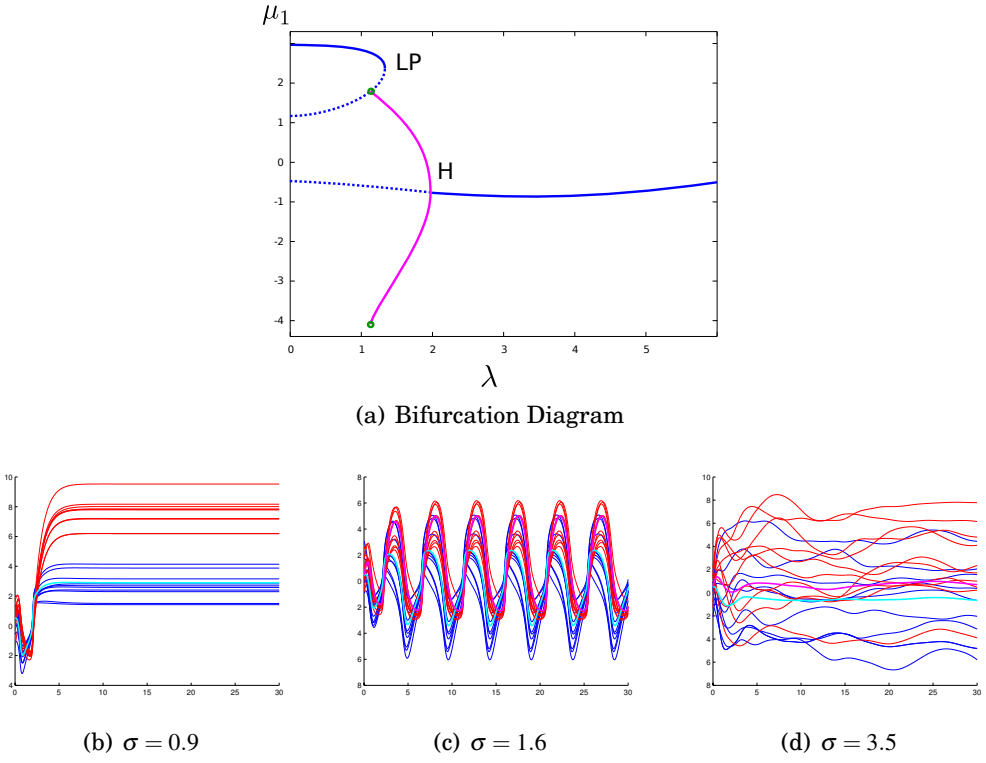


(a) Stationary States (b) Localized Chaos

Figure 6.5: Stationary and partially chaotic solutions of a two populations network with deterministic lateral around non-zero fixed points. $J_{12} = J_{21} = 4$, $\sigma_2 = 0.5$, (left): $\sigma_1 = 2$: both populations display a stationary behavior, (right): $\sigma_1 = 5$: only population 1 is chaotic and the irregularity is not transmitted to population 2.

6.3.3 Heterogeneity-induced oscillations in two-populations networks

Let us eventually discuss the effect of heterogeneities in a more biologically plausible neuronal network. Biologically realistic networks of the type of equation 6.1 analyzed in the present chapter involve at least two populations, one excitatory and one inhibitory, i.e. synaptic coefficients are centered on positive or negative values depending on the population they belong to. Moreover, these include strictly positive sigmoidal transforms (since these functions model the input to firing-rate transformation), that tend to zero at $-\infty$ and to 1 at ∞. The analysis of the present sections justifies a result presented in [238]. For such networks, the dynamics of the stationary solutions no more derive from a potential, and rigorous analysis of the dynamics of the mean-field equations are very involved. In order to uncover the dynamics of such networks, we artificially consider that the solutions achieve a stationary variance $\Gamma_\alpha(\sigma)$ where σ is the multidimensional variable corresponding to the heterogeneity parameters $(\sigma_{\alpha\gamma})_{(\alpha,\gamma)\in\{1\cdots P\}^2}$. Analyzing the dynamics of the average activity as a function of these variables reduces to the analysis of a closed set of ordinary differential equations:



(a) Bifurcation Diagram

(b) $\sigma = 0.9$ (c) $\sigma = 1.6$ (d) $\sigma = 3.5$

Figure 6.6: Heterogeneity-induced oscillations in a two populations excitatory-inhibitory network. (a) Bifurcation diagram of the system of ODEs (6.8) as a function of a presumably constant common value $\Gamma(\sigma)$. LP: saddle-node bifurcation, H: Hopf bifurcation, Sh: Saddle-homoclinic bifurcation, blue: fixed points (solid: stable, dashed: unstable), magenta: limit cycles. (b)-(d): simulations of the heterogeneous network with 2000 neurons per population, for increasing values of the heterogeneity parameter σ displays a transition from stationary to periodic phase-locked behaviors, and then to chaotic behavior. blue (resp. red): 30 arbitrary trajectories of population 1 (resp. 2), cyan (resp. magenta): average on all neurons of population 1 (resp. 2).

$$\dot{\mu}_\alpha(t) = -\frac{1}{\theta_\alpha}\mu_\alpha(t) + \sum_{\beta=1}^{P} \bar{J}_{\alpha\beta} f_{\alpha\beta}(\mu_\beta(t), \Gamma_\beta(\sigma)) + I_\alpha(t) \quad \alpha = 1 \dots P \tag{6.8}$$

Considering $P = 2$ populations, all sigmoids equal to $\mathrm{erf}(gx) = \int_{-\infty}^{gx} e^{-y^2/2}/\sqrt{2\pi}\,dy$ (yielding $f_{\alpha\beta}(x,v) = \mathrm{erf}(gx/\sqrt{1+g^2v})$), all time constants $\theta_\alpha = 1$, and the connectivity matrix, inspired from the seminal article of Wilson and Cowan [245]:

$$\bar{J} = \begin{pmatrix} 15 & -12 \\ 16 & -5 \end{pmatrix}.$$

These equations were studied as a function of the variance parameter $\Gamma_\alpha(\sigma)$ in [239] and displayed transitions to oscillations through a homoclinic bifurcation as the variance increases, oscillations that disappeared through a Hopf bifurcation when that parameter was further increased (see figure Fig. 6.6). Simulations of the network equations show that this phenomenon persists in our case when the variance parameter is induced by heterogeneous connections. Considering all $\sigma_{\alpha\beta}$ equal and denoting σ the common value, we observe that for small heterogeneity parameter σ, the network converges towards a stationary solution with non-zero mean. For intermediate values of the heterogeneity, phase-locked perfectly periodic behaviors appear at the network level, that disappear, as heterogeneity is further increased, through a SCS phase transition yielding chaotic activity.

This phenomenon relates the level of heterogeneity to the presence of synchronized oscillations in networks, an essential phenomenon, as we discuss in the conclusion.

6.4 Discussion

Our setting takes into account important features present in cortical networks: interconnection delays, multiple populations with non-zero average synaptic connection. All these refinements allowed going deeper into the understanding of the dynamics of neuronal networks. In particular, we showed that delays can induce oscillations in a one-population network modulated by the level of heterogeneity, and that non-zero average connectivity yields non-trivial dynamics that were not present in the original SCS model. Moreover, we showed that networks with multiple populations can show relative counter-intuitive phenomena such as the localization of chaos: a few populations can have a chaotic behavior which is not transmitted to the other populations, whatever the connection strength. Another phenomenon we illustrated was the apparition of heterogeneity-induced oscillations, phenomenon first presented in a very recent article [238]. As discussed, the latter phenomenon can be related to experimental studies that showed that the cortex of developing rats subject to absence seizures (abnormal synchronization of some cortical areas) was characterized by the same mean resting membrane potential, but an increased variance [13]. We further showed here that such oscillations were facilitated by the presence of delays. Let us eventually underline that the particular form of our system is not essential in the apparition of such phenomena, and in [238], it is shown that the transition to synchronized activity as a function of heterogeneity levels persists for realistic neuronal networks made of excitable cells, the Fitzhugh-Nagumo model.

An important observations is that in all the examples treated, the SCS phase transition to chaos is present as the heterogeneity is increased. This phenomenon seems relatively universal in this kind of randomly coupled neuronal networks. This chapter was restricted to the analysis of static random synaptic weights. As mentioned in the introduction, precise analysis of the synaptic transmission reveals stochastic variations of the synaptic efficacies that cannot be considered in the present framework. The analysis of networks with stochastic synapses is the subject of the second part of this paper.

The analysis of the present chapter underlines the fact that the structure of connectivity maps is essential to the function of the networks, and illustrated the fact that averaging effects do not cancel the structure into populations and allow serving functions such as oscillations. However, actual brain connectivity maps are not recurrent, and can display different topologies, with different computational capabilities. Moreover, our analysis did not take into account the plasticity mechanisms, resulting in the slow evolution of the synaptic weights as a function of the activity of neurons, which tends to correlate the synaptic weights to the voltage variables. Techniques to rigorously address the dynamics of neuronal networks with non-recurrent connectivity, with specific topologies, or with correlated synaptic weights, are deep questions that are still largely left unexplored, and we can expect that a wide range of novel phenomena will arise from the analysis of such networks.

CHAPTER 7

STUDY OF THE RANDOM KURAMOTO MODEL

7.1 INTRODUCTION

In South Asian forests, every night gives way to an astonishing phenomenon. After sunset, when darkness has already well spread over the country, a few dim lights coming from the trees and wild grasses begin to sparkle, each one at its own rhythm. Slowly, more and more twinkling join this luminous performance until, all of a sudden, hundreds flashes of light reach a perfect unison.

It was a great surprise to scientists of the beginning of the twentieth century that this uncommon spectacle was actually produced by fireflies. In fact, these insects are capable of lighting their bulb, and instinctively adjust their flashing in order to reduce the delays they have with siblings. As puzzling as it may seems, many other spontaneous examples of synchrony are to be found in nature, even involving nonliving entities. From Earth, achieving an entire rotation upon its axes in twenty four hours, to humans naturally syncing their sleeping time on the disappearance of the sun, our physical world is filled with such biological or physical temporal harmonies.

7.2 PRESENTATION OF THE MODEL

The Kuramoto model was first introduced by Yoshiki Kuramoto to describe the behavior of large assemblies of oscillators with heterogeneous natural frequencies [152, 153, 154, 247, 248]. It is now considered as the canonical model of coupled oscillators (see the excellent reviews [2, 226, 227]). Its study provides insightful results in such distinct domains as synchronization of pacemaker cells in the heart [189], synchronization of fireflies flashes [39], arrays of lasers [144], power networks [88], rhythmic applauses [181], and superconduction [244].

The Kuramoto consists of a large set of coupled nonlinear ODE in random environment. Here is its original form as introduced by Kuramoto:

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t}(t) = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i). \tag{7.1}$$

These equation describes N oscillators through there phase θ_i. As appears in the model, it is convenient to work on the torus, by considering that $\theta = (\theta_1, \dots, \theta_N) \in [0, 2\pi]^N$. In the classical model, the interactions are all-to-all, with same amplitude $\frac{K}{N}$. This kind of interaction is generally called *mean-field* (remark that other models with nearest neighbor interactions have also been studied [75, 210, 228]). The non-linear entrainment term is calculated through a sine function. It has a tendency to push the assembly towards synchronization. Furthermore, ω_i denote the natural frequency of oscillator i when it feels no interactions. The ω_i are independent identically distributed with density given by $g(\omega)$. This density is generally supposed centered, as we can always make the change of variable $\theta_i \to \theta_i - \Omega t$ with $\Omega = \mathscr{E}(\omega_i)$. Moreover, g is also often assumed even and unimodal.

This quenched disorder is motivated by the fact that a group of manufactured (metronomes, Josephson's junctions) or biological (fireflies, runners, persons applauding etc.) oscillators is always composed by slightly distinct individuals with different natural frequencies. For living individuals, this disparity might be important for the survival of the species. Hence, the Kuramoto model is a disordered system. The original question of Kuramoto was to understand in which extent synchronization was possible in a system composed of many distinct oscillators. To answer this question, he introduced a simple quantity describing the level of synchrony called *order parameter*:

$$r(t) \exp\{i\psi(t)\} := \frac{1}{N} \sum_{j=1}^{N} \exp\{i\theta_j(t)\}. \tag{7.2}$$

When $r(t) \simeq 0$, oscillators are well spread on the unit circle, so that the system is near incoherence. In comparison, if $r(t) \simeq 1$, the phase of the oscillators are almost equal, so that the system is in coherence. Strikingly, in the limit of an infinite number of oscillators, Kuramoto discovered a phase transition in the amplitude of interactions. For t big enough, the system reaches a steady state in which exists a critical value of this amplitude $K_c \in \mathbb{R}_+$ such that:

- when $K < K_c$, the phases of the oscillators are uniformly distributed on $[0, 2\pi]$, a state called *incoherence*. Moreover, this state is stable, and perturbations form this equilibrium are subject to a Landau damping ([108, 229])
- when $K > K_c$, partially synchrony emerges in which part of the oscillators is *locked* near the averaged phase ψ with $\dot{\theta}_i = 0$, while the other part evolves out of synchrony.

In this last phase, the locked oscillators are those whose natural frequency is small enough in order to be attracted by the synchronized group. In fact, rewriting equation (7.1) with the order parameter (7.2), we obtain:

$$\begin{aligned} \frac{\mathrm{d}\theta_i}{\mathrm{d}t}(t) &= \omega_i + K \operatorname{Im}\left(\exp\{-i\theta_i(t)\} \frac{1}{N} \sum_{j=1}^{N} \exp\{i\theta_j(t)\} \right) \\ &= \omega_i + K r(t) \sin\big(\psi(t) - \theta_i(t)\big). \end{aligned}$$

Hence, supposing that the order parameter has reached a steady state, $\dot{\theta}_i = 0$ implies

$$\omega_i = rK\sin\left(\theta_i - \psi\right)$$

equation that can be satisfied only if $|\omega_i| \leq rK$. Hence, in the partial synchrony regime, the locked oscillators all have the same frequency, but their phase are locked to slightly different positions distributed around the stationary averaged phase ψ, and given by

$$\theta_i = \psi + \arcsin\left(\frac{\omega_i}{rK}\right).$$

In comparison, oscillators with large individual frequency ω_i can never join this group. Moreover, their velocity is subject to changes along time: they are all the more rapid that their phase is near ψ. In the case of an even distribution of individual frequency $g(\omega)$, the critical value of the interaction amplitude K_c was found to follow the equation:

$$K_c = \frac{2}{\pi g(0)}.$$

Remark that this transition can also be seen to be dependent on the level of disparity of individuals. In fact, assuming that the ω_i are centered Gaussian variables with variance μ^2, and that K is fixed, one can see that exists a critical value of the heterogeneity parameter μ_c governing the transition from partial synchrony to incoherence:

$$\mu_c = \frac{K\sqrt{\pi}}{2\sqrt{2}}.$$

Despite the fact that important work that has been realized on the classical Kuramoto model, the dynamics of the randomly coupled Kuramoto system in the scaling considered in this thesis is still an open and debated issue [76, 77, 224, 225]. The results of chapter 5 provide a formal characterization of the limit, allowing to show how disorder competes with deterministic connectivity for the emergence of synchrony. We specifically apply these results to reveal how the disorder in the connections modifies the macroscopic synchronization properties of the system. A trade-off between the averaged connection amplitude and its fluctuations is quantitatively characterized using numerical simulations, allowing to exhibit how the synchronization transition is affected by the fluctuations of the interactions and intrinsic frequencies disorder of the oscillators.

7.3 KURAMOTO MODEL IN NEUROSCIENCE

Applying some of our analysis developed for neural networks to the Kuramoto model of coupled oscillators is not only anecdotal. Indeed, this model shares some interesting links with neuroscience, as the brain has been shown to display oscillatory behaviors. In fact, besides the well-known Greek rhythms observed by EEG recording, experimental studies have demonstrated that mammalian brain displayed many oscillatory patterns in neural assemblies found in distinct regions such as the olfactory, visual, and auditory cortex, the cerebellum and thalamus, as well as the olfactory bulb . Moreover, these patterns are also found in non-mammalian animals, as the optic tectum of pigeon, or olfactory bulbs of insects behaved in the same fashion [182].

One pioneering experimental works in this filed is that of Gray, Singer, Eckhorn and collaborators on the primary visual cortex of anesthetized cats [92, 123, 124]. Different electrodes were plugged to the animal's brain, and the visual stimuli the authors used were light bars traveling on a screen. They not only showed that stimuli induced synchronized oscillations of neurons within a cortical column, but also that - when the animal was adequately stimulated - coherence was found for neighboring hypercolumns (macrocolumns of the mammalian visual cortex), distant hypercolumns (up to 7mm), but also remote hypercolumn located in two different region of the cortex. These oscillatory responses were short-lasting (a few ms), with a frequency ranging from 35 to 60 Hz, and both the bars orientation, direction, velocity and position impacted the oscillations.

Because of this sensitivity to such characteristics, it was hypothesized in [123] that this collaboration between nearby or far away assemblies of neurons might permit to put into relation the different features of the presented stimulus. This would be a possible solution to the *binding problem*, which could be formulated as follow: "how can the brain bind together all the features of one object, and segregate them from those of another jointly perceived object?" [212, 217]. Under this hypothesis, also named the *labeling hypothesis* [17], neuronal structures oscillating with the same phase and/or frequency share the same label, and can be easily differentiated from other labels presenting different kind of oscillations, thus resolving this associative problem. Furthermore, other cognitive abilities such as memory, categorization, or attention, may also involve coherent neural oscillations [182]. Moreover, mathematical models - involving oscillators model and neural networks - have been proposed, that have accounted for several of these synchronization patterns [17, 128].

Ermentrout argued that Kuramoto model furnishes a useful framework to study the synchrony observed throughout the nervous system. He emphasized the fact that without stimulus, the natural observed patterns in the cortex are traveling waves. In [97], he propose some possible computational benefits - unconcerned with the binding effect - of these temporal patterns: while the waves could result in an better sensitivity to stimuli, oscillations might enhance synaptic plasticity. Moreover, in an analogy between syncing cortical columns, and the coupled oscillators from the Kuramoto model, Ermentrout argued that correlated noise could induce oscillations in cortical areas, for macrocolumns computing the same feature of a stimulus [96].

7.4 APPLICATION OF OUR RESULTS TO THE RANDOM KURAMOTO MODEL

We now apply the theoretical results obtain in chapter 5 to investigate the dynamics of the randomly connected Kuramoto model. This system describes the phase of one oscillator within its cycle, $X_t^{i,N} \in [0, 2\pi)$, rotating at its own natural frequency ω_i, and constantly adjusting to the others phase through a simple sine function:

$$dX_t^{i,N} = \left(\omega_i + \sum_{j=1}^{N} J_{ij} \sin(X_t^{i,N} - X_t^{j,N}) \right) dt + \lambda dW_t^i.$$

In recent years, important work was devoted to characterize the limits and fluctuations of the system around the mean-field regime (low levels of fluctuations) [27,

62, 63, 64, 118, 121, 136, 168, 174, 205].

The dynamics of the randomly connected Kuramoto model is much less understood. This question was raised twenty years ago by physicists who showed, using statistical physics methods and numerical simulations, a rich phenomenology [76, 77, 224, 225]. The model we studied here provides a mathematical approach to characterize the averaged and quenched behavior of coupled oscillators with independent non-centered random interactions with variance σ^2/N.

In detail, the randomly interacting Kuramoto model corresponds to the case where ω represents the intrinsic frequency of each oscillator, $\mathscr{E}(\omega) = 0$ (without loss of generality), $\mathscr{E}(\omega^2) = \mu^2$, $f(\omega,t,x) = \omega$, and $b(x,y) = \sin(y-x)$, i.e. the dynamic of the finite-size network reads:

$$dx_t^i = \Big(\omega_i + \sum_{j=1}^{N} J_{ij}\sin(x_t^j - x_t^i)\Big)dt + \lambda\, dW_t^i, \tag{7.3}$$

where the J_{ij}s are independent, identically distributed Gaussian variables:

$$J_{ij} \sim \mathscr{N}\Big(\frac{\bar{J}}{N}, \frac{\sigma^2}{N}\Big).$$

As proposed in the original work of Kuramoto [152], the dynamics can be well understood by recasting the Kuramoto equation in a complex form and introducing one complex variable, the *order parameter*. In the classical Kuramoto model (when $J_{ij} = \bar{J}/N$), the interaction term can indeed by expressed via a single complex variable $Z^N(t) := \frac{1}{N}\sum_{j=1}^{N} e^{ix_t^j} =: r^N(t)e^{i\psi^N(t)}$ (the order parameter):

$$\sum_{j=1}^{N} \frac{\bar{J}}{N}\sin(x_t^j - x_t^i) = \bar{J}r^N(t)\sin(\psi^N(t) - x_t^i). \tag{7.4}$$

As appears in the above expression, all particles are attracted towards the argument of Z^N with a strength proportional to its norm $r^N \in \{z \in \mathbb{C}, |z| \leq 1\}$. This last parameter provides a very convenient measure of the synchrony of the particles. From the disorder in the interaction strength of the random Kuramoto model, the order parameter $Z^N(t)$ is no more sufficient to describe the effect of other oscillators onto an arbitrarily chosen oscillator. Actually, each oscillator feels a distinct effect of the network which is due to the independence of the random connectivity coefficients. One can thus no more reduce the dynamics to the one of a single complex variable. It is convenient to introduce for each oscillator i its individuality parameter $Z^{N,i}(t)$, whose definition is analogous to the one of the order parameter, but that incorporates the unique way oscillator i interacts with the network:

$$Z^{N,i}(t) := \sum_{j=1}^{N}\Big(J_{ij} - \frac{\bar{J}}{N}\Big)e^{ix_t^j} =: \sigma\sum_{j=1}^{N}\tilde{J}_{ij}e^{ix_t^j} =: \sigma\rho_i^N(t)e^{i\phi_i^N(t)},$$

with $\tilde{J}_{ij} := \frac{1}{\sigma}(J_{ij} - \frac{\bar{J}}{N}) \sim \mathscr{N}\big(0, \frac{1}{N}\big)$ are independent standard Gaussian variables. We note that, in contrast to the modulus of the order parameter r_N, the modulus of the individuality parameter ρ_i^N may exceed 1.

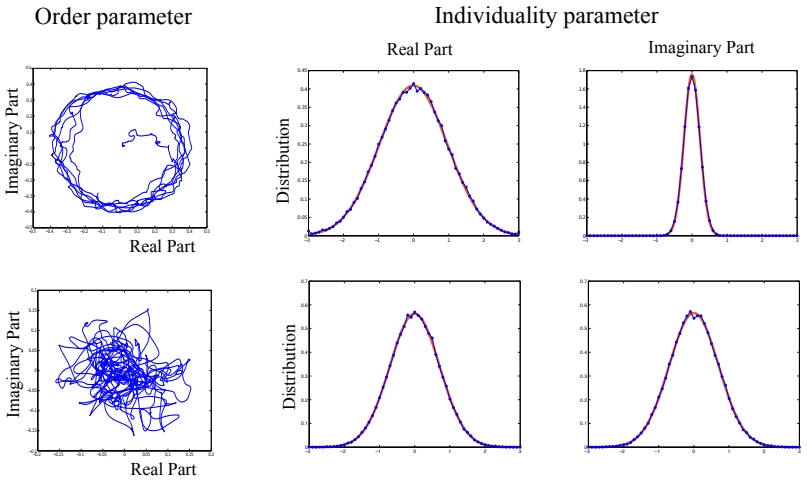


Figure 7.1: Trajectory of the order parameter in the complex plane, together with the distributions of the real part and imaginary part of the depolarized individuality parameters in a synchronized (top row) or desynchronized (bottom row) case. Parameters: $N = 2500$, $\sigma = 1$, $\mu = 1$, simulated with $dt = 0.01$ over a period of $T = 500$. Top row: $J = 5$, bottom row: $J = 1.5$. The statistics are taken over 30 independent realizations and across oscillators. The blue stars and curve correspond to the empirical distribution obtained with the numerical simulations, and the red curve is the Gaussian distribution with mean and variance evaluated on the same data. We observe that the empirical distribution is indeed well described by a Gaussian variable.

With these two notations, it is easy to rewrite the dynamics of the random Kuramoto model (7.3) in the following form:

$$dx_t^i = \Big(\omega_i + \bar{J} r^N(t)\sin(\psi^N(t) - x_t^i) + \sigma\rho_i^N(t)\sin(\phi_i^N(t) - x_t^i)\Big)dt + \lambda dW_t^i. \tag{7.5}$$

This microscopic equation reveals the interplay of coupling strength and disorder. The mean coupling strength has a synchronizing effect by driving all oscillators, through the order parameter $Z^N(t)$, to reach the phase $\psi^N(t)$ with an intensity $r^N(t) \leq 1$. In contrast, whereas the disorder in the coupling strengths creates a similar drive (individuality parameters $(Z^{N,i}(t))_{i=1\cdots N}$ attract oscillators $(x^i(t))_{i=1\cdots N}$ towards the distinct phases $(\phi_i^N(t))_{i=1\cdots N}$ with intensities $(\rho_i^N(t))_{i=1\cdots N}$), the heterogeneity of individuality parameters translates into a source of frustration for the synchronization of the network. We further note that, as opposed to the common drive, these intensities are unbounded.

The theory does not provide the correlation between the parameters $(Z^N(t), (Z^{N,i}(t))_{i=1\cdots N})$, but under the heuristic Boltzmann's molecular chaos hypothesis, one may expect the parameter $Z^N(t)$ to approach $\mathbb{E}\left[e^{ix_t^j}\right] =: \bar{r}(t)e^{i\bar{\psi}(t)}$. If synchronization occurs, all oscillators shall be polarized around a random phase $\bar{\psi}(t)$. Boltzmann's molecular chaos points towards the fact that the depolarized individuality parameters $\tilde{Z}^{N,i}(t) = e^{-i\psi^N(t)}Z^{N,i}(t)$ shall approach independent centered Gaussian variables. We confirmed this intuition using numerical simulations in Figure 7.1 in a case where the oscillators do not synchronize and in a case where they do.

Based on this independence, we conclude that the two interaction terms of equation (7.5) play opposite roles of synchronization and desynchronization. We quanti-

tatively study this competition using numerical simulations in section 7.4.2. Before we proceed to this analysis, we start by comparing our limit equation with previous works.

7.4.1 Comparison with previous results

The case treated previously in the physics literature corresponds either to $\sigma = 0$ for the classical Kuramoto model, or to $\bar{J} = 0$ in the random one that extends to the case where the J_{ij} may be dependent to the reciprocal connection J_{ji} with a correlation η. Using methods developed in the domain of the physical analysis of spin glasses, they obtain for this latter setting a mean-field equation for the thermodynamic limit that would write, in our notations, as:

$$dx_t(\omega) = \Big(\omega + \sigma\,\mathrm{Re}(e^{ix_t(\omega)}\zeta(t)) + \eta R(t)\Big)\,dt + \lambda\,dW_t \tag{7.6}$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion, ζ is a centered complex-valued Gaussian process with covariance given by a self-consistent formula:

$$\begin{cases} \mathbb{E}[\zeta(t)\zeta(t')] = K_+(t,t') = \mathbb{E}[e^{ix_t}e^{ix_{t'}}] \\ \mathbb{E}[\zeta(t)\zeta^*(t')] = K_-(t,t') = \mathbb{E}[e^{-ix_t}e^{ix_{t'}}] \end{cases}$$

and $R(t)$ is a term emerging from the symmetry of the interaction, made explicit in [224], but whose complex expression is not given here since our techniques do not allow to cover the case of partly symmetric interactions.

It is clear that the Kuramoto model enjoys the regularity hypotheses of the theoretical developments of the previous sections, in the case of independent non-symmetric disorder of the connections. Theorems 5.2.2 and 5.2.3 ensures that the empirical measure satisfies a weak LDP and converges towards a unique law Q characterized by Theorem 5.2.4. The object of this section is to derive more explicitly the limit of the disordered Kuramoto system and to check the consistency with previously derived limits.

The theory directly applies to the disordered Kuramoto equation (7.3). First of all, a direct application of theorem (5.4.1) ensures that there exists a unique minimizer Q to the associated good rate function. Furthermore, we know that this measure on $\mathscr{C}\times D$ can be decomposed as $dQ(x,\omega) = dQ_{\omega(x)}d\pi(\omega)$ where $Q_\omega \in \mathscr{M}_1^+(\mathscr{C})$ and $Q_\omega \ll P_\omega$. We will now provide a representation to this process as the solution of an implicit stochastic differential equation similar to (7.6). To this purpose, we fix $(e_j^Q)_{j\in\mathbb{N}}$ an orthonormal basis of the real Hilbert space $\mathbb{L}^2_Q(\mathscr{C}\times D)$. We draw a sequence $\big(J_j\big)_{j\in\mathbb{N}}$ of independent standard Gaussian $\big(\hat{\Omega},\hat{\mathscr{F}},\gamma\big)$-random variables with variance σ^2, and define the process (see (12)):

$$G_t^Q(x) := \sum_{j\in\mathbb{N}} J_j \int_{\mathscr{C}\times D} \sin(y_t - x_t)e_j^\mu(y,\omega)d\mu(y,\omega). \tag{7.7}$$

Theorem 7.4.1. *Consider the random Kuramoto system* (7.3) *on the time interval* $[0,T]$ *for any* $T < \frac{\lambda^2}{2\sigma^2}$. *We have the following convergence results*

(i) *Under* Q^N, *the law of each oscillator* $(x_t^i)_{t\leq T}$ *converges towards* Q,

(ii) *Let $\bar{X}(\omega)$ be the unique strong solution of the SDE restricted to the interval $[0,T]$:*

$$\begin{cases} d\bar{X}_t(\omega) = \Big(\omega + m_Q\big(t,\bar{X}(\omega)\big) + G_t^Q\big(\bar{X}(\omega)\big)\Big)\,dt + \lambda\,dW_t \\ \bar{X}_0(\omega) \overset{\mathscr{L}}{=} \mu_0(\omega) \end{cases} \tag{7.8}$$

where $(G_t^Q)_{t\leq T}$ is given by (7.7) *and $m_Q(t,x_t) := \frac{\bar{J}}{\lambda}\int_{\mathscr{C}\times D}\sin(y_t - x_t)dQ(y,\omega')$ for $x\in\mathscr{C}$ and $t\in[0,T]$. Then, the law of $(\bar{X}_t(\omega))_{t\leq T}$ averaged over the realization of the effective interactions G^Q, is equal to Q_ω.*

Proof. The convergence result (point 1.) is a direct consequence of Proposition (5.2.5). In order to demonstrate point 2., we first need to show that there exists a unique strong solution to equation (7.8). This is a simple result of the classical theory of stochastic differential equations, once noted that m_Q and G^Q are γ-almost surely Lipschitz-continuous in the variable x. Indeed, we have

$$\sin(y_t - x_t) - \sin(y_t - z_t) = \sin(y_t)\big(\cos(x_t) - \cos(z_t)\big) - \cos(y_t)\big(\sin(x_t) - \sin(z_t)\big).$$

The conclusion follows from Girsanov's theorem and the characterization of the solution (5.30). □

In order to confront our result to the formula (7.6) given in [224], we remark that $G_t^Q\big(\bar{X}(\omega)\big)$ can be written as the imaginary part of a complex process:

$$\begin{aligned} G_t^Q\big(\bar{X}(\omega)\big) &= \sum_{j\in\mathbb{N}} J_j \int_{\mathscr{C}\times D} \mathrm{Im}\Big(e^{i(y_t - \bar{X}_t(\omega))}\Big) e_j^Q(y,r')dQ(y,\omega') = \sum_{j\in\mathbb{N}} \mathrm{Im}\left(e^{-i\bar{X}_t(\omega)} J_j \int_{\mathscr{C}\times D} e^{iy_t} e_j^Q(y,\omega')dQ(y,\omega')\right) \\ &= \mathrm{Im}\left(e^{-i\bar{X}_t(\omega)} \underbrace{\sum_{j\in\mathbb{N}} J_j \int_{\mathscr{C}\times D} e^{iy_t} e_j^Q(y,\omega')dQ(y,\omega')}_{=:\zeta_t^Q}\right) \end{aligned}$$

where $\zeta_t^Q = a_t^Q + ib_t^Q$ is a well-defined non isotropic complex-valued process with correlated Gaussian imaginary and real parts:

$$\begin{aligned} a_t^Q &:= \sum_{j\in\mathbb{N}} J_j \int_{\mathscr{C}\times D} cos(y_t) e_j^Q(y,\omega')dQ(y,\omega') \sim \mathscr{N}\Big(0, \sigma^2 \int_{\mathscr{C}\times D} \cos(y_t)^2 dQ(y,\omega')\Big) \\ b_t^Q &:= \sum_{j\in\mathbb{N}} J_j \int_{\mathscr{C}\times D} sin(y_t) e_j^Q(y,\omega')dQ(y,\omega') \sim \mathscr{N}\Big(0, \sigma^2 \int_{\mathscr{C}\times D} \sin(y_t)^2 dQ(y,\omega')\Big) \\ \mathscr{E}\big[a_t^Q b_t^Q\big] &= \sigma^2 \int_{\mathscr{C}\times D} cos(y_t) sin(y_t) dQ(y,\omega'). \end{aligned}$$

Note that such processes with non isotropic distributions are not generally called complex Gaussian processes. Nevertheless, the statistics are exactly those given in [224]:

$$\mathscr{E}\Big(\zeta_t^Q \zeta_s^Q\Big) = \int_{\mathscr{C}\times D} e^{iy_t} e^{iy_s} dQ(y,\omega'), \qquad \mathscr{E}\Big((\zeta_t^Q)^* \zeta_s^Q\Big) = \int_{\mathscr{C}\times D} e^{-iy_t} e^{iy_s} dQ(y,\omega').$$

However, to the difference of the equation provided in [224], we find, in the centered case $\bar{J} = 0$, an interaction term involving the imaginary part of the term $e^{-ix_t}\zeta_t^Q$ rather than the real part of $e^{ix_t}\zeta_t^Q$. As explicit in the above calculations, the imaginary part

directly comes from the sine nature of the interaction here, and we expect that the same shall be true in [224].

When taking into account non-centered interactions, an additional term appears in the limit equation. Interestingly, this mean term $m_Q(t,\bar{X}_t(\omega))$ can be written as follow using the asymptotic order parameter

$$\bar{Z}(t) := \int_{\mathscr{C}\times D} e^{iy_t} dQ(y,\omega') = \bar{r}(t)e^{i\bar{\psi}(t)},$$

$$m_Q(t,\bar{X}_t(\omega)) = \bar{J}\,\mathrm{Im}\Big(e^{-i\bar{X}_t(\omega)}\bar{Z}(t)\Big) = \bar{J}\bar{r}(t)\sin(\bar{\psi}(t)-\bar{X}_t(\omega)).$$

Define now:

$$\check{Z}(t) := \sum_{j\in\mathbb{N}} J_j \int_{\mathscr{C}\times D} e^{iy_t} e_j^Q(y,\omega') dQ(y,\omega') = \bar{\rho}(t)e^{i\bar{\phi}(t)}.$$

Then, the general limit equation of the averaged network as the same law as the solution of

$$d\bar{X}_t(\omega) = \Big(\omega + \bar{J}\bar{r}(t)\sin\big(\bar{\psi}(t)-\bar{X}_t(\omega)\big) + \sigma\bar{\rho}(t)\sin\big(\bar{\phi}(t)-\bar{X}_t(\omega)\big)\Big)\,dt + \lambda\,dW_t \tag{7.9}$$

averaged over γ. Observe that it can be directly compared to equation (7.5). In that sense, $\check{Z}(t)$ provides an effective limit for the law of each individual order parameter $Z^{N,i}(t)$ for which no classic limit theorem applies.

7.4.2 Numerical analysis of the phase transition

The classical Kuramoto model ($J_{ij} = \frac{\bar{J}}{N}, \lambda = 0$) displays a celebrated phase transition as a function of the coupling strength $\bar{J}$: when $\bar{J}$ is smaller than a certain critical value $\bar{J}_c$, the system reaches a stable stationary state called incoherence. In this state, the phase of the oscillators are uniformly distributed across all possible phases, so that the order parameter tends towards 0 when N diverges. In contrast, when $\bar{J} > \bar{J}_c$, $r^N(t)$ remains non trivial in the thermodynamic limit and a partially synchronized state emerges locking all sufficiently slow oscillators at a common frequency. Moreover, when the distribution of the heterogeneity parameters $\omega_i \in D$ has an even density g, the critical value can be expressed explicitly as $\bar{J}_c = \frac{2}{\pi g(0)}$. In this section, we numerically investigate the effect of the heterogeneity parameter of the random connections σ on this phase transition.

In order to illustrate this dependence, we characterize the dynamics of the disordered Kuramoto model with centered Gaussian natural frequencies $\mathscr{N}(0,\mu^2)$, centered standard Gaussian initial conditions for the particles. To isolate the role of disorder, we consider similarly to the classical case a non-noisy network ($\lambda = 0$), for which the large-deviation results do not hold. We performed extensive numerical Monte-Carlo simulations to identify the possible transition from incoherence to partial synchrony for varying σ and for $\mu \in \{1,2\}$. The theory for $\sigma = 0$ predicts that the transition occurs at the critical coupling value $\bar{J}_c = 2\mu\sqrt{\frac{2}{\pi}}$. For all values of σ tested, we found a similar transition between incoherence and partial synchrony, but the value of the critical coupling $J_c^\mu(\sigma)$ increases with both μ and σ (see Fig. 7.2). In order to detect the transition, we have used the very different distributions of moduli of Z^N in the synchronized and incoherent regimes, owing to the fact that, in the incoherent phase, the order parameter fluctuates erratically around the origin, whereas

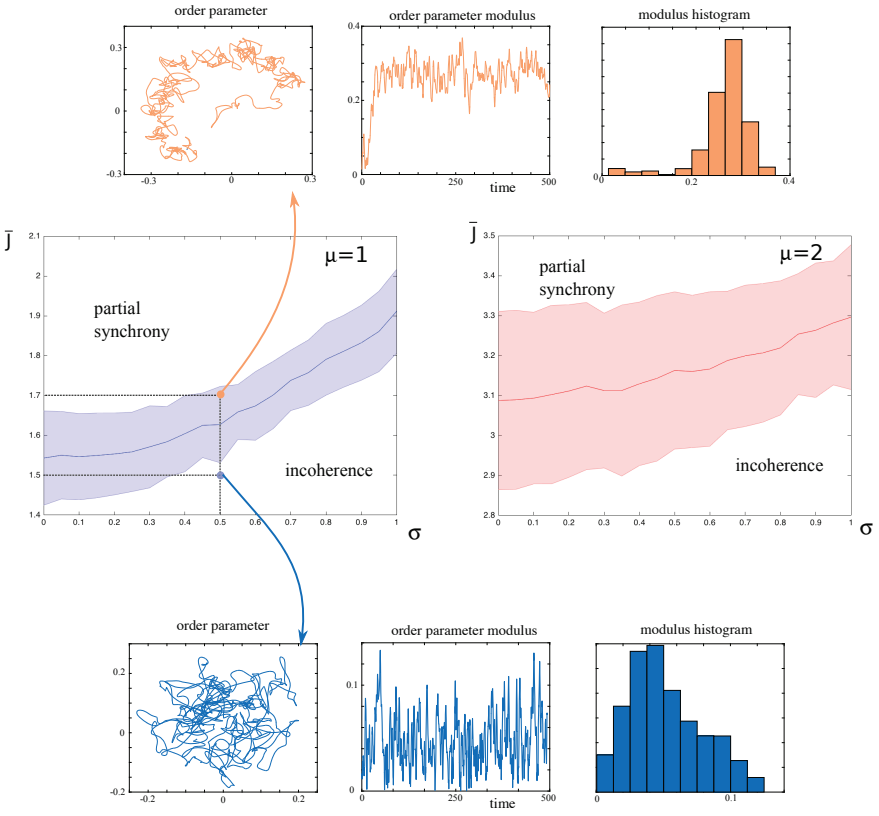


Figure 7.2: Transition from incoherent to partially synchronized activity in the random Kuramoto model. The curves have been obtained through extensive Monte-Carlo simulations for $N = 1500$, total time $T = 150$ and timestep 0.01. We computed, for 20 independent realizations of the disorder, a critical transition value using a dichotomized search. The transition curves are computed for $\mu = 1$ and $\mu = 2$: we observe the monotonic shape of the transition curve as a function of σ, and the fact that the transition occurs for strictly larger values of $\bar{J}$ for $\mu = 2$ compared to $\mu = 1$.

in the partially synchronized phase its modulus $r^N(t)$ accumulates at a specific positive value. Considering thus the distribution of the modulus of the order parameter allows distinguishing between the trajectories that accumulate around the origin and those being significantly peaked away from the origin.

The behavior of $J_c^{\mu}(\sigma)$ is consistent with the limit we have found, and is the expression of the competition between the entrainment of both $Z^N(t)$ and $Z^{N,i}(t)$ as appears in equation (7.5).

Conclusion and perspectives

In this thesis, we have analyzed neural networks presenting different level of heterogeneity, and obtained limit theorem involving the convergence of the empirical measure and propagation of chaos in the limit where the number of neurons tends toward infinity. These models notably include nontrivial random synaptic connections with variance scaling as the inverse of the number of neurons, case we have been able to address through the establishment of large deviations principles. To this end, we have generalized the large deviations approach developed in the context of spin glasses by Gérard Ben Arous and Alice Guionnet [22, 129] to the case of nonlinear interactions. Moreover, we have also further their results to incorporate: (i) non-centered synaptic weights [51, 52, 53], introducing additional deterministic terms in the coupling, (ii) transportation delays [51, 52], which projects the problem into infinite dimensions, (iii) several populations, which was handled showing that empirical measures on each populations simultaneously satisfy a large deviation principle [51], (iv) spatial extension, which we addressed by considering a two layered empirical measure [52, 53], (v) spatial dependence of transportation delays as well as mean and variance of interactions [52] bringing some complex dependence between time and space, (vi) non-Gaussian interactions with sub-Gaussian tails for which we revisited previous work developed in discrete time [178], and (vii) bi-dependent interactions [53] that we achieved through the use of sophisticated tools from Probability theory. To these ends, we have proved large deviations principle, exhibited and studied the associated good rate functions, and characterized their minima that we have shown to be unique. The proof proceeds by using a combination of Sanov's theorem and to extend Varadhan's lemma to a functional that does not directly satisfies the canonical assumptions. The limit of the system is a complex non-Markovian process whose dynamics is relatively hard to understand in the general case. The model we have worked with, the Wilson and Cowan system, is very popular in physics and neuroscience and is close of the famous Hopfield model, but as mentioned, can be easily generalized to nonlinear neuron models. Indeed, most of the proofs deal with a quantity which is related to the density of the coupled network with respect to the uncoupled dynamics, and this quantity is independent of the dynamics of individual cells. Eventually, let us note that this result provides large-deviations estimates on the convergence of deterministically coupled networks as studied in [239].

For the neuroscience viewpoint, this thesis has the interest to justify an approach introduced in the seminal paper of Sompolinsky and collaborators [221] and widely used in the analysis of large-scale networks. Let us mention that this approach is also implicit in the study of neural fields. Important, while many of the previous contribution was done in discrete time, our results holds in continuous time setting, resolving a long-standing problem in neuroscience. Moreover, extending the scope of firing-rate models, we addressed biologically relevant synapses depending on both the presynaptic and postsynaptic neurons. Furthermore, the limit equation we obtained fully takes into account various microscopic features such as the noise intensity, the mean and variance of synaptic connections, and delays that was notably lacking some mean-field descriptions of neural networks. Under the choice of Gaussian initial conditions and linear intrinsic dynamics, he have been able - through bifurcation diagram - to uncover interesting qualitative effect that was not known in the literature. These notably include mean driven pitch-fork bifurcation leading to stable up-and-down states, as well as heterogeneity-induced oscillations that may account for the results of Aradi and colleagues showing that increased heterogeneity

was related with epilepsy.

Applying this formalism, we gained insights on the role of connectivity disorder for random oscillators in the Kuramoto model. Our approach therefore justified in a mathematically rigorous way some results from the physics literature [76, 77, 224, 225], only in the case of non-symmetric connectivity coefficients, but in a slightly more general setting. This has led us to uncover the dependence of the synchronization transition in the disorder level. A number of results in the physics literature are not covered by the present analysis. These include the case of non-noisy oscillators, or of interacting oscillations with correlated or anti-correlated reciprocal interactions (i.e., symmetric or asymmetric connectivity matrices respectively). The refined large deviations techniques developed by Ben Arous and Guionnet in [23, 129] may provide a way to address these results. We have evidenced numerically a number of transitions due to disorder in the noiseless Kuramoto model.

However, the limits obtained are valid only in the presence of noise, since Girsanov's theorem is used to relate the dynamics of the coupled system to the uncoupled system. The limit of randomly connected systems in the absence of noise is a complex issue with numerous applications, and very little work have been done on this topic. One outstanding contribution that addresses a similar question is the work of Ben Arous, Dembo and Guionnet for spherical spin glass [14]. In that work, the authors characterize the thermodynamic limit of this system and analyze its long term behavior, providing a mathematical approach for aging. This approach uses the rotational symmetry of the Hamiltonian allowing, by a change of orthogonal basis, to rely on results on the eigenvalues of the coupling matrix. A similar approach seems unlikely to readily extend to the setting of the present manuscript.

From the mathematical viewpoint, it shall not be hard to combine the methods of the present article to those in [52, 53] and the specific methods developed here to extend the present results to spatially-dependent interactions with space dependent delays and bi-dependent synapses. Moreover, we expect that the limit obtained is universal with respect to the distribution of the connectivity coefficient as soon as their tails have a sufficiently fast decay, as demonstrated for a discrete-time neuronal network in [179] and for coefficients with spatial dependence in [52]. In addition, the results shall hold in cases where the intrinsic dynamics is not Lipschitz-continuous as soon as sufficient non-explosion estimates are obtained on the solutions of the uncoupled system, as was the case in [15, 129]. We however mention that in this case, the original fixed-point method developed in the present article to prove existence and uniqueness of solutions to the mean-field equations are no more valid and adequate methods needs to be used as the ones presented in [15, 129]. Furthermore, it preforming the study of the fluctuations of the convergence of the empirical measure by establishing a Central Limit Theorem is to be done. This kind of question was for example addressed in a similar simpler setting [74], using methods originally introduced by Bolthausen, and that consists in showing strict positivity and finiteness of the curvature of the good rate function at its unique minimum. From the dynamical point of view, it would also be very interesting to to confront our limit equations obtained in [52] with the phenomenology of heterogeneous neural fields, relying on the same approach as in 6. In fact, for the same choice of Gaussian initial conditions and linear dynamics, solutions are still Gaussian so that their study might provide many

interesting insights.

Part IV

General Appendix

CHAPTER 8

GENERALITIES

8.1 A FEW CLASSICAL THEOREMS FROM PROBABILITY THEORY.

Theorem 8.1.1 (PortManteau). *The following properties are equivalent:*

(i) $P_N \xrightarrow{\mathscr{L}} P$,

(ii) *for every closed set* $F \subset \Sigma$, $\limsup_N P_N(F) \leq P(F)$,

(iii) *for every open set* $O \subset \Sigma$, $P(O) \leq \liminf_N P_N(O)$,

(iv) *for every* $A \subset \Sigma$ *with* $P(\partial A) = 0$, $\lim_N P_N(A) = P(A)$,

(v) *for every bounded Lipschitz-continuous function* $f : \Sigma \to \mathbb{R}$, $\int_\Sigma f(x)dP_n(x) \to \int_\Sigma f(x)dP(x)$

Theorem 8.1.2 (Girsanov). *Let* $\big(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P}\big)$ *be a filtered probability space, and let* $(L_t, t \geq 0)$ *a continuous martingale such that* $L_0 = 0$. *Let* $D := E(L)$, *and suppose* $E(D_\infty) = 1$. *Let* Q *be the probability measure on* $(\Omega, \mathscr{F})$ *defined by* $Q := D_\infty \cdot P$. *Then, for every* P*-continuous local martingale* M, *the process* $M - \langle M, L\rangle$ *is a* Q*-continuous local martingale.*

8.2 GAUSSIAN CALCULUS

Definition 8.2.1 (Gaussian variables). Let X be a random real number with $\mathbb{E}\big[X\big] = m$, $Var(X) = \sigma^2 \in \mathbb{R}$. It is Gaussian iff

$$\mathbb{E}\big[\exp\{tX\}\big] = \exp\big\{\frac{t^2\sigma^2}{2} + tm\big\} \forall t \in \mathbb{C}.$$

Proposition 8.2.1 (Gaussian tail). Let $X \sim \mathscr{N}(0,1)$. Then for any $x > 0$,

$$\frac{1}{\sqrt{2\pi}}\big(\frac{1}{x} - \frac{1}{x^3}\big)e^{-\frac{x^2}{2}} \leq \mathbb{P}(X > x) \leq \frac{1}{\sqrt{2\pi}}\frac{1}{x}e^{-\frac{x^2}{2}}.$$

Proof.

$$\begin{aligned}\mathbb{P}(X>x) &= \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \leq \frac{1}{x\sqrt{2\pi}} \int_x^\infty u e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{x^2}{2}}, \\ &= \int_x^\infty \frac{u}{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \overset{\text{IPP}}{=} \left[-\frac{1}{u\sqrt{2\pi}} e^{-\frac{u^2}{2}} \right]_x^\infty - \int_x^\infty \frac{1}{u^2\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \\ &\geq \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}} - \frac{1}{x^3\sqrt{2\pi}} \int_x^\infty u e^{-\frac{u^2}{2}} du.\end{aligned}$$

□

Proposition 8.2.2. Let $\zeta \sim \mathcal{N}(\mu, \sigma^2)$ with $\sigma < 1, \mu \in \mathbb{R}$. Then

$$\mathbb{E}\Big[\exp\Big\{\frac{1}{2}\zeta^2\Big\}\Big] = \frac{1}{\sqrt{1-\sigma^2}} \exp\Big\{\frac{\mu^2}{2(1-\sigma^2)}\Big\} = \exp\Big\{\frac{1}{2}\Big(\frac{\mu^2}{1-\sigma^2} - \log(1-\sigma^2)\Big)\Big\}. \tag{8.1}$$

Proposition 8.2.3. For any $t \in \mathbb{R}_+$,

$$\mathbb{E}\Big[e^{\{t|\mathcal{N}(\mu,\sigma^2)|\}}\Big] \leq e^{t|\mu| + \frac{t^2\sigma^2}{2}} \big(1 + O(t\sigma)\big)$$

Proof.

$$\begin{aligned}\mathbb{E}\Big[e^{|\mathcal{N}(0,\sigma^2)|}\Big] &= 2\int_0^{+\infty} \frac{e^{x - \frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx = 2e^{\frac{\sigma^2}{2}} \int_0^{+\infty} \frac{e^{-\frac{(x-\sigma^2)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx = e^{\frac{\sigma^2}{2}} \left(1 + \int_{-\sigma^2}^{\sigma^2} \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx\right) \\ &= e^{\frac{\sigma^2}{2}} \big(1 + O(\sigma)\big)\end{aligned}$$

□

Definition 8.2.2 (Gaussian vector)**.** A random vector $X = (X_1, \ldots, X_k)$ is called Gaussian if for every $a \in \mathbb{R}^k$, the random variable $\langle X, a\rangle$ is Gaussian.

Definition 8.2.3 (Gaussian field, function)**.** (i) A Gaussian process (field, function) indexed by a set T is a collection of random variables $(X_t)_{t\in T}$ such that for any $t_1, \ldots, t_k \in T$, the random vector $(X_{t_1}, \ldots, X_{t_k})$ is Gaussian.

(ii) If X has value in $\mathbb{R}^N$, it is degenerate in $t \in T$ if $Var(X(t)) \notin GL_N(\mathbb{R})$.

Definition 8.2.4 (C^k random processes)**.** Let $(X_t)_{t\in U}$ be a random process on $U \subset \mathbb{R}^N, N > 0$. It is a C^k random process if exists $\Omega_0 \subset \Omega$ of probability 1, and such that $\forall \omega \in \Omega_0$, $t \to X_t(\omega)$ is a C^k function.

Theorem 8.2.4. *(Isserlis) Let $n \in \mathbb{N}^*$ and $(X_1, \ldots, X_{2n})$ be a centered Gaussian vector. Then*

$$\mathbb{E}\Big[X_1 X_2 \ldots X_{2n}\Big] = \sum_{E\in\mathscr{P}_n} \prod_{i=1}^n \mathbb{E}\Big[\prod_{k\in E_i} X_k\Big]$$

where $\mathscr{P}_n$ denote the set of partitions of $[\![1, 2n]\!]$ made of n pairs:

$$\mathscr{P}_n := \Big\{E = \{E_1, \ldots, E_n\} \mid \bigcup_{i=1}^n E_i = [\![1,2n]\!],\ \forall 1 \leq i \neq j \leq n, E_i \cap E_j = \emptyset\ Card(E_i) = 2\Big\}.$$

8.3 Stochastic Fubini Theorem

This subsection is devoted to give, in the most compact way, sufficient conditions in order to apply a Stochastic Fubini Theorem. It is inspired of [194]. We will restrict our scope to stochastic processes on a finite time horizon $[0,T]$ with $T>0$, on a a filtered probability space $\big(\Omega,\mathscr{F},(\mathscr{F}_t)_{0\leq t\leq T},\mathbb{P}\big)$. The filtration $(\mathscr{F}_t)_{0\leq t\leq T}$ will be assumed continuous, and satisfy the usual conditions. $\mathbb{L}$ will denote the space of adapted continuous processes, and $\mathscr{P}$ the predictable σ-algebra on $\mathbb{R}_+\times\Omega$ defined as the smallest σ-algebra such that every element of $\mathbb{L}$ is measurable:

$$\mathscr{P}:=\sigma\big\{H:H\in\mathbb{L}\big\}.$$

$\mathbb{E}$ will denote the expectation over $\mathbb{P}$. Moreover, we will consider another probability space $\big(\hat{\Omega},\hat{\mathscr{F}},\gamma\big)$ that will account for the second argument on which to integrate (note that γ needs only to be a finite measure on $\big(\hat{\Omega},\hat{\mathscr{F}}\big)$)

Definition 8.3.1 (Stopping-times)**.** A random variable $\tau:\Omega\to[0,\infty]$ is a stopping time if the event $\{T\leq t\}\in\mathscr{F}_t$, for every $0\leq t\leq\infty$.

Definition 8.3.2 (Martingale)**.** A real-valued, adapted process $X=(X_t)_{0\leq t\leq\infty}$ is called a martingale (resp. supermartingale, submartingale) with respect to the filtration $\mathscr{F}$ if

(i) $\mathbb{E}\big[|X_t|\big]<\infty$.

(ii) if $s\leq t$, then $\mathbb{E}\big[X_t|\mathscr{F}_s\big]=X_s$ a.s. (resp.$\mathbb{E}\big[X_t|\mathscr{F}_s\big]\leq X_s$, reps. $\mathbb{E}\big[X_t|\mathscr{F}_s\big]\geq X_s$).

Note that martingales are only defined on $[0,\infty[$; that is, for finite t and not $t=\infty$. It is often possible to extend the definition to $t=\infty$.

Definition 8.3.3 (Local Martingale)**.** An adapted continuous process X is a local martingale if there exists a sequence of increasing stopping times, T_n, with $\lim_{n\to\infty}T_n=\infty$ a.s. such that $X_{t\wedge T_n}\mathbf{1}_{T_n>0}$ is a uniformly integrable martingale for each n. Such a sequence (T_n) of stopping times is called a fundamental sequence.

Definition 8.3.4 (FV processes)**.** Let $A=(A_t)_{t\geq0}$ be a continuous process. A is a continuous increasing process if the paths of $A:t\to A_t(\omega)$ are non-decreasing for almost all ω. A is called a continuous finite variation process (FV) if almost all of the paths of A are of finite variation on each compact interval of $\mathbb{R}_+$.

Definition 8.3.5 (semimartingale, special semimartingale)**.** A continuous process X is a (continuous) semimartingale if it can be written

$$X=N+A \tag{8.2}$$

with N a continuous local martingale, and A a continuous adapted FV process. Equation (8.2) is called the canonical decomposition of X. Moreover, if A is predictable, them X is a special semimartingale.

Definition 8.3.6 (Quadratic variation of a semimartingale)**.** Let X be a continuous semimartingale. Its quadratic variation is defined by

$$\langle X\rangle=X^2-2\int XdX.$$

Definition 8.3.7. Let X be a semimartingale with canonical decomposition $X = N + A$. The $\mathscr{H}^2$-norm of X is defined as:

$$\|X\|_{\mathscr{H}^2} := \mathbb{E}\left[\langle N\rangle_T\right]^{\frac{1}{2}} + \mathbb{E}\left[\left(\int_0^T |dA_s| dt\right)^2\right]^{\frac{1}{2}}.$$

Moreover, the space $\mathscr{H}^2$ consists of every special semimartingale with finite $\mathscr{H}^2$-norm. It is a Banach space.

Definition 8.3.8. Let a semimartingale $X \in \mathscr{H}^2$ with canonical decomposition $X = N + A$, and a predictable process $H \in \mathscr{P}$. We say that H is $\left(\mathscr{H}^2, X\right)$-integrable if and only if

$$\mathbb{E}\left[\int_0^T H_s^2 d\langle N\rangle_s\right] + \mathbb{E}\left[\left(\int_0^T |H_t||dA_s|\right)^2\right] < +\infty.$$

Moreover, for a general semimartingale (not necessarily in $\mathscr{H}^2$), we will say that $H \in L(X)$, or that H is X integrable if exists an increasing sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ such that:

(i) $T_n \nearrow +\infty$ $\mathbb{P}$-a.s.

(ii) $\forall n \geq 1$, $X^{T_n} \in \mathscr{H}^2$ and H is $\left(\mathscr{H}^2, X^{T_n}\right)$-integrable.

The following results can be found in [194, IV.63-65]. Let X be a semimartingale with $X_0 = 0$ $\mathbb{P}$-a.s., and

$$H := \begin{cases} \Omega \times [0,T] \times \hat{\Omega} & \to \mathbb{R} \\ (\omega, t, \hat{\omega}) & \to H(\omega, t, \hat{\omega}) \end{cases}$$

be a $\hat{\mathscr{F}} \otimes \mathscr{P}$ measurable map.

Theorem 8.3.1. *Suppose that, for every $\hat{\omega} \in \hat{\Omega}$, $H(\cdot,\cdot,\hat{\omega}) \in L(X)$. Then, there exists a $\hat{\mathscr{F}} \otimes \mathscr{B}\left([0,T]\right) \otimes \mathscr{F}$-measurable function*

$$Z := \begin{cases} \Omega \times [0,T] \times \hat{\Omega} & \to \mathbb{R} \\ (\omega, t, \hat{\omega}) & \to Z(\omega, t, \hat{\omega}) \end{cases}$$

such that, for every $\hat{\omega} \in \hat{\Omega}$, $\left(Z(\omega,t,\hat{\omega})\right)_{0\leq t\leq T}$ is a continuous version of $\left(\int_0^t H(\omega,s,\hat{\omega})ds\right)_{0\leq t\leq T}$.

Theorem 8.3.2. *Stochastic Fubini Theorem Assume that*

$$\left(\int_{\hat{\Omega}} H(\omega,t,\hat{\omega})^2 d\gamma(\hat{\omega})\right)^{\frac{1}{2}} \in L^1(X).$$

Letting $(\omega,t,\hat{\omega}) \to Z(\omega,t,\hat{\omega})$ be as in the previous theorem. Then

$$\left(Y_t(\omega) := \int_{\hat{\Omega}} Z(\omega,t,\hat{\omega}) d\gamma(\hat{\omega})\right)_{0\leq t\leq T}$$

exists and is a continuous version (that is an indistinguishable process) of

$$\left(\int_0^t \int_{\hat{\Omega}} H(\omega,t,\hat{\omega}) d\gamma(\hat{\omega}) dX_s(\omega)\right)_{0\leq t\leq T}.$$

CHAPTER 9

LARGE DEVIATIONS

For their applications to interacting particles system, large deviations play a key role in this manuscript. They will constitute our main tool to derive mean-field equations for the dynamics of randomly connected neural networks, as they can contain a Law of Large Number (LLN). One notable advantages of large deviations is that they address dependent sequences of random variables much more readily than the Central Limit Theorem do. Large deviations have been the object of numerous previous compilations from which we will often borrow. Among these, notably appears a short monograph by Franck Den Hollander [85] that very clearly gives the quintessence of what large deviations are, the very complete book of A. Dembo and O. Zeitouni [84] that addresses all the most difficult and abstract results on the topic, and also the elegant book of Deuschel and Stroock [86].

Large deviations focus on the study of rare events. They furnish a set of useful tools to analyze the probability of these events, generally by establishing a large deviation principles (LDP). The framework for LDPs implies that we work on a measurable space $(\Sigma, \mathscr{B})$, where Σ is assumed to be topological, so that the open and closed sets that generates the Borel σ-field $\mathscr{B}$ are well-defined. For convenience, we will here restrict this short introduction to Polish spaces only, and refer to [84, 86] for broader results. The LDP characterizes the limiting behavior of a family of probability measures $(P_N)_{N\in\mathbb{N}}$ defined on $(\Sigma, \mathscr{B})$. It provides exponential asymptotic estimates for the probabilities of Borel sets $P_N(\Gamma)$, $\Gamma \in \mathscr{B}$ by bounding these quantities from above and below using a rate function. We will use the following notation s: if $\Gamma \subset \Sigma$, $\bar{\Gamma}$, $\mathring{\Gamma}$ and Γ^c will respectively denote its closure, interior and complement.

9.1 PRELIMINARIES

Let $(X_i)_{i\in\mathbb{N}^*}$ be a sequence of i.i.d real integrable random variables with $\mu := \mathbb{E}\big[X_1\big]$. Large deviations theory is typically interested in the asymptotic behavior of events of the type $\big\{S_N \geq aN\big\}$, where $S_N := \sum_{i=1}^N X_i$, and $a \in \mathbb{R}$. When $a > \mu$, this event can be understood as an excursion - or deviation - of the random walk S_N from its mean behavior μN. While the intuitive rarity of such excursions is a simple consequence of the Strong Law of Large Numbers (SLLN):

$$\mathbb{P}\big(S_N \geq aN\big) = \mathbb{P}\Big(\frac{S_N - \mu N}{N} \geq (a-\mu)\Big) \to_N 0, \tag{9.1}$$

large deviations theory goes further by finely estimating the speed of this convergence. We insist on the fact that the size of the deviation - N - is of importance here. Indeed, provided that $\sigma^2 := Var(X_1) < +\infty$, the Central Limit Theorem ensures that

deviations of order $\sqrt{N}$ are typical events:

$$\mathbb{P}\Big(\frac{S_N-\mu N}{\sqrt{N}} \geq (a-\mu)\Big) \to_N \mathbb{P}\big(\mathscr{N}\big(0,\sigma^2\big) \geq a-\mu\big) \neq 0.$$

This simple fact justifies the name of the theory: deviations of order N are called "large" as opposed to normal ones, of order $\sqrt{N}$, usually referred to as fluctuations. Let us build on CLT to estimate $\mathbb{P}\big(S_N \geq aN\big)$. For large N, a non-rigorous calculus gives

$$\begin{aligned}\mathbb{P}\big(S_N \geq aN\big) &= \mathbb{P}\Big(\frac{S_N-\mu N}{\sqrt{N}} \geq \sqrt{N}(a-\mu)\Big) \\ &\simeq \mathbb{P}\Big(\mathscr{N}\big(0,\sigma^2\big) \geq \sqrt{N}(a-\mu)\Big) \sim_N \frac{\sigma}{(a-\mu)\sqrt{N}} e^{-\frac{N(a-\mu)^2}{2\sigma^2}}, \end{aligned} \tag{9.2}$$

where the equivalence arises from Proposition (8.2.1). This suggests an exponential speed of convergence for (9.1), with rate $H(a) := \frac{(a-\mu)^2}{2\sigma^2}$ depending on a. While this heuristic is erroneous - the rate $H(a)$ will not be given by $\frac{(a-\mu)^2}{2\sigma^2}$ in general - it will appear in the theory that this exponential decreasing is generally valid as we will find:

$$\mathbb{P}\big(S_N \geq aN\big) \sim c_N(a) e^{-NH(a)} \tag{9.3}$$

for some "rate function" $H : \mathbb{R} \to [0,+\infty]$, and with $\frac{1}{N}\log(c_N(a)) \to_N 0$. [1]

To give us a more solid insight of this fact, let us assume that X_1 admits an exponential moment:

$$\exists t \in \mathbb{R}^*, \qquad \mathbb{E}\big[e^{tX_1}\big] < \infty,$$

and define the logarithmic moment generating function of X_1,

$$\forall t \in \mathbb{R}, \quad \phi(t) := \log \mathbb{E}\big[e^{tX_1}\big] \overset{\text{Jensen}}{\in}]-\infty,+\infty].$$

Then, an exponential Chebyshev's inequality ensures that, for any $t \in D_\phi := \big\{t \in \mathbb{R}, \phi(t) < \infty\big\}$,

$$\mathbb{P}\big(S_N \geq aN\big) \leq e^{-aNt}\mathbb{E}\big[e^{tX_1}\big]^N = e^{-N(at-\phi(t))},$$

so that

$$\mathbb{P}\big(S_N \geq aN\big) \leq \exp\Big(-N \sup_{t\in D_\phi}\big\{at-\phi(t)\big\}\Big).$$

From this inequality it appears that the rate function H depends on $\mathscr{L}(X_1)$, and will thus not be given by $H(a) = \frac{(a-\mu)^2}{2\sigma^2}$ in general [2]. For the case of i.i.d. random variables, H will turn out to precisely be given by $H(a) := \sup_{t\in D_\phi}\big\{at-\phi(t)\big\}$, also known as the Legendre transform of ϕ. This result is the consequence of the paradigmatic Cramér's Theorem (9.3.1), which will constitute our first step into large deviations theory. It will extend to much more abstract settings. Before entering the theory, we will first give a few useful preliminary results and definitions, among which particularly appears the Legendre transform that will ubiquitous in the theory.

[1] In their introduction [86] Deuschel and Stroock precisely define Large Deviations Theory as focusing on the study of "rare events" with exponential decay as in (9.3).

[2] Take for example $X_1 \overset{\mathscr{L}}{=} \log\big(\xi^{2p}\big)$, with $p \in \mathbb{R}_+^*$, and $\xi \sim \mathscr{N}(0,1)$. In this case, $\phi(t) = \log \mathbb{E}\big[\xi^{2pt}\big]$ so that $D_\phi = \mathbb{R}$, and for p small enough $\sup_{t\in\mathbb{R}}\big\{at-\phi(t)\big\} > \frac{(a-\mu)^2}{2\sigma^2}$ by the Dominated Convergence Theorem.

9.2 Preliminary results, and Legendre transform

This section is devoted to introduce some essential notions of probability, continuity and geometry, as well as proving a few results on which we will often rely. We will insist on the chief properties of the Legendre transform, and prove a version of the duality Fenchel-Legendre. Let us begin with a weak notion of continuity:

Definition 9.2.1. Let Σ be a topological space, and $f : \Sigma \to \bar{\mathbb{R}}$. Then f is lower semi-continuous on Σ iff one of the following conditions is verified:

(i) $\forall x \in \Sigma$, $\forall \varepsilon \geq 0, \exists U \in \mathscr{V}(x), \forall y \in U, f(y) \geq f(x) - \varepsilon$,

(ii) $\forall x \in \Sigma$, $\liminf_{y \to x} f(y) \geq f(x)$,

(iii) $\forall M \in \mathbb{R}$, $\{x \in \Sigma, f(x) \leq M\}$ is closed.

Moreover, f is upper semi-continuous iff $-f$ is lower semi-continuous.

It is very easy to see why the three conditions are equivalent. While 1. $\implies$ 2. $\implies$ 3. trivially holds, 3. $\implies$ 1. amounts at choosing $x \in \Sigma$ and $\varepsilon > 0$, setting $M = f(x) - \varepsilon$, and remarking that $x \in \{f \leq M\}^c$ which is open. With words, this definition basically means that the graph of a lower (upper resp.) semi-continuous function must satisfy this simple requirement: at a discontinuity the function remains below the points nearby (above resp.). Of course, this picture has its limits, as the notion exists for very abstract Σ. A consequence of having closed level sets is that lower semi-continuous function are measurable with respect to the Borel σ-field. In fact, the family $\big(]a, +\infty]\big)_{a \in \mathbb{R}}$ forms a π-system that generates the Borel σ-field on $\bar{\mathbb{R}}$, and

$$f^{-1}(]a, +\infty]) = f^{-1}(\bar{\mathbb{R}} \backslash [-\infty, a]) = f^{-1}(\bar{\mathbb{R}}) \backslash f^{-1}([-\infty, a]) = f^{-1}([-\infty, a])^c,$$

which is an open set of Σ.

Along with convexity, lower semi-continuity is stable by taking the supremum over an uncountable family:

Proposition 9.2.1. (i) If $\big(g_\alpha\big)_{\alpha \in I}$ is a family of convex functions from $\mathbb{R}$ to $\bar{\mathbb{R}}$, then $t \to \sup_\alpha g_\alpha(t)$ is convex from $\mathbb{R}$ to $\bar{\mathbb{R}}$.

(ii) If $\big(g_\alpha\big)_{\alpha \in I}$ is a family of lower semi-continuous function from $\mathbb{R}$ to $\bar{\mathbb{R}}$, then $t \to \sup_\alpha g_\alpha(t)$ is lower semi-continuous from $\mathbb{R}$ to $\bar{\mathbb{R}}$.

Proof. 1. For any $\alpha \in I$, $\theta \in [0,1]$, and $(x,y) \in \mathbb{R}^2$, one has $g_\alpha(\theta x + (1-\theta)y) \leq \theta g_\alpha(x) + (1-\theta) g_\alpha(y)$, so that

$$\sup_\alpha \big\{g_\alpha(\theta x + (1-\theta)y)\big\} \leq \sup_\alpha \big\{\theta g_\alpha(x) + (1-\theta) g_\alpha(y)\big\} \leq \theta \sup_\alpha \big\{g_\alpha(x)\big\} + (1-\theta) \sup_\alpha \big\{g_\alpha(y)\big\}.$$

2. Let $a \in \bar{\mathbb{R}}$, and remark that

$$\Big(\sup_{\alpha \in I} g_\alpha\Big)^{-1}([-\infty, a]) = \bigcap_{\alpha \in I} g_\alpha^{-1}([-\infty, a])$$

is a closed set. □

An interesting remark to link these two notions, is that a convex function f could be modified into a lower semi-continuous map only by changing its values at the edge of its domain of definition $D_f := \{x \in \Sigma,\ f(x) < \infty\}$. We now introduce the Legendre transform for the simple case of real functions.

Definition 9.2.2 (Legendre transform)**.** Let $f : \mathbb{R} \to \bar{\mathbb{R}}$. When it exists, its Legendre transform is the function $f^* : \mathbb{R} \to \bar{\mathbb{R}}$ defined by

$$f^*(z) := \sup_{t \in \mathbb{R}} \{zt - f(t)\}.$$

Applying Proposition (9.2.1) on the family of convex and continuous functions $z \to zt - f(t), t \in \mathbb{R}$, one immediately sees that the Legendre transform $f^* : \mathbb{R} \to \bar{\mathbb{R}}$ is convex and lower semi-continuous. To give us an idea of how the graph of f^* is connected with the one of f, let us prove this elementary proposition containing a first easy version of the Fenchel-Legendre duality.

Proposition 9.2.2. Let $f : \mathbb{R} \to \bar{\mathbb{R}}$ be strictly convex and $\mathscr{C}^2$ on the interior of $D_f := \{t \in \mathbb{R},\ f(t) < \infty\}$. Then

(i) If f satisfies the steepness condition $\lim_{t \to \partial D_f} |f'(t)| = +\infty$, then f^* is strictly convex and $\mathscr{C}^2$ on $\mathbb{R}$.

(ii) If $D_f = \mathbb{R}$, then f^* is strictly convex and $\mathscr{C}^2$ on $\overset{\circ}{D}_{f^*}$ and satisfies the steepness condition:
$$\lim_{t \to \partial D_{f^*}} |f^{*\prime}(t)| = +\infty.$$

(iii) In both cases $f^{**} = f$.

Proof. 1-2. In both cases, the strict monotony of f' on $\overset{\circ}{D}_f$ implies that it is in bijection from $\overset{\circ}{D}_f$ to $f'(\overset{\circ}{D}_f)$. Let $\psi := f'^{-1} : f'(\overset{\circ}{D}_f) \to \overset{\circ}{D}_f$. Then, as f is strictly convex and $\mathscr{C}^2$ on $\overset{\circ}{D}_f$, $f^{(2)}$ is strictly positive, and ψ is a $\mathscr{C}^1$ strictly increasing function. Moreover, by taking the derivative of $t \to zt - f(t)$, we find that for any $z \in f'(\overset{\circ}{D}_f)$, $f^*(z) = z\psi(z) - f(\psi(z)) < \infty$. In particular f^* is $\mathscr{C}^1$ on $f'(\overset{\circ}{D}_f) \subset \overset{\circ}{D}_{f^*}$, with derivative $f^{*\prime}(z) = z\psi'(z) + \psi(z) - f'(\psi(z))\psi'(z) = \psi(z)$, so that f^* is also strictly convex and $\mathscr{C}^2$ on $f'(\overset{\circ}{D}_f)$ with $f'^{-1} = f^{*\prime}$. Point 1. is proved by remarking that the steepness condition implies $f'(\overset{\circ}{D}_f) = \mathbb{R}$. For the second point, recall that $D_f = \mathbb{R}$ and suppose that $f'(\overset{\circ}{D}_f) \neq \mathbb{R}$, as there is nothing to prove otherwise. Then, remark that $\forall z \notin f'(\mathbb{R})$, $zt - f(t)$ becomes arbitrarily large for t going either to $-$ or $+\infty$, so that $\overset{\circ}{D}_{f^*} = f'(\overset{\circ}{D}_f)$.

3. Remark that if f satisfies conditions 1. (resp 2.), then f^* satisfies condition 2. (resp 1.), so that we can properly define f^{**} which is strictly convex and $\mathscr{C}^2$ on $\overset{\circ}{D}_{f^{**}}$. Moreover, if f is steep, then $D_{f^*} = \mathbb{R}$ and $\overset{\circ}{D}_{f^{**}} = f^{*\prime}(\mathbb{R}) = \psi(\mathbb{R}) = f'^{-1}(\mathbb{R}) = \overset{\circ}{D}_f$, whereas if $D_f = \mathbb{R}$, then f^* is steep and $D_{f^{**}} = \mathbb{R}$. Hence, in both cases $\forall t \in \overset{\circ}{D}_{f^{**}} = \overset{\circ}{D}_f$:

$$f^{**}(t) = tf'(t) - f^*(f'(t)) = tf'(t) - \Big(f'(t)\underbrace{\psi\big(f'(t)\big)}_{=t} - f\big(\psi(f'(t))\big)\Big) = f(t).$$

□

As the scope of large deviations will extend well beyond the case of real random variables, we mention without proof the much more abstract Fenchel-Legendre duality result [86, Theorem 2.2.15]:

Theorem 9.2.3 (Fenchel-Legendre Duality)**.** *Let X be a Hausdorff real topological vector space, let X^* be its dual, and let $f : X \to]-\infty,\infty]$ be convex and lower semi-continuous. If f is not identically equal to $+\infty$, then the function*

$$g := \left\{ \begin{array}{ll} X^* & \to [-\infty,+\infty] \\ x & \to \sup_{t\in X} \left\{ {}_{X^*}\langle x,t\rangle_X - f(t) \right\} \end{array} \right.$$

is never $-\infty$ and satisfies $f(t) = \sup_{x\in X^*} \left\{ {}_{X^*}\langle x,t\rangle_X - g(x) \right\}$, where ${}_{X^*}\langle\cdot,\cdot\rangle_X$ denote the duality bracket.

We now introduce an important notion from probability theory:

Definition 9.2.3. Let Σ be a Polish space, and let $\mathscr{B}$ be its Boral σ-field. A family of probability measures $(P_N)_{N\in\mathbb{N}}$ on $(\Sigma,\mathscr{B})$ is said to be tight (resp. exponentially tight) if for every $0 < M < +\infty$ there exists a compact set $K_M \subset \Sigma$ such that:

$$\limsup_{N\to+\infty} P_N\big(K_M^c\big) \leq \frac{1}{M},$$

respectively

$$\limsup_{N\to+\infty} \frac{1}{N} \log P_N\big(K_M^c\big) < -M.$$

Remark 15. On a Polish space, Ulam's Theorem allows one to change $\limsup_N$ into $\forall N \in \mathbb{N}$.

For a family of measures, to be tight (resp. exponentially tight) means that the probability mass that is not on compact sets decrease fast (resp. exponentially fast) when N goes to infinity. Hence, the most of the mass remains trapped on compact sets, and can't escape toward infinity.

The following lemma will also be very useful for the theory.

Lemma 9.2.4. Let $k \in \mathbb{N}^*$ sequences of strictly positive real numbers $\big(a_N^{(1)}\big)_{N\in\mathbb{N}}$, $\big(a_N^{(2)}\big)_{N\in\mathbb{N}}$, …, $\big(a_N^{(k)}\big)_{N\in\mathbb{N}}$. Then

$$\lim_N \frac{1}{N} \log\Big(\sum_{l=1}^k a_N^{(l)}\Big) - \frac{1}{N} \max_{1\leq l\leq k} \log\big(a_N^{(l)}\big) = 0.$$

Proof. It suffices to remark that $\max_{1\leq l\leq k} a_N^{(l)} \leq \sum_{l=1}^k a_N^{(l)} \leq k \max_{1\leq l\leq k} a_N^{(l)}$. □

Remark 16. Remark in particular that

$$\limsup_N \frac{1}{N} \log\Big(\sum_{l=1}^k a_N^{(l)}\Big) = \max_{1\leq l\leq k} \left\{ \limsup_N \frac{1}{N} \log\big(a_N^{(l)}\big) \right\},$$

and

$$\liminf_N \frac{1}{N} \log\Big(\sum_{l=1}^k a_N^{(l)}\Big) = \max_{1\leq l\leq k} \left\{ \liminf_N \frac{1}{N} \log\big(a_N^{(l)}\big) \right\}.$$

These inequalities will be particularly useful as large deviations theory makes a great use of $\liminf$ and $\limsup$.

9.3 CRAMÉR'S THEOREM

Let us now state and prove the Cramér's Theorem. While dealing with very simple settings - that of i.i.d. sequences of real random variables - this theorem condenses all the essential ingredients of Large Deviations, making it particularly appropriate to gain some insights on the theory.

Theorem 9.3.1 (Cramér's Theorem)**.** *Let $(X_i)_{i\in\mathbb{N}}$ be an i.i.d. sequence of real random variables such that*

$$\forall t \in \mathbb{R}, \qquad \phi(t) := \log\Big(\mathbb{E}\Big[\exp\{tX_1\}\Big]\Big) < +\infty,$$

and let $\mu := \mathbb{E}[X_1]$. Then

(i) *for every $a > \mu$, $\lim_{N\to\infty} \frac{1}{N}\log\mathbb{P}(S_N \geq aN) = -H(a)$,*

(ii) *for every $a < \mu$, $\lim_{N\to\infty} \frac{1}{N}\log\mathbb{P}(S_N \leq aN) = -H(a)$,*

where $H(z) := \sup_{t\in\mathbb{R}}\{zt - \phi(t)\}$ is the Legendre transform of ϕ.

Remark 17. Observe that the case $a = \mu$ is absent of the theorem, as it does not constitute an analysis of the large deviations of S_N. Indeed, as X_1 admits exponential moments, we can rely on the CLT to obtain:

$$\mathbb{P}(S_N \geq \mu N) = \mathbb{P}\Big(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(X_i - \mu) \geq 0\Big) \to_{N\to\infty} \frac{1}{2}.$$

In particular

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{P}(S_N \geq \mu N) = 0, \tag{9.4}$$

suggesting $H(\mu) = 0$. Moreover, following the idea of Proposition (9.2.2), the condition $D_\phi := \{t \in \mathbb{R}, \phi(t) < \infty\} = \mathbb{R}$ can been weakened in $0 \in \overset{\circ}{D}_\phi$ and $\lim_{t\to\partial D_\phi} |\phi'(t)| = \infty$, without having to change the proof too dramatically.

Let us, under the weaker assumptions of this remark, establish a few properties on the logarithmic moment generating function ϕ and its Legendre transform H before turning to the proof of the theorem.

Proposition 9.3.2. Let $D_\phi := \{t \in \mathbb{R},\ \phi(t) < \infty\} \subset \mathbb{R}$, $D_H := \{z \in \mathbb{R},\ H(z) < \infty\}$ and suppose that $0 \in \overset{\circ}{D}_\phi$. Then

(i) ϕ is convex on $\mathbb{R}$, $\mathscr{C}^\infty$ on $\overset{\circ}{D}_\phi$, and satisfies $\phi(0) = 0$, $\phi'(0) = \mu$, $\phi^{(2)}(0) = Var(X_1)$. Moreover, ϕ is strictly convex on $\overset{\circ}{D}_\phi$ iff $Var(X_1) \neq 0$.

(ii) $H : \mathbb{R} \to [0,\infty]$ is lower semi-continuous and convex on $\mathbb{R}$ with $H(\mu) = 0$,

(iii) H has compact level sets.

Moreover, if $Var(X_1) \neq 0$ and $\lim_{t\to\partial D_\phi} |\phi'(t)| = \infty$, then

(i) $D_H = \mathbb{R}$ and H is a smooth strictly convex function. In particular, it is strictly decreasing on $]-\infty,\mu]$, and strictly increasing on $[\mu,+\infty[$,

(ii) $H^{(2)}(\mu) = \frac{1}{Var(X_1)}$.

Proof. 1. Let $t_- < 0 < t_+ \in \overset{\circ}{D}_\phi$. As X_1 admits positive and negative exponential moments, Cauchy-Schwarz's inequality ensures that it admits moment of every order $k \in \mathbb{N}^*$

$$\begin{aligned}\mathbb{E}\left[X_1^k\right] &= \mathbb{E}\left[X_1^k \mathbb{1}_{\{X_1 \geq 0\}}\right] + \mathbb{E}\left[X_1^k \mathbb{1}_{\{X_1 < 0\}}\right] \\ &\overset{\text{C.S.}}{\leq} \mathbb{E}\left[\mathbb{1}_{\{X_1 \geq 0\}} X_1^{2k} e^{-t_+ X_1}\right]^{\frac{1}{2}} \mathbb{E}\left[e^{t_+ X_1}\right]^{\frac{1}{2}} + \mathbb{E}\left[\mathbb{1}_{\{X_1 < 0\}} X_1^{2k} e^{-t_- X_1}\right]^{\frac{1}{2}} \mathbb{E}\left[e^{t_- X_1}\right]^{\frac{1}{2}} < \infty\end{aligned}$$

Moreover, by Hölder inequality, the maps $(t,\omega) \to X_1^k(\omega)\exp\left\{tX_1(\omega)\right\} = \partial_t^k\left(\exp\left\{tX_1(\omega)\right\}\right)$ are well-defined, a.s. continuous in $t \in \overset{\circ}{D}_\phi$ and integrable in ω for all integer $k \in \mathbb{N}^*$. Thus, by classical theorems $\phi(t)$ is $\mathscr{C}^\infty$ on $\overset{\circ}{D}_\phi$. In particular,

$$\phi(t) = \log\left(\mathbb{E}\left[\exp\left\{tX_1\right\}\right]\right), \qquad \phi'(t) = \frac{\mathbb{E}\left[X_1 \exp\left\{tX_1\right\}\right]}{\mathbb{E}\left[\exp\left\{tX_1\right\}\right]}$$

$$\phi^{(2)}(t) = \frac{\mathbb{E}\left[X_1^2 \exp\left\{tX_1\right\}\right]\mathbb{E}\left[\exp\left\{tX_1\right\}\right] - \mathbb{E}\left[X_1 \exp\left\{tX_1\right\}\right]^2}{\mathbb{E}\left[\exp\left\{tX_1\right\}\right]^2} \overset{C.S.}{\geq} 0,$$

with equality iff $Var(X_1) = 0$.

2. H being the Legendre transform of ϕ, it is convex and lower semi-continuous by Proposition (9.2.2). Moreover, as $0 \in \overset{\circ}{D}_\phi$, $H(z) \geq z \times 0 - \phi(0) = 0$, whereas Jensen inequality ensures that $H(\mu) = \sup_{t \in D_\phi}\left\{-\log \mathbb{E}\left[e^{t(X_1-\mu)}\right]\right\} \leq \sup_{t \in D_\phi}\left\{-t\mathbb{E}\left[X_1 - \mu\right]\right\} = 0$, yielding $H(\mu) = 0$.

3. Let us show that $K_a := H^{-1}\left([0,a]\right)$ is compact for any $a \in \mathbb{R}^+$. As H is lower semi-continuous, K_a is a closed set. As we work in finite dimension, it is sufficient to prove it is bounded. Suppose that exists $(x_p)_{p\in\mathbb{N}}$ such that $x_p t - \phi(t) \leq a, \forall t \in D_\phi$ and $x_p \to \infty$. Let then $t_p := \frac{a+1}{x_p} \in \overset{\circ}{D}_\phi$ for p sufficiently large. Then $a + 1 - \phi(t_p) = x_p t_p - \phi(t_p) \leq a$, so that $\phi(t_p) \geq 1$ which contradicts the fact that $\phi(0) = 0$ as ϕ is continuous in 0 and $t_p \to_p 0$.

4-5. Suppose now that X_1 is not degenerate, and that ϕ is steep on $\overset{\circ}{D}_\phi$. Then ϕ is $\mathscr{C}^\infty$ and strictly convex so that Proposition (9.2.2) ensures that H is $\mathscr{C}^2$ and strictly convex on $D_H = \mathbb{R}$. Moreover, $H' = \phi'^{-1}$, so that H is also $\mathscr{C}^\infty$. As H is strictly convex and positive, $H^{-1}\left(\{0\}\right) = \{\mu\}$. Furthermore, $\forall z \in \mathbb{R}$

$$H^{(2)}(z) = \frac{1}{\phi^{(2)}\left(\phi'^{-1}(z)\right)}, \qquad \text{so that} \qquad H^{(2)}(\mu) = \frac{1}{Var(X_1)}$$

as $\phi'(0) = \mu$, and $\phi^{(2)}(0) = Var(X_1)$ according to 1. (Remark that, if $Var(X_1) = 0$, $D_\phi = \mathbb{R}$ and $H(z) = \sup_{t\in\mathbb{R}}\left\{t(z-\mu)\right\} = \infty \times \mathbb{1}_{\{z \neq \mu\}}$, so that $\overset{\circ}{D}_H = \emptyset$) □

Remark 18. From this Proposition and Proposition (9.2.2), we obtain by Fenchel-Duality that $\phi(t) := \sup_{z\in\mathbb{R}}\left\{zt - H(z)\right\}$. Hence, H generates the logarithmic generating function of X_1, and thus encodes its law. This highlights the fact, at least for the simple scope of Cramér's theorem, the rate function is in a one to one correspondence with $\mathscr{L}(X_1)$. Though this will not generalize to the whole theory, H will in general be

closely related to sequence of probability measures whom it controls the deviations. It will also turn out that, among the good properties displayed by rate function H in Proposition (9.3.2), some will be needed to properly construct the theory (see (9.4.1)).

We now prove the theorem.

Proof. Let $Y_i = X_i - a$. Then $\phi_Y(t) = \phi_X(t) - at$, so that $H_Y(0) = \sup_{t\in\mathbb{R}} \{at - \phi_X(t)\} = H_X(a)$. Hence, it is sufficient to prove the result for the case $a = 0$. Moreover, as $\phi_{-X}(t) = \phi_X(-t)$, points 1. and 2. are equivalent. Remark also that the result trivially holds when $Var(X_1) = 0$.

Let us then suppose that $\mu < 0$, $Var(X_1) \neq 0$, and define $\rho := H(0) = -\inf_{t\in\mathbb{R}} \phi(t)$. To make ρ more explicit, we are interested in the shape of the function ϕ. Proposition 9.3.2 ensures that it is a strictly convex smooth function on $\mathbb{R}$, whereas Jensen inequality gives $\phi(t) \geq t\mu \to_{t\to-\infty} +\infty$ as $\mu < 0$. Hence, by strict convexity, there are only two possible shapes for ϕ: L or U. In the first case, the infimum is attained at $+\infty$, whereas in the second case it is attained at a unique location. Remark that

$$\phi(t) = \log\left\{\mathbb{E}\left[\mathbb{1}_{\{X_1<0\}}e^{tX_1}\right] + \mathbb{E}\left[\mathbb{1}_{\{X_1>0\}}e^{tX_1}\right] + \mathbb{P}\left(X_1 = 0\right)\right\}$$

Hence, the following holds:

(i) If $\mathbb{P}\left(X_1 \leq 0\right) = 1$, then ϕ is strictly decreasing and the MCT ensures that $\rho = -\lim_{t\to+\infty}\phi(t) = -\log\left(\mathbb{P}\left(X_1 = 0\right)\right) \in [0, +\infty[$.

(ii) If $\mathbb{P}\left(X_1 > 0\right) > 0$, then $\lim_{t\to\pm\infty}\phi(t) = +\infty$ and there exists a unique $\tau \in \mathbb{R}$ such that $\rho = -\phi(\tau)$, and $\phi'(\tau) = 0$.

For the first case, $\mathbb{P}\left(S_N \geq 0\right) = \mathbb{P}\left(X_1 = \ldots = X_N = 0\right) = \mathbb{P}\left(X_1 = 0\right)^N$, so that

$$\frac{1}{N}\log\left(\mathbb{P}\left(S_N \geq 0\right)\right) = -\rho = -H(0),$$

and the result holds.

The second case is more tricky and necessitate to introduce a particular "tilted" transformation Y_i of the X_i. This new i.i.d. sequence will have the good property of being centered. The idea is that, for these variables, we have shifted the mean to zero without shifting the point $a = 0$ at which we are making the analysis, so that according to 9.4, we will be able to build on the CLT. Let then the "tilted" repartition function $F_Y(y) := e^{\rho}\int_{-\infty}^{y} e^{\tau x} dF_X(x)$, so that

$$\phi_Y(t) := \log\Big(\int_{\mathbb{R}} e^{ty} dF_Y(y)\Big) + \rho = \log\mathbb{E}\left[e^{(t+\tau)X}\right] + \rho = \phi(t+\tau) - \phi(\tau),$$

is $\mathscr{C}^\infty$ and strictly convex on $\mathbb{R}$ and attain its minimal value 0 at $t = 0$. Moreover, by Proposition 9.3.2, Y_1 is centered as $\mathbb{E}\left[Y_1\right] = \phi_Y'(0) = \phi'(\tau) = 0$, so that $H_Y(0) = 0$. Let then, $\left(Y_i\right)_{i\in\mathbb{N}}$ be a i.i.d sequence of random variables with repartition function F_Y, and define $\hat{S}_N := \sum_{i=1}^N Y_i$. Then

$$\mathbb{P}\left(S_N \geq 0\right) = \mathbb{E}\left[e^{-\rho N - \tau S_N}\Big(\prod_{i=1}^N e^{\rho}e^{\tau X_i}\Big)\mathbb{1}_{S_N\geq 0}\right] = \exp\left\{-\rho N\right\}\mathbb{E}\left[e^{-\tau\hat{S}_N}\mathbb{1}_{\hat{S}_N\geq 0}\right],$$

so that

$$\frac{1}{N}\log\mathbb{P}\Big(S_N\geq 0\Big)=-\rho+\frac{1}{N}\log\mathbb{E}\Big[e^{-\tau\hat{S}_N}\mathbb{1}_{\hat{S}_N\geq 0}\Big].$$

As the Y_i are independent and centered r.v., the event $\hat{S}_N\geq 0$ is typical whereas $S_N\geq 0>\mu$ is not. Let then $C>0$ be a fixed number, and $\hat{\sigma}^2$ be the variance of Y_1. The CLT ensures that

$$1\geq\mathbb{E}\Big[e^{-\tau\hat{S}_N}\mathbb{1}_{\hat{S}_N\geq 0}\Big]=\mathbb{E}\Big[e^{-\tau\hat{\sigma}\sqrt{N}\frac{\hat{S}_N}{\hat{\sigma}\sqrt{N}}}\mathbb{1}_{\left\{\frac{\hat{S}_N}{\hat{\sigma}\sqrt{N}}\geq 0\right\}}\Big]\geq e^{-\tau\hat{\sigma}\sqrt{N}C}\mathbb{P}\Big(\frac{\hat{S}_N}{\hat{\sigma}\sqrt{N}}\in[0,C]\Big),$$

so that we obtain the desired result. □

Remark 19. As $\inf_{z\in[a,\infty[}H(z)=0$ as soon as $a\leq\mu$ and as H is strictly increasing on $[\mu,+\infty[$, 1. of (9.3.1) is equivalent to:

$$\forall a\in\mathbb{R},\qquad\lim_N\frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in[a,+\infty[\Big)=-\inf_{z\in[a,+\infty[}H(z).\tag{9.5}$$

Remark that this limit equality also holds for interval of the form $[a,b[$ with $\mu<a<b$. Indeed

$$\mathbb{P}\Big(\frac{S_N}{N}\in[a,b[\Big)=\mathbb{P}\Big(\frac{S_N}{N}\in[a,+\infty[\Big)-\mathbb{P}\Big(\frac{S_N}{N}\in[b,+\infty[\Big),$$

so that

$$\frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in[a,b[\Big)=\frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in[a,+\infty[\Big)+\frac{1}{N}\log\Big(1-\frac{\mathbb{P}\big(S_N\geq bN\big)}{\mathbb{P}\big(S_N\geq aN\big)}\Big).$$

As $\mathbb{P}\big(S_N\geq bN\big)/\mathbb{P}\big(S_N\geq aN\big)=e^{-N(H(b)-H(a)+o(1))}$,

$$\lim_N\frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in[a,b[\Big)=-H(a)=-\inf_{z\in[a,b[}H(z).$$

This highlights an interesting fact: for any $\varepsilon>0$, $a>\mu$,

$$\begin{aligned}\lim_N\frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in[a,+\infty[\Big)&=\lim_N\frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in[a,a+\varepsilon[\Big)=-H(a)\\&>\lim_N\frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in[a+\varepsilon,+\infty[\Big)=-H(a+\varepsilon).\end{aligned}$$

This can be understood as follow: the larger a deviation is, the less likely it becomes. Hence, under the condition $\frac{S_N}{N}\geq a$, the most probable scenario is that $\frac{S_N}{N}$ has deviated no further than a.

Here is a first insights of an interesting property of large deviations: they contain a LLN.

Lemma 9.3.3. Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of r.v. such that exists $\mu\in\mathbb{R}$, a function $H:\mathbb{R}\to[0,+\infty]$ for which $H^{-1}(\{0\})=\{\mu\}$ and

(i) for every $a>\mu$, $\lim_N\frac{1}{N}\log\mathbb{P}\big(S_N\geq aN\big)=-H(a)$,

(ii) for every $a<\mu$, $\lim_N\frac{1}{N}\log\mathbb{P}\big(S_N\leq aN\big)=-H(a)$.

Then,

$$\frac{S_N}{N} \overset{\text{a.s.}}{\to} \mu.$$

Proof. Let $p > 0$ and $N_p \in \mathbb{N}$ such that $\forall N \geq N_p$,

$$\frac{1}{N}\log \mathbb{P}\Big(\frac{S_N - \mu N}{N} \geq \frac{1}{p}\Big) \leq -\frac{1}{2}H\big(\mu + \frac{1}{p}\big),$$

$$\frac{1}{N}\log \mathbb{P}\Big(\frac{S_N - \mu N}{N} \leq -\frac{1}{p}\Big) \leq -\frac{1}{2}H\big(\mu - \frac{1}{p}\big).$$

Hence

$$\mathbb{P}\Big(\Big|\frac{S_N - \mu N}{N}\Big| \geq \frac{1}{p}\Big) = \mathbb{P}\Big(\frac{S_N - \mu N}{N} \geq \frac{1}{p}\Big) + \mathbb{P}\Big(\frac{S_N - \mu N}{N} \leq -\frac{1}{p}\Big) \leq 2e^{-\frac{N}{2}\left(H\left(\mu+\frac{1}{p}\right)\wedge H\left(\mu-\frac{1}{p}\right)\right)},$$

which is summable as $H\big(\mu + \frac{1}{p}\big) \wedge H\big(\mu - \frac{1}{p}\big) > 0$. Hence, Borel-Cantelli Lemma ensures that

$$\mathbb{P}\Big(\bigcap_{N_0}\bigcup_{N\geq N_0}\Big\{\Big|\frac{S_N - \mu N}{N}\Big| \geq \frac{1}{p}\Big\}\Big) = 0,$$

so that the set $\bigcap_{p\geq 1}\bigcup_{N_0}\bigcap_{N\geq N_0}\Big\{\Big|\frac{S_N-\mu N}{N}\Big| < \frac{1}{p}\Big\}$ is of measure one and the assertion holds true. □

Remark that this theorem does not makes any assumption concerning the independence of the X_is, so that its conclusion does not fall in the classical SLLN. Moreover, one might expect that a CLT also holds provided that H is smooth on a neighborhood of μ with $H^{(2)}(\mu) = \frac{1}{\sigma^2} > 0$. In fact, this implies strict convexity in a neighborhood of μ, uniqueness of the minimum $H(\mu) = H'(\mu) = 0$, and

$$H(\mu \pm \varepsilon) = \frac{\varepsilon^2}{2\sigma^2} + o(\varepsilon^2),$$

for any $\varepsilon > 0$ small enough.

Now, letting $x > 0$, N large enough, and assuming that the converge (9.5) holds for $A := [\mu + \frac{x}{\sqrt{N}}, \mu + \frac{x+dx}{\sqrt{N}}[$ with an error term as $O(\frac{1}{N})$, we find

$$\begin{aligned}\mathbb{P}\Big(\frac{S_N - \mu N}{\sqrt{N}} \in [x, x+dx[\Big) &= \exp\Big\{-N\Big(H\big(\mu + \frac{x}{\sqrt{N}}\big) + O(N^{-1})\Big)\Big\} \\ &= \exp\Big\{-\frac{x^2}{2\sigma^2} + O(1)\Big\},\end{aligned}$$

so that, up to a multiplicative constant, we find the density of the centered Gaussian variable with variance σ^2.

9.4 General Theory

9.4.1 Basic definitions and properties

We now give the full formulation of large deviations principles that encompasses setting much more general than the case of i.i.d. sequence of real random variable. For convenience, we will restrict our attention to Polish spaces. More general settings in [84]. Let then Σ be a Polish space endowed with its Borel σ-field $\mathscr{B}(\Sigma)$.

The starting point of large deviations is to define what are the nice properties that we will require for our rate functions. We will distinguish between two notions, one stronger than the other.

Definition 9.4.1 (rate function)**.** (i) A rate function H is a lower semi-continuous function taking value in $[0,+\infty]$ and non uniformly infinite.

(ii) A good rate function is a rate function with compact level sets.

Remark 20. A good rate function always attains its minimal value on a non empty closed set F. In fact, if $\alpha = \inf_F H$, remark that the set

$$\bigcap_{k\geq 1} H^{-1}\Big(\Big[\alpha,\alpha+\frac{1}{k}\Big]\Big)\cap F = H^{-1}(\{\alpha\})\cap F$$

is compact and non-empty as the countable intersection of the sequence of decreasing non-empty compact sets. In contrast, a rate function may never attain its infimum, e.g. $x\to e^{-x}$ on $\mathbb{R}$.

By convention, the infimum (resp. supremum) of a function over an empty set will be equal to $+\infty$ (resp. $-\infty$). Moreover, as is traditional for large deviations theory, we will denote $H(\Gamma) := \inf_\Gamma H \neq \big\{x\in\Sigma, H(x)\big\}$ when H is a rate function.
We are now able to fully state the definition of a full or strong large deviation principle (FLDP):

Definition 9.4.2. Let $(P_N)_{N\in\mathbb{N}}$ be a sequence of probability measures on $\big(\Sigma,\mathscr{B}(\Sigma)\big)$. Then $(P_N)_{N\in\mathbb{N}}$ satisfies a large deviation principle with good rate function H iff for every Borel set $\Gamma\in\mathscr{B}(\Sigma)$,

$$-\inf_{x\in\overset{\circ}{\Gamma}} H(x) \leq \liminf_{N\to\infty}\frac{1}{N}\log P_N(\Gamma) \leq \limsup_{N\to\infty}\frac{1}{N}\log P_N(\Gamma) \leq -\inf_{x\in\bar{\Gamma}} H(x). \tag{9.6}$$

The right and left-hand sides of (9.6) are referred to as the upper and lower-bound of the FLDP.

As was expected by remark (18), the following proposition ensures that the good rate function associated with a FLDP is unique.

Proposition 9.4.1. Let $(P_N)_N$ be a sequence of probability measure on $\big(\Sigma,\mathscr{B}(\Sigma)\big)$, and two good rate functions H and I such that

$$\forall\Gamma\in\mathscr{B}(\Sigma),\quad \limsup_{N\to\infty}\frac{1}{N}\log P_N(\Gamma) \leq -\inf_{x\in\bar{\Gamma}} H(x),$$

$$\forall\Gamma\in\mathscr{B}(\Sigma),\quad -\inf_{x\in\overset{\circ}{\Gamma}} I(x) \leq \liminf_{N\to\infty}\frac{1}{N}\log P_N(\Gamma).$$

Then, $H\leq I$. In particular, if $(P_N)_N$ satisfies a FLDP, the rate function is unique.

Proof. Define the decreasing sequence of open balls $B_k := B\big(x, \frac{1}{k}\big), \forall k \in \mathbb{N}$. Then

$$\begin{aligned} -I(x) \leq -I(B_{k+1}) &\leq \liminf_N \frac{1}{N} \log P_N(B_{k+1}) \leq \limsup_N \frac{1}{N} \log P_N(B_{k+1}) \\ &\leq -H(\bar{B}_{k+1}) \leq -H(B_k). \end{aligned}$$

Letting $k \to \infty$, and using the lower semi-continuity of H, we get $\lim_k H(B_k) \geq H(x)$ so that $I(x) \geq H(x)$. □

A FLDP imposes further conditions to both the rate function and sequence of probability it is associated with it is associated with. Indeed, this proposition shows that $(P_N)_N$ must be exponentially tight.

Proposition 9.4.2. A sequence $(P_N)_{N\in\mathbb{N}}$ satisfying a FLDP on a Polish space is exponentially tight.

Proof. Let H denote the good rate function associated with the FLDP, and recall that it attains its minimum on every non-empty closed set. Σ being a Polish space, it is separable and exists a dense sequence of Σ $(x_i)_{i\in\mathbb{N}^*}$.

Let $M, \delta > 0$, and define the non-decreasing sequence of closed sets $\left(F_p := \bigcap_{i=1}^p B(x_i, \delta)^c\right)_{p\in\mathbb{N}^*}$. By density of the sequence (x_i), we have $H\Big(\bigcap_{p\geq 1} F_p\Big) = +\infty$. Moreover, $\{H \leq M + \delta^{-1}\}$ and $\{H \leq M + \delta^{-1}\} \cap F_p$ are compact sets with $\bigcap_{p\geq 1} \big(F_p \cap \{H \leq M\}\big) = \emptyset$. As Σ is a complete space, then exists $K \in \mathbb{N}^*$ such that $\Big(\bigcap_{p=1}^K F_p\Big) \cap \{H \leq M + \delta^{-1}\}$ is empty, implying that $\forall p \geq K, H(F_p) > M + \delta^{-1}$. Hence, the upper-bound of the FLDP ensures the existence of $n_0 \in \mathbb{N}^*$ such that $\forall N \geq n_0$ $\frac{1}{N} \log P_N\big(F_K\big) < -(M + \delta^{-1})$. Moreover, as for every $\mu \in M_1^+(\Sigma)$, $\mu(F_p) \to_p 0$, we can extend the previous inequality to every $N \in \mathbb{N}^*$. Similarly, for every $k \in \mathbb{N}^*$, we can find $K_k \in \mathbb{N}^*$ such that

$$\forall N \in \mathbb{N}^*, \frac{1}{N} \log P_N\big(\bigcap_{i=1}^{K_k} B(x_i, k^{-1})^c\big) < -(M + k).$$

Let

$$A := \bigcap_{k\geq 1} {F_{K_k}^k}^c = \bigcap_{k\geq 1} \bigcup_{p=1}^{K_k} B(x_i, k^{-1}),$$

which has compact closure as Σ is Polish (this classical result can be found in [29]), and remark that

$$P_N(\bar{A}^c) \leq P_N(A^c) \leq \sum_{k\geq 1} e^{-(M+k)N} = e^{-MN} \frac{1}{e^N - 1},$$

so that

$$\limsup_N \frac{1}{N} \log P_N(\bar{A}^c) < -M.$$

□

Moreover, remark that for any Borel set $\Gamma \in \mathscr{B}(\Sigma)$, $H(\bar{\Gamma}) > 0$ implies an exponential decay for $P_N(\Gamma)$ with rate lying between $H(\mathring{\bar{\Gamma}})$ and $H(\bar{\Gamma})$. P_N being a probability measure, we necessarily have $H(\Sigma) = 0$, so that $H^{-1}(\{0\})$ is a non-empty compact set by goodness of H (see remark 20).

In definition 9.4.2, one sees that for Borel sets Γ satisfying $H(\overset{\circ}{\Gamma}) = H(\bar{\Gamma})$ the two inequalities become an equality:

$$\lim_N \frac{1}{N} \log P_N(\Gamma) = -H(\Gamma). \tag{9.7}$$

It seems tempting to replace the upper and lower-bound by this simple limit that evokes Cramér's theorem (see remark 9.5). Nonetheless, it will appear that the difference $H(\overset{\circ}{\Gamma}) - H(\bar{\Gamma})$ is very relevant, in particular when dealing with Borel sets composed of atoms. Suppose for example that P_N has no atoms for every $N \in \mathbb{N}^*$. Then (9.7) would yield

$$\forall x \in \Sigma, \lim_N \frac{1}{N} \log P_N(\{x\}) = -\infty,$$

so that $\forall x \in \Sigma$, $H(x) = +\infty$, which is not an acceptable rate function according to definition (9.4.1). Moreover, take $\Gamma = \{x\} \cup [a, \infty[$, with $x < a$, and suppose that H is increasing on $[x, \infty[$ with $H(x) < H(a)$ (think for example of Cramér's case for $\mu < x < a$). Then, the FLDP satisfied by $(P_N)_N$ ensures that

$$\lim_N \frac{1}{N} \log P_N(\Gamma) = -\inf_{z \in [a,\infty[} H(z) = -H(a) \neq -H(x) = -\inf_{z \in \{x\} \cup [a,\infty[} H(z),$$

as $P_N(\Gamma) = P_N([a, \infty[)$, and $\{x\}$ seems not not to be seen by the FLDP.

Theorem 9.4.3. *Let $(X_i)_{i \in \mathbb{N}^*}$ be a sequence satisfying the hypothesis of Cramér's Theorem, and let $\phi(t) := \log \mathbb{E}\left[e^{tX_1}\right]$ the logarithmic moment generating function of X_1. Further suppose that X_1 admits not atoms.*

Then, $P_N := \mathscr{L}\left(\frac{S_N}{N}\right)$ satisfies a FLDP with good rate function $H : z \to \sup_{t \in D_\phi} \left\{zt - \phi(t)\right\}$ given by the Legendre transform of ϕ.

Remark 21. It is very easy to see why the FLDP fails when X_1 admits atoms. Indeed, suppose that $\mathbb{P}(X_1 = a) = p > 0$. Then

$$\mathbb{P}\left(\frac{S_N}{N} \in \{a\}\right) \geq \mathbb{P}(X_1 = \ldots = X_N = a) = p^N,$$

so that

$$\frac{1}{N} \log \mathbb{P}\left(\frac{S_N}{N} \in \{a\}\right) \geq \log p > -\inf_{\overset{\circ}{\{a\}}} H = -\infty,$$

as $\overset{\circ}{\{a\}} = \emptyset$.

Also remark that, provided the existence of the logarithmic moment generating function ϕ, the upper-bound for the Legendre transform naturally arises from an exponential Chebyshev's inequality and an optimization over the parameter in the exponential.

Proof. According to Proposition 9.3.2 we already that H is a good rate function. To prove that (9.6) holds for any $\Gamma \in \mathscr{B}(\mathbb{R})$, first remark that if $\mu \in \overset{\circ}{\Gamma}$ the equality is obvious relying on the SLLN and on the fact that $H(\mu) = 0$. Moreover, if $\mu \notin \overset{\circ}{\Gamma}$, we can find two Borel disjoint sets Γ_+ and Γ_- such that $\Gamma = \Gamma_+ \cup \Gamma_-$, $\Gamma_+ \subset [\mu, \infty[$, $\Gamma_- \subset]-\infty, \mu]$. Furthermore, supposing that (9.6) holds for $\Gamma_\pm$ Lemma (9.2.4) ensures that

$$\begin{aligned}\limsup_N \frac{1}{N} \log \mathbb{P}\left(\frac{S_N}{N} \in \Gamma\right) &= \max\left\{\limsup_N \frac{1}{N} \log \mathbb{P}\left(\frac{S_N}{N} \in \Gamma_-\right), \limsup_N \frac{1}{N} \log \mathbb{P}\left(\frac{S_N}{N} \in \Gamma_+\right)\right\} \\ &= \max\left\{-\inf_{\bar{\Gamma}_-} H, -\inf_{\bar{\Gamma}_+} H\right\} = -\inf_{\bar{\Gamma}} H,\end{aligned}$$

and

$$\begin{aligned}\liminf_N \frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in\Gamma\Big) &= \max\Big\{\liminf_N \frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in\Gamma_-\Big), \liminf_N \frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in\Gamma_+\Big)\Big\}\\ &= \max\big\{-\inf_{\overset{\circ}{\Gamma}_-} H, -\inf_{\overset{\circ}{\Gamma}_+} H\big\} = -\inf_{\overset{\circ}{\Gamma}} H,\end{aligned}$$

so that proving (9.6) for $\Gamma\in\mathscr{B}(\mathbb{R})$ with $\Gamma\subset[\mu,+[$ suffices by symmetry.

Choose then such a Γ, and let us prove the upper-bound by an exponential Chebyshev inequality. Let $\alpha := \inf(\Gamma)\geq\mu$. Then

$$\begin{aligned}\frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in\Gamma\Big) &\leq \frac{1}{N}\log\mathbb{E}\Big[\mathbb{1}_{\{\frac{S_N}{N}\in\Gamma\}}e^{tS_N-Nt\alpha}\Big]\\ &\leq -\alpha t+\phi(t).\end{aligned}$$

As this holds for any $t\in\mathbb{R}$, we find

$$\frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in\Gamma\Big)\leq\inf_{t\in\mathbb{R}}\big\{-\alpha t+\phi(t)\big\} = -H(\alpha) = -\inf_{\bar{\Gamma}} H, \tag{9.8}$$

by strictly monotony of H on $[\mu,\infty[$.

We now prove the lower bound for open sets only which will be sufficient as $P_N(\Gamma)\geq P_N(\overset{\circ}{\Gamma})$. Cramér's Theorem (9.3.1) and equality (9.5) already ensure that the lower-bound is satisfied for sets of the form $[a,b[$, with $\mu\leq a<b\leq+\infty$. As $\inf_{[a,b[}H = H(a) = \inf_{]a,b[}H$, and $\mathbb{P}\big(\frac{S_N}{N}\in]a,b[\big) = \mathbb{P}\big(\frac{S_N}{N}\in[a,b[\big)$, X_1 admitting no atom., the lower-bound also holds for sets of the form $]a,b[$, with $\mu\leq a<b\leq+\infty$. Now, Γ being an open set of $[\mu,+\infty[$, it can be decomposed in a countable union of disjoint sets of the form $I_p=]a_p,b_p[$, with $\mu\leq a_1<b_1\leq\ldots\leq a_p<b_p\leq+\infty, \forall p\in\mathbb{N}^*$. We then conclude by remarking that, for any $P\in\mathbb{N}^*$, Lemma (9.2.4) ensures

$$\begin{aligned}\liminf_N \frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in\Gamma\Big) &\geq \liminf_N \frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in\bigcup_{p=1}^{P} I_p\Big) = \max_{1\leq p\leq P}\Big\{\liminf_N \frac{1}{N}\log\mathbb{P}\Big(\frac{S_N}{N}\in I_p\Big)\Big\}\\ &= \max_{1\leq p\leq P}\big\{-H(a_p)\big\} = -H(a_1) = -\inf_{\Gamma} H.\end{aligned}$$

□

The main interest of LDP is that, under good conditions, they contain a convergence in law for the sequence of measures at stake. As in Lemma (9.3.3), such result can go well beyond the scope of classical theorems from probability theory, encompassing in particular nonlinear cases. Note that the links of large deviations theory with the weak convergence of probability measures has been the object of an intense investigation (see e.g. [185], [184], and [90]). In this context, the upper and lower-bound of (9.6) can be seen as exponential versions of that involved in PortManteau Theorem. Taken conjointly, they give an equality for Borel set Γ satisfying a regularity condition: $P(\partial\Gamma)=0$ in PortManteau theorem, and $H(\bar{\Gamma}) = H(\overset{\circ}{\Gamma})$ for LDP. The following proposition makes explicit the good conditions required for the obtaining of a convergence in law, and also stresses an important fact: when satisfying a FLDP, the mass of the sequence of probability measures P_N asymptotically concentrates around the zeros of the rate function H.

Proposition 9.4.4. Let $(P_N)_{N\in\mathbb{N}}$ satisfies a FLDP with good rate function H on $\big(\Sigma,\mathscr{B}\big)$. Then

(i) $\exists x^* \in \Sigma$, $H(x^*) = 0$,

(ii) if $P_N \xrightarrow{\mathscr{L}} P \in \mathscr{M}_1^+(\Sigma)$ and H is continuous, then $P\big(H^{-1}(\{0\})\big) = 1$,

(iii) if $H^{-1}(\{0\}) = \{x^*\}$, then $P_N \xrightarrow{\mathscr{L}}_N \delta_{x^*}$,

where δ_{x^*} is the probability measure of $\mathscr{M}_1^+(\Sigma)$ degenerated at x^*.

Proof. 1. As $\frac{1}{N}\log P_N(\Sigma) = 0 \leq -H(\Sigma)$ so that $H(\Sigma) = 0$. Moreover, by remark (20), H attains its infimum, so that the first point is proved, and $H^{-1}(\{0\}) \neq \emptyset$.

2. Let $\varepsilon > 0$, and define the closed set $\Gamma_\varepsilon := \{x \in \Sigma, H(x) \geq \varepsilon\}$. Then, the upper-bound of the FLDP brings

$$\limsup \frac{1}{N}\log P_N(\Gamma_\varepsilon) \leq -H(\Gamma_\varepsilon) \leq -\varepsilon < 0,$$

so that PortMantau theorem ensures that

$$0 = \liminf P_N(\Gamma_\varepsilon) \geq P(\Gamma_\varepsilon).$$

Taking $\varepsilon \to 0$, we conclude by the Monotone Convergence Theorem.

3. Suppose now that $H^{-1}(\{0\}) = \{x^*\}$, and take any open set $\Gamma \subset \Sigma$. We want to show that

$$\liminf_N P_N(\Gamma) \geq \delta_{x^*}(\Gamma)$$

and conclude by PortManteau Theorem. When $x^* \notin \Gamma$, the proof is immediate. Suppose that $x^* \in \Gamma$. Then, using the upper-bound for the closed set Γ^c:

$$\limsup \frac{1}{N}\log P_N(\Gamma^c) \leq -H(\Gamma^c) < 0,$$

so that necessarily, $P_N(\Gamma^c) \to 0$, and the proof is completed. □

Remark 22. In Proposition (9.4.4) point 2., the continuity hypothesis of H can be relaxed as follow: exists a decreasing positive sequence $(\varepsilon_p)_{p \in N^*}$ such that

$$\varepsilon_P \searrow 0, \text{ and } \forall p \in \mathbb{N}^*, \quad \Gamma_p := \big\{x \in \Sigma, H(x) \geq \varepsilon_p\big\} \text{ is closed.}$$

An alternative hypothesis could also be that the number of connected components of $H^{-1}(\{0\})$ is finite, so that the infimum of H over the closed set $\bigcap_{x \in H^{-1}(\{0\})} B(x,\varepsilon)^c$ is strictly positive.

As FLDP can be hard to obtain, there exists a weaker notion of LDP:

Definition 9.4.3. Let $(P_N)_{N \in \mathbb{N}^*}$ be a sequence of probability measures on $\big(\Sigma, \mathscr{B}\big)$. Then $(P_N)_{N \in \mathbb{N}^*}$ satisfies a weak large deviation principle (WLDP) with rate function H if and only if

- for any set $\Gamma \in \mathscr{B}$ with compact closure,

$$\limsup_{N \to \infty} \frac{1}{N}\log P_N(\Gamma) \leq -\inf_{x \in \bar{\Gamma}} H(x). \tag{9.9}$$

- $\forall \Gamma \in \mathscr{B}$,

$$-\inf_{x \in \overset{\circ}{\Gamma}} H(x) \leq \liminf_{N \to \infty} \frac{1}{N}\log P_N(\Gamma).$$

Remark 23. An alternative (and equivalent) formulation of the weak upper-bound (9.9) is: for any compact set $K \in \mathscr{B}(\Sigma)$

$$\limsup_{N\to\infty} \frac{1}{N} \log P_N(K) \leq -\inf_{x\in K} H(x).$$

This weaker formulation looses most of the good properties of the FLDP. Whereas the rate function associated with a WLDP is still unique (see [184]), it does not need to be good anymore, so that $(P_N)_N$ might not be exponentially tight, or even tight, allowing cases where P_N utterly vanish in the limit $N \to \infty$. Moreover the infimum of H needs not to be 0 in general as Σ is not necessarily a compact set.

Example. The sequence $\left(P_N := e^{-N}\delta_0 + (1-e^{-N})\delta_N\right) \in \mathscr{M}_1^+(\mathbb{R})^{\mathbb{N}^*}$ satisfies a WLDP with rate function $H : \mathbb{R} \to \bar{\mathbb{R}}$ defined by $H(0) = 1$, $H(x) = +\infty$ if $x \neq 0$.
In fact, for any Borel set $\Gamma \in \mathscr{B}(\mathbb{R})$, $H(\Gamma) = 1 + (+\infty) \times \mathbb{1}_{\Gamma^c}(0)$, so that the lower-bound

$$\forall N \in \mathbb{N}^*, -H(\Gamma) \leq \frac{1}{N} \log(P_N(\Gamma))$$

is trivial when $0 \notin \Gamma$, and still holds in the other case, as $P_N(\Gamma)$ is then greater than $\frac{1}{N}\log(e^{-N}) = -1$. Moreover, the upper-bound for any compact set K also holds as $\lim_N \frac{1}{N}\log(P_N(K)) = -1 + (-\infty) \times \mathbb{1}_{K^c}(0) = -H(K)$. Nevertheless, remark that $(P_N)_N$ is not tight and that $\inf H = 1 \neq 0$.

The following lemma highlights the links between weak and full LDP:

Lemma 9.4.5. Let $(P_N)_N$ be a sequence of probability measures on a Polish space Σ. We have the equivalence:

(i) $(P_N)_N$ is exponentially tight and satisfies a WLDP with rate function H,

(ii) $(P_N)_N$ satisfies a FLDP with good rate function H.

Proof. By Proposition (9.4.2), we only have to prove 1. $\implies$ 2.:
Let us prove first that H is a good rate function. Let $M > 0$, and consider the closed set $K_M := \{H \leq M\}$, and a compact K_M such that

$$\limsup_N \frac{1}{N} \log P_N\left(K^c\right) \leq -(M+1).$$

The lower-bound of the WLDP implies

$$-H(K^c) \leq \liminf_N \frac{1}{N} \log P_N\left(K^c\right) \leq -(M+1).$$

Hence, $H(K^c) \geq (M+1)$, and $K_M \subset K$. Then the set $K_M \cap K = K_M$ is compact so that H is a good rate function.
Let us now prove the upper-bound for closed sets. Let F be a closed set. Then $F \cap K$ is a compact set, and we have

$$\limsup_N \frac{1}{N} \log P_N\left(F \cap K\right) \leq -H(F \cap K) \leq -H(F).$$

Moreover,

$$\begin{aligned}\limsup_N \frac{1}{N} \log P_N(F) &\leq \limsup_N \frac{1}{N} \log\left(P_N(F \cap K) + P_N(K^c)\right) \\ &\leq \max\left(-H_F, -M\right),\end{aligned}$$

so that sending M towards infinity gives the result.

□

The advantage of WLDP is that they require very mild conditions to ensure a convergence in law for the sequence $(P_N)_N$.

Proposition 9.4.6. Suppose that $(P_N)_{N\in\mathbb{N}}$ satisfies an upper-bound for compact sets (9.9) with rate function H on $\left(\Sigma,\mathscr{B}\right)$. If $H^{-1}(\{0\})=\{x^*\}$ and $(P_N)_N$ is tight, then

$$P_N \overset{\mathscr{L}}{\to}_N \delta_{x^*}.$$

Remark 24. Remark here that we do not need a lower-bound for the convergence to hold. Whereas these upper-bound and uniqueness of the 0 of H put sufficient constraint on the sequence $(P_N)_N$ for ensuring its convergence, such a function H might not be unique and do not characterize the speed of the convergence. As suggested by Proposition (9.4.1) in the case of FLDP, it should nevertheless be smaller than the rate function provided its existence. In particular, it implies that this latter can only have one 0: the same as H.

Proof. Let $\varepsilon,\delta>0$, and choose a compact set K_ε such that

$$\forall N\in\mathbb{N}^*,\quad P_N(K_\varepsilon^c)<\varepsilon.$$

Then

$$P_N(B(x^*,\delta)^c)\leq P_N(B(x^*,\delta)^c\cap K_\varepsilon)+P_N(K_\varepsilon^c)\leq P_N(B(x^*,\delta)^c\cap K_\varepsilon)+\varepsilon.$$

The upper-bound for compact sets ensures that

$$\limsup_N \frac{1}{N}\log P_N(B(x^*,\delta)^c\cap K_\varepsilon)=-H(B(x^*,\delta)^c\cap K_\varepsilon)<0,$$

so that $\lim_N P_N(B(x^*,\delta)^c\cap K_\varepsilon)=0$, and

$$\limsup_N P_N(B(x^*,\delta)^c)\leq\varepsilon.$$

Sending $\varepsilon\to 0$, we find that $\lim_N P_N(B(x^*,\delta)^c)=0$. Take now $F\in\mathscr{B}(\Sigma)$ a closed set of Σ. If $x^*\in F$, then

$$\limsup_N P_N(F)\leq\delta_{x^*}(F)=1$$

trivially holds. If $x^*\notin F$, then, exists $\delta>0$ such that $F\subset B(x^*,\delta)^c$, so that $\limsup_N P_N(F)\leq\lim_N P_N(B(x^*,\delta)^c)=0=\delta_{x^*}(F)$, and we conclude by PortManteau theorem. □

Remark 25. Remark that the speed of the convergence might no more be exponential as in Proposition (9.4.4).

9.5 OBTAINING A LDP

Now we have an idea of the utility of large deviations principles, we would like to know how to obtain them. The answer dwells in two theorems. Sanov's theorem deals with i.i.d. Polish-valued sequences of random variables. It can be seen as an extension of the weak law of large numbers, and have the advantage of encompassing a wide class of settings. We furnish here the general formulation of the theorem, introduce important notions, and In order to give an idea of its meaning, we also give a full proof of the theorem in the case of finite spaces at the end of the appendix 9.7.1. On the other hand, Gartnër-Ellis theorem an be seen as a generalization of Cramér's theorem, and deals with dependent real random vectors. For readers interested in the full proofs of these theorem, we refer to [84, 85].

9.5.1 Sanov's theorem

In this section $P \in \mathscr{M}_1^+(\Sigma)$ is a probability measure on a Polish space, and $\mathbf{X} = (X_i)_{i\in\mathbb{N}^*}$ is an i.i.d. sequence with law P. We first define the empirical measure of $\mathbf{X}$:

Definition 9.5.1 (Empirical measure). The empirical measures of the sequence $\mathbf{X} = (X_i)_{i\in\mathbb{N}^*} \in \Sigma^{\mathbb{N}^*}$ is given, for any $N \in \mathbb{N}^*$, by

$$\hat{\mu}_N^{\mathbf{X}} := \frac{1}{N}\sum_{i=1}^{N} \delta_{X_i} \in \mathscr{M}_1^+(\Sigma),$$

where for any $x \in \Sigma$, $\delta_x \in \mathscr{M}_1^+(\Sigma)$ is the degenerate measure at x.

Remark in particular that if the X_is are random (which is the case here), the empirical measure is a random element of $\mathscr{M}_1^+(\Sigma)$. A way of understanding the empirical measure $\hat{\mu}_N^{\mathbf{X}} \in \mathscr{M}_1^+(\Sigma)$ is to see it as a version of the classical empirical mean

$$\frac{S_N^{\mathbf{X}}}{N} := \frac{\sum_{i=1}^N X_i}{N} \in \Sigma.$$

Of course, one immediately sees that

$$\hat{\mu}_N^{\mathbf{X}} = \frac{S_N^{\delta_{\mathbf{X}}}}{N}.$$

More interestingly, $S_N^{\mathbf{X}}/N$ might not be well-defined or lie in Σ, as Polish spaces are not generally real vector spaces, or even to be convex ensembles. In these problematic cases, $\hat{\mu}_N^{\mathbf{X}}$ still exists and has value in the convex Polish space $\mathscr{M}_1^+(\Sigma)$. It thus appears as the most natural quantity describing the averaged behavior of the distribution P, and as a good candidate to follows some kind of law of large number that Sanov's theorem will make explicit.

To avoid confusion, we insist on the fact that when $S_N^{\mathbf{X}}/N$ is a well-defined element of Σ, $\hat{\mu}_N^{\mathbf{X}} \neq \mathscr{L}\left(S_N^{\mathbf{X}}/N\right)$ in general. For example, taking i.i.d. Bernoulli variables, one sees that the empirical measure has support in $\{0,1\}$, whereas the empirical mean can be equal to $\frac{1}{2}$ with positive probability.

We now define what will be the good rate function of Sanov's theorem, namely the Kullback-Leibler divergence, or relative entropy:

Definition 9.5.2. Let $P, \mu \in \mathscr{M}_1^+(\Sigma)$, where Σ is a Polish space. Then, the relative entropy of μ with respect to P is defined as follow

$$I(\mu|P) = \begin{cases} \int_\Sigma \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}P}(x)\right) d\mu(x) & \text{if } \mu \ll P, \\ +\infty & else. \end{cases}$$

It is linked with information theory, and can be seen as a heuristic distance on the space $\mathscr{M}_1^+(\Sigma)$. Nevertheless, it is not symmetric nor does it follow axiom the triangular inequality. Before proving that it is indeed a rate-function, let us state now the theorem:

Theorem 9.5.1. *The sequence of probability measures* $\left(P^{\otimes N}\left(\hat{\mu}_N^X \in \cdot\right)\right)_N \in \mathscr{M}_1^+\left(\mathscr{M}_1^+(\Sigma)\right)^{\mathbb{N}^*}$ *satisfies a FLDP with good rate function* $I(\cdot|P)$.

Proposition 9.5.2. Let $D_I := \left\{\mu \in \mathscr{M}_1^+(\Sigma), I(\mu|P) < \infty\right\}$ and $\mu \in \mathscr{M}_1^+(\Sigma)$. Then

(i) $I(\cdot|P) \geq 0$,

(ii) $I(\cdot|P)$ is a lower semi-continuous and convex function,

(iii) $I(\mu|P) = 0$ iff $\mu = P$,

(iv) $I(\cdot|P)$ is strictly convex on $\mathring{D}_I$,

(v) Moreover,

$$I(\mu|P) = \Lambda^*(\mu) := \sup_{\phi \in \mathscr{C}_b(\Sigma)} \left\{ \int_\Sigma \phi(x) d\mu(x) - \Lambda(\phi) \right\}, \tag{9.10}$$

where $\Lambda(\phi) := \log \int_\Sigma e^{\phi(x)} dP(x)$.

(vi) $I(\cdot|P)$ is a good rate function.

Remark 26. (i) This lemma reveals the true nature of the relative entropy: it is the Legendre transform of a kind of logarithmic moment generating function of P. It is then not that surprising that it constitutes the good rate function associated with the FLDP of Sanov's Lemma.

(ii) By Proposition (9.4.4), $P^{\otimes N}\left(\hat{\mu}_N^X \in \cdot\right) \overset{\mathscr{L}}{\to} \delta_P$, as P is the unique minimum of $I(\cdot|P)$.

(iii) From the proof of this lemma appears a useful inequality for the relative entropy: for any bounded measurable function ϕ on Σ,

$$\int_\Sigma \phi d\mu \leq I(\mu|P) + \log \int_\Sigma \exp \phi dP. \tag{9.11}$$

The monotone convergence theorem ensures that this inequality also holds for positive measurable ϕ.

Proof. 1. Let $\mu \ll P$, and remark that

$$I(\mu|P) = \int_\Sigma \frac{\mathrm{d}\mu}{\mathrm{d}P} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}P}\right) dP.$$

Then, as $x \to x\log(x)$ is a convex function, Jensen inequality gives the result.

2. Remark that, for any $\phi \in \mathscr{C}_b(\Sigma)$, $\mu \to \int_\Sigma \phi d\mu$ is convex and continuous. We then rely on (9.2.1).

3-4. Remark that $I(P|P) = 0$. Uniqueness of the minimum and strict convexity on $\overset{\circ}{D}_I$ are then obtained relying on the strict convexity of $x \to x\log(x)$ and Jensen inequality.

5. Let $\mu \not\ll P$. If $\Lambda^*(\mu) \neq \infty$, we would have for any $\phi \in \mathscr{C}_b(\Sigma)$:

$$\int_\Sigma \phi d\mu - \log\Big(\int_\Sigma e^\phi dP\Big) \leq \Lambda^*(\mu) < \infty, \tag{9.12}$$

inequality that also holds for bounded measurable functions by the dominated convergence theorem. Let Γ such that $P(\Gamma) = 0$, and define $\phi := r\mathbb{1}_\Gamma$ for a given $r > 0$. Then, inequality (9.12) reads

$$r\mu(\Sigma) - \log\Big(\int_{\Gamma^c} e^0 dP\Big) = r\mu(\Sigma) \leq \Lambda^*(\mu) < \infty,$$

so that, as r can be chosen arbitrary large, we necessarily have $\mu(\Sigma) = 0$, which contradicts $\mu \not\ll P$. Suppose now that $\mu \ll P$, and define $f := \frac{\mathrm{d}\mu}{\mathrm{d}P}$. $\forall t \in [0,1]$, also define $\mu_t := tP + (1-t)\mu$. Then, $\mu_t \ll P$, with $f_t := \frac{\mathrm{d}\mu_t}{\mathrm{d}P} = t + (1-t)f$, so that by convexity of $x \to x\log(x)$ on the one hand, and as $x \to \log(x)$ is both non-decreasing and concave on the other hand, we find

$$I(\mu_t|P) = \int f_t \log f_t dP \leq (1-t)I(\mu|P),$$

and

$$\begin{aligned} I(\mu_t|P) &= t\int_\Sigma \log\Big(t + (1-t)f\Big)dP + (1-t)\int \log\Big(t\times 1 + (1-t)f\Big)\frac{\mathrm{d}\mu}{\mathrm{d}P}dP \\ &\geq t\log(t) + (1-t)^2 I(\mu|P) \end{aligned}$$

In particular, $\lim_{t\searrow 0} I(\mu_t|P) = I(\mu|P)$. Let us prove the same for Λ^*. It is a convex and lower semi-continuous by Proposition 9.2.1, also positive as $0 \in \mathscr{C}_b(\Sigma)$. Moreover, for any $\phi \in C_b$, Jensen inequality yields $\Lambda_\phi \geq \int_\Sigma \phi(x)dP(x)$, so that $\Lambda^*(P) = 0$, and

$$\Lambda^*(\mu_t) \leq (1-t)\Lambda^*(\mu).$$

Let $\varepsilon > 0$, and choose ϕ_ε such that $\Lambda^*(\mu) \leq \int_\Sigma \phi_\varepsilon(x)d\mu(x) - \Lambda_{\phi_\varepsilon} + \varepsilon$. Then

$$\Lambda^*(\mu_t) \geq (1-t)\big(\Lambda^*(\mu) - \varepsilon\big) + t\Big(\int_\Sigma \phi_\varepsilon(x)dP(x) - \Lambda_{\phi_\varepsilon}\Big).$$

As $\big(\int_\Sigma \phi_\varepsilon(x)dP(x) - \Lambda_{\phi_\varepsilon}\big)$ is bounded, taking the limit $t \to 0$ yields

$$\Lambda^*(\mu) - \varepsilon \leq \liminf_{t\searrow 0} \Lambda^*(\mu_t) \leq \limsup_{t\searrow 0} \Lambda^*(\mu_t) \leq \Lambda^*(\mu),$$

so that $\lim_{t\searrow 0}\Lambda^*(\mu_t) = \Lambda^*(\mu)$. It is now sufficient to prove the result equality for μ_t with $t \in]0,1[$ or, more generally, for $\mu \ll P$ such that their density is bounded away from 0. Suppose thus that exists $\eta > 0$ such that $f \geq \eta$, and let us show that $I(\mu|P) \geq \sup_{\phi\in\mathscr{C}_b(\Sigma)} \big\{\int_\Sigma \phi(x)d\mu(x) - \Lambda_f\big\}$. Let $\phi \in \mathscr{C}_b$ and remark that

$$\exp\Big\{\int_\Sigma \phi d\mu - I(\mu|P)\Big\} = \exp\Big\{\int_\Sigma \big(\phi - \log f\big)d\mu\Big\} \leq \int_\Sigma \frac{e^\phi}{f}d\mu = \int_\Sigma e^\phi dP,$$

so that $I(\mu|P) \geq \Lambda^*(\mu)$. Set $f_M = f \wedge M$ that is positive bounded and measurable so that

$$I(\mu|P) = \int_\Sigma f \log f dP = \int_\Sigma \log f d\mu \overset{\text{Fatou}}{\leq} \liminf_{M\to\infty} \int_\Sigma \log f_M d\mu$$
$$\overset{(9.12)}{\leq} \Lambda^*(\mu) + \liminf_{M\to\infty} \log\Big(\int_\Sigma f_M d\mu\Big) \overset{\text{MCT}}{=} \Lambda^*(\mu),$$

which complete the proof of equality (9.10).

6. The goodness of $I(\cdot|P)$ is somewhat technical. We admit it here, and refer to [86, Lemma 3.2.7]. □

9.5.2 Gartnër-Ellis Theorem

Let $d \in \mathbb{N}^*$, and (Z_N) be a sequence of random vectors of the probability space $\big(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d), \mathbb{P}\big)$, where $\mathscr{B}(\mathbb{R}^d)$ is the Borel σ-field on $\mathbb{R}^d$. Define for any $t \in \mathbb{R}^d$ the logarithmic moment generating function:

$$\Lambda_N(t) := \log \mathbb{E}\Big[e^{N\langle t, Z_N\rangle}\Big],$$

where $\langle .,. \rangle$ is the scalar product on $\mathbb{R}^d$, and let the probability measure $P_N := \mathbb{P}(Z_N \in .)$. The aim of Gartnër-Ellis theorem is to prove a FLDP for the sequence (P_N). Before stating the theorem, let us introduce a few definitions.

Definition 9.5.3 (Legendre transform and exposed points). Let $f : \mathbb{R}^d \to \bar{\mathbb{R}}$. Its Legendre transform is defined by f^*:

$$f^*(x) = \sup_{t\in\mathbb{R}^d} \big\{\langle x,t\rangle - f(t)\big\},\ \forall x \in \mathbb{R}^d.$$

Moreover, a point $x \in \mathbb{R}^d$ is called exposed for f^* iff exists $t \in \mathbb{R}^d$ such that

$$f^*(y) - f^*(x) > \langle y - x, t\rangle, \quad \forall y \neq x.$$

Such t is called an exposing hyperplane for x.

Remember that Proposition 9.2.1 ensures that f^* is a convex function. In words, an exposing point for f^* is a point at which f^* is strictly convex, while for an hyperplane exposing $t \in \mathbb{R}^d$ is such that exists an affine sub-space with direction $\{t\}^\perp$ that tangents f^* at x without touching any other point of the graph of f^*.

Remark that for $d = 1$, and $Z_N := \frac{1}{N}\sum_{i=1}^N X_i$ with X_i i.i.d. random variables with exponential moments, ϕ_N is stationary as we have

$$\forall t \in \mathbb{R},\ \Lambda_N(t) = \log \mathbb{E}\big[e^{t\sum_{i=1}^N X_i}\big] = N\phi(t),$$

with ϕ defined as in Cramér's theorem. In this simple setting, the rate function was given by the Legendre transform of ϕ. Nevertheless, for dependent sequence, or even more general Z_N, $\frac{1}{N}\phi_N$ generally depends on N. As a large deviations principle is only interested in asymptotic behaviors, the natural hope is that provided $\frac{1}{N}\phi_N \to \Lambda$ when $N \to \infty$, a FLDP will follows for $(P_N)_N$, with good rate function Λ^*. Gartnër-Ellis theorem specify that this is indeed the case when Λ is sufficiently regular, and Λ^* is strictly convex.

Theorem 9.5.3 (Gartnër-Ellis Theorem). *Assume that*

(a) $\forall t \in \mathbb{R}^d, \quad \Lambda(t) := \lim_N \frac{1}{N}\phi_N(t)$ *exists in* $\bar{\mathbb{R}}$,

(b) $0 \in \overset{\circ}{D_\Lambda}$, *with* $D_\Lambda := \{t \in \mathbb{R}^d : \Lambda(t) < \infty\}$.

Then, for any $\Gamma \in \mathscr{B}(\mathbb{R}^d)$

(i) Λ^* *is a good rate function,*

(ii) $\limsup_N \frac{1}{N} \log P_N(\Gamma) \leq -\Lambda^*(\Gamma)$,

(iii) $\liminf_N \frac{1}{N} \log P_N(\Gamma) \geq -\Lambda^*(\Gamma \cap E)$,

where E *is the set of exposed points of* Λ^* *whose exposing hyperplane belongs to* $\overset{\circ}{D_\Lambda}$. *Moreover, if* Λ *is also lower semi-continuous, differentiable on* $\overset{\circ}{D_\Lambda}$ *with either* $D_\Lambda = \mathbb{R}^d$ *or a steep condition at* ∂D_Λ *then,* $\overset{\circ}{\Gamma} \cap E$ *can be replaced by* $\overset{\circ}{\Gamma}$ *in 3., so that* (P_N) *satisfies a FLDP with good rate function* Λ^*.

Remark that Gartnër-Ellis theorem fails to establish a LDP when Λ^* is not strictly convex.

9.6 DERIVING A LDP

This section provides the general results of the theory that build a FLDP out of another one. With this arsenal, it will be very easy to obtain FLDP for complex cases presenting e.g. nonlinearities or dependent sequences. We begin with the contraction principle, that deals with the continuous image of a sequence of measures satisfying a FLDP.

Theorem 9.6.1 (Contraction Principle). *Let* $(P_N)_{N\in\mathbb{N}}$ *satisfies a FLDP with good rate function* H *on* $(\Sigma, \mathscr{B})$, $(\tilde{\Sigma}, \mathscr{B}_{\tilde{\Sigma}})$ *another Polish space endowed with its Borel* σ*-field, and* $f : \Sigma \to \tilde{\Sigma}$ *a continuous application. Then,* $\left(P_N \circ f^{-1}\right)_{N\in\mathbb{N}}$ *satisfies a FLDP on* $(\tilde{\Sigma}, \mathscr{B}_{\tilde{\Sigma}})$ *with good rate function:*

$$J := \begin{cases} \tilde{\Sigma} & \to [0,\infty] \\ y & \to H\left(f^{-1}(\{y\})\right), \end{cases}$$

with the convention $\inf \emptyset = +\infty$.

Proof. Let $\Gamma \subset \tilde{\Sigma}$, and remark that $\left(P_N \circ f^{-1}\right)(\Gamma) = \int_\Sigma \mathbb{1}_\Gamma(f(x)) dP_N(x) = P_N\left(f^{-1}(\Gamma)\right)$.

As the property of being open or closed are preserved by f^{-1}, and as $\Gamma \to \overset{\circ}{\Gamma}$, and $\bar{\Gamma}$ commute with $\Gamma \to f^{-1}(\Gamma)$, we easily obtain the upper and lower bounds building on the FLDP of $(P_N)_N$:

$$-H\left(f^{-1}(\overset{\circ}{\Gamma})\right) \leq \liminf_N \frac{1}{N} \log\left(P_N \circ f^{-1}\right)(\Gamma) \leq \limsup_N \frac{1}{N} \log\left(P_N \circ f^{-1}\right)(\Gamma) \leq -H\left(f^{-1}(\bar{\Gamma})\right).$$

We thus only need to prove that J is a good rate function. It is clear that it has value in $[0,\infty]$. Let $x \in \Sigma$ such that $H(x) = 0$. Then, $J(f(x)) \leq H(x) = 0$, so that J is not uniformly infinite. Moreover, let $K_a^J := \{J \leq a\}$ for $a \in \mathbb{R}$, and define K_a^H similarly. It is clear that $f(K_a^H) \subset K_a^J$. Let then $y \in K_a^J$, and remark that the closed set $f^{-1}(\{y\}) \neq \emptyset$ so that H attains its minimum on it. Let x be such a point. Then $J(y) = H(x) \leq a$, so that $x \in K_a^H$, and $f(K_a^H) = K_a^J$ is a compact set. □

Remark 27. We need to work more when dealing with WLDP as the reciprocal image of compact set by a continuous function is not necessarily compact, and we need the goodness of H to conclude on the lower semi-continuity of J.

The following theorem, known as Varadhan's Lemma, provides a control on exponential integrals.

Theorem 9.6.2. *Varadhan's lemma*
Let $(P_N)_{N\in\mathbb{N}}$ satisfies a FLDP with good rate function H on $\big(\Sigma,\mathscr{B}(\Sigma)\big)$. Let $f:\Sigma\to\mathbb{R}$. Then

(i) *if f is lower semi-continuous, one has for every open set $O\in\mathscr{B}(\Sigma)$*

$$-\inf_O\big\{H-f\big\}\le \liminf\frac{1}{N}\log\left\{\int_O e^{Nf(x)}dP_N(x)\right\}$$

(ii) *if f is upper semi-continuous and bounded from above, one has for every closed set $F\in\mathscr{B}(\Sigma)$*

$$\limsup\frac{1}{N}\log\left\{\int_F e^{Nf(x)}dP_N(x)\right\}\le -\inf_F\big\{H-f\big\}.$$

(iii) *if f is continuous and bounded from above, then*

$$\lim_N\frac{1}{N}\log\left\{\int_\Sigma e^{Nf(x)}dP_N(x)\right\}=\sup_\Sigma\big\{f-H\big\}.$$

Proof. Remark first that 3. is a consequence of 1. and 2. applied to the open and closed set Σ. Define, for any $\Gamma\in\mathscr{B}(\Sigma)$, the quantity

$$Z_N^f(\Gamma):=\int_\Gamma e^{Nf(x)}dP_N(x).$$

1. Let $O\in\mathscr{B}(\Sigma)$ be an open set, $x\in O$, and $\varepsilon>0$. As f is lower semi-continuous, exists $\delta>0$ such that $B(x,\delta)\subset O$ and $\inf_{y\in B(x,\delta)}f(y)\ge f(x)-\varepsilon$. Then,

$$Z_N^f(O)\ge Z_N^f(B(x,\delta))\ge e^{N(f(x)-\varepsilon)}P_N(B(x,\delta)).$$

As $(P_N)_N$ satisfies a FLDP, the latter implies that

$$\liminf\frac{1}{N}\log Z_N^f(O)\ge f(x)-\varepsilon-H\big(B(x,\delta)\big)\ge(f-H)(x)-\varepsilon,$$

which proves the result.

2. Let $F\in\mathscr{B}(\Sigma)$ be a closed set, $\varepsilon>0$ and define $F_a:=F\cap\{H\le a\}$ for a given $a\in\mathbb{R}$. As f is upper semi-continuous, and H is lower semi-continuous, for any $x\in\mathscr{F}_a$ exists $\delta_x>0$ such that $\inf_{y\in B(x,2\delta_x)}H\ge H(x)-\varepsilon$ and $\sup_{y\in B(x,2\delta_x)}f\le f(x)+\varepsilon$. Moreover, as $F_a:=F\cap\{H\le a\}$ is compact, it can be covered by a finite union of open balls. Let then $K\in\mathbb{N}^*$, $x_1,\dots,x_K\in F_a$, $\delta_{x_1},\dots,\delta_{x_K}\in\mathbb{R}_+^*$ be such that $F_a\subset O=\bigcup_{k=1}^K B(x_k,\delta_{x_k})$. As $F\subset O\cup(F\cap O^c)$, we have

$$\begin{aligned}Z_N^f(F)&\le Z_N^f(F\cap G^c)+\sum_{k=1}^K Z_N^f\Big(\overline{B(x_k,\delta_{x_k})}\Big)\\&\le e^{NM}P_N(F\cap G^c)+\sum_{k=1}^K e^{N(f(x_k)+\varepsilon)}P_N\Big(\overline{B(x_k,\delta_{x_k})}\Big),\end{aligned}$$

where $M \in \mathbb{R}$ is such that $f(x) \leq M$, $\forall x \in \Sigma$. As (P_N) follows a FLDP, we then have using Lemma 9.2.4

$$\begin{aligned}\limsup \frac{1}{N} \log\left(Z_N^f(F)\right) &\leq \Big(M - H(F \cap G^c)\Big) \vee \max_{1\leq k\leq K} \Big(f(x_k) + \varepsilon - H\Big(\overline{B(x_k, \delta_{x_k})}\Big)\\ &\leq \Big(M - H(F \cap G^c)\Big) \vee \max_{1\leq k\leq K} \Big(f(x_k) - H(x_k) + 2\varepsilon\Big).\end{aligned}$$

Moreover, as $F_a \subset G$, $H(F \cap G^c) \geq H(F \cap F_a^c) \geq a$, so that

$$\limsup \frac{1}{N} \log\left(Z_N^f(F)\right) \leq \big(M - a\big) \vee \sup_{y\in F_a} \Big(f(x) - H(x) + 2\varepsilon\Big).$$

Sending $\varepsilon \to 0$ and $a \to \infty$ concludes the proof. □

Remark 28. As stated in [86, Lemma 2.1.8] the above condition of boundedness for cases 2.-3. can readily be replaced by:

$$\lim_{M\to\infty} \limsup_{N\to\infty} \frac{1}{N} \log\left(\int_{f\geq M} e^{Nf(x)} dP_N(x)\right) = -\infty. \tag{9.13}$$

Indeed, working with $f_M = f \wedge M$, and remarking that under this condition, for any closed set F

$$\limsup_N \frac{1}{N} \log\left(\int_F e^{Nf(x)} dP_N(x)\right) = \limsup_N \frac{1}{N} \log\left(\int_F e^{Nf_M(x)} dP_N(x)\right) \leq \sup_F \big\{f_M - H\big\} \leq \sup_F \big\{f - H\big\}.$$

As the contraction principle, Varadhan's lemma furnishes a way of deriving a large deviation principle out of another one. It will only apply to "tilted" transformation of probability sequence $(P_N)_N$ satisfying a LDP, i.e. probability measures absolutely continuous with respect to $(P_N)_N$, with exponential Radon-Nykodym density. This corollary finds many applications in statistical physics as the Gibbs measure naturally appears as a "tilted" probability sequence of the independent system. It will also be of chief importance in the manuscript, as it furnishes the main intuition to our analysis on particle systems.

Theorem 9.6.3. *Let $(P_N)_{N\in\mathbb{N}}$ be a sequence of measure on $(\Sigma, \mathscr{B}(\Sigma))$ following a FLDP with good rate function H, and f be continuous and bounded from above. Denote $Z_N^f := Z_N^f(\Sigma)$. Then, the sequence of "tilted" probability measures*

$$Q_N(\Gamma) := \frac{Z_N^f(\Gamma)}{Z_N^f}, \ \forall \Gamma \in \mathscr{B}(\Sigma)$$

follows a FLDP with good rate function $H_Q(x) := sup_\Sigma\big\{f - H\big\} - \big(f - H\big)(x)$.

Proof. Let $M \in \mathbb{R}$ be such that $f(x) \leq M$, $\forall x \in \Sigma$. Then, for every $\Gamma \in \mathscr{B}(\Sigma)$, $0 \leq Z_N^f(\Gamma) \leq e^{NM}$, with strict positivity as soon as Γ is of non-empty interior. Hence, $Q_N(\Gamma)$ is well-defined. Moreover, Q_N is σ-additive by the monotone convergence theorem, so that it is a probability measure.

Let us show that H_Q is indeed a good rate function. First H_Q is positive, as $\sup_\Sigma\big\{f - H\big\} \geq f(x) - H(x)$, $\forall x \in \Sigma$. Moreover, as f is continuous, then, $H + (-f)$ is the sum of two lower semi-continuous functions so that H_Q is also lower semi-continuous. Moreover, as H is a good rate function, exists a $x \in \Sigma$ such that $H(x) < \infty$,

so that $\infty > M \geq \sup_\Sigma \{f-H\} \geq f(x)-H(x) > -\infty$ and $H_Q(x) < \infty$. Furthermore, for any $a \in \mathbb{R}$

$$\{H_Q \leq a\} = \Big\{H - f \leq a - \sup_\Sigma \{f-H\}\Big\} \subset \Big\{H \leq a + M - \sup_\Sigma \{f-H\}\Big\}.$$

Hence $\{H_Q \leq a\}$ is a closed set contain in a compact set and is thus compact. We now establish the lower and upper-bounds. Remark that, for any $\Gamma \in \mathscr{B}(\Sigma)$ we have

$$\frac{1}{N}\log Q_N(\Gamma) = \frac{1}{N}\log\Big(\int_\Gamma e^{Nf(x)} dP_N(x)\Big) - \frac{1}{N}\log(Z_N^f).$$

On the one hand, Σ is both open and closed, so that

$$-\inf_\Sigma\{H-f\} \leq \liminf \frac{1}{N}\log(Z_N^f) \leq \limsup \frac{1}{N}\log(Z_N^f) \leq -\inf_\Sigma\{H-f\},$$

Hence, $\lim_N \frac{1}{N}\log(Z_N^f) = sup_\Sigma\{f-H\}$. On the other hand

$$-\inf_{\overset{\circ}{\Gamma}}\{H-f\} \leq \liminf \frac{1}{N}\log\Big(\int_\Gamma e^{Nf(x)} dP_N(x)\Big) \leq \limsup \frac{1}{N}\log\Big(\int_\Gamma e^{Nf(x)} dP_N(x)\Big) \leq -\inf_{\bar{\Gamma}}\{H-f\}.$$

Subtracting these two equations yields

$$-\inf_{\overset{\circ}{\Gamma}} H_Q \leq \liminf \frac{1}{N}\log Q_N(\Gamma) \leq \limsup \frac{1}{N}\log Q_N(\Gamma) \leq -\inf_{\bar{\Gamma}} H_Q.$$

□

Remark 29. In the proof of Cramér's Theorem, we have precisely used a "tilted" version of our original i.i.d. sequence $(X_i)_{i\in\mathbb{N}^*}$. The function we used was $f : x \to \tau x$, where $\tau \in \mathbb{R}$ is such that $\phi(\tau) = \inf\phi = -\rho = -H(0)$ in order to conclude. In fact, in this case, $P_N = \mathscr{L}\big(\frac{S_N}{N}\big)$ so that

$$Z_N^\tau = \frac{1}{N}\log\mathbb{E}\Big[e^{N\tau\frac{S_N}{N}}\Big] = e^{\phi(\tau)} = e^{-N\rho}.$$

Hence

$$dQ_N(x) := \frac{e^{N\tau x}}{Z_N^\tau} dP_N(x)$$

which precisely corresponds to the law of a sum of N independent copies of variable Y of law with the repartition function

$$F_Y(y) = e^\rho \int_{-\infty}^y e^{\tau x} dF_X(x).$$

As appear in the proof, the goal of introducing such a tilted measure it to rely on the CLT arguments. In fact, we know that we can use a CLT at the minimum of the good rate function (see remark (9.4)), and Varadhan's lemma ensures that

$$H_Q(0) = \phi(\tau) - \tau\times 0 + H(0) = \phi(\tau) + \rho = 0.$$

Varadhan's Lemma (for continuous bounded functions) admits a reciprocal theorem (proof can be found e.g. in [84, Section 4.4.]):

Theorem 9.6.4 (Bryc's Lemma). *Let $(P_N)_N$ be an exponentially tight family of probability measures on $\big(\Sigma, \mathscr{B}(\Sigma)\big)$ such that, for any $f \in \mathscr{C}_b(\Sigma)$, the quantity*

$$\Lambda(f) := \lim_N \frac{1}{N} \log\Big(\int_\Sigma e^{Nf(x)} dP_N(x) \Big)$$

exists in $\mathbb{R}$. Then $(P_N)_N$ satisfies a FLDP with good rate function

$$H(x) = \sup_{f \in \mathscr{C}_b(\Sigma)} \big\{ f(x) - \Lambda(f) \big\}.$$

Moreover, for any $f \in \mathscr{C}_b(\Sigma)$ we have $\Lambda(f) = \sup_{x \in \Sigma} \big[f(x) - H(x) \big]$.

Bryc's theorem can be seen as an infinite dimensional version of Gartnër-Ellis theorem (see [185]). Remark that it proves an interesting Fenchel duality in a non-convex setting.

9.6.1 Remarks on Legendre transform and convex rate functions

Because of the ubiquity of the Legendre transform in the theory, it seems important to fully understand in what extent good rate functions are all Legendre transform of the logarithmic moment generating function. In this short section, inspired from [86, Section 2.2.], we will gather our knowledge on the theory to answer this question.

In order to encompass all the settings we have encountered so far, we will work, in this section, on X, a locally convex Hausdorff real topological vector space. We will denote its real topological dual space by X^*, and the associated duality bracket by ${}_{X^*}\langle \cdot, \cdot \rangle_X$. Moreover, as done in we will let $E \subset X$ be a closed convex subset, Polish with respect to the inherited topology, and consider sequence of probability measures $(P_N)_N$ lying in $\mathscr{M}_1^+(\Sigma)$. In this theoretical framework, we will cite, without proof, a few clarifying results from [86, Section 2.2.].

Here are the canonical examples one should have in mind:

Example. (i) $X = E = \mathbb{R}^d$, $X^* = \mathbb{R}^d$, ${}_{X^*}\langle \cdot, \cdot \rangle_X = \langle \cdot, \cdot \rangle_{\mathbb{R}^d}$ is the Euclidean scalar product on $\mathbb{R}^d$. $d \in \mathbb{N}^*$. This encompasses Cramér's and Gartnër-Ellis theorems.

(ii) $X = \mathscr{M}^+(\Sigma)$, $E = \mathscr{M}_1^+(\Sigma)$, $X^* = \mathscr{C}_b(\Sigma)$, and ${}_{X^*}\langle \phi, \mu \rangle_X := \int_\Sigma \phi(x) d\mu(x)$, where Σ is Polish space. This setting accounts for Sanov's theorem.

We first extend the definition of Legendre transform to real topological Hausdorff spaces:

Definition 9.6.1. Let Λ be a convex map from X^* to $\bar{R}$. Then, its Legendre transform is the map from X to $\bar{\mathbb{R}}$ defined by:

$$\Lambda^*(x) := \sup_{t \in X^*} \big\{ {}_{X^*}\langle t, x \rangle_X - \Lambda(t) \big\}, \forall x \in X.$$

We will denote Λ_E^* its restriction to the closed convex set E

Moreover, we introduce the logarithmic moment generating function of probability measure $\mu \in \mathscr{M}_1^+(X)$:

$$\Lambda_\mu(t) := \log\left(\int_X e^{{}_{X^*}\langle t, x \rangle_X} d\mu(x) \right) \in [-\infty, \infty], \forall t \in X^*. \tag{9.14}$$

With the many insights given by the theory on how to find the rate function, let us try to understand why and when it corresponds to a Legendre transform. First, stepping back on what we have done in the above, it appears that an upper-bound for closed sets involving the Legendre transform of the logarithmic moment generating function (when it exists), naturally arises from an exponential Chebyshev's inequality. Moreover, in both Cramér's and Sanov's theorems, the good rate function of the FLDP is precisely given by this Legendre transform. Because of the inconvenient dependency in N, Gartnër-Ellis theorems, suggests another version of asymptotic logarithmic moment generating function:

$$\forall t \in X^*, \quad \Lambda(t) := \lim_N \frac{1}{N} \Lambda_{P_N}(Nt),$$

that condenses the asymptotic Laplace moments of the sequence $(P_N)_N$, should they exist in $\bar{\mathbb{R}}$.

This intuition is also compatible with both Cramér's and Sanov's theorem.

Let $\mathbf{X} := (X_i)_{i \in \mathbb{N}^*}$ be a sequence of independent random variables of law P. In the first case $P \in \mathscr{M}_1^+(\mathbb{R})$, $P_N = \mathscr{L}(\frac{S_N}{N})$ and provided the existence of exponential moments, we find

$$\frac{1}{N} \Lambda_{P_N}(Nt) = \frac{1}{N} \log \mathbb{E}\left[e^{Nt \frac{S_N}{N}} \right] = \Lambda_P(t), \forall t \in \mathbb{R}.$$

In the second case, $P \in \mathscr{M}_1^+(\Sigma)$, Σ being a Polish space, $P_N = P^{\otimes N}\left(\hat{\mu}_N^{\mathbf{X}} \in \cdot\right) \in \mathscr{M}_1^+\left(\mathscr{M}_1^+(\Sigma)\right)$, and for any $\phi \in \mathscr{C}_b(\Sigma)$,

$$\begin{aligned} \frac{1}{N} \Lambda_{P_N}(N\phi) &= \frac{1}{N} \log \left(\int_{\mathscr{M}_1^+(\Sigma)} e^{N \int_\Sigma \phi(x) d\mu(x)} dP_N(\mu) \right) = \frac{1}{N} \log \mathbb{E}\left[e^{N \int_\Sigma \phi(x) d\hat{\mu}_N^{\mathbf{X}}(x)} \right] \\ &= \frac{1}{N} \log \mathbb{E}\left[e^{\sum_{i=1}^N \phi(X_i)} \right] = \Lambda_{P(\delta_X \in \cdot)}(\phi). \end{aligned}$$

Hence, in these two independent cases, the sequence $\left(\frac{1}{N} \Lambda_{P_N}(Nt)\right)_N$ are stationary, and equal the logarithmic moment generating function originally introduced in both theorems.

We now state some general results applying to the topological framework introduced in the section.

Theorem 9.6.5. *Let $(P_N)_N \in \mathscr{M}_1^+(X)^{\mathbb{N}^*}$, and assume that*

$$\Lambda(t) := \lim_N \frac{1}{N} \Lambda_{P_N}(Nt) \in [-\infty, \infty] \tag{9.15}$$

exists for every $t \in X^*$. Then, Λ is convex on X^*. Moreover, its Legendre transform Λ^* (in the sense of (9.6.1)*) satisfies the upper-bound for closed sets:*

$$\forall \Gamma \in \mathscr{B}(E), \quad \limsup \frac{1}{N} \log P_N(\Gamma) \leq -\inf_{\bar{\Gamma}} \left\{ \Lambda_E^* \right\}.$$

Proof. See [86, Section 2.2.]. □

This theorem only partially confirms our intuition. Its proof, fully performed in [86], relies classically on an exponential Chebyshev's inequality and optimization in order to obtain an upper-bound for closed sets (see (21)). Nevertheless, supposing that $(P_N)_N$ satisfies a FLDP with good rate function H, this upper-bound dominated by Λ_E^* only provides an inequality: $\Lambda_E^* \leq H$ (see Proposition (9.4.1)).

The lower-bound is more difficult to obtain, as illustrated by the demonstration of Cramér's Theorem. If it holds, the rate function must then be the Legendre transform: $\Lambda_E^* = H$, making H a convex function.

Nevertheless, the rate function has no particular reason for being convex!

While convex rate functions naturally arise for the large deviations of i.i.d. sequences of random variables (see [86, Section 3.] and [85, Section III.7]), and is preserved by the contraction principle for linear transform, dependent sequences of random variables generally produce non-convex rate functions. The contraction principle 9.6.1 and Theorem (9.6.3) furnish an opportune way of building counterexamples: we can easily pick an $f \in \mathscr{C}_b(\Sigma)$ in order for the rate functions $J(y) := H(f^{-1}(\{y\}))$ and $H_Q(x) := \sup_\Sigma \{f - H\} - (f(x) - H(x))$ to fail to be convex. Provided that $(P_N)_N$ satisfies a FLDP with a good rate function H that is convex, the following lemma ensures that a tail condition brings the identity $H = \Lambda_E^*$.

Lemma 9.6.6. Let $(P_N)_N$ satisfy a FLDP with convex good rate function H on a Polish space endowed with its Borel σ-field $\big(X, \mathscr{B}(X)\big)$, and suppose that

$$\lim_{M\to\infty} \limsup_{N\to\infty} \frac{1}{N} \log\left(\int_{\{X^* \langle t,x\rangle_X \geq M\}} e^{X^* \langle Nt,x\rangle_X} dP_N(x) \right) = -\infty, \forall t \in X^*.$$

Then the limit function $\Lambda(t)$ of (9.15) exists $\forall t \in X^*$, and satisfies $H = \Lambda_E^*$.

Proof. Let us first extend H to the whole X by setting it to $+\infty$ outside E. We will denote H_X this extension, which is also convex. We apply 3. of Varadhan's Lemma with the extended condition of remark 28 to the function $f_t : x \to_{\Sigma^*} \langle t,x\rangle_\Sigma$ to obtain that $\Lambda(t) = \sup_{x\in\Sigma} \{f_t(x) - H(x)\} = \sup_{x\in\Sigma} \{f_t(x) - H_X(x)\} = H_X^*(t)$. As H_X satisfies the hypothesis of Theorem (9.2.3), it follows the Fenchel-Legendre duality equation: $H_X = H_X^{**} = \Lambda^*$, so that the proof is completed. □

The case of Varadhan's lemma and Bryc's theorem is slightly more subtle.

In order to dwell in our topological framework, we would like to work with probability measures on $\mathscr{M}_1^+(\Sigma)$ rather than Σ.

A natural way to do so is to define $\hat{P}_N := P_N\big(\delta_x \in \cdot\big) = P_N\big(\{x \in \Sigma, \delta_x \in \cdot\}\big)$. Interestingly, for any $\phi \in \mathscr{C}_b(\Sigma)$:

$$\begin{aligned}\Lambda_{\hat{P}_N}(N\phi) &= \log\left(\int_{\mathscr{M}_1^+(\Sigma)} e^{N\int_\Sigma \phi(y) d\mu(y)} d\hat{P}_N(\mu) \right) = \log\left(\int_\Sigma e^{N\int_\Sigma \phi(y) d\delta_x(y)} dP_N(x) \right) \\ &= \log\left(\int_\Sigma e^{N\phi(x)} dP_N(x) \right),\end{aligned}$$

which is precisely the kind of integrals found in Varadhan's lemma. The following promising lemma ensures that this comparison makes some sense. It is also an illustration that good rate functions are not always Legendre transform.

Lemma 9.6.7. Let $(P_N)_N$ be a sequence of probability measure on a Polish space Σ, and define $\hat{P}_N := P_N(\delta_x \in \cdot) \in \mathscr{M}_1^+\big(\mathscr{M}_1^+(\Sigma)\big)$. Then, we have the equivalence between the two following assertions:

(i) $(P_N)_N$ satisfies a FLDP on $\big(\Sigma, \mathscr{B}(\Sigma)\big)$,

(ii) $(\hat{P}_N)_N$ satisfies a FLDP on $\Big(\mathscr{M}_1^+(\Sigma), \mathscr{B}\big(\mathscr{M}_1^+(\Sigma)\big)\Big)$.

If these hold, the respective good rate functions H and $\hat{H}$ satisfy

$$\forall \mu \in \mathscr{M}_1^+(\Sigma), \qquad \hat{H}(\mu) = \begin{cases} H(x) \text{ if exists } x \in \Sigma, \text{ such that } \mu = \delta_x, \\ +\infty, \text{ if not.} \end{cases} \tag{9.16}$$

Moreover, if $(\hat{P}_N)_N$ is such that the limits

$$\Lambda(\phi) := \lim_N \frac{1}{N} \Lambda_{\hat{P}_N}(N\phi) = \lim_N \frac{1}{N} \log \left(\int_\Sigma e^{N\phi(x)} dP_N(x) \right) \in [\min_\Sigma \phi, \max_\Sigma \phi],$$

exists in $\mathbb{R}$, then the Legendre transform $\Lambda^* : \mathscr{M}^+(\Sigma) \to \bar{\mathbb{R}}$ exists and coincide with $\hat{H}$ on Dirac masses $\delta_x, x \in \Sigma$. Furthermore, we have the equivalence

$$\hat{H} = \Lambda^*_{\mathscr{M}_1^+(\Sigma)} \iff \hat{H} \text{ is convex} \iff \exists x \in \Sigma,\ H^{-1}(\mathbb{R}) = H^{-1}(\{0\}) = \{x\}.$$

Remark 30. (i) We can show (see 9.7.2) that the map Λ^* is uniformly $+\infty$ on $\mathscr{M}^+(\Sigma) \backslash \mathscr{M}_1^+(\Sigma)$, and that, when H is sufficiently regular, we have for any $\mu \in \mathscr{M}_1^+(\Sigma)$

$$\Lambda^*(\mu) = \int_\Sigma H(x) d\mu(x).$$

Moreover, supposing that H is a Legendre transform of some kind (as suggested by Bryc's theorem) non infinite at at least two points, $\hat{H}$ would loose this good property.

(ii) Suppose that $(P_N)_N$ satisfies a FLDP with good rate function H, so that $(\hat{P}_N)_N$ satisfies a FLDP with good rate function $\hat{H}$. Then, if $\hat{H}$ is not a Legendre transform, there is no point iterating the process. In fact, let $\widehat{\hat{P}}_N := \hat{P}_N(\delta_x \in \cdot) = P_N\big(\{x \in \Sigma, \delta_{\delta_x} \in \cdot\}\big) \in \mathscr{M}_1^+\Big(\mathscr{M}_1^+\big(\mathscr{M}_1^+(\Sigma)\big)\Big)$, that satisfies a FLDP with good rate function $\widehat{\hat{H}}$. Observe that $\hat{H}$ is a Legendre transform if and only if $\widehat{\hat{H}}$ is, that is when $H^{-1}(\mathbb{R})$ admits only one element.

The relevance of this lemma, striving to include Varadhan's setting into the framework of locally convex Hausdorff topological vector spaces, is questionable. Firstly, its statement hide a powerlessness, as the condition for $\hat{H}$ to be convex is extremely restrictive: $H^{-1}(\mathbb{R}) = H^{-1}(\{0\}) = \{x\}$ for some $x \in \Sigma$! In this very particular hardly representative case, their is no need to make such efforts as to consider $(\hat{P}_N)_N$. We would better choose the closed convex set E to be equal to $\{\delta_x\}$. Provided the existence of the logarithmic moment generating function Λ, the inequality $0 = H \geq \Lambda_E^*$ turns into an equality $H = \Lambda_E^*$. But we know nothing of more general cases where $H^{-1}(\mathbb{R})$ is not a singleton. Moreover, supposing that H has good properties (for example is a non-trivial Legendre-transform), $\hat{H}$ would automatically fail to reproduce them.

Secondly, one striking characteristic of Varadhan's and Bryc's settings is that they deal with non-convex and non-linear functions on a Polish space Σ that has no reason to be a vector space. In this context, it seems hardly appropriate to plunge them into the framework of real Hausdorff topological vector spaces, or to even hope that the good rate function will be convex. One very interesting insights of Bryc's theorem nevertheless relates it with a form of Fenchel duality through the equalities:

$$\forall x \in \Sigma,\ H(x) = \sup_{\phi \in \mathscr{C}_b(\Sigma)} \big\{\phi(x) - \Lambda(\phi)\big\}, \qquad \forall \phi \in \mathscr{C}_b(\Sigma),\ \Lambda(\phi) = \sup_{x \in \Sigma} \big\{\phi(x) - H(x)\big\}.$$

This relation is investigated in the next section.

9.6.2 Non-convex Fenchel duality

In this section, X denotes a (complete) metric space. We first introduce a non-convex Fenchel duality for a class of functions from $X \to \bar{\mathbb{R}}$ that includes rate functions.

Definition 9.6.2. Let f be a function from X to $\mathbb{R} \cup \{+\infty\}$. Its conjugate is the function $f^\times$ from $\mathscr{C}_b(X)$ to $\bar{\mathbb{R}}$ defined by

$$f^\times(\phi) := \sup_{x \in X} \big\{ \phi(x) - f(x) \big\}$$

for any $\phi \in \mathscr{C}_b(X)$. The conjugate of any $\Lambda : \mathscr{C}_b(X) \to \bar{\mathbb{R}}$ is the function $\Lambda^\times : X \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\Lambda^\times(x) := \sup_{\phi \in \mathscr{C}_b(X)} \big\{ \phi(x) - \Lambda(\phi) \big\}.$$

Moreover, $f^{\times\times}$ is called the second conjugate of f.

Remark 31. It is also possible to restrict these definitions to any subspace $\mathscr{A}(X)$ of $\mathscr{C}_b(X)$ by considering only $\phi \in \mathscr{A}(X)$. The obtained conjugate functions of course depend on the choice of $\mathscr{A}(X)$.

Theorem 9.6.8. *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be bounded from below. Then*

(i) *$f^\times$ is a convex Lipschitz-continuous function of $\big(\mathscr{C}_b(\mathbb{R}), \|\cdot\|_\infty\big)$,*

(ii) *$f^{\times\times}$ is lower semi-continuous, bounded from below with $f^{\times\times} \leq f$,*

(iii) *$f^{\times\times} = f$ if and only if f is not uniformly infinite and lower semi-continuous.*

Proof. This corresponds to [20, Theorems 2.1-2] for the particular case $\mathscr{A}(X) = \mathscr{C}_b(X)$ □

Remark 32. (i) For any $f : X \to \mathbb{R} \cup \{+\infty\}$ lower semi-continuous, bounded from above and non uniformly infinite, [20, Theorems 2.8-9] also ensures that $f^\times$ satisfies a notion of differentiability on a G_δ dense $G \subset \mathscr{C}_b(X)$. More precisely $\forall \phi \in G, \exists x \in X$

$$\forall h \in \mathscr{C}_b(X), \lim_{t \to 0^+} \frac{1}{t} \big(f^\times(\phi + th) - f^\times(\phi) - {}_{C_b(X)}\langle h, \delta_x \rangle_{\mathscr{M}^+(X)} \big) = 0,$$

and

$$\phi(z) - f(z) = \sup_{y \in X} \big\{ \phi(y) - f(y) \big\} \iff z = x.$$

(ii) Similar results hold for a number of subspaces $\mathscr{A}(X) \subset \mathscr{C}_b(X)$. See [20] for further details.

In the case of a Polish space $\big(\Sigma, \mathscr{B}(\Sigma)\big)$ endowed with its Borel σ-field, theorem (9.6.8) ensures that virtually any rate function are some kind of non-convex Legendre transform:

Lemma 9.6.9. Let $(P_N)_N$ be a sequence of probability measures of $\big(\Sigma, \mathscr{B}(\Sigma)\big)$. Then, we have the following:

(i) Any rate function H satisfies:

$$\forall x \in \Sigma,\ H(x) = H^{\times\times}(x) = \sup_{\phi \in \mathscr{C}_b(\Sigma)} \left\{ \phi(x) - H^{\times}(\phi) \right\}.$$

(ii) If $(P_N)_N$ is exponentially tight and admits a logarithmic moment generating function:

$$\Lambda(\phi) := \lim_N \frac{1}{N} \Lambda_{P_N}(N\phi) := \lim_N \frac{1}{N} \log \left(\int_\Sigma e^{N\phi(x)} dP_N(x) \right) \in [\min_\Sigma \phi, \max_\Sigma \phi], \tag{9.17}$$

then $\Lambda^{\times\times} = \Lambda$ and $(P_N)_N$ satisfies a FLDP with good rate function $\Lambda^{\times}$.

Proof. These two assertions are easy consequences of Bryc's theorem along with Theorem (9.6.8). □

Remark 33. With this in mind and supposing (9.17) holds, the form of the inherited rate function H_Q for the sequence of "tilted" probability measures $(Q_N)_N$ in Varadhan's corollary, is elementary. In fact, let $f \in \mathscr{C}_b(\Sigma)$ and define the "tilted" probability measure on Σ:

$$dQ_N := \frac{e^{Nf}}{Z_N^f} dP_N.$$

We have, for any $\phi \in \mathscr{C}_b(\Sigma)$

$$\Lambda_{Q_N}(N\phi) = \log \left(\int_\Sigma e^{N(f(x)+\phi(x))} dP_N(x) \right) - \log \left(\int_\Sigma e^{Nf(x)} dP_N(x) \right),$$

so that the logarithmic moment generating function of Q_N also exists:

$$\Lambda_Q(\phi) := \lim_N \frac{1}{N} \Lambda_{Q_N}(N\phi) = \Lambda(f+\phi) - \Lambda(f),$$

and necessarily, $H_Q = \Lambda_Q^{\times}$. As $H^{\times} = \Lambda^{\times\times} = \Lambda$, remark that for any $x \in \Sigma$,

$$\begin{aligned} \Lambda_Q^{\times}(x) &= \sup_{\phi \in \mathscr{C}_b} \left\{ \phi(x) - \Lambda_Q(\phi) \right\} = \Lambda(f) - f(x) + \sup_{\phi \in \mathscr{C}_b} \left\{ (f+\phi)(x) - \Lambda(f+\phi) \right\} \\ &= \Lambda(f) + \Lambda^{\times}(x) - f(x) = \sup_{y \in \Sigma} \left\{ f(y) - H(y) \right\} - \left(f(x) - H(x) \right). \end{aligned}$$

9.7 Large Deviations Appendix

9.7.1 Sanov's theorem for finite sequences.

In this section we will a toy version of Sanov's Theorem in order to give the general idea underlying the theorem and highlight its links with information theory. It will apply to the very limited case of finite alphabet $\Sigma = \{a_1, a_2, \ldots, a_K\}, K \in \mathbb{N}$, for which the proof relies on combinatorial methods.

Let $\mathbf{X} := (X_1, \ldots, X_N, \ldots)$ be a sequence of Σ-valued i.i.d. random variables with law $P \in \mathscr{M}_1^+(\Sigma)$. When $\Sigma = \{a_1, a_2, \ldots, a_K\}$ is a finite alphabet, one can identify $\mathscr{M}_1^+(\Sigma)$ with the fragment of the hyperplane of $\mathbb{R}^K$ $\{y \in (\mathbb{R}_+)^K, \Sigma_{i=1}^K y_i = 1\}$. Define the support of P by $\Sigma_P := \{a_i : P(a_i) > 0\} \subset \Sigma$.

Definition 9.7.1 (Empirical measure). Let $\mathbf{y}$ be an element of Σ^d, $d \in \bar{\mathbb{N}}^*$. Its sequence of empirical measures $\left(\hat{\mu}_N\right)_{N\leq d} \in \mathscr{M}_1^+(\Sigma)^d$ is defined by

$$\hat{\mu}_N^{\mathbf{y}} := \frac{1}{N}\sum_{i=1}^{N} \delta_{y_i},$$

where $\delta_x \in \mathscr{M}_1^+(\Sigma)$ is the degenerate measure at $x \in \Sigma$.

Let $\mathscr{L}_N := \left\{\hat{\mu}_N^{\mathbf{y}},\ \mathbf{y} \in \Sigma^{\mathbb{N}}\right\}$ denote the set of all possible empirical measure sequence of length N, so that $\hat{\mu}_N^{\mathbf{X}}$ is a random element of $\mathscr{L}_N$.

Lemma 9.7.1. (i) $Card\left(\mathscr{L}_N\right) \leq (N+1)^K$,

(ii) $\forall \mu \in \mathscr{M}_1^+(\Sigma)$,

$$d_V(\mu, \mathscr{L}_N) := \inf_{\nu\in\mathscr{L}_N} d_V(\mu,\nu) \leq \frac{K}{2N},$$

where $d_V(\mu,\nu) := \sup_{A\subset\Sigma}$ is the variational distance between the measure μ and ν.

Proof. 1. It suffices to see that, given an element μ of $\mathscr{L}_N$ and a $a_i \in \Sigma$, the quantity $\mu(a_i)$ has at most $N+1$ possible values: $\frac{k}{N}$, $k \in [\![0,N]\!]$.
2. On a finite alphabet, we can get rid of the supremum in the expression of the variational distance

$$\forall \mu,\nu \in \mathscr{M}_1^+(\Sigma), \quad d_V(\nu,\mu) = \frac{1}{2}\sum_{i=1}^{K} |\mu(a_i) - \nu(a_i)|.$$

In fact, as the subsets of Σ are of finite number, the supremum is attained. Let $A \subset \Sigma$ in which it is attained. Then, for any $a_i \in A$, $\mu(a_i) \geq \nu(a_i)$, whereas for any $a_i \notin A$, $\mu(a_i) \leq \nu(a_i)$ by maximality. Remark now that

$$0 = \sum_{i=1}^{K} \mu(a_i) - \nu(a_i) = d_V(\mu,\nu) + \sum_{i:a_i\notin A} \mu(a_i) - \nu(a_i),$$

so that

$$d_V(\mu,\nu) = \sum_{i:a_i\notin A} |\mu(a_i) - \nu(a_i)| = \sum_{i:a_i\in A} |\mu(a_i) - \nu(a_i)|.$$

Fix then $\mu \in \mathscr{M}_1^+(\Sigma)$, and consider for any $i \in [\![1,K]\!]$ the quantity $p_i = \frac{1}{N}\lfloor N\mu(a_i)\rfloor \geq \mu(a_i) < p_i + \frac{1}{N}$. Let $k \in [\![0,N]\!]$ such that $\sum_{i=1}^K p_i = \frac{k}{N}$. Then, exists at least $N-k$ distinct i such that $p_i < \mu(a_i)$. For $N-k$ such i define $\tilde{p}_i = p_i + \frac{1}{N}$ and set $\tilde{p}_i = p_i$ for the others i. Then $\nu \in \mathscr{L}_N$ define by $\nu(a_i) = \tilde{p}_i$ satisfies our condition. □

Definition 9.7.2. Let $\mu \in \mathscr{L}_N$. Then its type class is defined by $T_N(\mu) := \left\{(y_1,\ldots,y_N),\ \hat{\mu}_N^{\mathbf{y}} = \mu,\ \mathbf{y} \in \Sigma^{\mathbb{N}}\right\} \subset \Sigma^N$.

We now define a central objects that will play the role of the good rate function in Sanov's Theorem:

Definition 9.7.3 (Entropy, relative Entropy). Let $\Sigma = \left\{a_1, a_2, \ldots, a_K\right\}$ be a finite set, and $\mu,\nu \in \mathscr{M}_1^+(\Sigma)$. Then, the entropy of a μ is defined by

$$I(\mu) := -\sum_{i=1}^{K} \mu(a_i)\log\mu(a_i).$$

Moreover, the relative entropy of ν with respect to μ is defined by

$$I(\nu|\mu) := \sum_{i=1}^{K} \nu(a_i) \log \frac{\nu(a_i)}{\mu(a_i)}.$$

A few insights of these notions are given in the next section.

Proposition 9.7.2. , Let $\mu, P \in \mathscr{M}_1^+(\Sigma)$, and define $D_{I(\cdot|P)} := \{\mu \in \mathscr{M}_1^+(\Sigma), I(\mu|P) < \infty\}$. Then

(i) $\mu \to I(\mu|P)$ is convex and lower semi-continuous,

(ii) $I(\mu|P) \geq 0$ with equality iff $\mu = P$.

(iii) $I(\mu|P) = \infty \iff \mu \not\ll P$,

(iv) $\mu \to I(\mu|P)$ is a good rate function.

Proof. If $\Sigma_\mu \not\subset \Sigma_P$, then exists $i \in [\![1,K]\!]$ such that $\mu(a_i) > 0 = P(a_i)$. This implies $\mu(a_i) \log \frac{\mu(a_i)}{P(a_i)} = +\infty$, which proves the third point. To prove the first point, first remark that $\mu \to I(\mu|P)$ is obviously continuous on $D_{I(\cdot|P)}$. Suppose that $\mu_N \to \mu$, which implies in particular $\mu_N(a_i) \to \mu(a_i)$ for every $i \in [\![1,K]\!]$. If $\Sigma_\mu \not\subset \Sigma_P$, this implies $\mu_N \notin D_{I(\cdot|P)}$ for N large enough, so that $I(\mu_N|P) \to I(\mu|P) = \infty$. Moreover, if $\Sigma_\mu \subset \Sigma_P$ either exists infinitely many N such that $\mu_N \in D_{I(\cdot|P)}$ so that lower semi-continuity hold by continuity of $I(\cdot|P)$ on $D_{I(\cdot|P)}$, either such Ns are finite, in which case lower semi-continuity trivially holds as $\liminf I(\mu_N|P) = +\infty$. Furthermore, strict convexity of $x \to x\log(x)$ yields strict convexity of $I(\cdot|P)$ on $D_{I(\cdot|P)}$, so that 1. is proved. We now suppose that $\Sigma_\mu \subset \Sigma_P$, that is $\mu \in D_{I(\cdot|P)}$. Jensen inequality yields

$$I(\mu|P) = \sum_{i=1}^{K} \frac{\nu(a_i)}{\mu(a_i)} \log \frac{\nu(a_i)}{\mu(a_i)} \mu(a_i) \geq \Big(\sum_{i=1}^{K} \frac{\nu(a_i)}{\mu(a_i)} \mu(a_i) \Big) \log \Big(\sum_{i=1}^{K} \frac{\nu(a_i)}{\mu(a_i)} \mu(a_i) \Big) = 0.$$

Moreover, if $\mu \neq P$, then

$$0 \leq I\big(\frac{1}{2}(\mu + P)|P\big) < \frac{1}{2} I(\mu|P)$$

which proves the second point. To prove the fourth point, it suffices to remark that for any $a > 0$, the set $K_a := \{\mu, I(\mu|P) \leq a\} \subset D_{I(\cdot|P)}$ is closed in finite dimension. □

Lemma 9.7.3. (i) Let $\mu \in \mathscr{L}_N$, and $\mathbf{y}_N := (y_1, \ldots, y_N) \in T_N(\mu)$. Then

$$\mathbb{P}\Big(\mathbf{X}_N = \mathbf{y}_N\Big) = e^{-N\left(I(\mu) + I(\mu|P)\right)}.$$

In particular, this probability does not depend on $\mathbf{y}_N \in T_N(\mu)$.

(ii) $(N+1)^{-K} e^{NI(\mu)} \leq Card\big(T_N(\mu)\big) \leq e^{NI(\mu)}$,

(iii) $(N+1)^{-K} e^{-NI(\mu|P)} \leq \mathbb{P}\big(\hat{\mu}_N^{\mathbf{X}_N} = \mu\big) \leq e^{-NI(\mu|P)}$

Proof. 1. If $\mu \not\ll P$, both sides equal 0 and the result trivially holds. Suppose that $\mu \ll P$, that is $\Sigma_\mu \subset \Sigma_P$. Then, as a_i appears exactly $N\mu(a_i)$ times in $\mathbf{y}_N$,

$$\mathbb{P}\Big(\mathbf{X}_N = \mathbf{y}_N\Big) = \prod_{i=1}^{K} P(a_i)^{N\mu(a_i)}.$$

Furthermore,

$$I(\mu)+I(\mu|P)=-\sum_{i=1}^{K}\mu(a_i)\log P(a_i),$$

so that the result holds.
2. Let $Y_1,\dots,Y_N$ be an i.i.d. sequence of law μ. Then, using 1. and $H(\mu|\mu)=0$,

$$1\geq\mathbb{P}\left(\hat{\mu}_N^{\mathbf{Y}_N}=\mu\right)=\mathbb{P}\left(\mathbf{Y}_N\in T_N(\mu)\right)=\sum_{\mathbf{y}_N\in T_N(\mu)}\mathbb{P}\left(\mathbf{Y}_N=\mathbf{y}_N\right)=e^{-NH(\mu)}Card\left(T_N(\mu)\right),$$

so that the upper bound holds. Let $\nu\in\mathscr{L}_N$ such that $\Sigma_\nu\subset\Sigma_\mu$, and observe that $Card\left(T_N(\nu)\right)=\frac{N!}{\prod_{i=1}^{K}(N\nu(a_i))!}$. Moreover,

$$\frac{\mathbb{P}\left(\hat{\mu}_N^{\mathbf{Y}_N}=\mu\right)}{\mathbb{P}\left(\hat{\mu}_N^{\mathbf{Y}_N}=\nu\right)}=\frac{Card\left(T_N(\mu)\right)\prod_{i=1}^{K}\mu(a_i)^{N\mu(a_i)}}{Card\left(T_N(\nu)\right)\prod_{i=1}^{K}\mu(a_i)^{N\nu(a_i)}}=\prod_{i=1}^{K}\frac{(N\nu(a_i))!}{(N\mu(a_i))!}\mu(a_i)^{N\left(\mu(a_i)-\nu(a_i)\right)}.$$

As $\frac{m!}{l!}\geq l^{m-l}$ for any $(m,l)\in\mathbb{Z}_+^2$, we have

$$\frac{\mathbb{P}\left(\hat{\mu}_N^{\mathbf{Y}_N}=\mu\right)}{\mathbb{P}\left(\hat{\mu}_N^{\mathbf{Y}_N}=\nu\right)}\geq\prod_{i=1}^{K}N^{N\left(\nu(a_i)-\mu(a_i)\right)}=N^{N\sum_{i=1}^{K}\left(\nu(a_i)-\mu(a_i)\right)}=1.$$

Hence, for every $\nu\in\mathscr{L}_N$,

$$\mathbb{P}\left(\hat{\mu}_N^{\mathbf{Y}_N}=\mu\right)\geq\mathbb{P}\left(\hat{\mu}_N^{\mathbf{Y}_N}=\nu\right),$$

so that

$$1=\sum_{\nu\in\mathscr{L}_N}\mathbb{P}\left(\hat{\mu}_N^{\mathbf{Y}_N}=\nu\right)\leq Card\Big(\mathscr{L}_N\Big)\mathbb{P}\left(\hat{\mu}_N^{\mathbf{Y}_N}=\mu\right)\leq(N+1)^K e^{-NH(\mu)}Card\left(T_N(\mu)\right).$$

3. By 1.,

$$\mathbb{P}\left(\hat{\mu}_N^{\mathbf{X}_N}=\mu\right)=Card(T_N(\mu))e^{-N\left(I(\mu)+I(\mu|P)\right)},$$

so that the result is an immediate consequence of 2. . □

We are now able to state Sanov's Theorem:

Theorem 9.7.4 (Sanov's Theorem for finite Alphabet). *For every Borel set* Γ *of* $\mathscr{M}_1^+(\Sigma)$,

$$-\inf_{\mu\in\mathring{\Gamma}}I(\mu|P)\leq\liminf\frac{1}{N}\log\mathbb{P}\left(\hat{\mu}_N^{\mathbf{X}_N}\in\Gamma\right)\leq\limsup\frac{1}{N}\log\mathbb{P}\left(\hat{\mu}_N^{\mathbf{X}_N}\in\Gamma\right)\leq-\inf_{\mu\in\bar{\Gamma}}I(\mu|P).$$

Proof. Using the result of the previous lemma, we have

$$\begin{aligned}\mathbb{P}\left(\hat{\mu}_N^{\mathbf{X}_N}\in\Gamma\right)&=\sum_{\mu\in\mathscr{L}_N\cap\Gamma}\mathbb{P}\left(\hat{\mu}_N^{\mathbf{X}_N}=\mu\right)\leq\sum_{\mu\in\mathscr{L}_N\cap\Gamma}e^{-NI(\mu|P)}\\&\geq Card(\mathscr{L}_N\cap\Gamma)e^{-N\inf_{\mu\in\mathscr{L}_N\cap\Gamma}I(\mu|P)}\leq(N+1)^K e^{-N\inf_{\mu\in\mathscr{L}_N\cap\Gamma}I(\mu|P)}.\end{aligned}\tag{9.18}$$

Similarly,

$$\mathbb{P}\left(\hat{\mu}_N^{\mathbf{X}_N}\in\Gamma\right)\geq\sum_{\mu\in\mathscr{L}_N\cap\Gamma}(N+1)^{-K}e^{-NI(\mu|P)}\geq(N+1)^{-K}e^{-N\inf_{\mu\in\mathscr{L}_N\cap\Gamma}I(\mu|P)}.$$

As $\frac{1}{N}\log\big((N+1)^K\big) \to 0$, and $\mathscr{L}_N \cap \Gamma \subset \bar{\Gamma}$, (9.18) yields, the upper bound of the LDP, whereas we only have

$$\liminf_{N\to\infty} \frac{1}{N}\log \mathbb{P}\big(\hat{\mu}_N^{\mathbf{X}_N} \in \Gamma\big) = -\limsup_{N\to\infty}\big\{\inf_{\mu\in\mathscr{L}_N\cap\Gamma} I(\mu|P)\big\}$$

for the lower bound. Nevertheless, as Γ is open, for any $\mu \in \overset{\circ}{\Gamma}, \exists \eta_\mu > 0$ such that $\big\{\nu, d_V(\mu,\nu) \le \eta_\mu\big\} \subset \Gamma$. By Lemma 9.7.1, exists $\mu_N \in \mathscr{L}_N \cap \Gamma$ such that $\mu_N \to_N \mu$. Choose now $\mu \in \overset{\circ}{\Gamma}$ such that $\inf_{\nu\in\overset{\circ}{\Gamma}} H(\nu|P) \ge H(\mu|P) - \varepsilon$. Then

$$\inf_{\nu\in\mathscr{L}_N\cap\Gamma} H(\nu|P) \le H(\mu_N|P) \to H(\mu|P)$$
$$\limsup_{N\to\infty}\big\{\inf_{\nu\in\mathscr{L}_N\cap\Gamma} H(\nu|P)\big\} \le \lim_N H(\mu_N|P) \le \inf_{\nu\in\overset{\circ}{\Gamma}} H(\nu|P) + \varepsilon,$$

so that sending $\varepsilon \searrow 0$ yields the lower bound. □

9.7.2 Proof of abstract Varadhan setting

Proof. of Lemma (9.6.7): Define the map:

$$\delta := \begin{cases} \Sigma & \to M_1^+(\Sigma) \\ x & \to \delta_x \end{cases} \tag{9.19}$$

which is continuous as $\delta_{x_p} \overset{\mathscr{L}}{\to}_p \delta_x \iff x_p \to_p x$. Remark that $\hat{P}_N = P_N \circ \delta^{-1}$. Then, 1. $\implies$ 2. is a consequence of the contraction principle. Let us now prove this is an equivalence. In fact, suppose that $\hat{P}_N$ satisfies a FLDP with good rate function $\hat{H}$:

$$-\hat{H}\big(\overset{\circ}{\Gamma}\big) \le \liminf \frac{1}{N}\log \hat{P}_N(\Gamma) \le \limsup \frac{1}{N}\log \hat{P}_N(\Gamma) \le -\hat{H}\big(\bar{\Gamma}\big),$$

and take a probability measure $\mu \neq \delta_x$, $\forall x \in \Sigma$. Then exists $\varepsilon > 0$ such that the open ball $B(\mu,\varepsilon) \in \mathscr{B}\big(\mathscr{M}_1^+(\Sigma)\big)$ contains no Dirac mass (indeed, if not we could find a sequence δ_{x_p} such that $\delta_{x_p} \overset{\mathscr{L}}{\to} \mu$ implying in particular $\exists x \in \Sigma, \mu = \delta_x$, and $x_p \to x$). As $\hat{P}_N(B(\mu,\varepsilon)) = 0$, the lower-bound ensures that $\hat{H}\big(B(\mu,\varepsilon)\big) = +\infty$, so that $\hat{H}(\mu) = +\infty$. Define

$$H := \begin{cases} \Sigma & \to [0,+\infty] \\ x & \to \hat{H}\circ\delta(x). \end{cases} \tag{9.20}$$

which is clearly lower semi-continuous and non uniformly infinite. Moreover, as $\delta : \Sigma \to \delta(\Sigma)$ is a bijective map with continuous inverse, and as the compact set $\big\{\hat{H} \le a\big\} \subset \delta(\Sigma)$ for any $a \in \mathbb{R}$, it follows that $\big\{x \in \Sigma,\, H \le a\big\} = \delta^{-1}\big(\big\{\hat{H} \le a\big\}\big)$ is also compact, so that H is a good rate function. The FLDP of $\hat{P}_N$ is easily converted into:

$$-H\big(\overset{\circ}{\Gamma}\big) \le \liminf_N \frac{1}{N}\log P_N(\Gamma) \le \limsup_N \frac{1}{N}\log P_N(\Gamma) \le -H\big(\bar{\Gamma}\big).$$

Observe that Λ is convex by Hölder inequality, so that it satisfies the assumptions of Theorem (9.2.3), and its Legendre transform Λ^* indeed exists and satisfies the equality of Fenchel-Legendre duality. Suppose now that $(\hat{P}_N)$ satisfies a FLDP. On the one hand, Lemma (9.6.7), Proposition (9.4.2) and Bryc's theorem respectively ensure that $(P_N)_N$ satisfies a FLDP, is exponentially tight, and that the associated good rate

is given by $H : x \in \Sigma \to \sup_{\phi \in \mathscr{C}_b(\Sigma)} \{\phi(x) - \Lambda(\phi)\}$. On the other hand, Lemma (9.6.7) ensures that the good rate function associated with the FLDP of $(\hat{P}_N)$ is given by:

$$\hat{H} := \begin{cases} \mathscr{M}_1^+(\Sigma) & \to \bar{\mathbb{R}} \\ \mu & \to \begin{cases} \sup_{\phi \in \mathscr{C}_b(\Sigma)} \{\phi(x) - \Lambda(\phi)\} & \text{if exists } x \in \Sigma \text{ such that } \mu = \delta_x, \\ +\infty & \text{if not.} \end{cases} \end{cases} \tag{9.21}$$

Remark that Λ^* and $\hat{H}$ indeed coincide on Dirac masses $\delta_x, x \in \Sigma$. Moreover, remark that if exists two distinct points $x, y \in \Sigma$ such that $\hat{H}(\delta_x) = H(x), \hat{H}(\delta_y) = H(y) < +\infty$ then $\hat{H}$ can not be convex, has

$$+\infty = \hat{H}\big(\frac{\delta_x + \delta_y}{2}\big) > \frac{H(x) + H(y)}{2},$$

so that it can not coincide with Λ^* which is convex by Proposition (9.2.1). Indeed, $\hat{H}$ is convex if and only if exists a unique $x \in \Sigma$ such that $H(x) = 0$, and $\forall y \in \Sigma,\ y \neq x,\ H(y) = +\infty$. In this case, Varadhan's lemma point 3.ensures that, for any $\phi \in \mathscr{C}_b(\Sigma)$, $\Lambda(\phi) = \sup_{y \in \Sigma} \{\phi(y) - H(y)\} = \phi(x)$, so that

$$\Lambda^*(\nu) = \sup_{\phi \in \mathscr{C}_b(\Sigma)} \Big\{ \int_\Sigma \big(\phi(y) - \phi(x)\big) d\nu(y) \Big\},$$

which equals $+\infty$ as soon as $\nu \not\ll \delta_x$. □

Proof of remark (30).1

Proof. 1. Suppose that $\nu(\Sigma) > 1$, and take the constant function $\phi = M \in \mathscr{C}_b(\Sigma)$, with $M \in \mathbb{R}_+^*$. Then

$$\Lambda^*(\nu) \geq \int_\Sigma M\nu(x) - \Lambda(M) = M(\nu(\Sigma) - 1),$$

so that sending $M \to \infty$, we conclude that $\Lambda^*(\nu) = +\infty$. The same holds if $\nu(\Sigma) < 1$, choosing $\phi = -M \in \mathscr{C}_b(\Sigma)$. Hence, $\Lambda^* = +\infty$ outside $\mathscr{M}_1^+(\Sigma)$.
2. Suppose that exists a point $y \in \Sigma$ with $\mu(\{y\}) > 0$ and $H(y) > 0$. Then, exists $\varepsilon > 0$ such that $H\big(B(y,\varepsilon)\big) > \frac{2-\mu(\{y\})}{2} H(y)$. For $p \in \mathbb{N}^*$, define a continuous positive function $\phi_p : \Sigma \to \mathbb{R}$ with support in $B(y, \frac{1}{p})$ and uniformly equal to its maximal value $H(y)$ on the ball $B(y, \frac{1}{2p})$. Then, for p large enough so that $B(y, \frac{1}{p}) \subset B(y, \varepsilon)$,

$$\begin{aligned} \Lambda^*(\mu) &\geq \int_\Sigma \phi_p(z) d\mu(z) - \sup_{x \in \Sigma} \{\phi_p(x) - H(x)\} = \int_{B(y, \frac{1}{2p})} \phi_p(z) d\mu(z) - \sup_{x \in B(y,\varepsilon)} \{\phi_p(x) - H(x)\} \\ &\geq H(y)\mu\big(B(y, \frac{1}{2p})\big) \int \phi_p(z) d\mu(z) - \sup_{x \in B(y,\varepsilon)} \{\phi_p(x)\} + H\big(B(y,\varepsilon)\big) \geq H(y) \frac{\mu(\{y\})}{2}. \end{aligned}$$

3. Suppose that H is continuous on $D_H := \{x \in \Sigma, H(x) < \infty\}$. Then, for any $M \in \mathbb{R}^+$, $H \wedge M$ is integrable with respect to μ, as it is continuous and μ is a Borel measure. Furthermore, $H \wedge M \in \mathscr{C}_b(\Sigma)$, $\sup_\Sigma \{H \wedge M - H\} = 0$, so that

$$\Lambda^*(\mu) \geq \int_\Sigma H(x) \wedge M d\mu(x).$$

The Monotone Convergence theorem then yields,

$$\Lambda^*(\mu) \geq \int_\Sigma H(x) d\mu(x).$$

Reversely, for any $\phi \in \mathscr{C}_b(\Sigma)$, remark that $\sup_\Sigma \{\phi - H\} \leq \sup_\Sigma \phi$ so that

$$\int_\Sigma H(x)d\mu(x) = \int_\Sigma \phi(x)d\mu(x) - \sup_\Sigma \{\phi - H\} + \underbrace{\left\{ \sup_\Sigma \{\phi - H\} - \int_\Sigma \phi(x) - H(x)d\mu(x) \right\}}_{\geq 0}$$

so that taking the supremum in $\phi \in \mathscr{C}_b(\Sigma)$ gives

$$\int_\Sigma H(x)d\mu(x) \geq \Lambda^*(\mu).$$

□

General Notations

LLN	Law of Large Numbers.
SLLN	Strong Law of Large Numbers.
CLT	Central Limit Theorem.
LDP	Large Deviations Principles.
DCT	Dominated Convergence Theorem.
MCT	Monotone Convergence Theorem.
B.D.G.	Burkholder-Davis-Gundy.
C.S.	Cauchy-Schwarz.
IBP	Integration By Part.
SDE	Stochastic differential equation.
ODE	Ordinary differential equation.
$:=$	Defined as.
$a \wedge b$	$\min(a,b)$.
$a \vee b$	$\max(a,b)$.
$\subset$	Subset of.
$\overset{\mathscr{L}}{=}$	Equal in law.
$\overset{\mathscr{L}}{\to}$	Convergence in law.
$Q \ll P$	Absolutely continuity of measures.
$Q \simeq P$	Equivalence of measures.
$\bigotimes$	Tensorial product of measures.
Σ	A Polish space.
$\mathscr{M}^+(\Sigma)$	The set of positive Borel measures on Σ.
$\mathscr{M}_1^+(\Sigma)$	The set of positive Borel probability measures on Σ.
$\mathscr{C}_{\mu,\nu}$	The set of measures $\xi \in \mathscr{M}_1^+\big(\Sigma \times \Sigma\big)$ with marginal μ and ν.
δ_x	The probability measure singular at $x \in \Sigma$.
$\bar{\Gamma}$	Closure of Γ.
$\mathring{\Gamma}$	Interior of Γ.
Γ^c	Complement of Γ.
$B(\nu,\delta)$	The open ball centered at ν with radius δ (various metrics).
$\|\cdot\|_{\infty,t}$	In Chapter 4, the supremum norm on $\mathscr{C}\big([-\tau,t],\mathbb{R}\big)$.
$\lim_N$	$\lim_{N\to+\infty}$.
$\mathscr{N}(m,v)$	The Gaussian distribution with mean m and variance v.
$\mathbf{1}_A$	Indicator function of A.
$\mathbf{x}$	Vector $(x^1,\ldots,x^N)$.
$\overset{\text{Fubini Fub.}}{\leq}$, $\leq$	Inequality obtained with Fubini's theorem (this also function for other theorem or inequalities and referenced equations).

List of Figures

Bibliography

[1] M Abeles. Firing rates and weil-timed events in the cerebral cortex. In *Models of neural networks*, pages 121–140. Springer, 1994.

[2] Juan A Acebrón, Luis L Bonilla, Conrad J Pérez Vicente, Félix Ritort, and Renato Spigler. The kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of modern physics*, 77(1):137, 2005.

[3] Edgar D Adrian. The impulses produced by sensory nerve endings. *The Journal of physiology*, 61(1):49–72, 1926.

[4] Edgar Douglas Adrian. The basis of sensation. 1928.

[5] S. Amari. Characteristics of random nets of analog neuron-like elements. *Syst. Man Cybernet. SMC-2*, 1972.

[6] S-I Amari. Characteristics of randomly connected threshold-element networks and network systems. *Proceedings of the IEEE*, 59(1):35–47, 1971.

[7] S.-I. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27(2):77–87, June 1977.

[8] Shun-ichi Amari. A method of statistical neurodynamics. *Kybernetik*, 14(4):201–215, 1974.

[9] Shun-Ichi Amari, Kiyonori Yoshida, and Ken-Ichi Kanatani. A mathematical foundation for statistical neurodynamics. *SIAM Journal on Applied Mathematics*, 33(1):95–126, 1977.

[10] Shunichi Amari. A theory of adaptive pattern classifiers. *IEEE Transactions on Electronic Computers*, (3):299–307, 1967.

[11] Per Andersen and Sven Anders Andersson. *Physiological basis of the alpha rhythm*, volume 1. Plenum Publishing Corporation, 1968.

[12] I. Aradi and I. Soltesz. Modulation of network behaviour by changes in variance in interneuronal properties. *The Journal of physiology*, 538(1):227, 2002.

[13] Ildiko Aradi and Ivan Soltesz. Modulation of network behaviour by changes in variance in interneuronal properties. *The Journal of physiology*, 538(1):227–251, 2002.

[14] G Ben Arous, Amir Dembo, and Alice Guionnet. Aging of spherical spin glasses. *Probability theory and related fields*, 120(1):1–67, 2001.

[15] G Ben Arous and A Guionnet. Large deviations for Langevin spin glass dynamics. *Probability Theory and Related Fields*, 102(4):455–509, 1995.

[16] Gerard Ben Arous and Alice Guionnet. Langevin dynamics for sherrington-kirkpatrick spin glasses. In *Mathematical Aspects of Spin Glasses and Neural Networks*, pages 323–353. Springer, 1998.

[17] Amir Atiya and Pierre Baldi. Oscillations and synchronizations in neural networks: an exploration of the labeling hypothesis. *International Journal of Neural Systems*, 1(02):103–124, 1989.

[18] A Babloyantz, JM Salazar, and C Nicolis. Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A*, 111(3):152–156, 1985.

[19] Agnessa Babloyantz and Alain Destexhe. Low-dimensional chaos in an instance of epilepsy. *Proceedings of the National Academy of Sciences*, 83(10):3513–3517, 1986.

[20] Mohammed Bachir. A non-convex analogue to fenchel duality. *Journal of Functional Analysis*, 181(2):300–312, 2001.

[21] MARTIENSTEN Bauer and W Martienssen. Quasi-periodicity route to chaos in neural networks. *EPL (Europhysics Letters)*, 10(5):427, 1989.

[22] G. Ben-Arous and A. Guionnet. Large deviations for langevin spin glass dynamics. *Probability Theory and Related Fields*, 102(4):455–509, 1995.

[23] G Ben Arous and A Guionnet. Symmetric langevin spin glass dynamics. *The Annals of Probability*, pages 1367–1422, 1997.

[24] G Ben Arous and O Zeitouni. Increasing propagation of chaos for mean field models. In *Annales de l'IHP Probabilités et statistiques*, volume 35, pages 85–102, 1999.

[25] R. Ben-Yishai, RL Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9):3844–3848, 1995.

[26] Michel Benaïm. *Dynamiques d'activation et dynamiques d'apprentissage des réseaux de neurones*. PhD thesis, 1992.

[27] Lorenzo Bertini, Giambattista Giacomin, and Khashayar Pakdaman. Dynamical aspects of mean field plane rotators and the kuramoto model. *Journal of Statistical Physics*, 138(1-3):270–290, 2010.

[28] Raymond L Beurle. Properties of a mass of cells capable of regenerating pulses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 240(669):55–94, 1956.

[29] Patrick Billingsley. *Convergence of Probability Measures*. Wiley series in probability and statistics, 1999.

[30] Sem Borst and Florian Simatos. A stochastic network with mobile users in heavy traffic. *Queueing Systems*, 74(1):1–40, 2013.

[31] William H Bosking, Ying Zhang, Brett Schofield, and David Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of neuroscience*, 17(6):2112–2127, 1997.

[32] Paul C Bressloff. Spatiotemporal dynamics of continuum neural fields. *Journal of Physics A: Mathematical and Theoretical*, 45(3):033001, 2011.

[33] Paul C Bressloff. Waves in neural media. *Lecture Notes on Mathematical Modelling in the Life Sciences, Springer, New York*, 2014.

[34] P.C. Bressloff. Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Physica D: Nonlinear Phenomena*, 155(1-2):83–100, 2001.

[35] P.C. Bressloff. Stochastic neural field theory and the system-size expansion. *SIAM J. on Applied Mathematics*, 70:1488–1521, 2009.

[36] P.C. Bressloff. Spatiotemporal dynamics of continuum neural fields. *Journal of Physics A: Mathematical and Theoretical*, 45:033001, 2012.

[37] P.C. Bressloff, J.D. Cowan, M. Golubitsky, P.J. Thomas, and M.C. Wiener. Geometric visual hallucinations, euclidean symmetry and the functional architecture of striate cortex. *Phil. Trans. R. Soc. Lond. B*, 306(1407):299–330, March 2001.

[38] Detlev W Bronk, Martin G Larrabee, et al. Chemical excitation of nerve. *Annals of the New York Academy of Sciences*, 47(4):457–485, 1946.

[39] John Buck. Synchronous rhythmic flashing of fireflies. ii. *Quarterly review of biology*, pages 265–289, 1988.

[40] MA Buice and JD Cowan. Field-theoretic approach to fluctuation effects in neural networks. *Physical Review E*, 75(5), 2007.

[41] MA Buice, JD Cowan, and CC Chow. Systematic fluctuation expansion for neural network activity equations. *Neural computation*, 22(2):377–426, 2010.

[42] Michael A Buice and Carson C Chow. Beyond mean field theory: statistical field theory for neural networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(03):P03003, 2013.

[43] Michael A Buice and Carson C Chow. Dynamic finite size effects in spiking neural networks. *PLoS Comput. Biol*, 9:e1002872, 2013.

[44] AJ Buller, JG Nicholls, and G Ström. Spontaneous fluctuations of excitability in the muscle spindle of the frog. *The Journal of physiology*, 122(2):409, 1953.

[45] P.L. Buono and M. Golubitsky. Models of central pattern generators for quadruped locomotion: I. primary gaits. *J. Math. Biol*, 42(4):291–326, 2001.

[46] B Delisle Burns. Some properties of isolated cerebral cortex in the unanaesthetized cat. *The Journal of physiology*, 112(1-2):156, 1951.

[47] B Delisle Burns, Woodburn Heron, and R Pritchard. Physiological excitation of visual cortex in cat's unanaesthetized isolated forebrain. *Journal of neurophysiology*, 1962.

[48] Benedict Delisle Burns. uncertain nervous system. 1968.

[49] D.P. Buxhoeveden and M.F. Casanova. The minicolumn hypothesis in neuroscience. *Brain*, 125:935–951, 2002.

[50] Gyorgy Buzsaki. *Rhythms of the brain*. Oxford University Press, USA, 2004.

[51] Tanguy Cabana and Jonathan Touboul. Large deviations and dynamics of large stochastic heterogeneous neural networks. *Journal of Statistical Physics*, 2013.

[52] Tanguy Cabana and Jonathan Touboul. Large deviations for spatially extended random neural networks. *(submitted)*, 2015.

[53] Tanguy Cabana and Jonathan Touboul. Large deviations of particle systems in random interaction. *arXiv preprint arXiv:1601.00985*, 2016.

[54] D Cai, L Tao, M Shelley, and DW McLaughlin. An effective kinetic representation of fluctuation-driven neuronal networks with application to simple and complex cells in visual cortex. *Proceedings of the National Academy of Sciences*, 101(20):7757–7762, 2004.

[55] M. Camperi and X.J. Wang. A model of visuospatial working memory in prefrontal cortex: Recurrent network and cellular bistability. *Journal of Computational Neuroscience*, 5:383–405, 1998.

[56] José A Carrillo, Massimo Fornasier, Giuseppe Toscani, and Francesco Vecil. Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical modeling of collective behavior in socio-economic and life sciences*, pages 297–336. Springer, 2010.

[57] B. Cessac. Increase in complexity in random neural networks. *Journal de Physique I (France)*, 5:409–432, 1995.

[58] Bruno Cessac, Bernard Doyon, Mathias Quoy, and Manuel Samuelides. Mean-field equations, bifurcation map and route to chaos in discrete time neural networks. *Physica D: Nonlinear Phenomena*, 74(1-2):24–44, 1994.

[59] Bruno Cessac, Olivier Mazet, Manuel Samuelides, and Hédi Soula. Mean field theory for random recurrent spiking neural networks. In *NOLTA*, volume 5, pages 18–21, 2005.

[60] Bruno Cessac and Manuel Samuelides. From neuron to neural networks dynamics. *EPJ Special topics: Topics in Dynamical Neural Networks*, 142(1):7–88, 2007.

[61] RD Chervin, PA Pierce, and BW Connors. Periodicity and directionality in the propagation of epileptiform discharges across neocortex. *Journal of Neurophysiology*, 60(5):1695–1713, 1988.

[62] Hayato Chiba. A proof of the kuramoto conjecture for a bifurcation structure of the infinite-dimensional kuramoto model. *Ergodic Theory and Dynamical Systems*, 35(03):762–834, 2015.

[63] MY Choi, HJ Kim, D Kim, and H Hong. Synchronization in a system of globally coupled oscillators with time delay. *Physical Review E*, 61(1):371, 2000.

[64] Nikhil Chopra and Mark W Spong. On exponential synchronization of kuramoto oscillators. *IEEE Transactions on Automatic Control*, 54(2):353–357, 2009.

[65] Eugene F Civillico and Diego Contreras. Integration of evoked responses in supragranular cortex studied with optical recordings in vivo. *Journal of neurophysiology*, 96(1):336–351, 2006.

[66] JD Clements. Transmitter timecourse in the synaptic cleft: its role in central synaptic function. *Trends in neurosciences*, 19(5):163–171, 1996.

[67] S. Coombes and M. R. Owen. Bumps, breathers, and waves in a neural network with spike frequency adaptation. *Phys. Rev. Lett.*, 94(14), 2005.

[68] Stephen Coombes. Waves, bumps, and patterns in neural fields theories. *Biological Cybernetics*, 93(2):91–108, 2005.

[69] Stephen Coombes, Peter beim Graben, Roland Potthast, and James Wright. *Neural Fields*. Springer, 2014.

[70] Stephen Coombes and Carlo Laing. Delays in activity based neural networks. *Submitted to the Royal Society*, 2011.

[71] Jack D Cowan. A statistical mechanics of nervous activity. Technical report, DTIC Document, 1969.

[72] JD Cowan. Statistical mechanics of nervous nets. In *Neural networks*, pages 181–188. Springer, 1968.

[73] G Da Prato and J Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge Univ Pr, 1992.

[74] Paolo Dai Pra and Frank den Hollander. Mckean-vlasov limit for interacting random processes in random media. *Journal of statistical physics*, 84(3-4):735–772, 1996.

[75] Hiroaki Daido. Lower critical dimension for populations of oscillators with randomly distributed frequencies: a renormalization-group analysis. *Physical review letters*, 61(2):231, 1988.

[76] Hiroaki Daido. Quasientrainment and slow relaxation in a population of oscillators with random and frustrated interactions. *Physical review letters*, 68(7):1073, 1992.

[77] Hiroaki Daido. Algebraic relaxation of an order parameter in randomly coupled limit-cycle oscillators. *Physical Review E*, 61(2):2145, 2000.

[78] Glenn Dallérac and Nathalie Rouach. Astrocytes as new targets to improve cognitive functions. *Progress in neurobiology*, 2016.

[79] E. Dauce, O. Moynot, O. Pinaud, and M. Samuelides. Mean-field theory and synchronization in random recurrent neural networks. *Neural Processing Letters*, 14(2):115–126, 2001.

[80] P. Dayan and L.F. Abbott. *Theoretical Neuroscience : Computational and Mathematical Modeling of Neural Systems*. MIT Press, 2001.

[81] C De Dominicis. Dynamics as a substitute for replicas in systems with quenched random impurities. *Physical Review B*, 18(9):4913, 1978.

[82] C De Dominicis. Systems with quenched random impurities, an overview of dynamics, replicas and frustration approaches. In *Dynamical Critical Phenomena and Related Topics*, pages 251–279. Springer, 1979.

[83] C De Dominicis and L Peliti. Field-theory renormalization and critical dynamics above t c: Helium, antiferromagnets, and liquid-gas systems. *Physical Review B*, 18(1):353, 1978.

[84] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Verlag, 2009.

[85] Frank Den Hollander. *Large deviations*, volume 14. American Mathematical Soc., 2008.

[86] J.D. Deuschel and D.W. Stroock. *Large deviations*, volume 137. Academic Pr, 1989.

[87] RW Ditchburn and BL Ginsborg. Vision with a stabilized retinal image. 1952.

[88] Florian Dorfler and Francesco Bullo. Synchronization and transient stability in power networks and nonuniform kuramoto oscillators. *SIAM Journal on Control and Optimization*, 50(3):1616–1642, 2012.

[89] B Doyon, B Cessac, M Quoy, and M Samuelides. Control of the transition to chaos in neural networks with random connectivity. *International Journal of Bifurcation and Chaos*, 3(02):279–291, 1993.

[90] Paul Dupuis and Richard S Ellis. *A weak convergence approach to the theory of large deviations*, volume 902. John Wiley & Sons, 2011.

[91] JC Eccles. Inhibition in thalamic and cortical neurones and its role in phasing neuronal discharges. *Epilepsia*, 6(2):89–115, 1965.

[92] Reinhard Eckhorn, Roman Bauer, Wolfgang Jordan, Michael Brosch, Wolfgang Kruse, Matthias Munk, and HJ Reitboeck. Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological cybernetics*, 60(2):121–130, 1988.

[93] Bard Ermentrout. Neural networks as spatio-temporal pattern-forming systems. *Reports on Progress in Physics*, 61:353–430, 1998.

[94] G Bard Ermentrout and Jack D Cowan. A mathematical theory of visual hallucination patterns. *Biological cybernetics*, 34(3):137–150, 1979.

[95] G Bard Ermentrout and Jack D Cowan. A mathematical theory of visual hallucination patterns. *Biological cybernetics*, 34(3):137–150, 1979.

[96] G Bard Ermentrout, Roberto F Galán, and Nathaniel N Urban. Reliability, synchrony and noise. *Trends in neurosciences*, 31(8):428–434, 2008.

[97] G Bard Ermentrout and David Kleinfeld. Traveling electrical waves in cortex: insights from phase dynamics and speculation on a computational role. *Neuron*, 29(1):33–44, 2001.

[98] G Bard Ermentrout and John Rinzel. Reflected waves in an inhomogeneous excitable medium. *SIAM Journal on Applied Mathematics*, 56(4):1107–1128, 1996.

[99] G. Bard Ermentrout and David Terman. *Foundations Of Mathematical Neuroscience*. Interdisciplinary Applied Mathematics. Springer, 2010.

[100] GB Ermentrout and DH Terman. Mathematical foundations of neuroscience, 2010.

[101] A Aldo Faisal, Luc PJ Selen, and Daniel M Wolpert. Noise in the nervous system. *Nature reviews neuroscience*, 9(4):292–303, 2008.

[102] BG Farley and WA Clark. Activity in networks of neuron-like elements. *Proc. 4th London Syrup. on Inf. Theory. Cherry, C, ed. London: Butterworths*, 1961.

[103] O. Faugeras, J. Touboul, and B. Cessac. A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in Neuroscience*, 3(1), 2009.

[104] Olivier Faugeras and James Maclaurin. Asymptotic description of stochastic neural networks. i. existence of a large deviation principleastic neural networks. ii. characterization of the limit law. *Comptes Rendus Mathematique*, 352(10):847–852, 2014.

[105] Olivier Faugeras and James Maclaurin. Asymptotic description of neural networks with correlated synaptic weights. *Entropy*, 17(7):4701–4743, 2015.

[106] Olivier Faugeras, Jonathan Touboul, and Bruno Cessac. A constructive mean-field analysis of multi population neural networks with random synaptic weights. In *COSYNE 09*, 2009.

[107] Isabelle Ferezou, Sonia Bolea, and Carl CH Petersen. Visualizing the cortical representation of whisker touch: voltage-sensitive dye imaging in freely moving mice. *Neuron*, 50(4):617–629, 2006.

[108] Bastien Fernandez, David Gérard-Varet, and Giambattista Giacomin. Landau damping in the kuramoto model. In *Annales Henri Poincaré*, pages 1–31. Springer, 2014.

[109] R FitzHugh. Fitzhugh-nagumo simplified cardiac action potential model. *Biophys. J*, 1:445–466, 1961.

[110] WALTER J Freeman. Relations between unit activity and evoked potentials in prepyriform cortex of cats. *Journal of neurophysiology*, 31(3):337–348, 1968.

[111] Walter J Freeman and John M Barrie. Analysis of spatial patterns of phase in neocortical gamma eegs in rabbit. *Journal of neurophysiology*, 84(3):1266–1278, 2000.

[112] Walter J Freeman, Yong Yao, and Brian Burke. Central pattern generating and recognizing in olfactory bulb: a correlation learning rule. *Neural Networks*, 1(4):277–288, 1988.

[113] W.J. Freeman. Simulation of chaotic eeg patterns with a dynamic model of the olfactory system. *Biological Cybernetics*, 56:139–150, 1987.

[114] Shintaro Funahashi, Charles J Bruce, and Patricia S Goldman-Rakic. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of neurophysiology*, 61(2):331–349, 1989.

[115] Stuart Geman. Almost sure stable oscillations in a large system of randomly coupled equations. *SIAM Journal on Applied Mathematics*, 42(4):695–703, 1982.

[116] Stuart Geman and Chii-Ruey Hwang. A chaos hypothesis for some large systems of random equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 60(3):291–314, 1982.

[117] W. Gerstner and W. M. Kistler. Mathematical formulations of hebbian learning. *Biological Cybernetics*, 87:404–415, 2002.

[118] Giambattista Giacomin, Khashayar Pakdaman, and Xavier Pellegrin. Global attractor and asymptotic dynamics in the kuramoto model for coupled noisy phase oscillators. *Nonlinearity*, 25(5):1247, 2012.

[119] Martin Giese. *Dynamic Neural Field Theory for Motion Perception*. Springer, 1998.

[120] Julijana Gjorgjieva, Guillaume Drion, and Eve Marder. Computational implications of biophysical diversity and multiple timescales in neurons and synapses for circuit performance. *Current opinion in neurobiology*, 37:44–52, 2016.

[121] D Golomb, David Hansel, B Shraiman, and Haim Sompolinsky. Clustering in globally coupled phase oscillators. *Physical Review A*, 45(6):3516, 1992.

[122] Carl Graham, Philippe Robert, et al. Interacting multi-class transmissions in large stochastic networks. *The Annals of Applied Probability*, 19(6):2334–2361, 2009.

[123] Charles M Gray, Peter König, Andreas K Engel, Wolf Singer, et al. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338(6213):334–337, 1989.

[124] Charles M Gray and Wolf Singer. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences*, 86(5):1698–1702, 1989.

[125] Henry Gray. *Anatomy of the human body*. Lea & Febiger, 1918.

[126] JS Griffith. On the stability of brain-like structures. *Biophysical journal*, 3(4):299, 1963.

[127] Charles G Gross. From imhotep to hubel and wiesel. In *Extrastriate Cortex in Primates*, pages 1–58. Springer, 1997.

[128] Stephen Grossberg and David Somers. Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*, 4(4):453–466, 1991.

[129] A Guionnet. Averaged and quenched propagation of chaos for spin glass dynamics. *Probability Theory and Related Fields*, 109(2):183–215, 1997.

[130] LD Harmon and ER Lewis. Neural modeling. *Physiol. Rev*, 46(3):513–591, 1966.

[131] Christopher M Harris and Daniel M Wolpert. Signal-dependent noise determines motor planning. *Nature*, 394(6695):780–784, 1998.

[132] HK Hartline and Floyd Ratliff. Spatial summation of inhibitory influences in the eye of limulus, and the mutual interaction of receptor units. *The Journal of general physiology*, 41(5):1049–1066, 1958.

[133] D.O. Hebb. *The organization of behavior: a neuropsychological theory.* Wiley, NY, 1949.

[134] Geoffroy Hermann and Jonathan Touboul. Heterogeneous connections induce oscillations in large-scale networks. *Physical review letters*, 109(1):018702, 2012.

[135] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952.

[136] Hyunsuk Hong, Moo-Young Choi, and Beom Jun Kim. Synchronization on small-world networks. *Physical Review E*, 65(2):026139, 2002.

[137] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. National Academy of Sciences, USA*, 79:2554–2558, 1982.

[138] Jonathan C Horton and Daniel L Adams. The cortical column: a structure without a function. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360(1456):837–862, 2005.

[139] E.M. Izhikevich. *Dynamical Systems in Neuroscience: The Geometry of Excitability And Bursting*. MIT Press, 2007.

[140] Eugene M Izhikevich and Richard FitzHugh. Fitzhugh-nagumo model. *Scholarpedia*, 1(9):1349, 2006.

[141] Pierre-Emmanuel Jabin. A review of the mean field limits for vlasov equations. *Kinet. Relat. Models*, 7(4):661–711, 2014.

[142] Dirk Jancke, Frédéric Chavane, Shmuel Naaman, and Amiram Grinvald. Imaging cortical correlates of illusion in early visual cortex. *Nature*, 428(6981):423–426, 2004.

[143] HK Janssen. Field-theoretic method applied to critical dynamics. In *Dynamical critical phenomena and related topics*, pages 25–47. Springer, 1979.

[144] Ziping Jiang and Martin McCall. Numerical simulation of a large number of coupled lasers. *JOSA B*, 10(1):155–163, 1993.

[145] Edward G Jones. Microcolumns in the cerebral cortex. *Proceedings of the National Academy of Sciences*, 97(10):5019–5021, 2000.

[146] Mark Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, volume 1955, pages 171–197, 1954.

[147] Jonathan Kadmon and Haim Sompolinsky. Transition to chaos in random neuronal networks. *Physical Review X*, 5(4):041030, 2015.

[148] Zachary P Kilpatrick. Interareal coupling reduces encoding variability in multi-area models of spatial working memory. *Frontiers in computational neuroscience*, 7, 2013.

[149] Z.P. Kilpatrick, S.E. Folias, and P.C. Bressloff. Traveling pulses and wave propagation failure in inhomogeneous neural media. *SIAM Journal on Applied Dynamical Systems*, 7(1):161–185, 2008.

[150] M Kokkinidis and P Anninos. Noisy neural nets exhibiting epileptic features. *Journal of theoretical biology*, 113(3):559–588, 1985.

[151] Leah Krubitzer. The organization of neocortex in mammals: are species differences really so different? *Trends in neurosciences*, 18(9):408–417, 1995.

[152] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, pages 420–422. Springer, 1975.

[153] Yoshiki Kuramoto. *Chemical oscillations, waves, and turbulence*, volume 19. Springer Science & Business Media, 2012.

[154] Yoshiki Kuramoto and Toshio Tsuzuki. Persistent propagation of concentration waves in dissipative media far from thermal equilibrium. *Progress of theoretical physics*, 55(2):356–369, 1976.

[155] Christian Kurrer and Klaus Schulten. Noise-induced synchronous neuronal oscillations. *Physical Review E*, 51(6):6213, 1995.

[156] C.L. Laing, W.C. Troy, B. Gutkin, and G.B. Ermentrout. Multiple bumps in a neuronal model of working memory. *SIAM J. Appl. Math.*, 63(1):62–97, 2002.

[157] Geneviève Leuba and Rudolf Kraftsik. Changes in volume, surface estimate, three-dimensional shape and total number of neurons of the human primary visual cortex from midgestation until old age. *Anatomy and embryology*, 190(4):351–366, 1994.

[158] B. Lindner, J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geier. Effects of noise in excitable systems. *Physics Reports*, 392(6):321–424

[159] A. Longtin, A. Bulsara, and F. Moss. Time-interval sequences in bistable systems and the noise-induced transmission of information by sensory neurons. *Physical review letters*, 67(5):656–659

[160] Eric Luçon. *Oscillateurs couplés, désordre et synchronisation*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 2012.

[161] Eric Luçon, Wilhelm Stannat, et al. Mean field limit for disordered diffusions with singular interactions. *The Annals of Applied Probability*, 24(5):1946–1993, 2014.

[162] Cheng Ly and Daniel Tranchina. Critical analysis of dimension reduction by a moment closure method in a population density approach to neural network modeling. *Neural Computation*, 19(8):2032–2092, Aug 2007.

[163] James Lynch and Jayaram Sethuraman. Large deviations for processes with independent increments. *The annals of probability*, pages 610–627, 1987.

[164] Zachary F Mainen and Terrence J Sejnowski. Reliability of spike timing in neocortical neurons. *Science*, 268(5216):1503, 1995.

[165] Xuerong Mao. *Stochastic Differential Equations and Applications*. Horwood publishing, 2008.

[166] Henry Markram, Maria Toledo-Rodriguez, Yun Wang, Anirudh Gupta, Gilad Silberberg, and Caizhi Wu. Interneurons of the neocortical inhibitory system. *Nature Reviews Neuroscience*, 5(10):793–807, 2004.

[167] Paul Cecil Martin, ED Siggia, and HA Rose. Statistical dynamics of classical systems. *Physical Review A*, 8(1):423, 1973.

[168] Paul C Matthews, Renato E Mirollo, and Steven H Strogatz. Dynamics of a large system of coupled nonlinear oscillators. *Physica D: Nonlinear Phenomena*, 52(2):293–331, 1991.

[169] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[170] Mark D McDonnell and Lawrence M Ward. The benefits of noise in neural systems: bridging theory and experiment. *Nature Reviews Neuroscience*, 12(7):415–426, 2011.

[171] Sylvie Méléard. Asymptotic behaviour of some interacting particle systems; mckean-vlasov and boltzmann models. In *Probabilistic models for nonlinear partial differential equations*, pages 42–95. Springer, 1996.

[172] John Milton and Peter Jung. *Epilepsy as a dynamic disease*. Springer Science & Business Media, 2013.

[173] Stéphane Mischler. Introduction aux limites de champ moyen pour des systèmes de particules. 2011.

[174] Yamir Moreno and Amalio F Pacheco. Synchronization of kuramoto oscillators in scale-free networks. *EPL (Europhysics Letters)*, 68(4):603, 2004.

[175] Catherine Morris and Harold Lecar. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical journal*, 35(1):193, 1981.

[176] V.B. Mountcastle. Modality and topographic properties of single neurons of cat's somatosensory cortex. *Journal of Neurophysiology*, 20:408–434, 1957.

[177] V.B. Mountcastle. The columnar organization of the neocortex. *Brain*, 120:701–722, 1997.

[178] O. Moynot and M. Samuelides. Large deviations and mean-field theory for asymmetric random recurrent neural networks. *Probability Theory and Related Fields*, 123(1):41–75, 2002.

[179] Olivier Moynot and Manuel Samuelides. Large deviations and mean-field theory for asymmetric random recurrent neural networks. *Probability Theory and Related Fields*, 123(1):41–75, 2002.

[180] Lyle Muller, Alexandre Reynaud, Frédéric Chavane, and Alain Destexhe. The stimulus-evoked population response in visual cortex of awake monkey is a propagating wave. *Nature communications*, 5, 2014.

[181] Zoltán Néda, Erzsébet Ravasz, Tamás Vicsek, Yves Brechet, and Albert-Lázló Barabási. Physics of the rhythmic applause. *Physical Review E*, 61(6):6987, 2000.

[182] SERGIO Neuenschwander, ANDREAS K Engel, PETER König, WOLF Singer, and FRANCISCO J Varela. Synchronization of neuronal responses in the optic tectum of awake pigeons. *Vis. Neurosci*, 13:575–584, 1996.

[183] Jacques Neveu. Processus aleatoires gaussiens. 1970.

[184] George L O'Brien. Sequences of capacities, with connections to large-deviation theory. *Journal of Theoretical Probability*, 9(1):19–35, 1996.

[185] George L O'Brien and Wim Vervaat. Compactness in the theory of large deviations. *Stochastic processes and their applications*, 57(1):1–10, 1995.

[186] Srdjan Ostojic. Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. *Nature neuroscience*, 17(4):594–600, 2014.

[187] Ulrike Pannasch, Dominik Freche, Glenn Dallérac, Grégory Ghézali, Carole Escartin, Pascal Ezan, Martine Cohen-Salmon, Karim Benchenane, Veronica Abudara, Amandine Dufour, et al. Connexin 30 sets synaptic strength by controlling astroglial synapse invasion. *Nature neuroscience*, 17(4):549–558, 2014.

[188] Lorenzo Pareschi and Giuseppe Toscani. *Interacting multiagent systems: kinetic equations and Monte Carlo methods*. Oxford University Press, 2013.

[189] Charles S Peskin. *Mathematical aspects of heart physiology*. Courant Institute of Mathematical Sciences, New York University, 1975.

[190] Carl CH Petersen, Amiram Grinvald, and Bert Sakmann. Spatiotemporal dynamics of sensory responses in layer 2/3 of rat barrel cortex measured in vivo by voltage-sensitive dye imaging combined with whole-cell voltage recordings and neuron reconstructions. *The Journal of Neuroscience*, 23(4):1298–1309, 2003.

[191] J Pham, K Pakdaman, and J-F Vibert. Noise-induced coherent oscillations in randomly connected neural networks. *Physical Review E*, 58(3):3610, 1998.

[192] Gian F Poggio and Lawrence J Viernstein. Time series analysis of impulse sequences of thalamic somatic sensory neurons. *Journal of Neurophysiology*, 27(4):517–545, 1964.

[193] Astrid A Prinz, Dirk Bucher, and Eve Marder. Similar network activity from disparate circuit parameters. *Nature neuroscience*, 7(12):1345–1352, 2004.

[194] P.E. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, 2004.

[195] Philip E Protter. *Stochastic Differential Equations*. Springer, 2005.

[196] Anatolii Anatol'evich Pukhalskii. *On functional principle of large deviation*, volume New Trends in Probability and Statistics, pages 198–218. VSP Moks'las, Moskva, 1991.

[197] Aaditya V Rangan, David Cai, and Louis Tao. Numerical methods for solving moment equations in kinetic theory of neuronal network dynamics. *Journal of Computational Physics*, 221(2):781–798, 2007.

[198] AV Rangan, G Kova i, and D Cai. Kinetic theory for neuronal networks with fast and slow excitatory conductances driven by the same spike train. *Physical Review E*, 77(4), 2008.

[199] Floyd Ratliff. *Mach bands: quantitative studies on neural networks*. Holden-Day, San Francisco London Amsterdam, 1965.

[200] D Revuz and M Yor. *Continuous martingales and Brownian motion*. Springer Verlag, 1999.

[201] Kristen A Richardson, Steven J Schiff, and Bruce J Gluckman. Control of traveling waves in the mammalian cortex. *Physical Review Letters*, 94(2):028103, 2005.

[202] Fred Rieke, David Warland, Rob de Ruyter van Steveninck, and William Bialek. Spikes: Exploring the neural code.

[203] Philippe Robert and Jonathan D Touboul. On the dynamics of random neuronal networks. *arXiv preprint arXiv:1410.4072*, 2014.

[204] Nathaniel Rochester, J Holland, L Haibt, and W Duda. Tests on a cell assembly theory of the action of the brain, using a large digital computer. *IRE Transactions on information Theory*, 2(3):80–93, 1956.

[205] Francisco A Rodrigues, Thomas K DM Peron, Peng Ji, and Jürgen Kurths. The kuramoto model in complex networks. *Physics Reports*, 610:1–98, 2016.

[206] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

[207] LI Rozonoer. Concerning stochastic logical nets. i. *Avtomat. i Telemekh*, (5):137–147, 1969.

[208] LI Rozonoer. Random logical nets, ii. *Avtomat. Telemekh*, 6:99–109, 1969.

[209] LI Rozonoer. Random logical nets, iii. *Avtomatika i telemekhanika*, 7:127–136, 1969.

[210] Hidetsugu Sakaguchi, Shigeru Shinomoto, and Yoshiki Kuramoto. Local and grobal self-entrainments in oscillator lattices. *Progress of Theoretical Physics*, 77(5):1005–1010, 1987.

[211] M Samuelides and Bruno Cessac. Random recurrent neural networks dynamics. *The European Physical Journal Special Topics*, 142(1):89–122, 2007.

[212] Terrence J Sejnowski. Open questions about computation in cerebral cortex. *Parallel distributed processing*, 2:372–389, 1986.

[213] P. Series, S. Georges, J. Lorenceau, and Y. Frégnac. Orientation dependent modulation of apparent speed: a model based on the dynamics of feed-forward and horizontal connectivity in v1 cortex. *Vision research*, 42(25):2781–2797, 2002.

[214] Sebastian Seung. *Connectome: How the brain's wiring makes us who we are*. Houghton Mifflin Harcourt, 2012.

[215] Gordon M Shepherd. *Creating modern neuroscience: the revolutionary 1950s*. Oxford University Press, 2009.

[216] D. Sherrington and S. Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792–1796

[217] Wolf Singer. The formation of cooperative cell assemblies in the visual cortex. *Journal of experimental Biology*, 153(1):177–197, 1990.

[218] Christine A Skarda and Walter J Freeman. How brains make chaos in order to make sense of the world. *Behavioral and brain sciences*, 10(02):161–173, 1987.

[219] William R Softky. Simple codes versus efficient codes. *Current opinion in neurobiology*, 5(2):239–247, 1995.

[220] D.C. Somers, S.B. Nelson, and M. Sur. An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience*, 15(8):5448, 1995.

[221] H. Sompolinsky, A. Crisanti, and HJ Sommers. Chaos in Random Neural Networks. *Physical Review Letters*, 61(3):259–262, 1988.

[222] H. Sompolinsky and A. Zippelius. Relaxational dynamics of the Edwards-Anderson model and the mean-field theory of spin-glasses. *Physical Review B*, 25(11):6860–6875, 1982.

[223] Richard B Stein, E Roderich Gossen, and Kelvin E Jones. Neuronal variability: noise or part of the signal? *Nature Reviews Neuroscience*, 6(5):389–397, 2005.

[224] JC Stiller and G Radons. Dynamics of nonlinear oscillators with random interactions. *Physical Review E*, 58(2):1789, 1998.

[225] JC Stiller and G Radons. Self-averaging of an order parameter in randomly coupled limit-cycle oscillators. *Physical Review E*, 61(2):2148, 2000.

[226] Steven Strogatz. *Sync: The emerging science of spontaneous order*. Hyperion, 2003.

[227] Steven H Strogatz. From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1):1–20, 2000.

[228] Steven H Strogatz and Renato E Mirollo. Phase-locking and critical phenomena in lattices of coupled nonlinear oscillators with random intrinsic frequencies. *Physica D: Nonlinear Phenomena*, 31(2):143–168, 1988.

[229] Steven H Strogatz, Renato E Mirollo, and Paul C Matthews. Coupled nonlinear oscillators below the synchronization threshold: relaxation by generalized landau damping. *Physical review letters*, 68(18):2730, 1992.

[230] AS Sznitman. Equations de type de boltzmann, spatialement homogenes. *Probability Theory and Related Fields*, 66(4):559–592, 1984.

[231] AS Sznitman. Topics in propagation of chaos. *Ecole d'Eté de Probabilités de Saint-Flour XIX*, pages 165–251, 1989.

[232] Michel Talagrand. *Mean field models for spin glasses: Volume I: Basic examples*, volume 54. Springer Science & Business Media, 2010.

[233] Michel Talagrand. *Mean field models for spin glasses: Volume II: Advanced Replica Symmetry and Low Temperature*, volume 54. Springer Science & Business Media, 2010.

[234] Jonathan Touboul. Limits and dynamics of stochastic neuronal networks with random heterogeneous delays. *Journal of Statistical Physics*, 149(4):569–597, 2012.

[235] Jonathan Touboul. On the dynamics of mean-field equations for stochastic neural fields with delays. *Physica D: Nonlinear Phenomena*, 241(15):1223–1244, 2012.

[236] Jonathan Touboul. Spatially extended networks with singular multi-scale connectivity patterns. *Journal of Statistical Physics*, 156(3):546–573, 2014.

[237] Jonathan Touboul et al. Propagation of chaos in neural fields. *The Annals of Applied Probability*, 24(3):1298–1328, 2014.

[238] Jonathan Touboul and Geoffroy Hermann. Heterogeneous connections induce oscillations in large-scale networks. *Physical Review Letters*, (in press), 2012.

[239] Jonathan Touboul, Geoffroy Hermann, and Olivier Faugeras. Noise-induced behaviors in neural mean field dynamics. *SIAM J. on Dynamical Systems*, 11(49–81), 2011.

[240] P.J. Uhlhaas and W. Singer. Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron*, 52(1):155–168

[241] C Villani. Limite de champ moyen. *Cours de DEA*, 2002:49, 2001.

[242] J Von Neumann. Automata studies, eds shannon ce, mccarthy j, ashby wr, 1956.

[243] Gilles Wainrib and Jonathan Touboul. Topological and dynamical complexity of random neural networks. *Physical review letters*, 110(11):118101, 2013.

[244] Kurt Wiesenfeld, Pere Colet, and Steven H Strogatz. Synchronization transitions in a disordered josephson series array. *Physical review letters*, 76(3):404, 1996.

[245] H.R. Wilson and J.D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12:1–24, 1972.

[246] H.R. Wilson and J.D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biological Cybernetics*, 13(2):55–80, September 1973.

[247] Arthur T Winfree. Biological rhythms and the behavior of populations of coupled oscillators. *Journal of theoretical biology*, 16(1):15–42, 1967.

[248] Arthur T Winfree. *The geometry of biological time*, volume 12. Springer Science & Business Media, 2001.

[249] Jian-Young Wu, Xiaoying Huang, and Chuan Zhang. Propagating waves of activity in the neocortex: what they are, what they do. *The Neuroscientist*, 14(5):487–502, 2008.

[250] Weifeng Xu, Xiaoying Huang, Kentaroh Takagaki, and Jian-young Wu. Compression and reflection of visually evoked cortical waves. *Neuron*, 55(1):119–129, 2007.

[251] Kechen Zhang. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *The journal of neuroscience*, 16(6):2112–2126, 1996.