Attacking the brain with neuroscience : mean-field theory, finite size effects and encoding capability of stochastic neural networks

Diego Fasoli

UNIVERSITÉ DE NICE-SOPHIA ANTIPOLIS

ÉCOLE DOCTORALE STIC

SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

T H È S E

pour l'obtention du grade de

Docteur en Sciences

de l'Université de Nice-Sophia Antipolis

Mention: Informatique

présentée et soutenue par

Diego FASOLI

English Title:

Attacking the Brain with Neuroscience: Mean-Field Theory, Finite Size Effects and Encoding Capability of Stochastic Neural Networks

Titre français:

Traiter le Cerveau avec les Neurosciences: Théorie de Champ-Moyen, Effets de Taille Finie et Capacité de Codage des Réseaux de Neurones Stochastiques

Thèse dirigée par *Olivier FAUGERAS*

soutenue le *25 septembre 2013*

Jury:

Président du Jury	Dr. Bruno CESSAC	INRIA Sophia-Antipolis
Rapporteurs	Dr. Alain DESTEXHE	CNRS Gif-sur-Yvette
	Dr. Stefano PANZERI	IIT@UNITN Rovereto
Examinateur	Dr. Gustavo DECO	UPF Barcelona
Directeur de thèse	Dr. Olivier FAUGERAS	INRIA Sophia-Antipolis

Contents

List of Figures

List of Tables

Abstract

The brain is the most complex system in the known universe. Its nested structure with small-world properties determines its function and behavior. The analysis of its structure requires sophisticated mathematical and statistical techniques. In this thesis we shed new light on neural networks, attacking the problem from different points of view, in the spirit of the *Theory of Complexity* and in terms of their information processing capabilities. In particular, we quantify the Fisher information of the system, which is a measure of its encoding capability. The first technique developed in this work is the mean-field theory of rate and FitzHugh-Nagumo networks without correlations in the thermodynamic limit, through both mathematical and numerical analysis. The second technique, the Mayer's cluster expansion, is taken from the physics of plasma, and allows us to determine numerically the finite size effects of rate neurons, as well as the relationship of the Fisher information to the size of the network for independent Brownian motions. The third technique is a perturbative expansion, which allows us to determine the correlation structure of the rate network for a variety of different types of connectivity matrices and for different values of the correlation between the sources of randomness in the system. With this method we can also quantify numerically the Fisher information not only as a function of the network size, but also for different correlation structures of the system. The fourth technique is a slightly different type of perturbative expansion, with which we can study the behavior of completely generic connectivity matrices with random topologies. Moreover this method provides an analytic formula for the Fisher information, which is

in qualitative agreement with the other results in this thesis. Finally, the fifth technique is purely numerical, and uses an Expectation-Maximization algorithm and Monte Carlo integration in order to evaluate the Fisher information of the FitzHugh-Nagumo network. In summary, this thesis provides an analysis of the dynamics and the correlation structure of the neural networks, confirms this through numerical simulation and makes two key counterintuitive predictions. The first is the formation of a perfect correlation between the neurons for particular values of the parameters of the system, a phenomenon that we term *stochastic synchronization*. The second, which is somewhat contrary to received opinion, is the explosion of the Fisher information and therefore of the encoding capability of the network for highly correlated neurons. The techniques developed in this thesis can be used also for a complete quantification of the information processing capabilities of the network in terms of information storage, transmission and modification, but this would need to be performed in the future.

Keywords: Computational Neuroscience, Stochastic Neural Networks, Theory of Complexity, Finite Size Effects, Connectome, Fisher Information.

Dedication

I dedicate this thesis to:

- my mother, from whom I inherited my interest in the deeper questions of life and existence, and my father, who transmitted to me the love of physics and mathematics;
- my friends, who are still asking me to invent the clear pill from the film *Limitless* [1];
- the patients with brain diseases or injuries that hopefully will benefit from the results of my research or from the future work inspired by my discoveries;
- the transhumanist movement and ideology, which is my strong motivation for doing scientific research.

Acknowledgements

I would like to express my thanks to:

- my parents, for everything;
- my friends, simply because without them this journey is not feasible;
- Bruno Magalini, for helping my family in my absence;
- my supervisor Olivier Faugeras, for the hours spent in front of the whiteboard trying to reveal the secrets of the brain;
- Guillaume Masson, Frédéric Chavane and Laurent Perrinet, for their cooperation on the study of the visual cortex;
- the FACETS-ITN community [2], for the incredible opportunity they have offered me;
- Bruno Cessac, for his help and encouragement;
- Pierre Kornprobst and Marie-Cécile Lafont, for their technical support and assistance;
- my colleagues, for time spent together and their contribution to science. A special thanks goes to Louis Capietto, James Maclaurin and Sylvain Merlet, for their help.

This work was partially supported by the ERC grant #227747 NerVi, the FACETS-ITN Marie-Curie Initial Training Network #237955 and the IP project BrainScaleS #269921.

Résumé (en Français)

Ce travail a été développé dans le cadre du projet européen FACETS-ITN, dans le domaine des Neurosciences Computationnelles. Son but est d'améliorer la compréhension des réseaux de neurones stochastiques de taille finie, pour des sources corrélées à caractère aléatoire et pour des matrices de connectivité biologiquement réalistes. Ce résultat est obtenu par l'analyse de la matrice de corrélation du réseau et la quantification de la capacité de codage du système en termes de son information de Fisher. Les méthodes comprennent diverses techniques mathématiques, statistiques et numériques, dont certaines ont été importés d'autres domaines scientifiques, comme la physique et la théorie de l'estimation. Ce travail étend de précédents résultats fondées sur des hypothèses simplifiées qui ne sont pas réaliste d'un point de vue biologique et qui peuvent être pertinents pour la compréhension des principes de travail liés cerveau. De plus, ce travail fournit les outils nécessaire à une analyse complète de la capacité de traitement de l'information des réseaux de neurones, qui sont toujours manquante dans la communauté scientifique.

Le Chapitre 1 présente la structure et les fonctions du cerveau, en particulier celle du cortex cérébral, en mettant l'accent sur sa structure imbriquée. Nous expliquons comment a évolué le cerveau au cours de l'histoire, selon la théorie de P. D. MacLean, ainsi que ses aires les plus importants. Nous clarifions aussi le rôle des différentes parties du cerveau, parmi l'incroyable collection de fonctions cognitives de l'être humain. Nous mettons l'accent sur la partie la plus externe du cerveau, appelée le *cortex cérébral*, en

décrivant sa subdivision en lobes, aires sensorielles primaires et uni modale ainsi qu'en aires associatives multimodales. Pour conclure, nous soulignons l'existence d'une structure imbriquée des connexions synaptiques dans chacune de ces aires, qui sera décrite mathématiquement dans le Chapitre 6.

Le Chapitre 2 explique deux points de vue différents qui sont couramment utilisés pour décrire les réseaux de neurones, à savoir la Théorie de la Complexité et l'approche computationnelle. La Théorie de la Complexité concerne les systèmes constitués de nombreuses particules qui interagissent pour d'atteindre un but. C'est donc est un candidat idéal pour décrire l'interaction entre les neurones du cerveau. En particulier, selon la Théorie de la Complexité, il est plus pertinent d'étudier le comportement du cerveau en utilisant des connexions synaptiques très réalistes plutôt que d'utiliser des modèles réalistes de la soma des neurones. Pour cette raison, nous mettons l'accent sur la nécessité de développer des modèles mathématiques de réseaux de neurones avec des matrices de connectivité biologiquement réalistes. Ce problème est abordé dans les Chapitres 5 et 6. Dans le Chapitre 2, nous expliquons egalement la différence entre le calcul et le traitement de l'information, en montrant comment le cerveau accomplit les deux à la fois. Dans le détail, nous allons présenter les différents types de calculs, c'est à dire quantique et classique, déterministe et probabiliste, numérique et analogique, et sémantique et non-sémantique, en expliquant leurs intérêts dans le contexte du cerveau. Nous présenterons également les différentes notions de l'information, en particulier l'information de Shannon, l'information de Fisher et l'information sémantique, avec un accent sur le second type d'information. L'information de Fisher du système, permettant de quantifier la capacité de codage du réseau de neurones, est l'un des principaux sujets de cette thèse, et sera quantifiée dans les chapitres suivants.

Le Chapitre 3 présente la théorie de champ moyen d'un réseau de neurones, à la fois pour les modèles stochastiques de type "rate" et pour ceux de FitzHugh-Nagumo, accompagnée de solutions analytiques et numériques des équations de champ moyen. Il s'agit de la pre-

mière étape de la thèse, à savoir l'analyse d'un réseau de neurones, dans les conditions les plus simples possible, c'est à dire en considérant des mouvements browniens indépendants et des neurones en nombre infinis (que l'on appelle la *limite thermodynamique*). Ce sont les conditions nécessaires pour l'émergence du phénomène connu sous le nom de *propagation du chaos*, sur laquelle la théorie de champ moyen est basée. Pour le modèle de type "rate", les équations de champ moyen peuvent être résolues analytiquement, tandis que pour le modèle FitzHugh-Nagumo nous avons résolu les équations numériquement en utilisant la méthode des lignes. Cependant, cette théorie ne peut pas décrire les effets de taille finie du réseau et son comportement pour des mouvements browniens corrélés.

Le Chapitre 4 présente la première des techniques utilisées dans cette thèse pour décrire les effets de taille finie du modèle de type "rate", à savoir l'expansion du groupe de Mayer, qui étend les résultats trouvés dans le Chapitre 3. De plus, il contient des simulations numériques de la densité de probabilité du réseau, avec une approche perturbative pour la quantification de l'information de Fisher du système. Selon les simulations, la capacité d'encodage du système est amélioré lorsque les neurones deviennent de plus en plus indépendants. Ce résultat sera confirmé dans les Chapitres 5 et 6 avec des approches différentes.

Le Chapitre 5 présente une approche perturbative qui sera utilisé pour déterminer la structure de corrélation d'un réseau de taille finie de type "rate" avec des sources corrélées à caractère aléatoire et pour différents types de matrices de connectivité. Cette approche étend les résultats du Chapitre 4, explorant la possibilité d'appliquer la théorie de champ moyen à la description d'un réseau de neurones. Dans le détail, nous allons prouver qu'en général, l'indépendance des neurones, sur quoi la théorie de champ moyen est basée, ne se produit pas pour les sources corrélées à caractère aléatoire, ni pour les matrices synaptiques qui ont un nombre insuffisant de connexions, ou pour des valeurs spécifiques des paramètres du système, qui génèrent un phénomène que nous avons appelé *synchronisation stochastique*. Par conséquent, en général, ces résultats ne valide pas l'utilisation

de la théorie de champ moyen pour la description du réseau de neurones. Le Chapitre 5 fournit également un algorithme pour évaluer numériquement l'information de Fisher du système, ce qui confirme les résultats du Chapitre 4 dans le cas de neurones indépendants. Par ailleurs, cet algorithme montre que la capacité de codage du réseau de neurones est beaucoup plus élevé lorsque les neurones sont fortement corrélés, un résultat qui sera prouvé analytiquement au Chapitre 6.

Le Chapitre 6 présente une autre approche perturbative pour les réseaux de taille finie de type "rate" avec des sources corrélées a caractère aléatoire, et il étend l'analyse de la structure de corrélation (qui ont été abordées au Chapitre 5) au cas de réseaux caractérisés par une topologie général et aléatoire. En particulier, il applique cette technique dans le cas d'une matrice de connectivité fractale qui estime la structure imbriquée des connections synaptiques analysées au Chapitre 1. Le Chapitre 6 présente également une formule analytique pour le calcul de l'information de Fisher, et qui explique de façon qualitative les résultats du Chapitre 5.

Le Chapitre 7 explore une méthode numérique basée sur l'algorithme Espérance - Maximisation et l'intégration de Monte Carlo qui nous permet de déterminer le comportement qualitatif de l'information de Fisher d'un réseau de FitzHugh-Nagumo, soutenant les résultats trouvés dans les Chapitres 5 et 6. Dans le détail, nous avons déterminé l'information de Fisher en fonction du nombre de neurones et de la corrélation entre les mouvements browniens, dans deux cas différents, à savoir lorsque les neurones génèrent des spikes et quand ils sont aux repos. Dans les deux cas, l'information de Fisher est plus élevée lorsque les potentiels de membrane sont fortement corrélés.

Le Chapitre 8 conclut la thèse, avec une discussion des résultats, leurs implications et leurs pertinence. Les techniques analytiques présentées dans les chapitres précédents ont été en mesure de quantifier avec succès les effets de taille finie de la structure de corrélation du réseau, pour des sources corrélées à caractère aléatoire et pour beaucoup de types de matrices de connectivité. Les simulations numériques en plus de ses études an-

alytiques, ont prouvé que la capacité d'encodage d'un réseau de neurones est plus élevée pour les potentiels de membrane fortement corrélés, contrairement à notre première intuition. Avec l'utilisation des les approches perturbatives vues dans les Chapitres 5 et 6, l'analyse du réseau peut être facilement étendu au-delà de l'information de Fisher. En fait, en utilisant la théorie de l'information de Shannon, il est possible de calculer le transfert, stockage et modification de l'information, caractérisant complétement la capacité des réseaux de neurones stochastiques.

Mots clés: Neurosciences Computationnelles, Réseaux Neuronaux Stochastiques, Théorie de la Complexité, Effets de Taille Finie, Connectome, Information de Fisher.

Thesis Organization

The structure of the thesis is as follows:

Chapter 1 provides an introduction to the structure and functions of the brain, in particular of the cerebral cortex, underlining its nested structure.

Chapter 2 explains two different points of view that are commonly used to describe the neural networks, namely the Theory of Complexity and the computational approach. Here we will also explain the difference between computation and information processing, showing how the brain accomplishes both.

Chapter 3 introduces the mean-field theory of a neural network, for both rate and FitzHugh-Nagumo neurons with independent Brownian motions, accompanied by analytic and numerical solutions of the mean-field equations.

Chapter 4 provides the first of the techniques used in this thesis to describe the finite size effects of the rate model, namely the Mayer's cluster expansion, and thereby extending the results found in Chapter 3. Moreover, it contains numerical simulations of the probability density of the network, together with a perturbative approach for the quantification of the Fisher information of the system.

Chapter 5 introduces a perturbative approach that will be used to determine the correlation structure of a finite rate network with correlated sources of randomness and for

different kinds of connectivity matrices. This approach extends the results of Chapter 4, exploring the possibility of applying the mean-field theory for the description of a neural network in the thermodynamic limit. This chapter provides also an algorithm to evaluate numerically the Fisher information of the system.

Chapter 6 contains another perturbative approach for finite rate networks with correlated sources of randomness, and it extends the analysis of the correlation structure started in Chapter 5 to the case of networks with general random topologies. In particular, it shows this technique applied to the case of a fractal connectivity matrix, which approximates the nested structure of the synaptic connections analyzed in Chapter 1. It provides also an analytic formula for the Fisher information of the system, which explains qualitatively the results found in Chapters 4 and 5.

Chapter 7 explores a numerical method based on the Expectation-Maximization algorithm and the Monte Carlo integration which allows us to determine the qualitative behavior of the Fisher information of a FitzHugh-Nagumo network, supporting the results found in Chapters 5 and 6.

Chapter 8 concludes the thesis, with a discussion of its findings, implications and relevance.

Software and Hardware

All the numerical results and figures in this thesis, with only the exception of Chapter 7, have been generated by Python Release 2.6 [3] and the libraries NumPy [4], SciPy [5], NetworkX [6] and matplotlib [7]. Instead the numerical results and figures of Chapter 7 have been generated by MATLAB and Statistics Toolbox Release $2008b$ [8] and the Fast EM_GM Algorithm [9].

All the simulations have been run on a laptop with an Intel®Core™2 Duo CPU $P9700$ at 2.80 GHz with 3.48 GB of RAM.

Présentation générale (en Français)

Le cerveau est le système connu le plus complexe de l'univers. Il est composé de milliards de neurones, dont les interactions et la coopération produisent les hautes fonctions cognitives de l'être humain, comme la pensée, le raisonnement, le discours, la reconnaissance de visage, ou encore les mouvements corporels complexes. La région du cerveau responsable des principales fonctions cognitives est appelée le *cortex cérébral*, que l'on peut diviser en aires spécialisées. Une des plus étudiées est le cortex visuel, chargé de la vision. De plus, chaque aire a une structure imbriquée: elle peut être fragmentée en régions de plus en plus petites exécutant des tâches plus spécifiques. Par exemple, le cortex visuel est formé de sous-aires consacrées à l'analyse de caractéristiques indépendantes des images, telles que le traitement des formes, des ombres, ou de la luminosité. Cette structure imbriquée a été observée dans les connexions synaptiques entre les neurones et représente la clé pour la compréhension des fonctions cognitives du cerveau humain. De plus le comportement d'un réseau neuronal peut varier fortement selon les connexions synaptiques en présence. Pour cette raison il est extrêmement important de décrire précisément leurs propriétés principales, telles que le nombre moyen de connexions par neurone et la distribution statistique de leur poids, pour reproduire des résultats biologiquement réalistes. Cet vision est conforme à la *Théorie de la Complexité*, établissant que pour un système composé de beaucoup de constituants interagissant (dans notre cas les neurones), il est

plus pertinent de décrire précisément les interactions que de définir d'une façon précise les propriétés d'un constituant. La modélisation de systèmes complexes requiert de nouvelles théories mathématiques, permettant d'analyser la coopération entre les constituants qui interagissent dans un but précis. Les colonies de fourmis, coopérant dans la construction de la fourmilière ou dans la recherche de nourriture, en est un exemple typique.

Une autre méthode importante qui est généralement utilisée afin d'analyser le cerveau est l'approche computationnel. Selon une des interprétations du cerveau, les réseaux neuronaux peuvent être pensés comme des systèmes exécutant une sorte de calcul, à la manière d'un ordinateur moderne. Cependant, ceci n'est pas la seule opération qui est exécutée par le cerveau. À vrai dire il est abondamment prouvé que le cerveau est lié avec le traitement de l'information. Les mots *calcul* et *traitement de l'information* sont souvent utilisés comme synonymes mais représentent en réalité deux concepts différents. Si le calcul est défini comme la manipulation de symboles, tel le code binaire utilisé par les ordinateurs modernes, le traitement de l'information est définie comme la modification d'informations, qui peuvent être de plusieurs types. Par exemple, dans le cas du cerveau les types les plus importants d'informations sont:

- l'information de Shannon, utilisée pour quantifier l'efficacité de la transmission d'informations entre des domaines cérébraux éloignés;
- l'information de Fisher, qui quantifie la capacité de codage du cerveau;
- l'information sémantique, liée à la signification des concepts qui sont stockés dans le cerveau.

Dans cette thèse nous présentons diverses techniques qui nous permettent de quantifier la structure corrélée de réseaux complexes de neurones, en mettant l'accent sur l'importance des connexions synaptiques. Nous utiliserons ensuite ces méthodes mathématiques afin de calculer analytiquement et numériquement l'information de Fisher du réseau, quantifiant la capacité de codage des modèles de neurone "rate" et "spiking". Malheureusement

une caractérisation largement acceptée par la communauté scientifique de l'information sémantique n'est pas disponible, donc la capacité du cerveau ne peut pas encore être étudiée en termes de traitement de l'information sémantique. A l'inverse la théorie de l'information de Shannon a des bases solides et a déjà été utilisé avec succès dans des domaines scientifiques variés. Toutes les quantités d'informations définies dans le contexte de la théorie de Shannon, comme les informations mutuelles et le transfert d'entropie, peuvent être facilement calculées avec les techniques mathématiques développées dans cette thèse, mais cette analyse est réservée pour une étude future.

Chapter 1

Introduction

THE human brain is the most complex system in the known universe. Its complexity, achieved by means of the cooperation of billions of neurons interconnected through the synapses, makes us able to think, to control our body, to reason with numbers, to use language, to memorize and recognize shapes and faces, and many other impressive skills which are unique to the animal realm. Moreover, within the animal kingdom, the human brain has the most neurons both in its whole volume and in the cerebral cortex. It is the latter, as we will see, which is responsible for the higher cognitive function of human beings. From the evolutionary point of view, both the complexity and size of the brain have increased throughout history in order to guarantee skills that are fundamental for the survival of the species. According to the *Triune Brain Theory* of P. D. MacLean [10][11], our current brain is the product of three different and cohabiting parts that emerged in the course of evolution: the *reptilian complex*, the *paleomammalian complex* (also known as the *limbic system*) and the *neomammalian complex* (namely the *neocortex*). The reptilian complex is the most ancient and primitive of the three: it is made up of the *brain stem* (composed of the superior part of the *spinal cord*, the *medulla oblongata*, the *pons* and the *midbrain*), the *diencephalon*, the *basal ganglia* and the *cerebellum*. It is the core of

the brain and derives from the *therapsids*, a form of mammal-like reptile that coexisted with dinosaurs and provided the evolutionary link between them and mammals at the end of the Triassic period. All the vertebrate animals, from reptiles to mammals, have the reptilian complex in common. The reptilian brain controls vital body functions that are fundamental for its maintenance, such as breathing, digestion and regulation of the body temperature, heart rate and blood pressure. It is also responsible for instinctive survival.

The second oldest part of the human brain is the paleomammalian complex, which emerged through evolution in the first mammals. It can be found in all mammals and is layered over the reptilian brain. It is composed of the *hippocampus*, the *septum*, the *limbic cortex* and the *amygdalae*. It is involved in emotions that ensure self-preservation, motivational functioning, behavior and biological drives, sense of smell, long term memory and many other functions.

Finally the neomammalian complex, the newest and most external part of the human brain, is represented by the cerebral neocortex, a structure that can be found only in mammals. It is responsible for our highest functions like language, abstraction, reasoning, imagination, conscious thought, sensory perception, etc.

All the three complexes are coordinated and communicate between them and depending on the different situations each one can dominate the others.

The neocortex is the most recent part of the cerebral cortex, and it is divided into six layers. The oldest part of the cerebral cortex is the *archicortex*, and has at most three layers. Moreover, together with the basal ganglia and the limbic system, the cerebral cortex forms the *cerebrum*, which is divided into two parts, known as *left and right hemispheres*, divided by the *corpus callosum*. The two hemispheres perform different functions. So for example the left hemisphere is responsible for speech and language, calculation, vigilance, writing and sequential processing. By contrast, the right hemisphere is devoted to spatial orientation and integration, facial and sound recognition and self-awareness. The cerebral

cortex is a layer of neural tissue that represents the outermost part of the cerebrum. It is conventionally divided into different parts with specific functions, according to alternative kinds of classification. An important example is the subdivision into lobes:

- *frontal lobe*: responsible for conscious thought, the planning of movements and problem solving;
- *parietal lobe*: involved in the integration of sensory information, speech and reading and object recognition;
- *temporal lobe*: engaged for speech (*Wernicke's area*), long term memory (hippocampus), but also auditory, visual and olfactory functions;
- *occipital lobe*: dedicated to visual processing.

Another important classification describes the cortex as structured into the following areas:

- the *primary sensory cortex*;
- the *primary motor cortex*;
- the *sensory association cortex*;
- the *multimodal association cortex*;
- the *motor association cortex*.

The primary sensory cortex is in turn divided into the following areas, dedicated to the perception of the senses:

- the *primary visual cortex*;
- the *primary auditory cortex*;

- the *primary olfactory cortex*;
- the *primary gustatory cortex*;
- the *primary somatosensory cortex*.

Even if *touch* is defined as one of the five senses, its perception is formed by the interplay of different sensory experiences, like pain, temperature, pressure, movement etc. These are called *somatic senses*, so this explains the origin of the name "somatosensory cortex". All these subareas are called "primary" because they represent the part of the cortex where the sensory information first arrives for a preliminary elaboration. Instead the primary motor cortex is the part dedicated to the execution of body movements through the muscles. The sensory association cortex is the collection of all the unimodal association areas, also called *secondary sensory areas*, namely the different parts of the cortex where information is further elaborated after the first processing that is performed in the primary sensory areas. In other terms, each unimodal association area performs a higher-order integration of the information that comes from the primary sensory areas. Moreover each unimodal association area is adjacent to its corresponding primary sensory area. The term "unimodal" refers to the fact that each association area further elaborates the information that comes from a *single* primary area and therefore it is devoted to the perception of a single stimulus (or sensory) modality, namely a physical phenomenon like smell, sound, touch or temperature. A simple example is represented by the primary visual cortex and its corresponding association area. In fact the primary visual cortex first receives the visual information from the eyes and extracts from it basic and independent features like edges, colors and shadows, while the corresponding association area combines these features in order to build representations of complex objects, like colored shapes, or to recognize faces.

Instead the multimodal association cortex integrates different unimodal sources of information creating coherent representations of complex objects and higher-order cognitive experiences. So while for example the visual association area performs the integration of

different kinds of visual information like color and orientation, the multimodal association area combines all the different kinds of sensory information, for example smell, color and sound. Therefore if we can perceive a complex object like a fire, it is due to the activity of the multimodal association cortex that integrates the visual information of the light intensity and the color of the fire, the olfactory information of the burnt smell, the auditory information of the crackling sound and the somatosensory information of the high temperature. All this information is combined together forming the coherent representation of what we know as *fire*. In other terms, the multimodal association cortex performs a higher-order integration of the information that comes from the sensory association cortex. The whole multimodal association cortex is then divided in three separated multimodal association areas: the *anterior*, *posterior* and *limbic cortices*. The anterior association area is situated in the *prefrontal cortex* (namely the anterior part of the frontal lobe), and is responsible for planning, reasoning, creating higher-order abstract concepts and also for the working memory. Instead the posterior association area is situated at the connection of the temporal, parietal and occipital lobes, and is responsible for written and spoken language (Wernicke's area), but also for perception, since it integrates the visual, auditory and somatosensory information that comes from the corresponding unimodal association areas. Finally, the limbic association area is found at the connection with the anterior part of the temporal lobe and the inferior part of the frontal lobe (this part of the brain is sometimes called *limbic lobe*), and is responsible for learning, emotions and memory.

To conclude, the motor association cortex is the part of the cortex devoted to the coordination of complex body movements. In contrast to the sensory areas, here the information flows from the motor association area (in particular its subpart known as *premotor cortex*) to the primary motor area. In turn, the premotor cortex receives information from the anterior association area. So, in summary, the cortex is a structured system that makes us able to receive sensory information from the environment, to modify it and then to use the result of this operation in order to move parts of our body accordingly. The schematic representation of this process is shown in Figure 1.1.

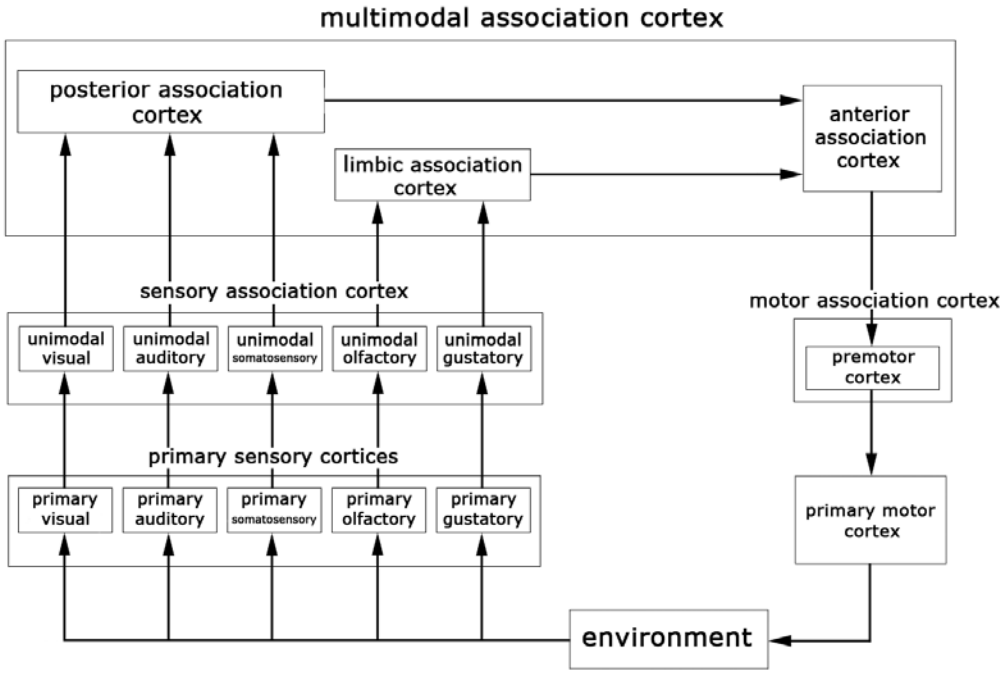

Figure 1.1: Simplified structure of the cerebral cortex. It may change slightly according to some authors, since there is no commonly accepted definition of the limbic association cortex.

Up to now we have described the subdivision of the brain into different macroscopic areas, that accomplish specific functions. However, if we want to understand the working principles of these areas, we have to analyze them more in detail, namely at smaller scales (in terms of spatial extension or number of neurons), finding out their building blocks. The macroscopic areas of the brain are typically of the order of $10^6 \div 10^9$ neurons and can be thought as made of many cooperating subsystems at an intermediate scale of $10^1 \div 10^5$ neurons. This is known as *mesoscopic scale* and represents the transition between the *microscopic scale* of the single neurons to the *macroscopic scale* of wide brain areas. In the cerebral cortex, at the mesoscopic level the neurons are vertically organized in structures known as *cortical columns* [12], in turn formed by many interconnected populations of neurons known as *neural masses* [13]. Even if the function of the cortical columns is not known yet, it is tempting to think about them as elementary computational blocks that cooperates in a parallel processing task.

So, to conclude this chapter, we underline the lesson that we have learnt from this analysis, namely that the brain has a *nested structure*, where each area contains other smaller areas with specific functions. This finding will be described from the mathematical point of view in Chapter 6.

Chapter 2

How can we describe the brain?

In this chapter we introduce two popular approaches that are commonly used in the description of the brain. The first is the Theory of Complexity, described in Section 2.1, which interprets the higher cognitive functions of the brain as emergent properties generated by the cooperative interactions of billions of neurons. Instead the second approach is introduced in Section 2.2, and is based on computation and information processing. The two words are often used as synonyms, even if they represent two distinct concepts. Their differences are explained in detail, underlining their corresponding roles and importance in the context of the brain.

2.1 The Theory of Complexity

As we have seen in Chapter 1, we can think about the brain as an incredible amount of neurons connected with each other in order to form structures at different scales with characteristic spatial extensions and specific computational functions, that are used as elementary blocks of higher-order structures with larger spatial extensions and that accomplish more complicated tasks. Increasing the complexity of the connections, starting from

the microscopic scale, passing through the mesoscopic scale and arriving up to the macroscopic scale, the computational power of the system increases at levels that let us acquire cognitive abilities. However, in nature several other systems are composed by thousands, millions or billions of interacting elements. Typical examples are matter, ant colonies, market economies and social structures, climate, ecosystems, living cells, genomes and the Internet. All these systems are studied in the field of the *Theory of Complexity* [14][15][16]. If the building elements of these systems interact in a random uncorrelated way, without accomplishing a specific result, according to W. Weaver [14] they are called *problems of disorganized complexity*, and this is the realm of chaos. A typical example is represented by gas molecules in a container. This kind of systems are studied with the tools of statistical mechanics and probability theory, developed from the 17th century to the early 20th century. Other kinds of systems, instead, interact with the aim of reaching a purpose, beyond the capabilities of the single element: these are the so called *problems of organized complexity* [14], and they are usually said to work at the transition point between order and chaos, the so called *edge of chaos* [17][18][19][20]. A characteristic example here is represented by ant colonies, where the ants collaborate in order to build the anthill, find the food, attack enemies, etc. Every ant seems to know what the others are doing and why. So the building elements of this kind of systems achieve collaboration or coordination through exchange of information, and the product is correlated behavior. For this reason the old techniques of statistical mechanics may not work anymore in this context. In this kind of systems there are phenomena which emerge from the interaction of their elementary blocks, or in other words very simple local interaction rules can generate highly complex global behaviors, that cannot be deduced easily from the properties of the individual parts. This is known today as *emergence*, a property usually described through the famous sentence "The whole is greater than the sum of its parts". A typical example of emergence is represented by flocks of birds. In fact, assuming simple interaction rules, i.e. flying in the same directions of the neighbors, with the same speed and avoiding obstacles or to bump into other birds, it is possible to recreate the flocks' ability to form stable

and complicated patterns and to rejoin when the group is split [21]. These properties are emergent because they are neither directly implemented in the system nor obvious consequences of the interaction rules. The formation of these patterns is one of the signatures of complex systems. Actually these systems seem the product of an intelligent design, and an important part of modern science is trying to understand how the organization they exhibit could be reached in an autonomous way, without any external intervention. This is the phenomenon known today with the name of *self-organization*. According to I. Prigogine [22], self-organization is achieved in open systems far from equilibrium through a reduction of the internal entropy of the system, at the expense of an increase of the entropy of the external environment. Therefore this process does not violate the second law of thermodynamics and allows the system to increase its internal order and degree of cooperation. One of the most studied self-organizing complex systems is the brain [23], where a number between a few hundreds (in the Caenorhabditis elegans) and a hundred of billions (in humans) of neurons collaborate through exchange of electric signals. Here the formation of spatio-temporal patterns, like for flocks of birds, can be observed and described theoretically, usually by means of the so called *neural field* models [24][25][26]. These patterns are emergent properties of the brain, a non-obvious consequence of the simple additive interactions realized by the synapses. Here self-organization is achieved through synaptic plasticity or learning, a phenomenon used by the brain to acquire, create and store new information from the environment. Learning is really fundamental for the brain since in the DNA of living beings there is not enough information to specify all the synaptic connections required for its correct functioning. Therefore the connections are molded using synaptic plasticity and the information that comes from the interaction with the environment. Actually, the behavior of a neural network can be completely different if we change its connectivity, and unfortunately the rules of synaptic plasticity are not always the same in the brain. In fact, for example, one of the main learning rules of the brain, called *Spike-Timing Dependent Plasticity* (or simply STDP) [27][28][29], changes enormously across different kinds of synapses and the regions of the brain [30]. How-

ever, in a mathematical model the synaptic connections and their strength could also be considered as static free parameters, and then it would be interesting to determine the differences in the behavior of the network corresponding to variations of the connectivity. The mapping of all the connections in the brain is called *connectome*, and it has been already completed for the Caenorhabditis elegans [31][32], and partially determined for the mouse [33][34][35], the rat [36][37], the cat [38][39] and the monkey [40]. Recently the project has been started also for humans [41][42][43]. Now, the basic topological properties of the brain of humans and other animal species have been quantified, in terms of clustering coefficients, degree distributions, path lengths, modularity, efficiency, connection densities, robustness, etc. [44][45][46], using the modern approach of *Graph Theory*. According to all these properties, the brain is a so called *small-world* network of neurons, namely a system that lies between the two extremes of completely regular and completely random topologies. The important lesson to learn is that the properties of a network are determined by its connectivity topology, and some of them, like the speed of propagation of the information, the wiring costs, the synchronizability and the computational power are enhanced in a small-world network [44][47]. Therefore it seems plausible that also the cognitive abilities are related to the topological structure of the brain and, according to the Theory of Complexity, they are emergent properties of the system. With this, we want to underline the importance of the synaptic connectivity in a neural network, therefore we believe that it could be more relevant to use simplified neural models (like the so called rate model [48][24][49][50][51]) with complex connectivities, than to use more biologically plausible neural models (like the Hodgkin-Huxley model [52]) with simple connectivities. This is in the spirit of Theory of Complexity, that says that the interactions are more important than the elementary components [53] in order to obtain an emergent behavior.

Up to now we have used the term “connectivity” to indicate the specification of all the pairs of neurons connected through the axons/synapses structure and of all the synaptic strengths. This is called more precisely *structural* or *anatomical connectivity*. However, in the literature there is also an intensive use of the term *functional connectivity*, that repre-

sents the specification of all the statistical dependencies (in terms of coherence or correlation) among the different regions of the brain. Today a big effort is devoted to the comprehension of the relation between these two kinds of connectivity [54][55][56][57][58][59][60], and in this thesis we will shed new light on the clarification of their link.

The number of neurons in the biggest of the brains is however much smaller than, for example, the number of water molecules contained in an ice cube. Nevertheless, in many species it is big enough to allow the use of special techniques already developed in physics, in order to study their behavior. For example, the so called *mean-field theory* has been used several times in the context of the brain [51][61][62][63][64][65]. The basic idea behind it, as we will see in Chapter 3, is the fact that, under appropriate assumptions, in the so called *thermodynamic limit* (namely when the number of neurons in the system becomes ideally infinite) the neurons interacting in a network become independent, a phenomenon called *propagation of chaos* (in order to avoid confusion, we clarify that here "chaos" simply means "independence", so it is not related to the definition of chaos in terms of the Lyapunov exponents). This allows a drastic reduction of the number of equations that describe the neural network and therefore represents a big simplification of the problem, but it does not let us determine the correlation structure (and therefore the functional connectivity) of the system since all the neurons become independent. For this reason, in Chapters 4, 5 and 6 we will develop new techniques that allow us to relax the hypothesis of infinite neurons and to study the neural networks under more general conditions.

2.2 Is the brain a computer?

So one approach toward the comprehension of the brain is represented by the Theory of Complexity, that through the interaction of many elementary units tries to explain higher-order cognitive functions as emergent properties of the system. Another approach is to study the brain in terms of *computation* and *information processing*. In fact there is plenty

of evidence that the brain is able to perform both the operations. These two words are often used as synonyms, but they represent actually two distinct concepts [66], whose differences are explained in the following sections. Often the brain is compared to a computer, and the mind to the software it executes. The subsequent analysis will help us to understand the accuracy of this analogy.

2.2.1 Computation

Computation is a difficult concept to characterize. However, it could be defined, to some extent, as *manipulation of symbols*. But what kind of computations is the brain able to perform? Computation can be classified into different categories, the most important of which are quantum and classical computation, deterministic and probabilistic, digital and analog, and semantic and non-semantic computation. In this section we introduce all these categories, explaining their main features and their relevance in the context of the brain.

Quantum computation

Every system is made of atoms and molecules, that are described by the laws of quantum mechanics. However, when a system has a macroscopic size, it is almost always possible to describe it using the laws of classical physics. There are special cases of macroscopic systems that exhibit quantum properties, like superfluidity and superconductivity, but these represent special exceptions, that emerge only at extremely low temperatures artificially created. Not only the brain, but also single neurons are made of so many atoms that they can be considered as macroscopic objects. Even if the evaluation of the brain's temperature is a poorly studied issue, its value is plausibly around the mean body's temperature, namely 37°C. Therefore it seems that there is no possibility to see any kind of quantum effect in its behavior [67]. Nevertheless, a minority group of researchers is trying to describe the working principles of the brain in terms of quantum mechanics. The most

famous quantum theories of the brain are the so called *Quantum Brain Dynamics* [68][69] and the *Orchestrated Objective Reduction* (Orch-OR) theory developed by R. Penrose and S. Hameroff [70][71][72], based on microtubules. The main motivation to use quantum mechanics in the description of the higher cognitive functions of the brain is the so called *mind-body problem*, namely the role that the observer's consciousness covers in the process of measurement according to some interpretations of quantum mechanics. However the idea of a quantum brain has been recently revalued due to the discovery of quantum effects in biological systems [73][74][75], and this has led to the birth of a new research field known as *quantum biology* [76]. The preservation of quantum effects in macroscopic systems at high temperatures can be explained through the so called *quantum Zeno effect* [77]. If the brain could use the laws of quantum mechanics, it would be possible to describe it as a sort of parallel *quantum computer*, namely a machine that performs computations on quantum states, that in this context are known as *quantum bits* or *qubits* [78]. However, even if quantum computers have been proved to be more powerful than classical computers for some kinds of tasks, for most problems they are only modestly faster than their classical counterpart. Efficient quantum algorithms have been found only for specific tasks like integer factorization [79], and these kinds of computations are not enough for accomplishing the complex functions of the brain. Moreover P. Shor has observed that only a few efficient quantum algorithms have been found [80][81]. A possible explanation may lie in their counter intuitiveness, that makes them hard to discover. However this raises also the second possibility, namely the fact that maybe quantum algorithms exist *only* for tasks which are relatively simple or not relevant in the context of neuroscience . In this case the quantum brain hypothesis should be abandoned definitely.

Classical computation

As we said previously, even if every system obeys the laws of quantum mechanics, it is possible to describe them using classical physics if they have a macroscopic size and a

high enough temperature. Therefore classical computation refers simply to the manipulation of symbols that are physically realized by classical quantities, namely computations performed on states that are not quantum superposed. Every personal computer is an example of machine that executes classical computations, manipulating (binary) classical symbols known as *bits* (namely the classical version of the qubits). In this thesis we will describe neurons and synapses using the laws of classical physics, therefore we suppose that neglecting every quantum effect inherited from their underlying atomic structure will not affect the validity of the results obtained with our models.

Deterministic and probabilistic computation

Given infinitely many repetitions of the same computation performed by a machine starting from the same initial condition and receiving the same input, we say that the computation is deterministic if the result is always the same for every repetition. Instead, if the result may change, the computation is probabilistic or non-deterministic. Typically, sufficiently strong sources of noise can transform a deterministic computation into a probabilistic one.

Digital computation

Digital computation is the manipulation of symbols represented by digital quantities, namely discrete variables. The modern computers are digital calculators based on the binary code, namely the manipulations of symbols that consist in combinations of zeros and ones. These numbers are discrete variables, that are usually implemented through physical quantities, like electrical currents or voltages. One of the main advantages of digital computation is its error tolerance against noise. In fact here the manipulated symbols are discrete, therefore the "distance" or difference between two of them is finite. Therefore a small amount of noise in general is not able to convert one symbol into another one. This

problem instead affects analog computers, since they work with continuous quantities and therefore the difference between two "adjacent" symbols is infinitesimal.

The most powerful device that performs digital computations is the so called *Turing machine* [82] (or its equivalent models). In this context, the famous *Church-Turing thesis* [83], in one of its various equivalent formulations, is a statement about the computational power of the Turing machine. In fact, according to this conjecture, any real-world computation (i.e. any computation that can be described by a step-by-step procedure, known as *algorithm*) can be carried out by a Turing machine. Notwithstanding, scientists have tried to build calculators even more powerful than the Turing machine (in the sense that they can compute quantities that a Turing machine cannot compute), and this defines the field of *hyper-computation* or *super-computation*. In other words, they have tried to invalidate the Church-Turing thesis. However, after 70 years of research from the definition of the Turing machine, it has not yet been discovered a physically implementable machine with more computational power than the Turing machine itself.

Moreover, a Turing machine can be deterministic or probabilistic. However, it can be proved that a probabilistic machine can be simulated by a deterministic machine, therefore they are equivalent, in terms of what can be computed. This does not mean that the time they need in order to perform a given task is the same.

In the theory of computation there is a branch known as *Computational Complexity Theory*, which must not be confused with the Theory of Complexity introduced in Section 2.1. In this context, computational scientists have defined different classes of complexity for the decision problems, and one of the most fundamental is called *non-deterministic polynomial time* class, abbreviated as NP. A decision problem belongs to the NP class if it can be solved by a *non-deterministic* Turing machine in polynomial time. If the decision problem is described by an algorithm whose input has size n, and defining $T(n)$ as the maximum amount of time taken by the algorithm on any n, saying that the decision problem is solved in polynomial time means that $T(n) = O\left(n^k\right)$, for some constant natural

number k. Another fundamental class in the Computational Complexity Theory is known as *deterministic polynomial time* class, abbreviated as P, which contains all decision problems that can be solved by a *deterministic* Turing machine in polynomial time. Another important class is that called *non-deterministic polynomial time hard* class, abbreviated as NP-hard, which informally is defined as the class containing the decision problems that are at least as hard as the hardest problems in NP, and that do not necessarily belong to NP itself. We conclude by defining the *non-deterministic polynomial time complete* class, abbreviated as NPC, namely the set of the decision problems which belong to both the NP and NP-hard classes. Therefore both P and NPC are contained in NP. Now, since it can be proved that the NPC problems can be reduced to each other in polynomial time, and that all the NP problems can be reduced to NPC problems in polynomial time, then given any NP-hard problem, all the NP problems can be reduced to it in polynomial time. Therefore, if we found a polynomial time solution for any NP-hard problem, it could be used to solve all the NP problems, proving that $P = NP$. However it is widely believed that $P \neq NP$.

To conclude, other kinds of computing devices, known as *finite state machines*, even if less powerful than the Turing machine, are successfully used to study important cognitive functions of the human brain, like natural languages [84].

Analog computation

Analog computation refers to the manipulation of symbols represented by analog quantities, namely continuous variables. The first computers ever invented were actually analog and they manipulated continuous mechanical, electrical, or hydraulic quantities. They were used especially during the World War II, but one of their main problems was the low performance in computations performed with high levels of noise. In fact analog quantities can vary considerably if the intensity of the noise is not negligible and therefore this affects the precision of the computation. However recently analog computers have

been revalued due to the work of L. Rubel, who in 1993 introduced the so called *Extended Analog Computer* (EAC) [85]. The Rubel's idea was to expand the capabilities of another analog computer, known as *General Purpose Analog Computer* (GPAC), invented by C. Shannon in 1941 [86]. According to Rubel, the GPAC was not able to solve problems considered fundamental for the simulation of the human brain, and therefore an extension was required. So the importance of the EAC is its attempt to introduce new paradigms of computing, that may shed new light on the working principles of the brain, beyond digital computers. The EAC has been physically realized with different primary computing elements, like gelatin [87], a sheet of conductive plastic foam [88] or slime mold [89]. It performs computations using the electrical, mechanical, chemical and heat properties of these materials, operating by *analogy* with the system that represents the problem to solve. In other words, the problem to solve is described in terms of ordinary or partial differential equations, and then the EAC is brought into congruence with this problem by mapping the variables of the differential equations into its hardware. Therefore at a first level the EAC is configured (not programmed, as for a digital computer) in order to copy the mathematical description of the problem to solve in terms of differential equations, and then it solves these equations at a second level using other analog devices, like adders, differentiators, multipliers, inverters etc. So the basic idea of the EAC is to solve a problem using physical processes that obey the same differential equations of the problem itself. For the EAC, analogy covers the same role that the algorithm has in the context of digital computers. Even if it has not been proved yet if it is more or less powerful than digital computers (and this is true in general for every kind of analog machine), it has some interesting advantages over them [90]:

- It is not affected by the so called *von Neumann bottleneck* [91], also known as *memory wall* [92], namely the increasing difference of speed between the digital CPU and the memory outside it. In fact the EAC does not use any memory, since it performs its tasks using directly the matter of its computing element.

- The *Moore's law* [93], namely the observation that the number of transistors on integrated circuits has doubled every two years since their invention, does not apply on it since the logic units of the EAC are atoms and molecules.
- Electromagnetic interferences generate sequences of random numbers, while digital computers can only create pseudo-random numbers (in general periodic sequences with very large periods).

Another example of unconventional analog machines are the so called *chemical computers*, that performs computations using reaction-diffusion processes [94][95][96]. This is related to the field of *collision-based computation* [97]. To conclude, also artificial neurons of second and third generation are analog computing devices [98]. The first generation of artificial neurons was introduced in 1943 by W. McCulloch and W. Pitts [99]. They have a binary output therefore they are classified as digital units. So if the output is 0, this is interpreted as if the neuron were not spiking, while a 1 as if it were spiking. Instead the second generation of artificial neurons is characterized by neural models with an analog output, generated by a so called *activation function*. This function converts the state of a neuron into an output that can be interpreted as the firing rate or frequency of the spikes generated by the neuron itself. Therefore the output of a neuron of second generation provides more information than a neuron of first generation, because it tells us not only if the neuron is spiking or not, but also at which frequency. Finally we have the neurons of third generation, which are the most realistic from the biological point of view. In fact their output is a spike train, namely a sequence of spikes at increasing time instants, as in biological neurons. Therefore the output of a neuron of third generation provides more information than a neuron of second generation, because it tells us the time instants of the spikes, from which we can calculate the firing rate, while the inverse operation is not possible, namely to obtain the time instants from the firing rate. A single spike is usually described as a continuous variable (the so called *presynaptic potential*), therefore a neuron of third generation is another example of model with analog output.

For a complete review of analog computation, see [100].

Now we open a parenthesis about the super-computational power of analog computers. In fact in this field scientists have discovered devices that can perform super-computations, but these are still based on idealized assumptions, that are not physically realistic. For example, in [101][102], namely in the context of recurrent analog neural networks, the use of dynamic rational-valued or, more generally, static or dynamic real-valued synaptic weights gives super-computational power to the system. This seems to be due to the infinite precision of the synaptic weights, but such a system cannot be realized at the physical level due to its intrinsic noise [103]. In general, computation on infinite-precision real numbers is called *real computation* [104][105]. It is important to observe that if real computation were physically realizable, NPC problems could be solved in polynomial time, therefore we would conclude that $P = NP$ [106]. However, up to now scientists have not found a polynomial-time algorithm for any of the known NPC problems, answering probably negatively the $P = NP$ question. Other examples of super-computation require infinite time [107] or infinite neurons [108][109], therefore there is always a requirement that is not physically realizable. Therefore the conclusion seems to be that there is no physically realizable classical machine with more computational power than the Turing machine. The same conclusion can be obtained for quantum computers. In fact they can be simulated on a Turing machine, therefore they are not computationally more powerful than the latter.

Semantic and non-semantic computation

Semantic computation is the manipulation of symbols (digital or analog, classical or quantum) based on their meaning. A typical example is natural language, where the manipulated symbols are the words taken from the vocabulary. So while for example a digital computer manipulates binary strings like "011001101" regardless their eventual meaning (this is an example of non-semantic computation), a person manipulates symbols like "cat" according to the meaning that these words have in his/her head. The concept of *meaning* of

a symbol has been the topic of a long philosophical discussion, which has been renewed after the introduction of the *Chinese room argument* by J. Searle [110]. In fact it is not clear how the words acquire this sort of "metaphysical" property in our minds: this is called *symbol grounding problem* [111]. So the symbol grounding is the process through which the words acquire their meaning, and only recently it has been declared as solved [112].

In the context of the brain, the part which is responsible for the storage of meanings and concept-based knowledge is called *semantic memory*, which is often modeled by means of *semantic networks* and *spreading-activation theory* [113][114]. In a semantic network, the nodes represent *concepts*, while the edges are the *semantic relations* between them. There is an historical debate about the definition of "concept", which started with the classical theory of Aristotle. However often concepts are described through *categorization* and implemented in *attractor neural networks* [115][116]. To conclude, according to E. Tulving, semantic memory is responsible for the cognitive ability that he has called *noetic consciousness* [117], underlining its importance among the higher cognitive functions of the brain.

2.2.2 Information processing

Information processing can be defined as the *modification of information* performed by a physical system. However the word "information" has different meanings, therefore it is not possible to give a single definition that accounts for all the cases. In this section we introduce the three main kinds of information involved in the brain, namely Shannon, Fisher, and semantic information.

Shannon information

This is probably the most studied kind of information, due to the rigorous definition of information given by C. Shannon in the context of telecommunications [118]. So we have

to consider a discrete memoryless source of symbols, described by a random variable X, which can be equal only to the values x_i taken from the set $\mathbb{X} = \{x_i\}_{i=0,1,\ldots,m-1}$ with probability $P(x_i)$. According to Shannon, the *uncertainty* of X, also called the *entropy* of the source, is

$$H(X) = \mathbb{E}_{P(x)}[-log_b P(x)] = -\sum_{i=0}^{m-1} P(x_i)\, log_b P(x_i) \tag{2.1}$$

The base b of the logarithm determines the unit of the entropy. In particular, the unit is called *bit* for $b = 2$, *trit* for $b = 3$, *nat* for $b = e$ (the Napier's constant), and *ban* for $b = 10$. If the communication channel is noisy, even if the source is X, the receiver observes a signal Y, due to the interaction of X with the noise of the channel. We suppose that Y can be equal only to the values y_j taken from the set $\mathbb{Y} = \{y_j\}_{j=0,1,\ldots,n-1}$ with probability $P(y_j)$. So the initial uncertainty of the source is $H(X)$, but after a channel output y_j has been observed by the receiver, the new value of uncertainty of the source is:

$$H(X|Y = y_j) = -\sum_{i=0}^{m-1} P(x_i|y_j)\, log_b P(x_i|y_j)$$

Therefore the amount by which the uncertainty of the source has been reduced, after the channel output y_j has been observed, is:

$$I(X, Y = y_j) = H(X) - H(X|Y = y_j)$$

If we take the mean of $H(X|Y = y_j)$ over all the possible values of Y, we obtain the so called *conditional entropy* $H(X|Y)$:

$$H(X|Y) = \mathbb{E}_{P(y)}[H(X|Y = y)] = -\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} P(x_i, y_j)\, log_b P(x_i|y_j) \tag{2.2}$$

having used the relation:

$$P(x_i, y_j) = P(x_i|y_j)\, P(y_j)$$

Now, in the same way, if we want to calculate the mean reduction of uncertainty, we have to take the average of $I(X, Y = y_j)$ over all the possible values of Y, obtaining:

$$I(X, Y) = \mathbb{E}_{P(y)}\left[I(X, Y = y)\right] = H(X) - H(X|Y) \tag{2.3}$$

$I(X, Y)$ is known as *mutual information*, and is given by the difference of two entropies. In other terms, here information is interpreted as *reduction of uncertainty*, namely the reduction of $H(X)$ by the amount $H(X|Y)$. Formula 2.3 has been used in [17] in order to quantify the amount of cooperation between two cells of a cellular automaton, therefore it could be used with the the same purpose in the context of the brain. Using Shannon's theory of information, it is possible to define other quantities, that determine three different kinds of information processing, namely *information storage*, *transfer* and *modification*. This kind of analysis has already been performed for cellular automata [119][120][121][122], but not in the case of neural networks.

All the quantities defined above can also be extended to the case of continuous systems. In this case the probabilities $P(x_i)$, $P(y_j)$ and $P(x_i, y_j)$ must be replaced with probability densities $p(x)$, $p(y)$ and $p(x, y)$, respectively. Therefore the entropy and the conditional entropy become:

$$h(X) = -\int_{\mathbb{X}} p(x)\, log_b p(x)\, dx \tag{2.4}$$

$$h(X|Y) = -\int_{\mathbb{X}}\int_{\mathbb{Y}} p(x, y)\, log_b p(x|y)\, dxdy \tag{2.5}$$

In particular, $h(X)$ is known as *differential entropy*. However, even if formula 2.5 is well defined, unfortunately formula 2.4 is not [123]. This can be seen clearly with an example. In the discrete case, for a deterministic source which selects always the symbol $X = x_{\overline{i}}$, we have $P(x_{\overline{i}}) = 1$, while all the other symbols have probability zero. From formula 2.1 we obtain therefore that $H(X) = 0$, as suggested by intuition. Instead, in the continuous case, a deterministic source which selects always a symbol $\overline{x}$, has probability density $p(x) = \delta(x - \overline{x})$, where $\delta(\cdot)$ is the Dirac delta function. If we see $p(x) = \delta(x - \overline{x})$ as the limit for $\sigma \to 0^+$ of $p(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\overline{x})^2}{2\sigma^2}}$, we obtain:

$$h(X) = \lim_{\sigma \to 0^+} \left[-\int_{-\infty}^{+\infty} p(x, \sigma)\, log_b p(x, \sigma)\, dx \right] = \lim_{\sigma \to 0^+} \left[log_b \left(\sqrt{2\pi}\sigma \right) + \frac{1}{2} log_b e \right] = -\infty$$

for $b > 1$, therefore clearly formula 2.4 cannot be correct. For this reason in some cases the differential entropy is replaced by the so called *entropy power* $N(X)$ (see for example [124]), defined as:

$$N(X) = \frac{1}{2\pi e} e^{2h(X)}$$

All these quantities can also be defined for a general number of dimensions. To conclude, we observe that they are function only of the probability of the process, regardless the meaning of the symbols. For this reason Shannon's theory is able to quantify only syntactic information, not the semantic one.

Fisher information

If Shannon's theory is able to quantify information storage, transfer and modification, the theory developed by R. Fisher in the 1920s deals with the quantification of *information encoding*. The value of a stimulus could be encoded in a neural network through the so

called *population coding* [125]. According to this hypothesis, the same stimulus is provided to many neurons, which form a neural population, and due to different sources of variability, like noise [126] or differences in the synaptic weights, the cells exhibit different behaviors. From this collection of inhomogeneous activities, it is possible to estimate the value of the stimulus, using the joint probability density of the population. The Fisher information measures the amount of information that an observable variable $\overrightarrow{X}$ (in our case the collection of the neural activities of the population), carries about a parameter θ of the system (in our case the stimulus provided to the population). In detail, the Fisher information for a continuous system is defined as:

$$\mathcal{I}(\theta) = \int_{\mathbb{X}} \left(\frac{\partial}{\partial \theta} \log p(\overrightarrow{x}, \theta) \right)^2 p(\overrightarrow{x}, \theta) \, d\overrightarrow{x} \tag{2.6}$$

where $\mathbb{X}$ is the set of all the possible values $\overrightarrow{x}$ of the variable $\overrightarrow{X}$. Formula 2.6 can also be extended to the case of multiple parameters θ_i, but in this thesis we will consider only a single parameter, represented by the external input current of the network. Given any *unbiased estimator* $\widehat{\theta}\left(\overrightarrow{X}\right)$ of the stimulus θ through the samples of $\overrightarrow{X}$, namely any estimator such that:

$$\int_{\mathbb{X}} \widehat{\theta}(\overrightarrow{x}) \, p(\overrightarrow{x}|\theta) \, d\overrightarrow{x} = \theta, \qquad \forall \theta,$$

the quantity $\mathcal{I}(\theta)$ is an information in the sense that:

$$Var\left(\widehat{\theta}\right) \geq \frac{1}{\mathcal{I}(\theta)}$$

This inequality is known as *Cramér–Rao bound*, and proves that the precision to which we can estimate the stimulus θ is limited by the inverse of $\mathcal{I}(\theta)$. So the higher the Fisher information is, the better the stimulus θ is encoded by $\overrightarrow{X}$.

In this thesis we deal only with the Fisher information, reserving the other kinds of information processing for a future study.

Semantic information

If we define semantic information processing as the modification of semantic information, we are faced to a problem: semantic information has not been quantified yet in a rigorous way [127][128][129][130]. For this reason semantic information processing cannot be defined yet in an unambiguous way. However this thesis does not deal with semantics, therefore this lack will not affect the analysis of the neural networks performed in the next chapters.

2.2.3 The human mind in terms of computation

In the previous sections we have seen that computation and information processing are two distinct concepts, and that they are both required for a complete comprehension of the brain. Notwithstanding, in the course of history, scientists have usually associated the highest cognitive functions of the brain to computation. In fact, in the field of cognitive science, namely the scientific study of the mind, the prevailing idea was that the mind is originated by a serial manipulation of symbols. This is the so called *computer metaphor of the mind* [131], while this current of thought is called *symbolism*. In philosophy of artificial intelligence, this point of view is called *physical symbol system hypothesis*, formulated by A. Newell and H. A. Simon in [132], while in philosophy it could be roughly identified with the *computational theory of mind* of H. Putnam and J. A. Fodor [133][134][135]. A typical example that supports the symbolist current of thought is human language, because finite state machines have been successfully applied to its modelization [84]. Language is considered one of the most important human cognitive abilities, and since it can be described

by finite state machines, which perform serial computations, then this is a strong evidence in favor of symbolism. Historically, this point of view has been criticized especially by *connectionism*, that supports the idea of the mind as an emergent property generated by interconnected neurons, or in other terms by parallel *subsymbolic* computations [136]. Moreover symbolism has opted for discrete representations of the symbols, while connectionism prefers analog distributed representations. It is also important to observe that the two hemispheres of the brain seem to behave in a serial or parallel way according to the task performed [137][138][139]. Therefore these two currents are in contradiction with each other, but the theory developed by P. Smolensky [140] seems able to incorporate the symbolic approach into the connectionist one, without losing the qualities of the latter. This idea has been used by P. beim Graben and R. Potthast in [141] to define a theory they have called *dynamic cognitive modeling*. Therefore dynamic cognitive modeling consists in the implementation of symbolic cognitive serial operations into a neural network or a neural field equation. A similar idea is expressed also in [142], in the context of the Turing machine, while the work of beim Graben and Potthast is based on less powerful devices (like the pushdown automaton and the finite state machine), since according to them "the Turing machine is of only marginal interest in cognitive psychology and psycholinguistics" [141]. So dynamic cognitive modeling is an implementation of ideas from cognitive science into a neural network or a neural field model, and therefore this is a strong cooperation between cognitive science and neuroscience. In other words, dynamic cognitive modeling uses the ideas of cognitive science about the nature of the processes involved in cognition, and it implements these ideas in a biologically realistic substrate. Since in this approach the connectionist ideas are not neglected, this theory may be able, at least in principle, to explain how the cognitive functions generated by the serial manipulation of symbols arise as emergent properties of interconnected neurons.

The concept of a super-Turing brain can be rejected not only according to the explanation provided in Section 2.2.1, but also according to the following idea, essentially expressed in [143][144]. Quantum mechanics can be simulated by a classical Turing machine up

to a desired level of precision. Moreover quantum mechanics describes every physical phenomenon down to the atomic and molecular level, so we can think to simulate in a very big Turing machine all the atoms and molecules that compose all the neurons, axons and synapses in the brain. This simulation will describe and reproduce the brain at an extremely high level of precision, and therefore also its working principles, or in other words the way it performs computations. So, according to this idea, the brain can be computed by a Turing machine, therefore it cannot be more powerful than the Turing machine itself. Notwithstanding, there is still some belief that the incredible power of the brain could come from some super-Turing computation or also from non-algorithmic processes. This is due to the famous halting problem, namely the inability of the Turing machine to determine in general if a program will stop running or if otherwise it will run forever. It is important to clarify that the Turing machine is able to determine if some algorithms will halt or not, but it cannot establish it *for all* the possible algorithms that can be written. For this reason we say that it is not able to solve the halting problem. However there is some belief that this task can be accomplished by the brain, even if it is actually still unclear. In fact a programmer usually can look at a program or algorithm and tell if it will halt or not, but this may be not always true. For example, nowadays nobody is able to say if the following program will ever halt:

```
define_function: Is_the_number_a_sum_of_2_primes(integer)
 {
  ... // Body of the function
 }

n = 4
result = yes

while(result = yes)
 {
  n = n + 2
  result = Is_the_number_a_sum_of_2_primes(n)
 }

exit
```

This algorithm is the brute force implementation of the famous *Goldbach's conjecture*, one of the open problems in number theory. It states that every even integer greater than 2 can be decomposed as the sum of two prime numbers. Since this problem is still unsolved, nobody can say if the program above will ever halt or if it will loop forever. Maybe one day someone will prove this conjecture, or maybe this task is simply beyond the capabilities of the human brain. The same reasoning can be applied for example to any of the currently unsolved conjectures in number theory. As we said, a computational system can solve the halting problem if it looks at any program and is always able to tell if it will halt or not. In other words, if there is at least one case where the system cannot give the answer, then this system is not able to solve the halting problem. Now, since currently no person can look at the brute force implementation of the Goldbach's conjecture and tell us if it will ever halt, it is not clear if the human brain can solve the halting problem. Therefore we conclude that there is no proof of a super-Turing computational power in our brains. Actually this is unsurprising, because if there were a mechanical procedure to establish if a generic program would halt or not, then many hard mathematical problems would be easily solvable.

2.2.4 Information processing in the brain

So much effort has been devoted to the comprehension of the human brain and its higher emerging product, the mind, in terms of computation. Instead a detailed analysis of the information processing capability of the brain is missing. Up to now, information analysis has been limited only to the case of the Shannon and Fisher information, because, as we said, there is still a debate on the quantification of the semantic information. Shannon information has been used in the context of neural learning [145][146][147], for the quantification of the efficiency of the neural code [148][149][150][151][152][153][154][155], and

in visual perception [156][157]. Instead Fisher information has been evaluated in networks with specific correlation structures and its behavior could be completely different from case to case [158]. A complete analysis of the behavior of these information quantities for different kinds of connectivity matrices has not been performed yet, even if the theory of complexity stresses the importance of the synaptic connections (see Section 2.1). In this thesis we develop a mathematical formalism to fill the gap, and we apply it explicitly to the calculation of the Fisher information. Because of space and time requirements, we have not used these techniques for evaluating the Shannon information, but this calculation is straightforward and left for a future study.

To conclude this section, we observe that Shannon and Fisher information quantities are functions of the probability density of the system. An important theory that can be used to determine these probability densities and that is of particular relevance in the context of the brain is the *Bayesian Theory* [159][160]. This theory is based on the Bayes formula, according to which, given two random variables X and Y, their conditional probabilities are related as follows:

$$p\left(x|y\right) = \frac{p\left(y|x\right)p\left(x\right)}{p\left(y\right)} = \frac{p\left(y|x\right)p\left(x\right)}{\int_{\mathbb{X}} p\left(y|x\right)p\left(x\right)dx}$$

supposing that $p\left(y\right) \neq 0$. Interpreting x as an *hypothesis* to test, for example that "the distance of an object is 10 meters", and y as the *data* (also called *evidence* in this context) about the object that we receive from our senses, like its image through the visual cortex and its sound through the auditory cortex, then the Bayesian Theory tells us how to update our previous belief about the hypothesis, namely $p\left(x\right)$, also known as *prior probability* (namely the knowledge we have about x, before receiving any y), into the probability $p\left(x|y\right)$, also known as *posterior probability* (namely our knowledge of x after, or given, the reception of y). Instead the probability $p\left(y|x\right)$ is called the *likelihood* of sensing y given the hypothesis x. To conclude, $p\left(y\right)$ is called *evidence* (or *marginal likelihood*), because it represents the prior probability of the evidence y. Now, since y represents the data arriving

from different areas of the brain, the Bayesian Theory describes the sensory integration performed by the multimodal association cortex introduced in Chapter 1. This underlines its importance in the context of the brain.

2.2.5 Partial conclusion

After this complicated analysis, the conclusion is that the question "Is the brain a computer?" does not have a univocal answer. Clearly the brain is not like a digital computer with a von Neumann architecture, but it has some similarities with analog computers. Moreover, the brain is involved with information processing, which does not necessarily take place in a digital computer (if we define information processing as in Section 2.2.2). Notwithstanding, if we follow the point of view of *materialism*, since the brain is made up of atoms, then it has to follow the laws of physics, therefore it can be interpreted as some sort of machine, even if probably a modern computer does not represent the best basis for comparison. Therefore there must be no mystery in the formation of the cognitive functions of the brain, because they must emerge at some point in a natural way when the complexity of the synaptic connections is high enough.

Chapter 3

Mean-field analysis of some neural networks

THIS chapter is devoted to the mean-field theory of neural networks. In particular, in Section 3.1 we describe the system using rate equations and in Sections 3.1.1 and 3.1.3 we derive heuristically the mean-field limit in two different ways. We also obtain the mean-field Fokker-Planck equation of the network, and in Section 3.1.4 we show how to calculate its solution analytically. In Section 3.2 we describe the network using the FitzHugh-Nagumo model. In Section 3.2.1 we show its corresponding mean-field Fokker-Planck equation and in Section 3.2.2 we propose an algorithm to solve it numerically. To conclude, in Section 3.3 we discuss briefly the analysis of the Fisher information in the thermodynamic limit.

3.1 The rate model

In this section we describe a neural network using the following system of stochastic differential equations:

$$dV_i(t) = \left[-\frac{1}{\tau} V_i(t) + \sum_{j=0}^{N-1} J_{ij}(t) S(V_j(t)) + I(t) \right] dt + \sigma_1 dB_i^V(t) \tag{3.1}$$

$$V_i(0) \sim \mathcal{N}(\mu, \sigma_2) \tag{3.2}$$

with $i = 0, 1, ..., N-1$, where:

- N is the number of neurons in the network;
- $V_i(t)$ is the membrane potential of the i-th neuron;
- τ is a time constant that describes the speed of convergence to a stationary state;
- $I(t)$ is the deterministic external input current, which is supposed to be the same for all the neurons (because we want to avoid inhomogeneities which are difficult to describe analytically);
- $B_i^V(t)$ is the Brownian motion that describes the background noise of the i-th neuron (or equivalently the stochastic part of the external input current);
- σ_1 is the standard deviation of the Brownian motions, which is supposed to be the same for all the neurons and time-independent;
- μ is the mean of the initial conditions;
- σ_2 is the standard deviation of the initial conditions;
- $J_{ij}(t)$ is the random synaptic weight from the j-th neuron to the i-th neuron;
- $S(\cdot)$ is a function that converts the membrane potential of a neuron into the rate or frequency of the spikes it generates.

Usually in neuroscience $S(\cdot)$ is a sigmoid function, defined as:

$$S(V) = \frac{T_{MAX}}{1 + e^{-\lambda(V - V_T)}} \tag{3.3}$$

or an integral function, defined as:

$$\begin{aligned} S(V) =& T_{MAX} E(\lambda(V - V_T)) \\ E(V) =& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{V} e^{-\frac{V'^2}{2}} dV' = \frac{1}{2}\left[1 + erf\left(\frac{V}{\sqrt{2}}\right)\right] \end{aligned} \tag{3.4}$$

where T_{MAX} is the maximum amplitude of the function (which is reached for $V \to +\infty$), λ is the parameter that determine its slope for T_{MAX} fixed, while V_T represents the horizontal shift of the function along the V axis. Now, many different kinds of connections can be used, which typically are studied by the branch of mathematics known as *Graph Theory*. In this context, a graph whose adjacency matrix is circulant is called *circulant graph*. It is usually represented by the notation $C_N(1, 2, ..., q)$, which means that the i-th node in the graph (for $i = 0, 1, ..., N-1$) is connected to the $(i-j)$-th and $(i+j)$-th nodes (mod N) for each j in the list $(1, 2, ..., q)$. Special cases are the *cycle graph* $Cy_N = C_N(1, 0, 0, ..., 0)$, where every neuron is connected only with other two (therefore its connectivity matrix is tridiagonal with corner elements) and the *complete graph* $K_N = C_N\left(1, 2, ..., \left\lfloor \frac{N}{2} \right\rfloor\right)$ (this is the case of the fully connected network). Usually in graph theory they are represented as undirected unweighted graphs (see Figure 3.1), namely with non-directional connections (their graphs are drawn using connections without arrows) and identical weights. This means that their connectivity matrix is supposed to be symmetric. However in this section we can consider directed graphs as well, therefore the connectivity matrix in general is not symmetric.

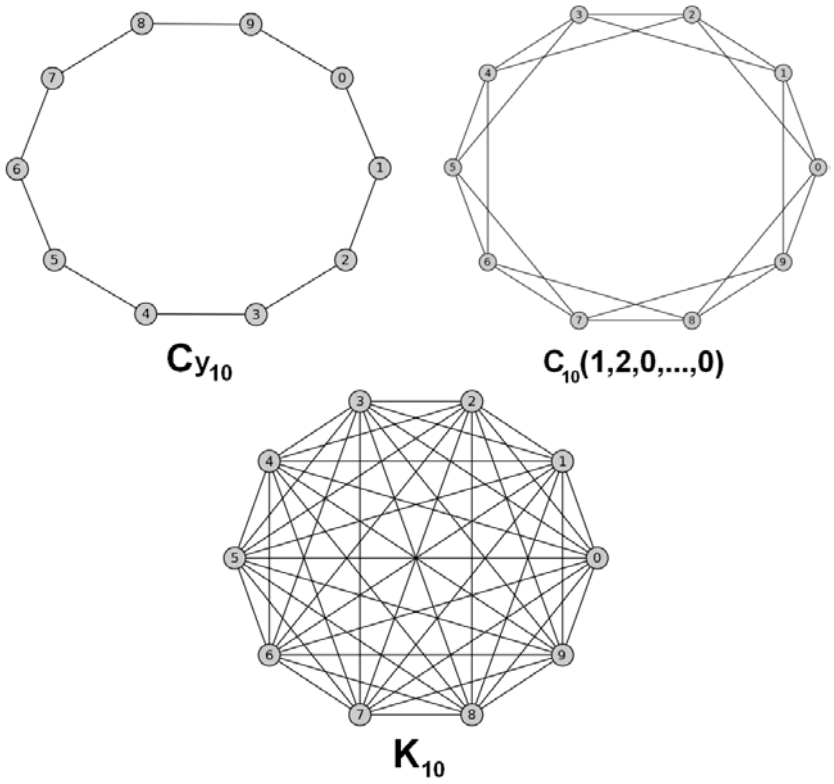

Figure 3.1: Three examples of circulant graphs: $Cy_N = C_N(1,0,0,...,0)$ (top-left), also known as *cycle graph*, $C_N(1,2,0,...,0)$ (top-right) and $K_N = C_N\left(1,2,...,\left\lfloor\frac{N}{2}\right\rfloor\right)$ (bottom), also known as *complete graph* or *fully connected network*.

Here we suppose that *all the neurons have the same number of incoming connections*, that we call M. Therefore circulant graphs are ideal candidates, even if many other graphs have this property. If there is no connection from the j-th neuron to the i-th neuron, we set $J_{ij} = 0$. Otherwise, if there is a connection, we set:

$$J_{ij}(t) = \frac{1}{M}\left(\Lambda + \sigma_3 \frac{dB_i^J(t)}{dt}\right) \tag{3.5}$$

where Λ is a free parameter, while $B_i^J(t)$ are new Brownian motions that describe the stochastic fluctuations of the synaptic weights. Instead σ_3 represents the amplitude of these fluctuations around the mean Λ. In order to avoid biologically unrealistic changes of the sign of $J_{ij}(t)$, we have to set $\sigma_3 \ll \Lambda$. Moreover, here the derivative $\frac{dB_i^J(t)}{dt}$ is meant in the weak sense of the distributions, and can be interpreted as white noise. The division by M is required in order to ensure the existence of a well-defined thermodynamic limit of the system. In fact when M increases, each neuron receives a larger and larger input from the remainder of the network through the term $\sum_{j=0}^{N-1} J_{ij}(t) S(V_j(t))$ in equation 3.1, therefore in order to fix this divergence the normalization of the synaptic weights is required. This technique has been already used in [51][64][65]. Moreover, a fundamental assumption of this chapter, is the *independence of the Brownian motions* $B_i^V(t)$ *and* $B_i^J(t)$ *and of the initial conditions* $V_i(0)$, whose importance will be clarified in Chapters 5 and 6.

To conclude, we want to underline that the assumption of having the same quantities σ_1, σ_2, σ_3, μ, $I(t)$, M and Λ for all the neurons, makes the system invariant under exchange of the neural indices. As we already said, this decision is driven by the necessity to avoid complicated inhomogeneities in the system, but the real advantage of this assumption will be clarified in Sections 3.1.1 and 3.1.2.

The key observation of this chapter is the emergence of the phenomenon known as *propagation of chaos*, namely the increase of the independence of the neurons when the network

Neuron	Input	Synaptic Weights	Sigmoid Function
$\tau = 1$	$I = 0.2$	$\Lambda = 1$	$T_{MAX} = 1$
$\sigma_2 = 0.1$	$\sigma_1 = 0.1$	$\sigma_3 = 0.1$	$\lambda = 1$
$\mu = 0.5$			$V_T = 0$

Table 3.1: Values of the parameters of the system 3.1 and of the initial conditions 3.2, used to obtain Figures 3.2 - 3.8.

grows larger in size, starting from independent initial conditions. Up to now propagation of chaos has been studied only for fully connected networks, and in [51] the reader can find the rigorous proof of the occurrence of this phenomenon for $N \to \infty$. In this thesis we do not provide such a proof, but we show its emergence numerically. Figure 3.2 shows the behavior of the correlation between a chosen pair of neurons for a fully connected network, described by equations 3.1, starting from independent initial conditions.

These results have been obtained using the Euler-Maruyama scheme [161] with integration time step $\Delta t = 0.1$ and calculating the statistics with $10,000$ Monte Carlo simulations. The values of the parameters are shown in Table 3.1.

Clearly Figure 3.2 shows that in a fully connected network the correlation decreases with N. However, since the membrane potentials of the neurons described by rate equations are not normally distributed (due to the non-linearity introduced by the function $S(\cdot)$), decorrelation is not equivalent to independence. Therefore, in order to show the increase of their independence when the network grows larger in size, we have to show also that their moments do factorize, namely:

$$\mathbb{E}\left[V_i^m(t)\, V_j^n(t)\right] = \mathbb{E}\left[V_i^m(t)\right] \mathbb{E}\left[V_j^n(t)\right] \tag{3.6}$$

for every pair of integers m, n. Obviously decorrelation corresponds to the case $m = n = 1$. Some examples are shown in Figures 3.3 - 3.6, from which it is possible to see that this factorization actually does occur.

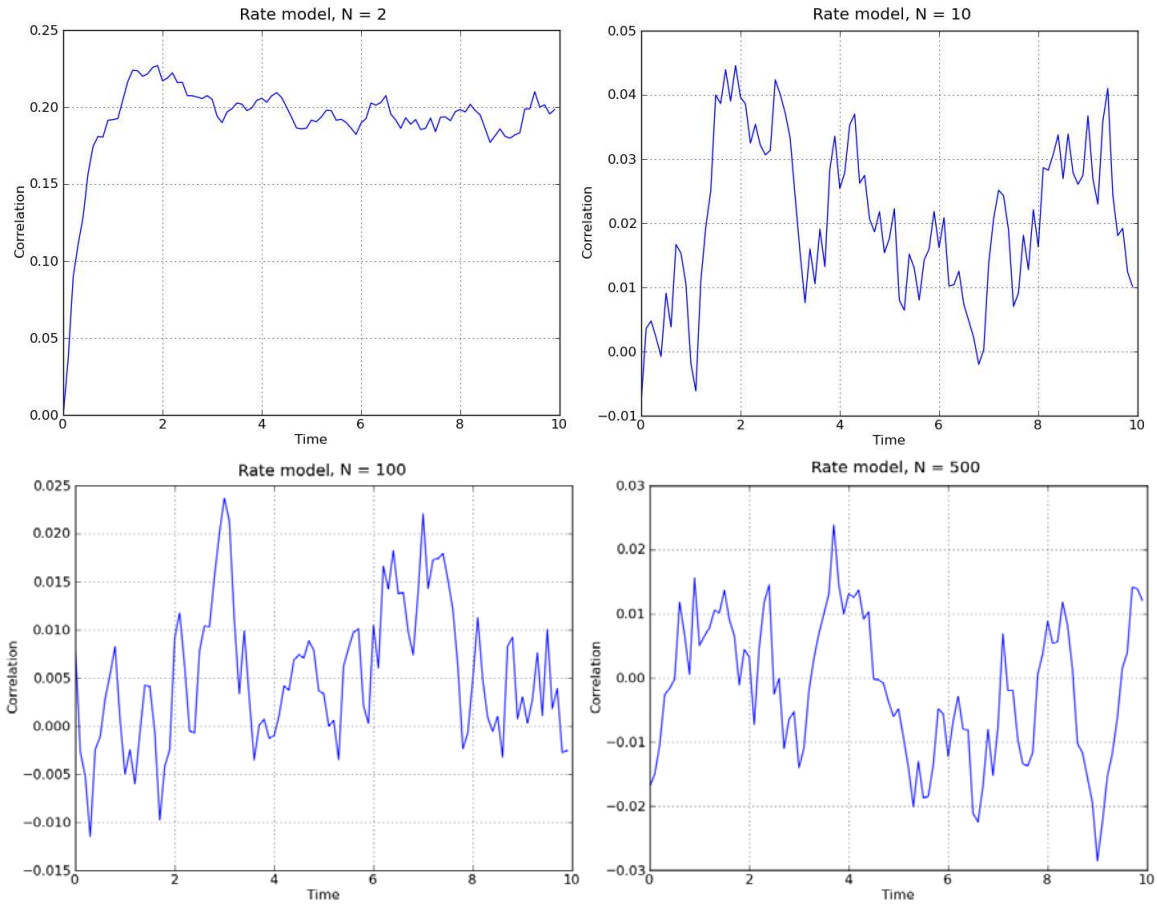

Figure 3.2: Correlation for $N = 2$ (top-left), $N = 10$ (top-right), $N = 100$ (bottom-left) and $N = 500$ (bottom-right) in a fully connected network, starting from independent initial conditions. These results have been obtained for $10,000$ Monte Carlo simulations and with the parameters of Table 3.1. Clearly correlation decreases as a function of N, but for large numbers of neurons ($N > 10$) correlation is so small that we need to run much more Monte Carlo simulations in order to approximate its real value.

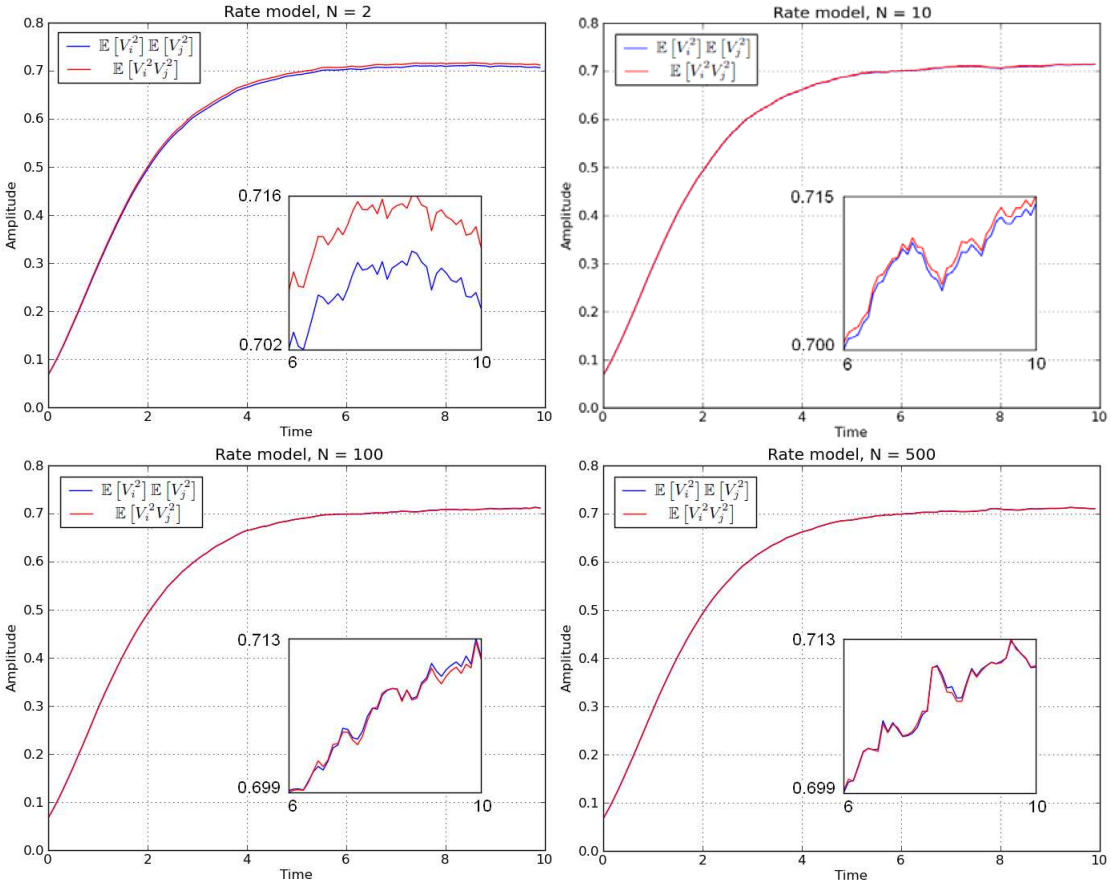

Figure 3.3: Factorization of the second-order moments of two different neurons for $N = 2$ (top-left), $N = 10$ (top-right), $N = 100$ (bottom-left) and $N = 500$ (bottom-right) in a fully connected network described by equations 3.1 and with independent initial conditions. These results have been obtained for $10,000$ Monte Carlo simulations and with the parameters of Table 3.1. This figure clearly shows that increasing N the equality 3.6 is satisfied for $m = n = 2$, providing a numerical evidence for the independence of the neurons.

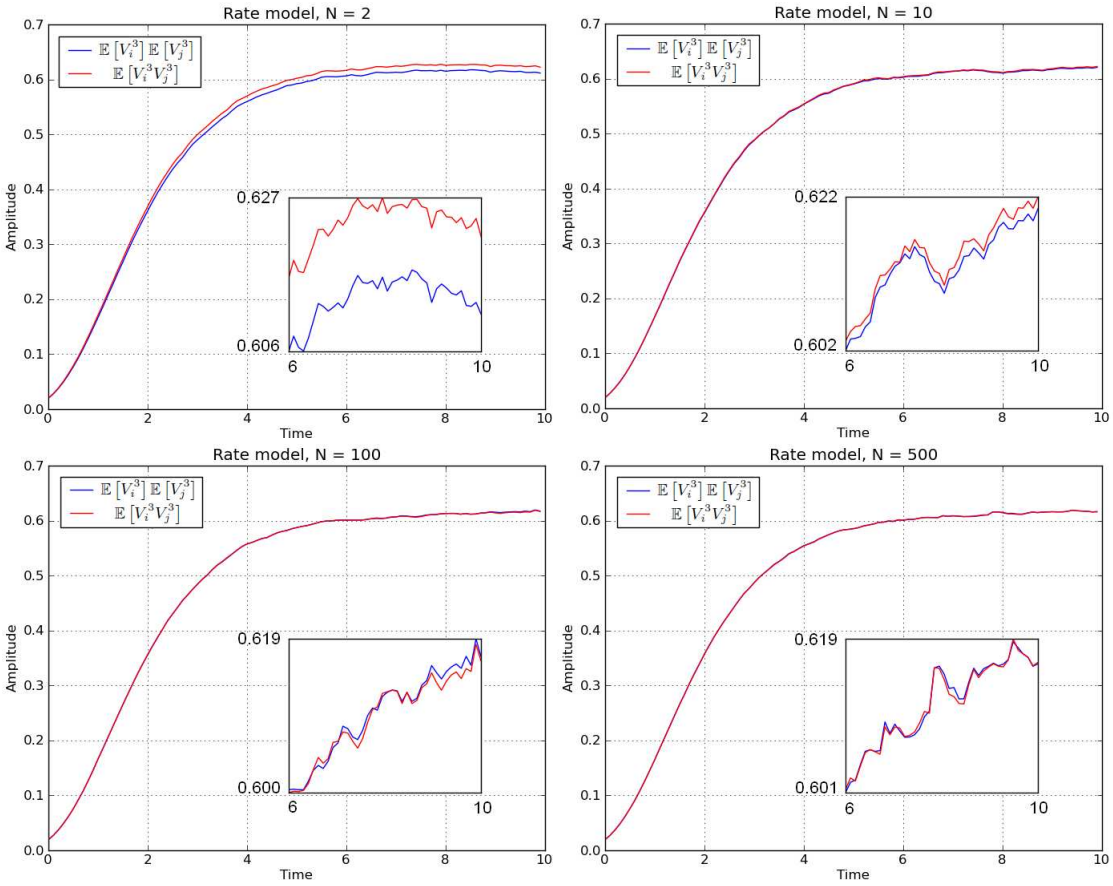

Figure 3.4: Factorization of the moments for $m = n = 3$ in a fully connected network described by equations 3.1 and with independent initial conditions.

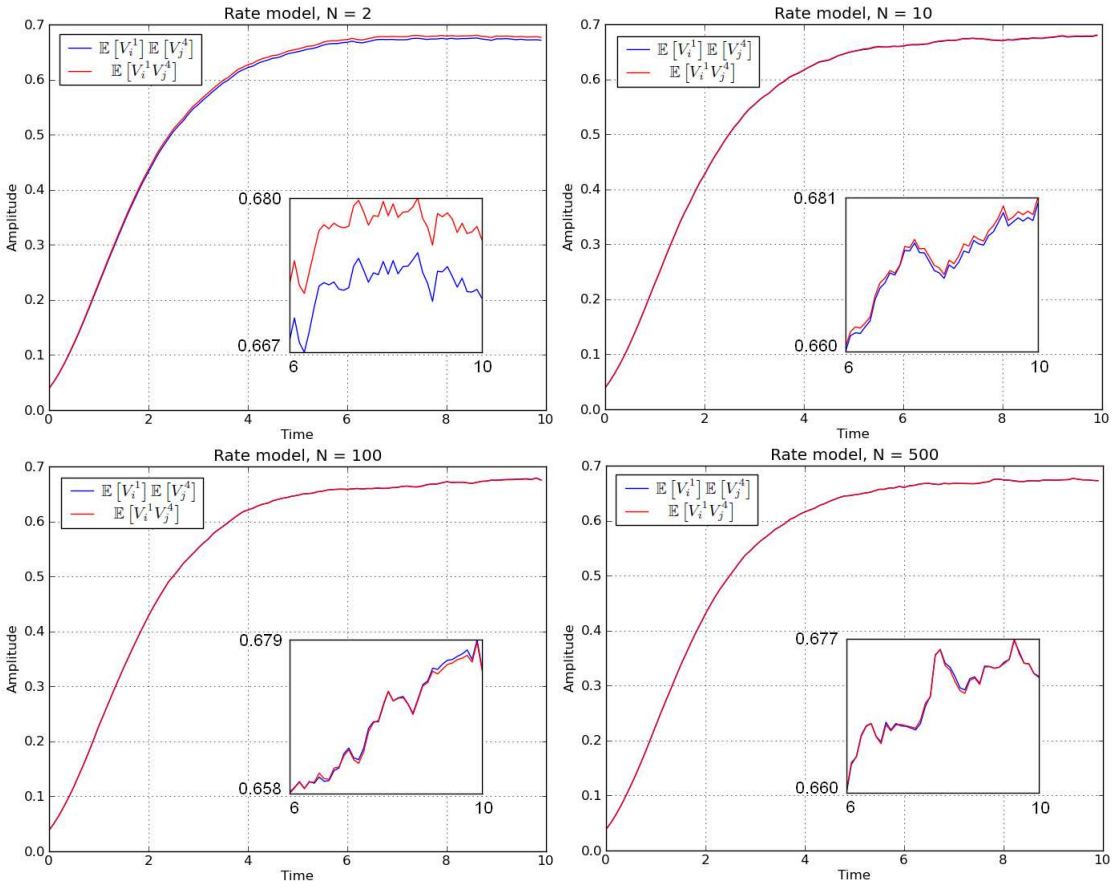

Figure 3.5: Factorization of the moments for $m = 1$ and $n = 4$ in a fully connected network described by equations 3.1 and with independent initial conditions.

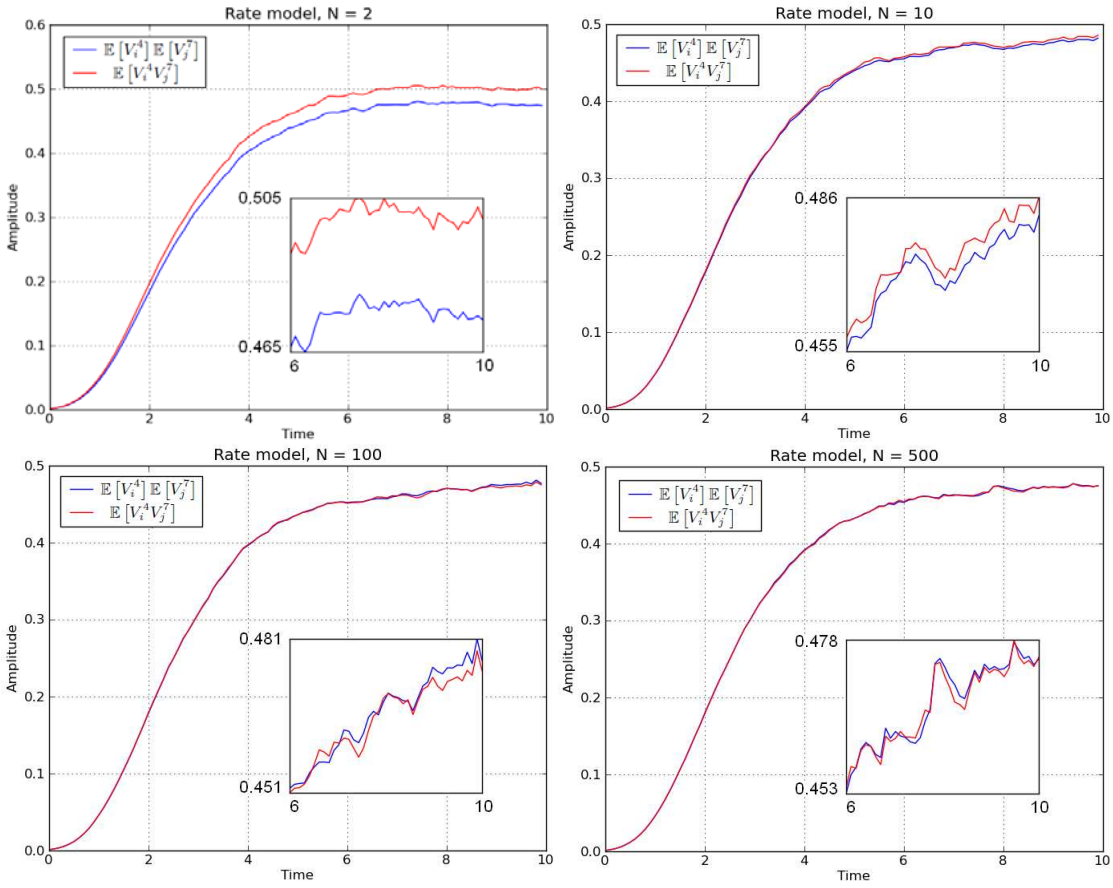

Figure 3.6: Factorization of the moments for $m = 4$ and $n = 7$ in a fully connected network described by equations 3.1 and with independent initial conditions.

This is not a rigorous proof of independence, just a numerical evidence. However, does propagation of chaos still occur when the network is not fully connected? According to the numerical simulations, the answer in general is negative. In fact, Figures 3.7 and 3.8 show the behavior of the correlation and of the higher order moments when the connectivity matrix is given by the cycle graph Cy_N. From them it is possible to see that propagation of chaos does not occur anymore.

Obviously, the only difference from the fully connected case is in the number of incoming connections per neuron, which is $M = 2$ for the cycle graph and $M = N-1$ for the complete graph. Therefore in the fully connected network not only the size of the network grows to infinity, but also the number of incoming connections, while the latter is finite for the cycle network. This seems to show that the real necessary condition for the emergence of propagation of chaos is the explosion of the number of incoming connections per neuron for $N \to \infty$. So for example we have propagation of chaos when $M = N - 1$, $M = \sqrt{N}$, $M = \log N$, etc. This result will be proved analytically in Chapter 5, for weak sources of noise. We will also prove that propagation of chaos is a consequence of the assumption of independence of the Brownian motions and of the initial conditions. Now, in Section 3.1.1 we provide a first method to determine the mean-field equation of the neural network. Instead, in Section 3.1.2, we show a preliminary manipulation of the Fokker-Planck equation of the system that will allow us to obtain the mean-field equation in a different way, which is explained in Section 3.1.3.

3.1.1 The mean-field equation: method #1

From 3.5 we obtain that:

$$\sum_{j=0}^{N-1} J_{ij}(t)\, S(V_j(t)) = \frac{1}{M(N)} \left(\Lambda + \sigma_3 \frac{dB_i^J(t)}{dt} \right) \sum_{j \in G_i(N)} S(V_j(t)) \tag{3.7}$$

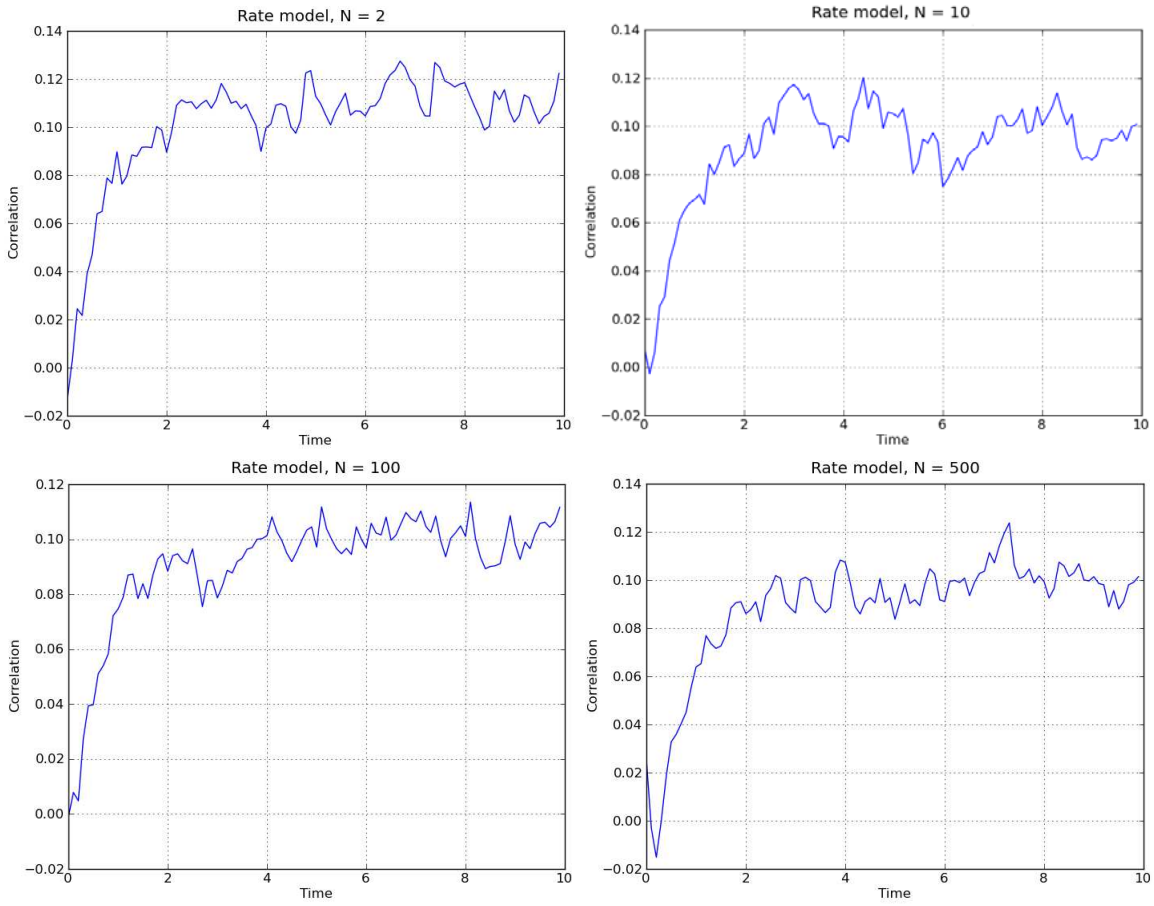

Figure 3.7: Correlation for $N = 2$ (top-left), $N = 10$ (top-right), $N = 100$ (bottom-left) and $N = 500$ (bottom-right) in a cycle network described by equations 3.1 and with independent initial conditions. These results have been obtained for $10,000$ Monte Carlo simulations and with the parameters of Table 3.1. Correlation does not decrease anymore with N, therefore the neurons do not become independent in thermodynamic limit.

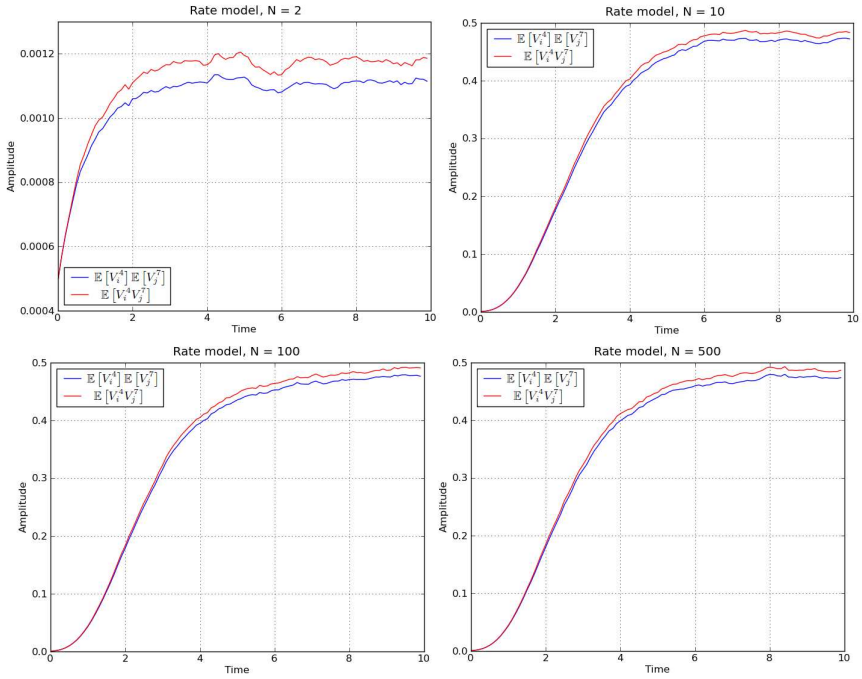

Figure 3.8: No factorization of the moments for $m = 4$ and $n = 7$ in a cycle network described by equations 3.1 and with independent initial conditions.

because $J_{ij}(t)$ depends only on the index i. Here $G_i(N)$, for a given value of the index i, is the set of all the values of the index j such that $J_{ij} \neq 0$, in a network of size N. So clearly $i \notin G_i(N)$, due to the absence of self-connections ($J_{ii} = 0$). We have previously stated that if the number of incoming connections per neuron grows to infinity in the thermodynamic limit, then the neurons become independent. In other terms, in order to have propagation of chaos, we have to assume that:

$$\lim_{N \to \infty} M(N) = \infty \tag{3.8}$$

Due to the invariance of the system under exchange of the neural indices, in the thermodynamic limit all the neurons have the same marginal probability density $p(V, t)$. Moreover, since for $N \to \infty$ they become independent under assumption 3.8 (this is the so called *local chaos hypothesis*), the infinite set of all the $V_i(t)$ can be seen as a collection of samples generated by the single distribution $p(V, t)$. For this reason we can write:

$$\lim_{N\to\infty}\frac{1}{M\left(N\right)}\sum_{j\in G_i(N)}S\left(V_j\left(t\right)\right)=\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]=\int_{-\infty}^{+\infty}S\left(V'\right)p\left(V',t\right)dV' \tag{3.9}$$

Therefore combining 3.7 and 3.9, we obtain that in the thermodynamic limit the system 3.1 becomes:

$$dV_i\left(t\right)=\left[-\frac{1}{\tau}V_i\left(t\right)+\Lambda\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]+I\left(t\right)\right]dt+\sigma_1 dB_i^V\left(t\right)+\sigma_3\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]dB_i^J\left(t\right) \tag{3.10}$$

This is the *mean-field equation* of the network, also known as the *McKean-Vlasov equation*, after the work of H. P. McKean and A. A. Vlasov on similar kinds of systems [51][162][163][164]. Now, since $B_i^V\left(t\right)$ and $B_i^J\left(t\right)$ are independent, we can combine them generating a new total Brownian motion $B_i\left(t\right)$ such that:

$$\sigma_1 dB_i^V\left(t\right)+\sigma_3\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]dB_i^J\left(t\right)=\sqrt{\sigma_1^2+\sigma_3^2\left(\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]\right)^2}dB_i\left(t\right)$$

Therefore finally the mean-field equation of the system can be equivalently rewritten as:

$$dV_i\left(t\right)=\left[-\frac{1}{\tau}V_i\left(t\right)+\Lambda\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]+I\left(t\right)\right]dt+\sqrt{\sigma_1^2+\sigma_3^2\left(\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]\right)^2}dB_i\left(t\right) \tag{3.11}$$

This stochastic differential equation (SDE) is not easily solvable since it is a function of the probability density $p\left(V,t\right)$, namely the law of the solution $V_i\left(t\right)$, which is not known a priori. Therefore we have to transform it in its corresponding Fokker-Planck equation (FPE), which is a function only of the unknown $p\left(V,t\right)$:

$$\frac{\partial}{\partial t}p\left(V,t\right)=-\frac{\partial}{\partial V}\left[\left(-\frac{V}{\tau}+\Lambda\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]+I\left(t\right)\right)p\left(V,t\right)\right]+\frac{1}{2}\left[\sigma_1^2+\sigma_3^2\left(\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]\right)^2\right]\frac{\partial^2}{\partial V^2}p\left(V,t\right) \tag{3.12}$$

Once the FPE is solved, we can replace $p(V,t)$ in the equation 3.11, transforming it in a common SDE which can be solved with the standard techniques.

To conclude, we want to show the advantage of using the thermodynamic limit of the neural network. If we want to determine the behavior of the network for a finite number of neurons N, we have to solve the system of N stochastic differential equations 3.1. The probability density of the system can be obtained by solving the corresponding FPE, which is extremely complicated (see formula 3.14 in Section 3.1.2) since the SDEs are coupled . Instead, in the thermodynamic limit, we have to solve the system 3.11, where the equations are decoupled. Therefore now the FPE is much simpler (see equation 3.12), and moreover it is the same for all the neurons. Now, choosing a finite sub-system made up of η neurons, whose indices belong to the subset $\mathcal{U} = \{i_0, i_1, ..., i_{\eta-1}\}$, and supposing to know the probability density $p(V,t)$ of a single neuron, we can calculate the joint probability density of the sub-system as follows:

$$p\left(V_{i_0}, V_{i_1}, ..., V_{i_{\eta-1}}, t\right) = \prod_{i \in \mathcal{U}} p\left(V_i, t\right) \tag{3.13}$$

since in the thermodynamic limit the neurons are independent.

However, in order to obtain this simplification of the FPE, we have to pay a price, namely the introduction of the term $\mathbb{E}_{p(V,t)}\left[S\left(V\right)\right]$, which transforms the FPE from a partial differential equation (PDE) into a partial integro-differential equation (PIDE). PIDEs can be equivalently seen as a special subset of the kind of equations known as *partial functional differential equations*.

3.1.2 Fokker-Planck equation of a finite neural network

In this section we introduce the Fokker-Planck equation of a finite neural network. This will allow us to obtain straightforwardly, in Section 3.1.3 (with a method that is basically

different from that shown in Section 3.1.1), the mean-field equation of the system in the thermodynamic limit. Moreover the results of this section will be used later in Chapter 4, since they allow us to study the finite size effects of the neural network.

So, the Fokker-Planck equation corresponding to 3.1 and 3.5 for a finite number of neurons is:

$$\frac{\partial}{\partial t}p\left(V_0,...,V_{N-1},t\right)=\sum_{i=0}^{N-1}\left\{-\frac{\partial}{\partial V_i}\left[\left(-\frac{V_i}{\tau}+\frac{\Lambda}{M}\sum_{j\in G_i}S\left(V_j\right)+I\left(t\right)\right)p\left(V_0,...,V_{N-1},t\right)\right]\right.$$

$$\left.+\frac{1}{2}\frac{\partial^2}{\partial V_i^2}\left[\left(\sigma_1^2+\sigma_3^2\left(\frac{1}{M}\sum_{j\in G_i}S\left(V_j\right)\right)^2\right)p\left(V_0,...,V_{N-1},t\right)\right]\right\} \tag{3.14}$$

which can be rewritten in a more compact form as:

$$\frac{\partial}{\partial t}p\left(\overrightarrow{V},t\right)=-\overrightarrow{\nabla}\cdot\left[\overrightarrow{f}\left(\overrightarrow{V},t\right)p\left(\overrightarrow{V},t\right)\right]+\frac{1}{2}\nabla^2\left[\overrightarrow{g}\left(\overrightarrow{V},t\right)p\left(\overrightarrow{V},t\right)\right] \tag{3.15}$$

where $\overrightarrow{V}=\left(V_0,...,V_{N-1}\right)$ and:

$$\overrightarrow{f}\left(\overrightarrow{V},t\right)=\begin{bmatrix}f_0\left(\overrightarrow{V},t\right)\\f_1\left(\overrightarrow{V},t\right)\\\vdots\\f_{N-1}\left(\overrightarrow{V},t\right)\end{bmatrix},\quad\overrightarrow{g}\left(\overrightarrow{V},t\right)=\begin{bmatrix}g_0\left(\overrightarrow{V},t\right)\\g_1\left(\overrightarrow{V},t\right)\\\vdots\\g_{N-1}\left(\overrightarrow{V},t\right)\end{bmatrix} \tag{3.16}$$

$$f_i\left(\overrightarrow{V},t\right)=-\frac{V_i}{\tau}+\frac{\Lambda}{M}\sum_{j\in G_i}S\left(V_j\right)+I\left(t\right),\quad g_i\left(\overrightarrow{V},t\right)=\sigma_1^2+\sigma_3^2\left(\frac{1}{M}\sum_{j\in G_i}S\left(V_j\right)\right)^2$$

$\overrightarrow{f}\left(\overrightarrow{V},t\right)$ and $\overrightarrow{g}\left(\overrightarrow{V},t\right)$ are known as *drift and diffusion functions*, respectively. Now, given a set of integers $m_0, m_1, ..., m_{N-1}$ with values between 0 and $N-1$, we define the marginal probability densities:

$$p_{m_0}\left(V_{m_0},t\right) = \int_{\mathbb{R}^{N-1}} p\left(V_0, V_1, ..., V_{m_0}, ..., V_{N-1}, t\right) \prod_{k \neq m_0} dV_k$$

$$p_{m_0,m_1}\left(V_{m_0},V_{m_1},t\right) = \int_{\mathbb{R}^{N-2}} p\left(V_0, V_1, ..., V_{m_o}, ..., V_{m_1}, ..., V_{N-1}, t\right) \prod_{k \neq m_0,m_1} dV_k \qquad \text{for} \qquad m_0 \neq m_1 \tag{3.17}$$

$$\vdots$$

Moreover, given a generic m_0, we integrate the Fokker-Planck equation with respect to $\prod_{k \neq m_0} dV_k$ on the domain $\mathbb{R}^{N-1}$. Therefore the partial derivative with respect to time on the left-hand side of the Fokker-Planck equation (3.15) becomes:

$$\int_{\mathbb{R}^{N-1}} \frac{\partial}{\partial t} p\left(\overrightarrow{V},t\right) \prod_{k \neq m_0} dV_k = \frac{\partial}{\partial t} \int_{\mathbb{R}^{N-1}} p\left(\overrightarrow{V},t\right) \prod_{k \neq m_0} dV_k = \frac{\partial}{\partial t} p_{m_0}\left(V_{m_0},t\right)$$

Now we have to see what happens to the drift and diffusion terms on the right-hand side of the Fokker-Planck equation. For the drift term we have:

$$\int_{\mathbb{R}^{N-1}} \left\{ -\overrightarrow{\nabla} \cdot \left[\overrightarrow{f}\left(\overrightarrow{V},t\right) p\left(\overrightarrow{V},t\right) \right] \right\} \prod_{k \neq m_0} dV_k$$

$$= -\int_{\mathbb{R}^{N-1}} \frac{\partial}{\partial V_{m_0}} \left[f_{m_0}\left(\overrightarrow{V},t\right) p\left(\overrightarrow{V},t\right) \right] \prod_{k \neq m_0} dV_k - \int_{\mathbb{R}^{N-1}} \overrightarrow{\nabla}_{N-1} \cdot \left[\overrightarrow{f}_{N-1}\left(\overrightarrow{V},t\right) p\left(\overrightarrow{V},t\right) \right] \prod_{k \neq m_0} dV_k$$

$$= -\frac{\partial}{\partial V_{m_0}} \left[\int_{\mathbb{R}^{N-1}} f_{m_0} \left(\overrightarrow{V}, t \right) p \left(\overrightarrow{V}, t \right) \right] \prod_{k \neq m_0} dV_k$$

$$= -\frac{\partial}{\partial V_{m_0}} \left\{ \left[-\frac{V_{m_0}}{\tau} + I(t) \right] \int_{\mathbb{R}^{N-1}} p \left(\overrightarrow{V}, t \right) \prod_{k \neq m_0} dV_k \right\} - \frac{\Lambda}{M} \frac{\partial}{\partial V_{m_0}} \sum_{j \in G_{m_0}} \int_{\mathbb{R}^{N-1}} p \left(\overrightarrow{V}, t \right) S(V_j) \prod_{k \neq m_0} dV_k$$

where $\overrightarrow{f}_{N-1} \left(\overrightarrow{V}, t \right)$ is the vector $\overrightarrow{f} \left(\overrightarrow{V}, t \right)$ without the m_0-th component, and $\overrightarrow{\nabla}_{N-1} \cdot = \sum_{l \neq m_0} \frac{\partial}{\partial V_l}$. This result is due to the divergence theorem:

$$\int_{\mathbb{R}^{N-1}} \overrightarrow{\nabla}_{N-1} \cdot \left[\overrightarrow{f}_{N-1} \left(\overrightarrow{V}, t \right) p \left(\overrightarrow{V}, t \right) \right] \prod_{k \neq m_0} dV_k = \int_{\partial \mathbb{R}^{N-1}} \left[\overrightarrow{f}_{N-1} \left(\overrightarrow{V}, t \right) p \left(\overrightarrow{V}, t \right) \right] \cdot \widehat{n} dS = 0$$

having used the boundary conditions:

$$p \left(\overrightarrow{V}, t \right) = 0 \qquad \text{on} \qquad \partial \mathbb{R}^{N-1}$$

Here $\widehat{n}$ is the outward pointing unit normal field of the boundary $\partial \mathbb{R}^{N-1}$. Now we can see that:

$$-\frac{\Lambda}{M} \frac{\partial}{\partial V_{m_0}} \sum_{j \in G_{m_0}} \int_{\mathbb{R}^{N-1}} p \left(\overrightarrow{V}, t \right) S(V_j) \prod_{k \neq m_0} dV_k$$

$$= -\frac{\Lambda}{M} \frac{\partial}{\partial V_{m_0}} \sum_{j \in G_{m_0}} \int_{\mathbb{R}} S(V_j) \left[\int_{\mathbb{R}^{N-2}} p \left(\overrightarrow{V}, t \right) \prod_{k \neq m_0, j} dV_k \right] dV_j$$

$$= -\frac{\Lambda}{M} \frac{\partial}{\partial V_{m_0}} \sum_{j \in G_{m_0}} \int_{\mathbb{R}} S(V_j) \, p_{m_0, j} (V_{m_0}, V_j, t) \, dV_j$$

Since the system is invariant under exchange of the neural indices and since all the neurons with index j are connected to the neuron with index m_0, the function $p_{m_0,j}(\cdot,\cdot,t)$ does not depend anymore on the indices m_0 and j. In other terms, we can define a new function $p^{(2)}(\cdot,\cdot,t)$ such that:

$$p_{m_0,j}(\cdot,\cdot,t) = p^{(2)}(\cdot,\cdot,t), \qquad \forall m_0, j \in G_{m_0} \tag{3.18}$$

In the same way the function $p_{m_0}(\cdot,t)$ does not depend anymore on the index m_0, so we can define a new function $p^{(1)}(\cdot,t)$ such that:

$$p_{m_0}(\cdot,t) = p^{(1)}(\cdot,t), \qquad \forall m_0 \tag{3.19}$$

So from these definitions we obtain:

$$-\frac{\Lambda}{M}\frac{\partial}{\partial V_{m_0}}\sum_{j\in G_{m_0}}\int_{\mathbb{R}} S(V_j)\, p_{m_0,m_j}(V_{m_0},V_j,t)\, dV_j = -\Lambda\frac{\partial}{\partial V_{m_0}}\int_{\mathbb{R}} S(V')\, p^{(2)}(V_{m_0},V',t)\, dV'$$

Therefore the drift term has been fully developed:

$$\int_{\mathbb{R}^{N-1}}\left\{-\overrightarrow{\nabla}\cdot\left[\overrightarrow{f}\left(\overrightarrow{V},t\right)p\left(\overrightarrow{V},t\right)\right]\right\}\prod_{k\neq m_0} dV_k$$

$$= -\frac{\partial}{\partial V_{m_0}}\left\{\left[-\frac{V_{m_0}}{\tau}+I(t)\right]p^{(1)}(V_{m_0},t)\right\} - \Lambda\frac{\partial}{\partial V_{m_0}}\int_{\mathbb{R}} S(V')\, p^{(2)}(V_{m_0},V',t)\, dV'$$

The same trick can be repeated with the diffusion term of the Fokker-Planck equation, since $\nabla^2 = \overrightarrow{\nabla}\cdot\overrightarrow{\nabla}$ and therefore we can apply again the divergence theorem. The result is:

$$\int_{\mathbb{R}^{N-1}} \left\{ \frac{1}{2} \nabla^2 \left[\overrightarrow{g} \left(\overrightarrow{V}, t \right) p \left(\overrightarrow{V}, t \right) \right] \right\} \prod_{k \neq m_0} dV_k$$

$$= \frac{\sigma_1^2}{2} \frac{\partial^2}{\partial V_{m_0}^2} p^{(1)} \left(V_{m_0}, t \right) + \frac{\sigma_3^2}{2} \frac{\partial^2}{\partial V_{m_0}^2} \left[\int_{\mathbb{R}^2} S\left(V' \right) S\left(V'' \right) p^{(3)} \left(V_{m_0}, V', V'', t \right) dV' dV'' \right]$$

having used the boundary conditions $\overrightarrow{\nabla} p \left(\overrightarrow{V}, t \right) = \overrightarrow{0}$ on the boundary $\partial \mathbb{R}^{N-1}$. Putting everything together, we obtain that our original Fokker-Planck equation has become:

$$\frac{\partial}{\partial t} p^{(1)} \left(V, t \right) = - \frac{\partial}{\partial V} \left[\left(-\frac{V}{\tau} + I\left(t \right) \right) p^{(1)} \left(V, t \right) \right] - \Lambda \frac{\partial}{\partial V} \int_{\mathbb{R}} S\left(V' \right) p^{(2)} \left(V, V', t \right) dV'$$

$$+ \frac{\sigma_1^2}{2} \frac{\partial^2}{\partial V^2} p^{(1)} \left(V, t \right) + \frac{\sigma_3^2}{2} \frac{\partial^2}{\partial V^2} \left[\int_{\mathbb{R}^2} S\left(V' \right) S\left(V'' \right) p^{(3)} \left(V, V', V'', t \right) dV' dV'' \right] \tag{3.20}$$

In order to solve equation 3.20, we need to know $p^{(2)} \left(V, V', t \right)$ and $p^{(3)} \left(V, V', V'', t \right)$. So we could think to obtain the PDEs satisfied by these functions using the same trick we used before, namely integrating equation 3.15 with respect to $\prod_{k \neq m_0, m_1} dV_k$ (for $p^{(2)} \left(V, V', t \right)$) and $\prod_{k \neq m_0, m_1, m_2} dV_k$ (for $p^{(3)} \left(V, V', V'', t \right)$). However, the PDEs for these two functions will depend on higher order marginal distributions, namely $p^{(4)}$ and $p^{(5)}$, and so on and so forth. This means that the partial integration of 3.15 generates a sequence of PDEs for the marginal densities, which does not form a closed system of equations. This is known as *Bogoliubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy*. In Section 3.1.3 and in Chapter 4 we will show how to deal with this problem.

3.1.3 The mean-field equation: method #2

Now we are ready to use the results of Section 3.1.2, in order to obtain in a different way the mean-field equation of the system. Since in the thermodynamic limit the neurons become independent, the joint probability density of any finite sub-system can be factorized as the product of the marginal probability densities, by definition of independence. This has been already shown in formula 3.13. Due to this factorization, for $N \to \infty$ we obtain:

$$p_{m_0,m_1}\left(V_{m_0},V_{m_1},t\right) = p^{(1)}\left(V_{m_0},t\right) p^{(1)}\left(V_{m_1},t\right)$$

$$p_{m_0,m_1,m_2}\left(V_{m_0},V_{m_1},V_{m_2},t\right) = p^{(1)}\left(V_{m_0},t\right) p^{(1)}\left(V_{m_1},t\right) p^{(1)}\left(V_{m_2},t\right)$$

therefore equation 3.20 can be rewritten as:

$$\begin{aligned}\frac{\partial}{\partial t}p^{(1)}\left(V,t\right) = &-\frac{\partial}{\partial V}\left[\left(-\frac{V}{\tau}+\Lambda\int_{\mathbb{R}} S\left(V'\right)p^{(1)}\left(V',t\right)dV'+I\left(t\right)\right)p^{(1)}\left(V,t\right)\right] \\ &+\frac{1}{2}\left[\sigma_1^2+\sigma_3^2\left(\int_{\mathbb{R}} S\left(V'\right)p^{(1)}\left(V',t\right)dV'\right)^2\right]\frac{\partial^2}{\partial V^2}p^{(1)}\left(V,t\right) \qquad (3.21)\end{aligned}$$

This is the FPE in the thermodynamic limit, and in fact we can see that it is perfectly equivalent to 3.12. Therefore the two methods that we have used in order to derive the mean-field equation of the system give the same result, as it must be. Obviously, the method developed in this section is more complicated, but unlike the technique of Section 3.1.1, it can be used to study the finite size effects of the network, as we will show in Chapter 4.

3.1.4 Solution of the mean-field Fokker-Planck equation

In this section we show how to calculate the solution of the mean-field FPE 3.12 or 3.21, using an alternative method to that of [51]. If we denote:

$$R_p(t) = \int_{-\infty}^{+\infty} S(V')\, p(V', t)\, dV' \tag{3.22}$$

that represents the *mean firing rate* (namely the mean frequency of the spikes produced by a single neuron) and:

$$\sigma_p^{tot}(t) = \sqrt{\sigma_1^2 + \sigma_3^2 R_p^2(t)}$$

$$I_p^{tot}(t) = \Lambda R_p(t) + I(t)$$

then the mean-field FPE can be rewritten in the following equivalent way:

$$\frac{\partial}{\partial t} p(V,t) = -\frac{\partial}{\partial V}\left[\left(-\frac{V}{\tau} + I_p^{tot}(t)\right) p(V,t)\right] + \frac{1}{2}\left(\sigma_p^{tot}(t)\right)^2 \frac{\partial^2}{\partial V^2} p(V,t)$$

Now we apply the "partial" Fourier transform, namely the transform only with respect to the potential V:

$$\widetilde{p}(\omega, t) = \mathcal{F}_V[p(V,t)] = \int_{-\infty}^{+\infty} p(V,t)\, e^{-\iota\omega V} dV$$

where $\iota = \sqrt{-1}$. In this way, using the following properties of the Fourier transform:

$$\mathcal{F}\left[\frac{d^n y(x)}{dx^n}\right] = (\iota\omega)^n \widetilde{y}(\omega)$$

$$\mathcal{F}\left[x^n y(x)\right] = \iota^n \frac{d^n \widetilde{y}(\omega)}{d\omega^n}$$

the FPE becomes:

$$\frac{\partial}{\partial t}\widetilde{p}(\omega,t) = -\iota\omega\left[\left(-\frac{\iota}{\tau}\frac{\partial}{\partial\omega} + I_{\widetilde{p}}^{tot}(t)\right)\widetilde{p}(\omega,t)\right] - \frac{1}{2}\left(\sigma_{\widetilde{p}}^{tot}(t)\right)^2\omega^2\widetilde{p}(\omega,t)$$

As we can see, this partial differential equation is of the first order, so we can solve it analytically using the *method of characteristics*. In order to use this technique, we have to write our equation in the following equivalent way:

$$\frac{\partial}{\partial t}\widetilde{p}(\omega,t) + \frac{\omega}{\tau}\frac{\partial}{\partial\omega}\widetilde{p}(\omega,t) = -\left[\iota\omega I_{\widetilde{p}}^{tot}(t) + \frac{1}{2}\left(\sigma_{\widetilde{p}}^{tot}(t)\right)^2\omega^2\right]\widetilde{p}(\omega,t)$$

So according to the method of characteristics we obtain:

$$\frac{dt}{1} = \tau\frac{d\omega}{\omega} = \frac{d\widetilde{p}}{-\left[\iota\omega I_{\widetilde{p}}^{tot}(t) + \frac{1}{2}\left(\sigma_{\widetilde{p}}^{tot}(t)\right)^2\omega^2\right]\widetilde{p}} \tag{3.23}$$

Using the first two differentials, we obtain $\frac{d\omega}{\omega} = \frac{dt}{\tau}$, whose solution is:

$$\omega(t) = \omega(0)\, e^{\frac{t}{\tau}} \tag{3.24}$$

Now we have to solve the equation generated by the first and the third differentials in formula 3.23, that remembering 3.24 becomes:

$$\frac{dt}{1} = \frac{d\widetilde{p}}{-\left[\iota\omega\left(0\right)e^{\frac{t}{\tau}}I_{\widetilde{p}}^{tot}\left(t\right) + \frac{1}{2}\left(\sigma_{\widetilde{p}}^{tot}\left(t\right)\right)^{2}\omega^{2}\left(0\right)e^{2\frac{t}{\tau}}\right]\widetilde{p}}$$

Its solution is:

$$\widetilde{p} = Ke^{-\left(\frac{1}{2}\omega^{2}(0)\int_{0}^{t}\left(\sigma_{\widetilde{p}}^{tot}(s)\right)^{2}e^{2\frac{s}{\tau}}ds + \iota\omega(0)\int_{0}^{t}I_{\widetilde{p}}^{tot}(s)e^{\frac{s}{\tau}}ds\right)} \tag{3.25}$$

where K is the integration constant (in general complex valued). If now we suppose to have the following initial condition:

$$\widetilde{p}\left(\omega, t=0\right) = \widetilde{f}\left(\omega\right)$$

where $\widetilde{f}\left(\cdot\right)$ is the partial Fourier transform of a probability density function $f\left(\cdot\right)$, then we have also:

$$\widetilde{p}\left(\omega\left(0\right), t=0\right) = \widetilde{f}\left(\omega\left(0\right)\right) \tag{3.26}$$

But from 3.25 we obtain that:

$$\widetilde{p}\left(t=0\right) = K \tag{3.27}$$

therefore, comparing 3.26 and 3.27, we have $K = \widetilde{f}\left(\omega\left(0\right)\right)$. However $\omega\left(0\right) = \omega\left(t\right)e^{-\frac{t}{\tau}}$ according to 3.24, therefore finally:

$$K = \widetilde{f}\left(\omega\left(t\right)e^{-\frac{t}{\tau}}\right)$$

Replacing this expression of K inside 3.25, we obtain the final result:

$$\widetilde{p}\left(\omega,t\right)=\widetilde{f}\left(\omega\left(t\right)e^{-\frac{t}{\tau}}\right)e^{-\left(\frac{1}{2}\omega^{2}(0)\int_{0}^{t}\left(\sigma_{\widetilde{p}}^{tot}(s)\right)^{2}e^{2\frac{s}{\tau}}ds+\iota\omega(0)\int_{0}^{t}I_{\widetilde{p}}^{tot}(s)e^{\frac{s}{\tau}}ds\right)}$$

and therefore, remembering again that $\omega\left(0\right)=\omega\left(t\right)e^{-\frac{t}{\tau}}$, we obtain:

$$\widetilde{p}\left(\omega,t\right)=\widetilde{f}\left(\omega e^{-\frac{t}{\tau}}\right)e^{-\left(\frac{1}{2}\omega^{2}e^{-2\frac{t}{\tau}}\int_{0}^{t}\left(\sigma_{\widetilde{p}}^{tot}(s)\right)^{2}e^{2\frac{s}{\tau}}ds+\iota\omega e^{-\frac{t}{\tau}}\int_{0}^{t}I_{\widetilde{p}}^{tot}(s)e^{\frac{s}{\tau}}ds\right)}$$

So, finally, the solution of the FPE is:

$$p\left(V,t\right)=\mathcal{F}_{V}^{-1}\left[\widetilde{p}\left(\omega,t\right)\right]=\int_{-\infty}^{+\infty}\widetilde{f}\left(\omega e^{-\frac{t}{\tau}}\right)e^{-\left(\frac{1}{2}\omega^{2}e^{-2\frac{t}{\tau}}\int_{0}^{t}\left(\sigma_{\widetilde{p}}^{tot}(s)\right)^{2}e^{2\frac{s}{\tau}}ds+\iota\omega e^{-\frac{t}{\tau}}\int_{0}^{t}I_{\widetilde{p}}^{tot}(s)e^{\frac{s}{\tau}}ds\right)}e^{\iota\omega V}d\omega \tag{3.28}$$

If now, according to 3.2, we use Gaussian initial conditions:

$$f\left(V\right)=p\left(V,t=0\right)=\frac{1}{\sqrt{2\pi}\sigma_{2}}e^{-\frac{(V-\mu)^{2}}{2\sigma_{2}^{2}}}$$

that is equivalent to say that:

$$\widetilde{f}\left(\omega\right)=\widetilde{p}\left(\omega,t=0\right)=\mathcal{F}_{V}\left[\frac{1}{\sqrt{2\pi}\sigma_{2}}e^{-\frac{(V-\mu)^{2}}{2\sigma_{2}^{2}}}\right]=e^{-\left(\iota\mu\omega+\frac{1}{2}\sigma_{2}^{2}\omega^{2}\right)}$$

then we can write:

$$\widetilde{f}\left(\omega e^{-\frac{t}{\tau}}\right)=e^{-\left(\iota\mu\omega e^{-\frac{t}{\tau}}+\frac{1}{2}\sigma_{2}^{2}\omega^{2}e^{-2\frac{t}{\tau}}\right)}$$

Now, substituting this expression of $\widetilde{f}\left(\omega e^{-\frac{t}{\tau}}\right)$ inside 3.28, we obtain:

$$p\left(V,t\right)=\int_{-\infty}^{+\infty}e^{-\left(\iota\mu\omega e^{-\frac{t}{\tau}}+\frac{1}{2}\sigma_{2}^{2}\omega^{2}e^{-2\frac{t}{\tau}}\right)}e^{-\left(\frac{1}{2}\omega^{2}e^{-2\frac{t}{\tau}}\int_{0}^{t}\left(\sigma_{\widetilde{p}}^{tot}(s)\right)^{2}e^{2\frac{s}{\tau}}ds+\iota\omega e^{-\frac{t}{\tau}}\int_{0}^{t}I_{\widetilde{p}}^{tot}(s)e^{\frac{s}{\tau}}ds\right)}e^{\iota\omega V}d\omega$$

and therefore finally:

$$p\left(V,t\right)=\frac{1}{\sqrt{2\pi\left(\sigma_2^2e^{-2\frac{t}{\tau}}+e^{-2\frac{t}{\tau}}\int_0^t\left(\sigma_p^{tot}\left(s\right)\right)^2e^{2\frac{s}{\tau}}ds\right)}}e^{-\frac{\left(V-e^{-\frac{t}{\tau}}\left(\mu+\int_0^t I_p^{tot}(s)e^{\frac{s}{\tau}}ds\right)\right)^2}{2\left(\sigma_2^2e^{-2\frac{t}{\tau}}+e^{-2\frac{t}{\tau}}\int_0^t\left(\sigma_p^{tot}(s)\right)^2e^{2\frac{s}{\tau}}ds\right)}} \tag{3.29}$$

Therefore the solution of the FPE in the thermodynamic limit is a Gaussian distribution with mean:

$$\mu_V\left(t\right)=\mathbb{E}_{p(V,t)}\left[V\right]=e^{-\frac{t}{\tau}}\left(\mu+\int_0^t I_p^{tot}\left(s\right)e^{\frac{s}{\tau}}ds\right) \tag{3.30}$$

and variance:

$$\Sigma_V\left(t\right)=Var\left(V\left(t\right)\right)=\sigma_2^2e^{-2\frac{t}{\tau}}+e^{-2\frac{t}{\tau}}\int_0^t\left(\sigma_p^{tot}\left(s\right)\right)^2e^{2\frac{s}{\tau}}ds \tag{3.31}$$

So we have obtained the exact solution $p\left(V,t\right)$ as a function of the mean firing rate $R_p\left(t\right)$. But of course $R_p\left(t\right)$ is a function of $p\left(V,t\right)$ itself, therefore using formulae 3.22 and 3.29, we obtain:

$$R_p\left(t\right)=\frac{1}{\sqrt{2\pi\left(\sigma_2^2e^{-2\frac{t}{\tau}}+e^{-2\frac{t}{\tau}}\int_0^t\left(\sigma_1^2+\sigma_3^2R_p^2\left(s\right)\right)e^{2\frac{s}{\tau}}ds\right)}}\int_{-\infty}^{+\infty}S\left(V'\right)e^{-\frac{\left(V'-e^{-\frac{t}{\tau}}\left(\mu+\int_0^t[I(s)+\Lambda R_p(s)]e^{\frac{s}{\tau}}ds\right)\right)^2}{2\left(\sigma_2^2e^{-2\frac{t}{\tau}}+e^{-2\frac{t}{\tau}}\int_0^t\left(\sigma_1^2+\sigma_3^2R_p^2(s)\right)e^{2\frac{s}{\tau}}ds\right)}}dV' \tag{3.32}$$

This is a self-consistency constraint, expressed as an integral equation in the unknown $R_p\left(t\right)$. Now, in the special case when the activation function is given by 3.4 (and only in this case), following [51] we obtain:

$$\int_{-\infty}^{+\infty}S\left(V'\right)e^{-\frac{(V'-\alpha)^2}{2\beta^2}}dV'=\sqrt{2\pi}\beta T_{MAX}E\left(\frac{\lambda\left(\alpha-V_T\right)}{\sqrt{1+\left(\lambda\beta\right)^2}}\right)$$

(while for general forms of the function $S(\cdot)$ this integrals cannot be evaluated straightforwardly, therefore we have opted for formula 3.4). Moreover, if we define two new functions:

$$U(t) = e^{-\frac{t}{\tau}} \int_0^t R_p(s)\, e^{\frac{s}{\tau}} ds$$

$$Q(t) = e^{-2\frac{t}{\tau}} \int_0^t R_p^2(s)\, e^{2\frac{s}{\tau}} ds$$

we obtain:

$$\begin{cases} \frac{dU(t)}{dt} = & -\frac{1}{\tau} U(t) + R_p(t) \\ \\ \frac{dQ(t)}{dt} = & -\frac{2}{\tau} Q(t) + R_p^2(t) \end{cases}$$

with:

$$R_p(t) = T_{MAX} E \left(\frac{\lambda \left[e^{-\frac{t}{\tau}} \left(\mu + \int_0^t I(s)\, e^{\frac{s}{\tau}} ds \right) + \Lambda U(t) - V_T \right]}{\sqrt{1 + \lambda^2 \left[\frac{\sigma_1^2 \tau}{2} \left(1 - e^{-2\frac{t}{\tau}} \right) + \sigma_2^2 e^{-2\frac{t}{\tau}} + \sigma_3^2 Q(t) \right]}} \right)$$

Therefore the integral equation 3.32 is equivalent to a system of two Wilson and Cowan ordinary differential equations (ODEs) in the unknowns $U(t)$ and $Q(t)$. For a more intuitive interpretation of this result, we can observe that 3.30 and 3.31 can be rewritten as:

$$\mu_V(t) = e^{-\frac{t}{\tau}} \left(\mu + \int_0^t I(s)\, e^{\frac{s}{\tau}} ds \right) + \Lambda U(t)$$

$$\Sigma_V(t) = \frac{\sigma_1^2 \tau}{2} \left(1 - e^{-2\frac{t}{\tau}} \right) + \sigma_2^2 e^{-2\frac{t}{\tau}} + \sigma_3^2 Q(t)$$

Taking the derivative with respect to time of these expressions and after some algebra, we obtain:

$$\begin{cases} \frac{d\mu_V(t)}{dt} = & -\frac{1}{\tau}\mu_V(t) + I(t) + \Lambda T_{MAX} E\left(\frac{\lambda(\mu_V(t)-V_T)}{\sqrt{1+\lambda^2\Sigma_V(t)}}\right) \\ \\ \frac{d\Sigma_V(t)}{dt} = & -\frac{2}{\tau}\Sigma_V(t) + \sigma_1^2 + \sigma_3^2 T_{MAX}^2 E^2\left(\frac{\lambda(\mu_V(t)-V_T)}{\sqrt{1+\lambda^2\Sigma_V(t)}}\right) \end{cases} \tag{3.33}$$

with initial conditions $\mu_V(0) = \mu$ and $\Sigma_V(0) = \sigma_2^2$. So for Gaussian initial conditions the probability density of the network in the thermodynamic limit is always Gaussian, whose mean and variance satisfy the ODE system 3.33.

Instead, if the initial conditions are not Gaussian, the system converges exponentially fast (with time constant $\frac{1}{\tau}$) to a Gaussian distribution for $t \to +\infty$. This can be easily observed from 3.28, using the fact that:

$$\lim_{t\to+\infty} \widetilde{f}\left(\omega e^{-\frac{t}{\tau}}\right) = \lim_{t\to+\infty} \int_{-\infty}^{+\infty} p(V, t=0)\, e^{-\iota\omega e^{-\frac{t}{\tau}} V} dV = 1$$

due to the normalization condition of the probability density.

To conclude, it is possible to extend these results to the case of an arbitrary number P of neural populations. In general, we will obtain a system of P coupled FPEs, one for each population. Moreover, if the initial conditions are Gaussian, the probability density of each population is always Gaussian, and their means and variances are given by a system of $2P$ coupled ODEs, which is the natural generalization of 3.33.

3.2 The FitzHugh-Nagumo model

The methods developed in Sections 3.1.1 and 3.1.3 can be used to determine the mean-field equation and its corresponding FPE for every kind of networks. For example, the so

called *FitzHugh-Nagumo model* with chemical synapses is often used, due to the relative simplicity of the equations (see for example [64]):

$$\begin{cases} dV_i(t) = & \left[V_i(t) - \frac{V_i^3(t)}{3} - w_i(t) - \sum_{j=0}^{N-1} J_{ij}(t)\left(V_i(t) - V_{rev}\right) y_j(t) + I(t) \right] dt + \sigma_1 dB_i^V(t) \\ dw_i(t) = & c\left(V_i(t) + a - b w_i(t)\right) dt \\ dy_i(t) = & \left[\alpha S\left(V_i(t)\right)\left(1 - y_i(t)\right) - \beta y_i(t)\right] dt + \sigma_4\left(V_i(t), y_i(t)\right) dB_i^y(t) \end{cases} \tag{3.34}$$

where:

- $V_i(t)$, $w_i(t)$ and $y_i(t)$ are respectively the membrane potential, the adaptation (also known as recovery function) and the fraction of open ion channels (we will refer to it as the conductance) of the i-th neuron;
- $I(t)$ is the deterministic external input current;
- V_{rev} is the reversal potential of the chemical synapses;
- $B_i^V(t)$ and $B_i^y(t)$ are the Brownian motions that describe respectively the background noise and the fluctuations in the synaptic conductance of the i-th neuron;
- σ_1 and $\sigma_4\left(V_i(t), y_i(t)\right)$ are the standard deviations of the Brownian motions in the background and in the ion channels respectively;
- $J_{ij}(t)$ is the maximum conductance of the synapse from the j-th neuron to the i-th neuron, which as usual is given by 3.5;
- a, b and c are the parameters that describe the kinetics of the adaptation;
- α and β determine respectively the rise and decay rates of the synaptic conductance;

- $S(\cdot)$ is a sigmoid function.

In particular, the noise intensity $\sigma_4(V_i(t), y_i(t))$ is given by:

$$\sigma_4(V_i(t), y_i(t)) = \sqrt{\alpha S(V_i(t))(1 - y_i(t)) + \beta y_i(t)}\chi(y_i(t))$$

where $\chi(\cdot)$ is a function that vanishes outside the range $(0, 1)$. Following [64], we use:

$$\chi(y) = \begin{cases} \Gamma e^{-\frac{\Upsilon}{1-(2y-1)^2}} & \text{if} \quad y \in (0,1) \\ \\ 0 & \text{otherwise} \end{cases}$$

This particular choice of $\sigma_4(V_i(t), y_i(t))$ guarantees that the function $y_i(t)$ is always in the range $[0, 1]$, as required by the definition of the synaptic conductance. The first and the second equation of the system 3.34 (excluding the interaction term) represent the FitzHugh-Nagumo model of a single neuron [165][166], while the third equation (without the noise term, which was first introduced in [167]) is a kinetic model that quantifies the synaptic transmission [168][169]. For a detailed description of the synaptic noise in cortical neurons, the reader is referred to [170][171][172].

3.2.1 The mean-field equation

Again, for a fully connected network in the limit $N \to \infty$, it can be shown numerically the emergence of propagation of chaos, while the exact proof can be found in [64]. Therefore the neurons become independent and can be described by a single probability density $p(V, w, y, t)$. Using the same methods developed for the rate model, we obtain the following mean-field system of equations for the network:

$$\begin{cases} dV_i(t) = & \left[V_i(t) - \frac{V_i^3(t)}{3} - w_i(t) - \Lambda\left(V_i(t) - V_{rev}\right)\mathbb{E}_{p(V,w,y,t)}[y] + I(t)\right]dt \\ & +\sigma_1 dB_i^V(t) - \sigma_3\left(V_i(t) - V_{rev}\right)\mathbb{E}_{p(V,w,y,t)}[y]\, dB_i^J(t) \\ dw_i(t) = & c\left(V_i(t) + a - bw_i(t)\right)dt \\ dy_i(t) = & \left[\alpha S\left(V_i(t)\right)\left(1 - y_i(t)\right) - \beta y_i(t)\right]dt + \sigma_4\left(V_i(t), y_i(t)\right)dB_i^y(t) \end{cases} \tag{3.35}$$

where:

$$\mathbb{E}_{p(V,w,y,t)}[y] = \int_{\mathbb{R}^2\times[0,1]} y' p\left(V', w', y', t\right) dV' dw' dy' \tag{3.36}$$

while its corresponding FPE is:

$$\begin{aligned} \frac{\partial p(V,w,y,t)}{\partial t} = & -\frac{\partial}{\partial V}\left\{\left[V - \frac{V^3}{3} - w - \Lambda\left(V - V_{rev}\right)\mathbb{E}_{p(V,w,y,t)}[y] + I(t)\right]p(V,w,y,t)\right\} \\ & -\frac{\partial}{\partial w}\left[c\left(V + a - bw\right)p(V,w,y,t)\right] - \frac{\partial}{\partial y}\left\{\left[\alpha S(V)(1-y) - \beta y\right]p(V,w,y,t)\right\} \\ & +\frac{1}{2}\frac{\partial^2}{\partial V^2}\left\{\left[\sigma_1^2 + \sigma_3^2\left(V - V_{rev}\right)^2\mathbb{E}^2_{p(V,w,y,t)}[y]\right]p(V,w,y,t)\right\} \\ & +\frac{1}{2}\frac{\partial^2}{\partial y^2}\left\{\left[\alpha S(V)(1-y) + \beta y\right]\chi^2(y)\, p(V,w,y,t)\right\} \end{aligned} \tag{3.37}$$

The analytic solution of 3.37 is not known, therefore we have to solve this equation using numerical methods, as explained in the next section.

3.2.2 Numerical solution of the Fokker-Planck equation

The numerical scheme that we have decided to implement is the so called *method of lines* [173][174]. This numerical technique is based on a discretization of the phase-space, which converts the FPE into a system of ODEs with continuous time. In other terms, since the domain of the probability density of the single neuron is $D = \mathbb{R}^2 \times [0,1]$, we can suppose that the state of the system is always in a sufficiently large but finite subset $A \subseteq D$. Now A can be discretized dividing the axes V, w and y respectively into n_V, n_w and n_y segments with lengths ΔV, Δw and Δy. Therefore, if $A = [V_{min}, V_{MAX}] \times [w_{min}, w_{MAX}] \times [0,1]$, we have:

$$n_V = \frac{V_{MAX} - V_{min}}{\Delta V}, \qquad n_V = \frac{w_{MAX} - w_{min}}{\Delta w}, \qquad n_V = \frac{1}{\Delta y}$$

This discretization of the phase-space generates a grid of points at which the probability density will be evaluated. Since now the phase-space is discrete, the partial derivatives of first and second order with respect to V, w and y which appear in 3.37 can be calculated numerically using a finite difference scheme. According to [175], the central difference expansions:

$$\begin{aligned} \frac{df(x)}{dx} \approx & \frac{1}{\Delta x} \sum_{\substack{k=-n \\ k \neq 0}}^{n} (-1)^{k+1} \frac{(n!)^2}{k\,(n-k)!\,(n+k)!} f(x + k\Delta x) \\ \frac{d^2 f(x)}{dx^2} \approx & \frac{1}{\Delta x^2} \left[g_{0,2n}^{C,2} f(x) + \sum_{\substack{k=-n \\ k \neq 0}}^{n} (-1)^{k+1} \frac{2}{k^2} \frac{(n!)^2}{(n-k)!\,(n+k)!} f(x + k\Delta x) \right] \end{aligned} \tag{3.38}$$

with:

$$
\begin{aligned}
g_{0,2n}^{C,2} &= -2\sum_{k=1}^{n} g_{k,2n}^{C,2} \\
g_{k,2n}^{C,2} &= (-1)^{k+1}\,\frac{2}{k^2}\,\frac{(n!)^2}{(n-k)!\,(n+k)!}
\end{aligned}
\tag{3.39}
$$

are approximations of order $2n$ of the first and second order derivatives. Therefore with these expansions we transform the FPE 3.37 into a system of ODEs, one for each point of the grid, where the unknowns are the functions $p\left(V_{min}+k_V\Delta V, w_{min}+k_w\Delta w, k_y\Delta y, t\right)$, for $k_V=2,...,n_V-2$, $k_w=2,...,n_w-2$ and $k_y=2,...,n_y-2$. These ODEs have continuous time, therefore they can be solved with standard numerical techniques, like the Runge-Kutta method. Instead, for $k_V=0,1,n_V-1,n_V$, $k_w=0,1,n_w-1,n_w$ and $k_y=0,1,n_y-1,n_y$, we set $p\left(V_{min}+k_V\Delta V, w_{min}+k_w\Delta w, k_y\Delta y, t\right)=0$. In this way we have implemented the boundary conditions $p=0$ and $\frac{\partial p}{\partial V}=\frac{\partial p}{\partial w}=\frac{\partial p}{\partial y}=0$ at infinity and $\forall t$. However we need also the initial conditions of the probability density. For simplicity we choose a Gaussian distribution, namely:

$$
p\left(V,w,y,t=0\right)=\frac{1}{(2\pi)^{\frac{3}{2}}\,\sigma_2^V\sigma_2^w\sigma_2^y}e^{-\frac{1}{2}\left[\left(\frac{V-\mu^V}{\sigma_2^V}\right)^2+\left(\frac{V-\mu^w}{\sigma_2^w}\right)^2+\left(\frac{V-\mu^y}{\sigma_2^y}\right)^2\right]}
\tag{3.40}
$$

where μ^V, μ^w, μ^y and σ_2^V, σ_2^w, σ_2^y are the means and the standard deviations of the initial conditions for V, w and y respectively. To conclude, we need an integration scheme in order to calculate the term 3.36. Since the dimensionality of the phase-space is large, in order to decrease the numerical error it is preferable to use a highly precise integration scheme. In particular, we have opted for the so called *Newton-Cotes method* of order 6, defined below:

$$\int_{x_1}^{x_2} f(x)\,dx \approx \frac{5}{288}\Delta x \sum_{i=1}^{n_x/5} [19f(x_1+(5i-5)\Delta x)+75f(x_1+(5i-4)\Delta x)$$

$$+50f(x_1+(5i-3)\Delta x)+50f(x_1+(5i-2)\Delta x)$$

$$+75f(x_1+(5i-1)\Delta x)+19f(x_1+5i\Delta x)]$$

$$n_x = \frac{x_2-x_1}{\Delta x}$$

The only drawback with this technique is the necessity to choose n_x as a multiple of 5. Now, for the method of lines applied to PDEs with second order derivatives, discretized with a step Δx, when we numerically solve the system of ODEs with respect to time using an integration time step Δt, it is well known that the effective integration step is $\Delta = \frac{\Delta t}{\Delta x^2}$. Therefore if we decide to use a very dense grid, namely a very small Δx, then Δ will increase, and this can cause the instability of the algorithm. This problem can be fixed by decreasing Δt, at the cost of a lower execution speed of the numerical scheme. Not surprisingly, we have found that the problem persists if we keep Δx fixed while increasing the approximation order $2n$ in formulae 3.38 and 3.39. Therefore we have opted for $n = 2$, namely approximations of order 4 of the first and second order derivatives. In this case formulae 3.38 and 3.39 give:

$$\frac{df(x)}{dx} \approx \frac{1}{12\Delta x}[f(x-2\Delta x)-8f(x-\Delta x)+8f(x+\Delta x)-f(x+2\Delta x)]$$

$$\frac{d^2f(x)}{dx^2} \approx \frac{1}{12\Delta x^2}[-f(x-2\Delta x)+16f(x-\Delta x)-30f(x)+16f(x+\Delta x)-f(x+2\Delta x)] \tag{3.41}$$

Initial conditions	Phase-space	FitzHugh-Nagumo	Synaptic weights	Synapses	Other
$\mu^V = 0$	$V_{min} = -3$	$a = 0.7$	$\Lambda = 1$	$V_{rev} = 1$	$\Delta t = 0.01$
$\mu^w = 0.5$	$V_{MAX} = 3$	$b = 0.8$	$\sigma_3 = 0.2$	$\alpha = 1$	
$\mu^y = 0.3$	$\Delta V = 0.1$	$c = 0.08$		$\beta = 1$	
$\sigma_2^V = 0.4$	$w_{min} = -2$	$I = 0.4$		$T_{MAX} = 1$	
$\sigma_2^w = 0.4$	$w_{MAX} = 2$	$\sigma_1 = 0$		$\lambda = 0.2$	
$\sigma_2^y = 0.05$	$\Delta w = 0.1$			$V_T = 2$	
	$\Delta y = 0.06$			$\Gamma = 0.1$	
				$\Upsilon = 0.5$	

Table 3.2: Values of the parameters of equation 3.37 and of the initial conditions 3.40, used to obtain Figures 3.9-3.12.

So finally, using a Runge-Kutta method of order 2 and the parameters of Table 3.2, we have obtained the results shown in Figures 3.9 and 3.10 for the marginal probability densities, performing the simulation on a regular laptop (see Software and Hardware for the hardware specifications).

Intuitively, we have chosen a high enough value of the input current I in order to generate a spiking activity in the network. Moreover, we have chosen the parameters μ^V, μ^w, μ^y of the initial probability density in such a way that the Gaussian peak of the distribution is located inside the limit cycle of the spikes. Due to the nullclines of the mean-field system 3.35, the phase-space and the volume inside the limit cycle are divided in two parts. For this reason we observe a split of the trajectories of the neurons in the phase space, as shown in Figure 3.11.

This split corresponds to the two peaks of the probability density observed in Figures 3.9 and 3.10, and it clearly shows that the system becomes highly non-Gaussian. This result will be used in Chapter 7 for the numerical calculation of the Fisher information of the system. Instead Figure 3.12 shows the formation of a rest state, where most of the neurons are not spiking, since they are forced to quiescence by an external input current $I = -0.8$.

Unfortunately the quality of these results is very poor, because smaller values of ΔV,

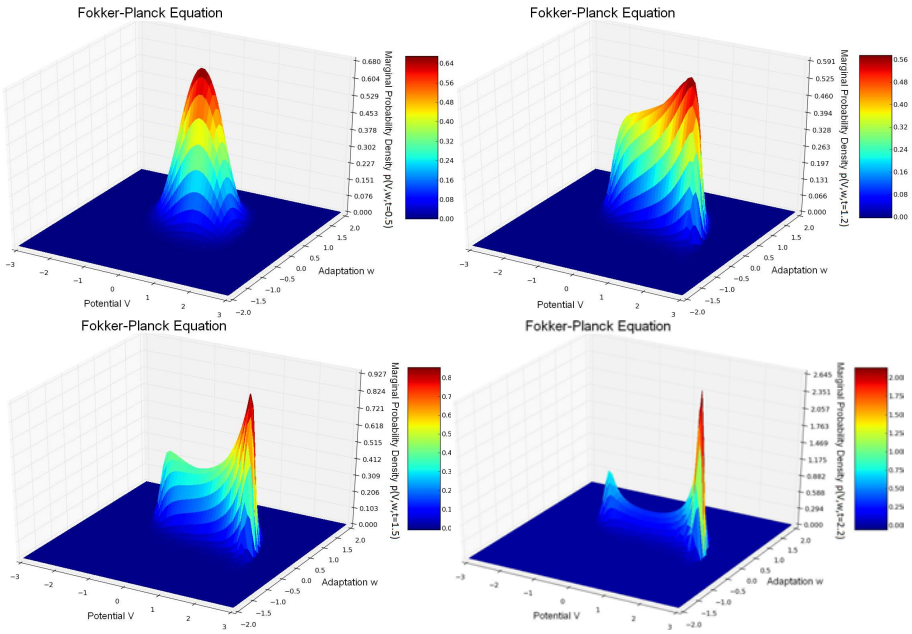

Figure 3.9: Marginal probability density $p(V, w, t)$ for the FitzHugh-Nagumo system, obtained from the simulation of equation 3.37 for the values of the parameters reported in Table 3.2. The probability density has been evaluated for $t = 0.5$ (top-left), 1.2 (top-right), 1.5 (bottom-left) and 2.2 (bottom-right), showing the split of the initial Gaussian peak centered inside the limit cycle.

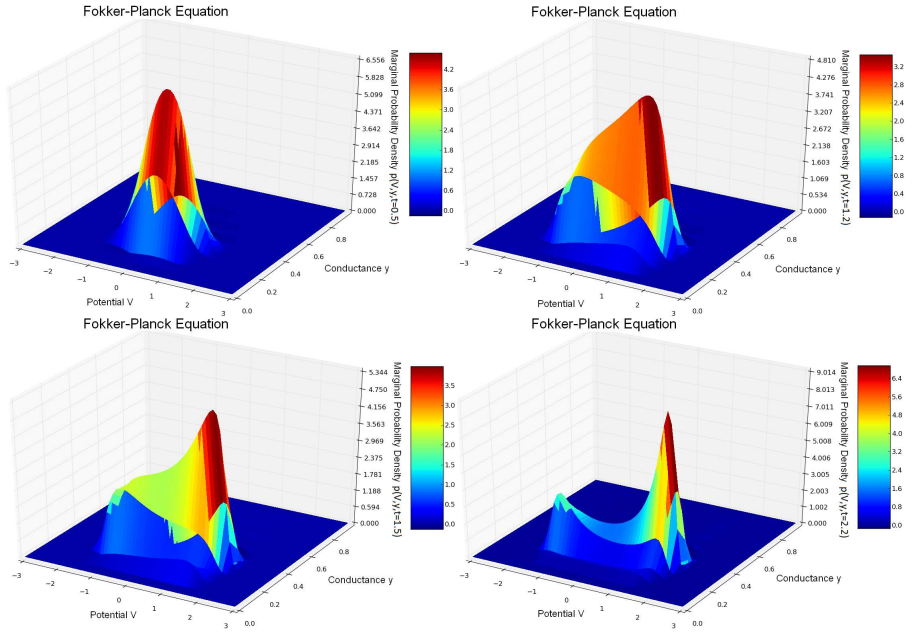

Figure 3.10: Marginal probability density $p(V, y, t)$ for the FitzHugh-Nagumo system, obtained from the simulation of equation 3.37 for the values of the parameters reported in Table 3.2. Again, the probability density has been evaluated for $t = 0.5$ (top-left), 1.2 (top-right), 1.5 (bottom-left) and 2.2 (bottom-right), showing the split of the initial Gaussian peak centered inside the limit cycle.

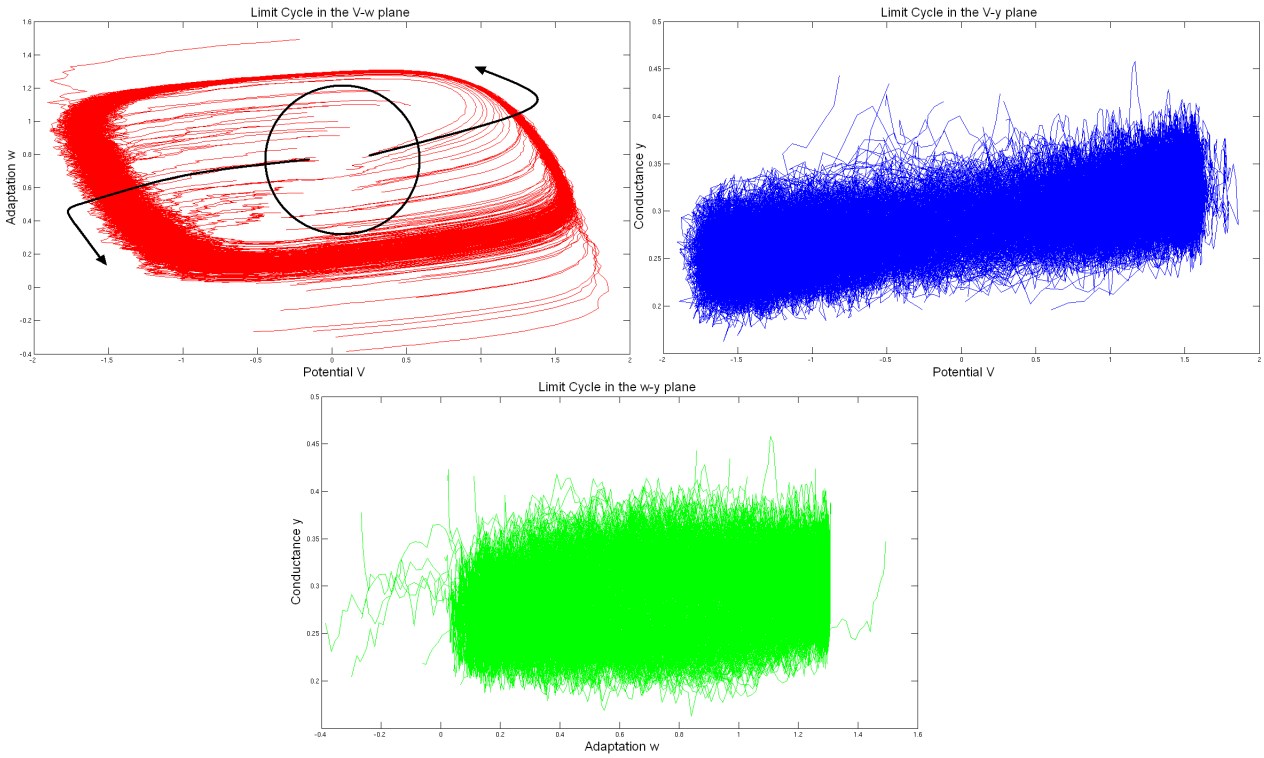

Figure 3.11: Projection of 100 trajectories in the (V, w) (top-left), (V, y) (top-right) and (w, y) (bottom) planes, obtained from the simulation of a large network for the values of the parameters reported in Table 3.2. The split of the trajectories is particularly visible in the (V, w) plane.

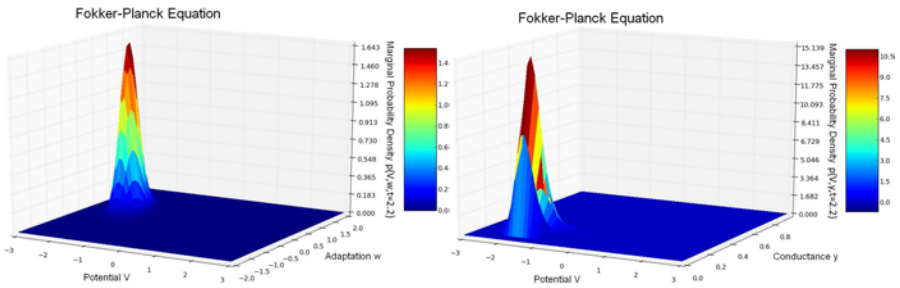

Figure 3.12: Marginal probability densities $p(V, w, t)$ (left-hand side) and $p(V, y, t)$ (right-hand side) for the FitzHugh-Nagumo system, obtained from the simulation of equation 3.37 for the values of the parameters reported in Table 3.2, with only the exception of the external input and the intensity of the background noise, which have been set respectively to $I = -0.8$ and $\sigma_1 = 0.45$. The probability density has been evaluated at $t = 2.2$, showing the formation of a rest state where most of the neurons are not spiking.

Δw, Δy and therefore also of Δt are required, making the simulation impossible to run in practice on a regular laptop. For this reason we have implemented the same algorithm on GPUs, improving the quality of the results and the execution speed of the numerical scheme. For more details, the interested reader is referred to [64][176][177].

3.3 Fisher information

According to 2.6, the Fisher information of a network with joint probability density $p\left(\overrightarrow{x}, \theta, t\right)$ is quantified by the following formula:

$$\mathcal{I}\left(\theta, t\right) = \int_{\mathbb{R}^{kN}} \left(\frac{\partial}{\partial \theta} \log p\left(\overrightarrow{x}, \theta, t\right) \right)^2 p\left(\overrightarrow{x}, \theta, t\right) d\overrightarrow{x} \tag{3.42}$$

where θ is the parameter with respect to which the Fisher information is calculated (in this thesis we will use $\theta = I$, supposing that the external input current is constant in time), while $\overrightarrow{x} = (x_0, x_1, ..., x_{N-1})$, and x_i is the set of the values of all the k variables that are used to describe the i-th neuron (so for example $x_i = V_i$ and $k = 1$ for the rate network, while $x_i = (V_i, w_i, y_i)$ and $k = 3$ for the FitzHugh-Nagumo network). It is well known that if the variables x_i are independent, then formula 3.42 for $\theta = I$ becomes:

$$\begin{aligned} \mathcal{I}\left(I, t\right) &= \sum_{i=0}^{N-1} \mathcal{I}_i\left(I, t\right) \\ \mathcal{I}_i\left(I, t\right) &= \int_{\mathbb{R}^k} \left(\frac{\partial}{\partial I} \log p\left(x_i, I, t\right) \right)^2 p\left(x_i, I, t\right) dx_i \end{aligned} \tag{3.43}$$

where $\mathcal{I}_i\left(I, t\right)$ is the Fisher information of the i-th neuron. In the cases studied in this chapter, all the neurons have common parameters and therefore the same marginal probability density, so formula 3.43 gives N times the Fisher information of a single neuron.

Therefore in the thermodynamic limit, when the neurons become independent if the condition 3.8 is satisfied, the total Fisher information of the system diverges, so the normalized Fisher information $\frac{\mathcal{I}(I,t)}{N}$ is a more useful quantity. However for $N \to \infty$ it is not possible to study the Fisher information per neuron as a function of N, because in this case the value of N is fixed, being infinite. Moreover, in the cases we have analyzed up to now, the correlation between the Brownian motions and that between the membrane potentials are always zero, so the Fisher information cannot be studied in terms of these quantities. Clearly the only possibility is to determine $\mathcal{I}(I,t)$ as a function of the input current. However, for the sake of coherence, this would imply, in the next chapters of this thesis, a study of the Fisher information as a function of the network size, the correlation of the neurons and the input, which is computationally prohibitive. In fact, as we said, all the simulations have been run on a regular laptop, see Software and Hardware. This means that, depending on the number of Monte Carlo simulations through which the statistics of the system are evaluated, some of the simulations shown in the next chapters can last for hours, therefore it is not possible, due to time requirements, to study the behavior of $\mathcal{I}(I,t)$ as a function of three variables (namely N, the correlation of the neurons and I) with this hardware. Therefore, considering also the fact that in this thesis we are interested mainly in the quantification of the finite size effects and of the correlation between the neurons, it is natural to give up the study of the Fisher information as a function of I. This explains why in the current section we do not show the behavior of $\frac{\mathcal{I}(I,t)}{N}$ as a function of I, because otherwise this would be in contradiction with the rest of the thesis. The reader is asked to wait patiently until Chapter 4, where finally we can start to analyze the Fisher information realistically due to the quantification of the finite size effects of the network, while its behavior in terms of the correlation between the neurons will be determined in Chapters 5, 6 and 7.

3.4 Partial conclusion

In this chapter we have seen how to use the mean-field theory in order to determine the activity of rate or spiking neurons in the thermodynamic limit. The FPE of the system has been calculated in two different ways. In Section 3.1.1 it has been obtained through the relation 3.9, which is a consequence of the law of large numbers for $N \to \infty$. Instead in Section 3.1.3 we have followed a longer path, since we have calculated the FPE in the thermodynamic limit passing through the FPE of the corresponding finite size network. In other words, the first approach is straightforward, since it "jumps" directly to the thermodynamic limit, but it does not allow us to determine the behavior of the probability density for a finite N. Instead the second method is more complex, but prepares the ground for the analysis of the finite size effects that will be performed in Chapter 4. The advantage of the mean-field technique is in the fact that it allows us to describe the whole network in a very economical way, using the independence of the neurons. For the rate model we have proved analytically that the probability density of the system is Gaussian for $N \to \infty$, if the initial conditions are Gaussian, while for the FitzHugh-Nagumo network it is highly non-Gaussian, therefore its probability density has been calculated numerically. However the mean-field theory does not allow us to study the finite size effects of the system, since it predicts the behavior of the network only in the thermodynamic limit. Moreover it can be applied only if the sources of noise in the system are independent, because otherwise the phenomenon of propagation of chaos does not occur anymore. For this reason it cannot be used to study the Fisher information as a function of the correlation of the noise, or of the correlation between the membrane potentials of the neurons. In Chapter 4 we show a possible way to deal with the problem of the finite size, using again independent sources of noise. Instead, in Chapters 5 and 6, we introduce a more efficient technique that allows us to study both the finite size effects and the behavior of the network with correlated noise, with special emphasis on their consequences for the Fisher information.

Chapter 4

Mayer's cluster expansion and finite size effects of the neural networks

In this chapter we develop a first approach for quantifying the finite size effects in a stochastic neural network. This technique is based on the Mayer's cluster expansion already developed for the physics of plasmas. Even if this methodology can be applied to every kind of neural equations, for the sake of simplicity in this chapter we consider only the case of the rate model, as shown in Sections 4.1 and 4.2. Due to the complexity of the resulting equations, they have been solved numerically for different network sizes, as shown in Section 4.3.Moreover, from the Mayer's cluster expansion we have developed an approximate series expansion for the Fisher information of the system, whose numerical results are given in Section 4.4.

4.1 Mayer's cluster expansion

In this section we describe the neural network using again formulae 3.1, 3.2 and 3.5. Due to the high complexity of the computations, we assume for simplicity that the synaptic weights are deterministic, namely $\sigma_3 = 0$. The Brownian motions $B_i^V(t)$ and the initial conditions $V_i(0)$ are still supposed to be independent, and the system has a finite size N. Moreover, we suppose again that the system is invariant under exchange of the neural indices. The basic idea of this chapter is to use the similarity between this neural network and a *plasma* that comes from particle physics. A plasma, namely the fourth state of matter, is made up of ionized gas, which is created using high temperatures or electromagnetic fields. Since the particles are ionized, their mean distances are very large, therefore they show very small levels of correlation. In other terms, the density of plasma is very small, and in this way the particles interact weakly through the Coulomb's force. In the case of our neural network, we are in a similar situation. In effect, for every finite N, if the number of connections is high enough, the correlation between the neurons is small, due to a partial propagation of chaos. Therefore, it is natural to study the neural network using the same mathematical formalism as the one developed for the physics of plasmas (this is not surprising, since the original work of Vlasov [164] introduced in Chapter 3 was actually on plasmas). In this context, a well known method is the so called *Mayer's cluster expansion*, which is based on the decomposition of the joint probability density of the system. In more detail, it consists in analyzing the contributions to the joint probability density that come from the interactions between pairs, triplets, quadruplets etc of particles (in our case of neurons). Since it involves the interactions between the single units of the system, clearly there is an important difference between plasmas and neural networks. In fact, in plasmas every particle is interacting with all the others, since the Coulomb's force is always active if all the particles have a non-zero electric charge. Instead, in the case of a neural network, the neurons are not always fully connected, since some synapses can be missing. Therefore we start to define the Mayer's cluster expansion in the simplest case of a fully

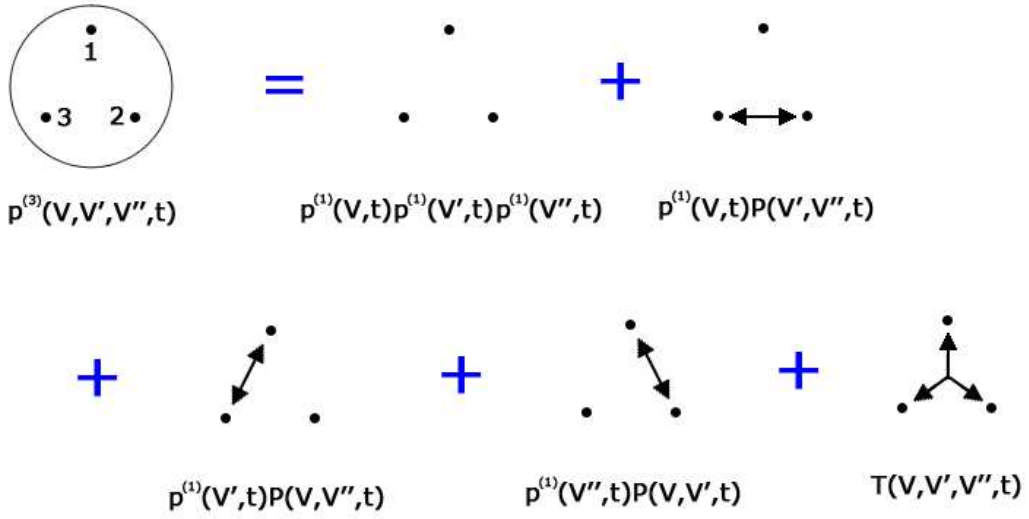

Figure 4.1: Mayer's cluster expansion for the marginal probability density $p^{(3)}(V, V', V'', t)$. The arrows represent the interactions between the neurons, while their absence means that the neurons are independent.

connected network, then we extend it to the general case. So, remembering the definitions 3.17, 3.18 and 3.19, the expansion in the fully connected case is:

$$
\begin{aligned}
p^{(2)}(V, V', t) =& p^{(1)}(V, t)\, p^{(1)}(V', t) + P(V, V', t) \\
p^{(3)}(V, V', V'', t) =& p^{(1)}(V, t)\, p^{(1)}(V', t)\, p^{(1)}(V'', t) + p^{(1)}(V, t)\, P(V', V'', t) \\
&+ p^{(1)}(V', t)\, P(V, V'', t) + p^{(1)}(V'', t)\, P(V, V', t) + T(V, V', V'', t)
\end{aligned}
\tag{4.1}
$$

where $P(V, V', t)$ and $T(V, V', V'', t)$ are the so called *pair and triplet correlation functions*, respectively. See also Figure 4.1 for an intuitive interpretation of this expansion.

The expansion can be extended to all orders, so for example $p^{(4)}(V, V', V'', V''', t)$ depends also on the *quadruplet correlation function* $Q(V, V', V'', V''', t)$, and so on and so forth. In the Appendix A we give the conditions that must be satisfied by the correlation functions

in order to have a self-consistent Mayer's cluster expansion. It is easy to understand why, for example, $P(V, V', t)$ is called pair correlation function. In fact, from 4.1, it is defined as $P(V, V', t) = p^{(2)}(V, V', t) - p^{(1)}(V, t)\, p^{(1)}(V', t)$, therefore it represents the "part" of the joint probability density of two neurons which is not taken into account by the independence term $p^{(1)}(V, t)\, p^{(1)}(V', t)$. So we conclude that intuitively $P(V, V', t)$ describes the correlation between the pair of neurons. More rigorously, if we quantify the correlation with the Pearson's coefficient, namely:

$$Corr(V_i(t), V_j(t)) = \frac{Cov(V_i(t), V_j(t))}{\sqrt{Var(V_i(t))\, Var(V_j(t))}}$$

$$Var(V_i(t)) = Cov(V_i(t), V_i(t))$$

where:

$$\begin{aligned} Cov(V_i(t), V_j(t)) &= \mathbb{E}[V_i(t)\, V_j(t)] - \mathbb{E}[V_i(t)]\, \mathbb{E}[V_j(t)] \\ &= \int_{\mathbb{R}^2} p^{(2)}(V, V', t)\, VV' dV dV' - \left[\int_{\mathbb{R}} p^{(1)}(V, t)\, V dV\right]^2 \\ &= \int_{\mathbb{R}^2} P(V, V', t)\, VV' dV dV \end{aligned}$$

and:

$$Var\left(V_i\left(t\right)\right) = \mathbb{E}\left[V_i^2\left(t\right)\right] - \mathbb{E}^2\left[V_i\left(t\right)\right]$$

$$= \int_{\mathbb{R}} p^{(1)}\left(V,t\right) V^2 dV - \left[\int_{\mathbb{R}} p^{(1)}\left(V,t\right) V dV\right]^2$$

then we can clearly see that the correlation is "proportional" to $P\left(V,V',t\right)$, which explains the choice of its name. At this point we open a short parenthesis about the quantification of the correlation function in this thesis: we always calculate the correlation between the membrane potentials, while many authors prefer to calculate the correlation between the spike-counts or the firing rates (see for example [158][178][179]). However, at least for relatively small values of the fluctuations of the membrane potentials, these two quantities are equivalent, as proved for the rate model in the Appendix B.

It is important to observe that in a fully connected network the functions $P\left(V,V',t\right)$, $T\left(V,V',V'',t\right)$ etc are the same for all the neurons, as a consequence of the invariance of the system under exchange of the neural indices and of the all-to-all connectivity matrix. We are now ready for the definition of the Mayer's cluster expansion in a general network. The difference with a fully connected network is that the function $P\left(V,V',t\right)$ between two neurons depends on the existence or absence of a connection between them, while in a fully connected network there is no distinction since all the neurons are connected. The same problem affects all the higher order correlation functions. So, for example, given three neurons i, j and k, the function $T\left(V,V',V'',t\right)$ has $2^3 = 8$ different forms, because in general we have eight different kinds of connections between these neurons (for example, i is connected with j but not with k and j is connected with k, or they are all connected, etc). The situation is even more complicated since in general the networks do not have symmetric connections, therefore the case with connection from i to j is different from the case with the connection in the opposite direction, namely from j to i. For simplicity, we consider only the case of symmetric connections, namely if i is connected to j, then j is

connected to i. The Mayer's cluster expansion is therefore defined as:

$$
\begin{aligned}
p^{(2)}_{C_\alpha}(V,V',t) =& p^{(1)}(V,t)\,p^{(1)}(V',t) + P_{C_\alpha}(V,V',t) \\
p^{(3)}_{C_\alpha,C_\beta,C_\gamma}(V,V',V'',t) =& p^{(1)}(V,t)\,p^{(1)}(V',t)\,p^{(1)}(V'',t) + p^{(1)}(V,t)\,P_{C_\beta}(V',V'',t) \\
&+ p^{(1)}(V',t)\,P_{C_\gamma}(V,V'',t) + p^{(1)}(V'',t)\,P_{C_\alpha}(V,V',t) + T_{C_\alpha,C_\beta,C_\gamma}(V,V',V'',t) \\
&\vdots
\end{aligned}
\tag{4.2}
$$

where $\alpha,\beta,\gamma \in \{0,1\}$. Here C_0 means that there is no connection between the two corresponding neurons, while C_1 means that the connection is present. So for example $P_{C_1}(V_i,V_j,t)$ is the pair correlation function between the i-th and the j-th neuron, that are supposed to be connected. Instead $T_{C_0,C_1,C_0}(V_i,V_j,V_k,t)$ is the triplet correlation function between these three neurons, where only the j-th and the k-th neuron are connected. In other words, the sequence $C_\alpha,C_\beta,C_\gamma$ represents the presence or absence of the connections between the neurons in positions $(i,j)\,(j,k)\,(k,i)$. We observe that we have not used the subscript C_x for the marginal density $p^{(1)}$. In fact this function does not depend on the presence or absence of the connections, because it describes a single neuron, whose behavior is always the same, due to the invariance of the system under exchange of the neural indices.

Usually correlations of higher and higher order are more and more negligible (we show an explicit example in Appendix E), therefore we can suppose that $Q \ll Tp^{(1)} \ll Pp^{(1)}p^{(1)} \ll p^{(1)}p^{(1)}p^{(1)}p^{(1)}$ etc. Therefore the higher order correlation functions can be neglected. In this chapter we take into account only the correlations between pairs of neurons, therefore we set $T_{C_\alpha,C_\beta,C_\gamma} = 0$ and the same for all the higher order correlation functions. This approximation allows us to break the infinite BBGKY hierarchy (see Chapter 3), trans-

forming it into a closed system of equations. Now, from Section 3.1.2 we know that:

$$\frac{\partial}{\partial t}p_{m_0}\left(V_{m_0},t\right)=-\frac{\partial}{\partial V_{m_0}}\left[\left(-\frac{V_{m_0}}{\tau}+I\left(t\right)\right)p_{m_0}\left(V_{m_0},t\right)\right]+\frac{1}{2}\sigma_1^2\frac{\partial^2}{\partial V_{m_0}^2}p_{m_0}\left(V_{m_0},t\right)$$

$$-\frac{\Lambda}{M}\frac{\partial}{\partial V_{m_0}}\sum_{j\in G_{m_0}}\int_{\mathbb{R}}S\left(V_j\right)p_{m_0,j}\left(V_{m_0},V_j,t\right)dV_j \tag{4.3}$$

which, using 4.2, can be equivalently rewritten in the following way:

$$\frac{\partial}{\partial t}p^{(1)}\left(V,t\right)=-\frac{\partial}{\partial V}\left[\left(-\frac{V}{\tau}+I\left(t\right)\right)p^{(1)}\left(V,t\right)\right]+\frac{1}{2}\sigma_1^2\frac{\partial^2}{\partial V^2}p^{(1)}\left(V,t\right)$$

$$-\Lambda\frac{\partial}{\partial V}\left\{\left[\int_{\mathbb{R}}S\left(V'\right)p^{(1)}\left(V',t\right)dV'\right]p^{(1)}\left(V,t\right)+\int_{\mathbb{R}}S\left(V'\right)P_{C_1}\left(V,V',t\right)dV'\right\} \tag{4.4}$$

Now we have to find the PDEs satisfied by the functions $P_{C_0}\left(V,V',t\right)$ and $P_{C_1}\left(V,V',t\right)$, in order to close the system of equations. This can be done using a similar trick to that used in Section 3.1.2, as shown in the next section.

4.2 Second equation

In order to obtain the PDEs satisfied by $P_{C_0}\left(V,V',t\right)$ and $P_{C_1}\left(V,V',t\right)$, we integrate the FPE 3.15 with respect to $\prod_{k\neq m_0,m_1}dV_k$ on the domain $\mathbb{R}^{N-2}$. So we obtain that:

$$\int_{\mathbb{R}^{N-2}}\frac{\partial}{\partial t}p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1}dV_k=\frac{\partial}{\partial t}p_{m_0,m_1}\left(V_{m_0},V_{m_1},t\right)$$

while for the drift term defined by 3.16 we obtain:

$$\int_{\mathbb{R}^{N-2}} \sum_{i=0}^{N-1} \left\{ -\frac{\partial}{\partial V_i} \left[f_i\left(\overrightarrow{V},t\right) p\left(\overrightarrow{V},t\right) \right] \right\} \prod_{k \neq m_0, m_1} dV_k$$

$$= -\int_{\mathbb{R}^{N-2}} \frac{\partial}{\partial V_{m_0}} \left[f_{m_0}\left(\overrightarrow{V},t\right) p\left(\overrightarrow{V},t\right) \right] \prod_{k \neq m_0, m_1} dV_k - \int_{\mathbb{R}^{N-2}} \frac{\partial}{\partial V_{m_1}} \left[f_{m_1}\left(\overrightarrow{V},t\right) p\left(\overrightarrow{V},t\right) \right] \prod_{k \neq m_0, m_1} dV_k$$

$$- \sum_{i \neq m_0, m_1} \int_{\mathbb{R}^{N-2}} \frac{\partial}{\partial V_i} \left[f_i\left(\overrightarrow{V},t\right) p\left(\overrightarrow{V},t\right) \right] \prod_{k \neq m_0, m_1} dV_k \tag{4.5}$$

The last term is equal to zero for the divergence theorem and the boundary conditions at infinity. Now consider the first term of 4.5. It writes:

$$-\int_{\mathbb{R}^{N-2}} \frac{\partial}{\partial V_{m_0}} \left[f_{m_0}\left(\overrightarrow{V},t\right) p\left(\overrightarrow{V},t\right) \right] \prod_{k \neq m_0, m_1} dV_k$$

$$= -\int_{\mathbb{R}^{N-2}} \frac{\partial}{\partial V_{m_0}} \left[\left(-\frac{V_{m_0}}{\tau} + I\left(t\right) \right) p\left(\overrightarrow{V},t\right) \right] \prod_{k \neq m_0, m_1} dV_k$$

$$- \frac{\Lambda}{M} \int_{\mathbb{R}^{N-2}} \frac{\partial}{\partial V_{m_0}} \left[\left(\sum_{j \in G_{m_0}} S\left(V_j\right) \right) p\left(\overrightarrow{V},t\right) \right] \prod_{k \neq m_0, m_1} dV_k \tag{4.6}$$

The first term of the right-hand side of 4.6 is:

$$-\frac{\partial}{\partial V_{m_0}} \left[\left(-\frac{V_{m_0}}{\tau} + I\left(t\right) \right) \int_{\mathbb{R}^{N-2}} p\left(\overrightarrow{V},t\right) \prod_{k \neq m_0, m_1} dV_k \right] = -\frac{\partial}{\partial V_{m_0}} \left[\left(-\frac{V_{m_0}}{\tau} + I\left(t\right) \right) p_{m_0,m_1}\left(V_{m_0}, V_{m_1}, t\right) \right]$$

The analysis of the second term of the right-hand side of 4.6 illuminates the difference in the behavior between connected and disconnected neurons. In effect, if $m_1 \notin G_{m_0}$, namely

if the m_0-th and the m_1-th neurons are not connected, this second term writes:

$$
\begin{aligned}
&-\frac{\Lambda}{M}\int_{\mathbb{R}^{N-2}}\frac{\partial}{\partial V_{m_0}}\left[\left(\sum_{j\in G_{m_0}}S\left(V_j\right)\right)p\left(\overrightarrow{V},t\right)\right]\prod_{k\neq m_0,m_1}dV_k\\
&=-\frac{\Lambda}{M}\frac{\partial}{\partial V_{m_0}}\sum_{j\in G_{m_0}}\int_{\mathbb{R}}S\left(V_j\right)\left[\int_{\mathbb{R}^{N-3}}p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1,j}dV_k\right]dV_j\\
&=-\frac{\Lambda}{M}\frac{\partial}{\partial V_{m_0}}\sum_{j\in G_{m_0}}\int_{\mathbb{R}}S\left(V_j\right)p_{m_0,m_1,j}\left(V_{m_0},V_{m_1},V_j,t\right)dV_j
\end{aligned}
\tag{4.7}
$$

Instead, if $m_1\in G_{m_0}$, namely if the m_0-th and the m_1-th neurons are connected, the second term of the right-hand side of 4.6 gives:

$$
\begin{aligned}
&-\frac{\Lambda}{M}\int_{\mathbb{R}^{N-2}}\frac{\partial}{\partial V_{m_0}}\left[\left(\sum_{j\in G_{m_0}}S\left(V_j\right)\right)p\left(\overrightarrow{V},t\right)\right]\prod_{k\neq m_0,m_1}dV_k\\
&=-\frac{\Lambda}{M}\frac{\partial}{\partial V_{m_0}}\sum_{j\in G_{m_0}}\int_{\mathbb{R}^{N-2}}S\left(V_j\right)p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1}dV_k\\
&=-\frac{\Lambda}{M}\frac{\partial}{\partial V_{m_0}}\left[\int_{\mathbb{R}^{N-2}}S\left(V_{m_1}\right)p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1}dV_k+\sum_{\substack{j\in G_{m_0}\\ j\neq m_1}}\int_{\mathbb{R}^{N-2}}S\left(V_j\right)p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1}dV_k\right]\\
&=-\frac{\Lambda}{M}\frac{\partial}{\partial V_{m_0}}\left\{S\left(V_{m_1}\right)\int_{\mathbb{R}^{N-2}}p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1}dV_k+\sum_{\substack{j\in G_{m_0}\\ j\neq m_1}}\int_{\mathbb{R}}S\left(V_j\right)\left[\int_{\mathbb{R}^{N-3}}p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1,j}dV_k\right]dV_j\right\}
\end{aligned}
$$

$$= -\frac{\Lambda}{M}\frac{\partial}{\partial V_{m_0}}\left[S\left(V_{m_1}\right)p_{m_0,m_1}\left(V_{m_0},V_{m_1},t\right)+\sum_{\substack{j\in G_{m_0}\\ j\neq m_1}}\int_{\mathbb{R}}S\left(V_j\right)p_{m_0,m_1,j}\left(V_{m_0},V_{m_1},V_j,t\right)dV_j\right] \tag{4.8}$$

To finish, the diffusion term in the FPE gives:

$$\frac{1}{2}\sigma_1^2\int_{\mathbb{R}^{N-2}}\left[\sum_{i=1}^{N}\frac{\partial^2}{\partial V_i^2}p\left(\overrightarrow{V},t\right)\right]\prod_{k\neq m_0,m_1}dV_k$$

$$= \frac{1}{2}\sigma_1^2\left[\int_{\mathbb{R}^{N-2}}\frac{\partial^2}{\partial V_{m_0}^2}p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1}dV_k+\int_{\mathbb{R}^{N-2}}\frac{\partial^2}{\partial V_{m_1}^2}p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1}dV_k+\sum_{i\neq m_0,m_1}\int_{\mathbb{R}^{N-2}}\frac{\partial^2}{\partial V_i^2}p\left(\overrightarrow{V},t\right)\prod_{k\neq m_0,m_1}dV_k\right]$$

$$= \frac{1}{2}\sigma_1^2\left[\frac{\partial^2}{\partial V_{m_0}^2}p_{m_0,m_1}\left(V_{m_0},V_{m_1},t\right)+\frac{\partial^2}{\partial V_{m_1}^2}p_{m_0,m_1}\left(V_{m_0},V_{m_1},t\right)\right]$$

having used again the divergence theorem and the boundary conditions at infinity. Putting everything together and remembering the following definitions (see 3.18):

$$p_{m_0,m_1}\left(\cdot,\cdot,t\right) = p_{C_\alpha}^{(2)}\left(\cdot,\cdot,t\right)$$

$$p_{m_0,m_1,j}\left(\cdot,\cdot,\cdot,t\right) = p_{C_\alpha,C_\beta,C_\gamma}^{(3)}\left(\cdot,\cdot,\cdot,t\right)$$

we obtain that the marginal probability density of two chosen neurons satisfies the following PDE:

$$\frac{\partial}{\partial t}p^{(2)}_{C_\alpha}(V,V',t) = -\frac{\partial}{\partial V}\left[\left(-\frac{V}{\tau}+I(t)\right)p^{(2)}_{C_\alpha}(V,V',t)\right] - \frac{\partial}{\partial V'}\left[\left(-\frac{V'}{\tau}+I(t)\right)p^{(2)}_{C_\alpha}(V,V',t)\right]$$

$$+\frac{1}{2}\sigma_1^2\left[\frac{\partial^2}{\partial V^2}p^{(2)}_{C_\alpha}(V,V',t)+\frac{\partial^2}{\partial V'^2}p^{(2)}_{C_\alpha}(V,V',t)\right]$$

$$-\frac{\Lambda}{M}\frac{\partial}{\partial V}\left[\alpha S(V')\,p^{(2)}_{C_\alpha}(V,V',t)+\sum_{\beta,\gamma=0}^{1}\zeta_{\alpha\beta\gamma}\int_{\mathbb{R}}S(V'')\,p^{(3)}_{C_\alpha,C_\beta,C_\gamma}(V,V',V'',t)\,dV''\right]$$

$$-\frac{\Lambda}{M}\frac{\partial}{\partial V'}\left[\alpha S(V)\,p^{(2)}_{C_\alpha}(V,V',t)+\sum_{\beta,\gamma=0}^{1}\zeta_{\alpha\beta\gamma}\int_{\mathbb{R}}S(V'')\,p^{(3)}_{C_\alpha,C_\beta,C_\gamma}(V,V',V'',t)\,dV''\right] \tag{4.9}$$

where $\zeta_{\alpha\beta\gamma}$ represents the number of times the sequence $C_\alpha, C_\beta, C_\gamma$ is generated by the sum over j in the terms 4.7 and 4.8. In general $\zeta_{\alpha\beta\gamma}$ depends on the connectivity matrix, and can be evaluated with combinatorial calculus. Now, equation 4.9 is the PDE satisfied by $p^{(2)}_{C_\alpha}(V,V',t)$, which is a function of $p^{(1)}(V,t)$ and $P_{C_\alpha}(V,V',t)$. The PDE for $p^{(1)}(V,t)$ is already known (see formula 4.4), so we need to transform 4.9 into the corresponding PDE satisfied by $P_{C_\alpha}(V,V',t)$. To this purpose, we observe that:

$$\frac{\partial}{\partial t}P_{C_\alpha}(V,V',t) = \frac{\partial}{\partial t}p^{(2)}_{C_\alpha}(V,V',t) - \left[\frac{\partial}{\partial t}p^{(1)}(V,t)\right]p^{(1)}(V',t) - p^{(1)}(V,t)\,\frac{\partial}{\partial t}p^{(1)}(V',t) \tag{4.10}$$

therefore substituting 4.9 into 4.10 and after some algebra we obtain:

$$\frac{\partial}{\partial t}P_{C_\alpha}\left(V,V',t\right) = -\frac{\partial}{\partial V}\left[\left(-\frac{V}{\tau}+I\left(t\right)\right)P_{C_\alpha}\left(V,V',t\right)\right] - \frac{\partial}{\partial V'}\left[\left(-\frac{V'}{\tau}+I\left(t\right)\right)P_{C_\alpha}\left(V,V',t\right)\right]$$

$$+\frac{1}{2}\sigma_1^2\left\{\left(\frac{\partial^2}{\partial V^2}+\frac{\partial^2}{\partial V'^2}\right)P_{C_\alpha}\left(V,V',t\right)\right\}$$

$$-\frac{\Lambda}{M}\left\{\alpha\left[S\left(V'\right)\frac{\partial}{\partial V}+S\left(V\right)\frac{\partial}{\partial V'}\right]\left[p^{(1)}\left(V,t\right)p^{(1)}\left(V',t\right)+P_{C_\alpha}\left(V,V',t\right)\right]\right.$$

$$+\sum_{\beta,\gamma=0}^{1}\zeta_{\alpha\beta\gamma}\left(\frac{\partial}{\partial V}+\frac{\partial}{\partial V'}\right)\left[R_{p^{(1)}}\left(t\right)p^{(1)}\left(V,t\right)p^{(1)}\left(V',t\right)\right.$$

$$\left.\left.+R_{p^{(1)}}\left(t\right)P_{C_\alpha}\left(V,V',t\right)+p^{(1)}\left(V,t\right)U_{P,C_\beta}\left(V',t\right)+p^{(1)}\left(V',t\right)U_{P,C_\gamma}\left(V,t\right)\right]\right\}$$

$$+\Lambda p^{(1)}\left(V',t\right)\frac{\partial}{\partial V}\left[R_{p^{(1)}}\left(t\right)p^{(1)}\left(V,t\right)+U_{P,C_1}\left(V,t\right)\right]$$

$$+\Lambda p^{(1)}\left(V,t\right)\frac{\partial}{\partial V'}\left[R_{p^{(1)}}\left(t\right)p^{(1)}\left(V',t\right)+U_{P,C_1}\left(V',t\right)\right] \tag{4.11}$$

where we have defined:

$$R_{p^{(1)}}\left(t\right) = \int_{\mathbb{R}} S\left(V\right)p^{(1)}\left(V,t\right)dV$$

$$U_{P,C_\alpha}\left(V,t\right) = \int_{\mathbb{R}} S\left(V'\right)P_{C_\alpha}\left(V,V',t\right)dV'$$

Therefore now we have a closed system of three coupled PDEs of the McKean-Vlasov type for $p^{(1)}\left(V,t\right)$, $P_{C_0}\left(V,V',t\right)$ and $P_{C_1}\left(V,V',t\right)$. Equation 4.11 can be considerably simplified neglecting all the terms multiplied by $\frac{\Lambda}{M}$, provided that M is large enough compared to

Λ (however, in the numerical simulations of Section 4.3 they will be always taken into account).

At this point, an important observation must be made. According to the Mayer's cluster expansion, in the case for example of a fully connected network, the joint probability density of the whole network is given by:

$$p\left(\overrightarrow{V},t\right)=\prod_{i=0}^{N-1}p^{(1)}\left(V_i,t\right)+\sum_{\substack{a,b\\a<b}}\left\{\left[\prod_{c\neq a,b}p^{(1)}\left(V_c,t\right)\right]P\left(V_a,V_b,t\right)\right\} \tag{4.12}$$

having neglected all the higher order correlation functions, like Q, PP, PT, and so on. Now, according to the numerical simulations (see Figures 4.2, 4.3 and 4.4), the function $p^{(1)}\left(V,t\right)$ depends weakly on N, while $P\left(V,V',t\right)$ is approximately proportional to $\frac{1}{N}$ (which confirms the phenomenon of propagation of chaos discussed in Chapter 3). Instead, the sum over a and b in formula 4.12 generates $\frac{N(N-1)}{2}$ terms, therefore roughly speaking the second term of the right-hand side of 4.12 is proportional to $\frac{N(N-1)}{2N}$. Since $p^{(1)}\left(V,t\right)$ depends weakly on N, for large network sizes the right-hand side of 4.12 dominates $\prod_{i=0}^{N-1}p^{(1)}\left(V_i,t\right)$. Moreover, in the Appendix E we prove that for a fully connected network, at least for weak Brownian motions, the 3-neurons correlation is zero. This means that we can neglect the contribution of the triplet correlation function T. According to the same analysis, the 4-neurons correlation is equal to the square of the 2-neurons correlation, which is proportional to $\frac{1}{N}$. This means that $Q\propto\frac{1}{N^2}$. Given a network with N neurons, the number of quadruplets that we can create is $\frac{N(N-1)(N-2)(N-3)}{24}$. This means that the total contribution of the quadruplet correlation function to the joint probability density is roughly $\frac{N(N-1)(N-2)(N-3)}{24N^2}$. This term grows quadratically with N, therefore for large networks it dominates the term generated by the pair correlation function, which is proportional to $\frac{N(N-1)}{2N}$. This proves that in general we cannot use 4.12 to evaluate the joint probability density of the whole network, because higher order contributions are not negligible. Notwithstanding, this formula can be used to evaluate the probability density of a small subset of the network, which con-

Neuron	Input	Synaptic Weights	Sigmoid Function	Phase-space	Other
$\tau = 1$	$I = 0.2$	$\Lambda = 1$	$T_{MAX} = 1$	$V_{min} = -3.5$	$\Delta t = 0.01$
$\sigma_2 = 0.5$	$\sigma_1 = 0.5$	$\sigma_3 = 0$	$\lambda = 1$	$V_{MAX} = 3.5$	
$\mu = 0.5$			$V_T = 0$	$\Delta V = 0.1$	

Table 4.1: Parameters used to generate Figures 4.2 - 4.6.

tains a number of neurons $\eta \ll N$. In effect, in this case the contributions of the functions P and Q are $\frac{\eta(\eta-1)}{2N}$ and $\frac{\eta(\eta-1)(\eta-2)(\eta-3)}{24N^2}$ respectively. These terms are much smaller than 1 if η is sufficiently smaller than N. The same happens to all the higher order terms that we have neglected in formula 4.12, therefore now this formula can be safely used. It is important to observe that for $N \to \infty$ and η finite, 4.12 gives formula 3.13, as it must be. This explains intuitively why in formula 3.13 we restricted the calculation of the joint probability density only to a finite number of neurons. For an alternative analysis of the contribution of the higher order terms, the reader is referred to [180].

4.3 Numerical simulations of the probability density

In this section we show the results obtained from the numerical simulation of formulae 4.4 and 4.11. For simplicity we consider only the case of a fully connected network (for which $\alpha = \beta = \gamma = 1$ and $\zeta_{111} = N - 2$), for the values of the parameters reported in Table 4.1. All results are shown at $t = 2$.

As in Section 3.2.2, these equations are solved on a regular laptop using the method of lines and the Newton-Cotes integration scheme of order 6. Figure 4.2 shows the results obtained for $p^{(1)}(V, t)$, while Figures 4.3 and 4.4 show the behavior of the function $P(V, V', t)$. The hypothesis $P \ll p^{(1)}p^{(1)}$ is seen to be satisfied.

From these results, we can see that $p^{(1)}(V, t)$ depends weakly on N and is approximately Gaussian. Moreover, also the shape of the function $P(V, V', t)$ does not change considerably

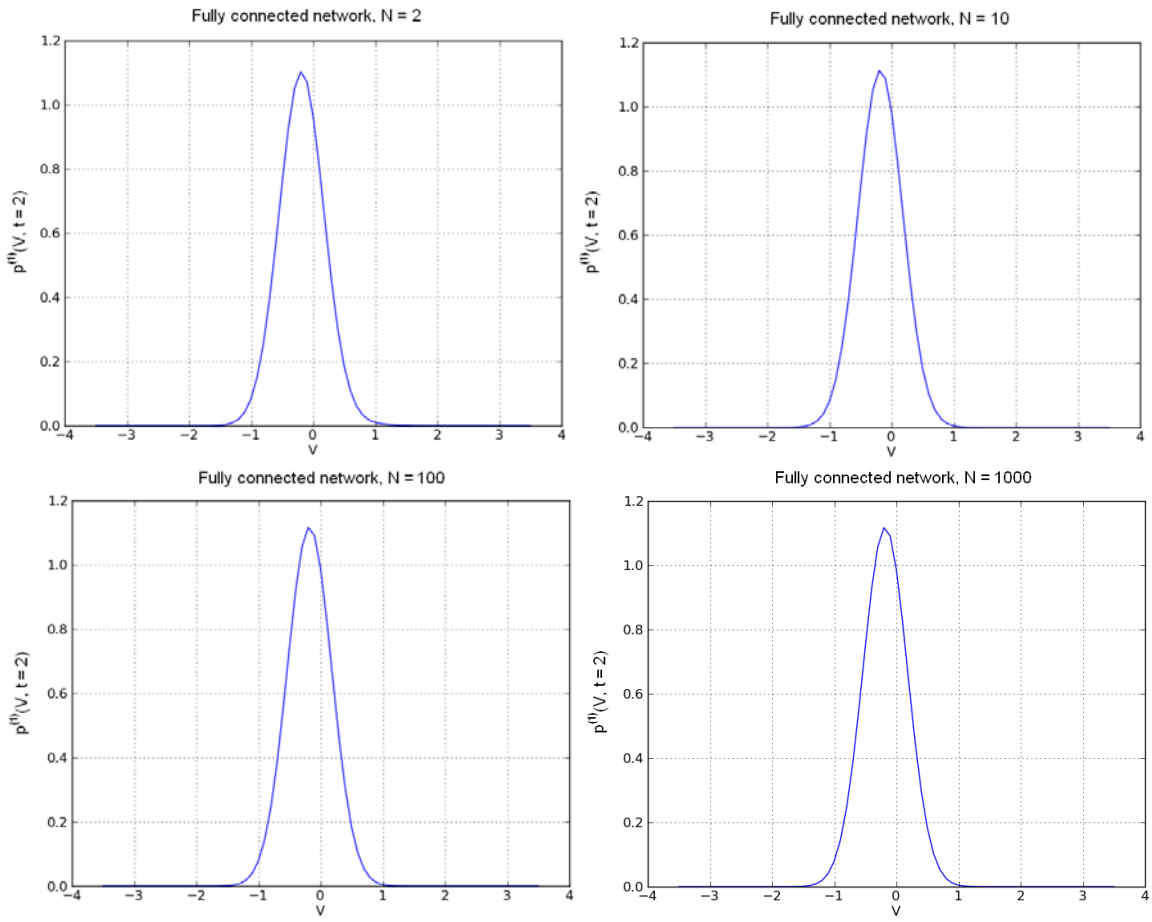

Figure 4.2: Function $p^{(1)}(V,t)$ in a fully connected network, obtained at $t = 2$ and for the values of the parameters reported in Table 4.1. This figure shows the marginal probability density of a single neuron for $N = 2$ (top-left), 10 (top-right), 100 (bottom-left) and 1000 (bottom-right), and its extremely weak dependence on the size of the network.

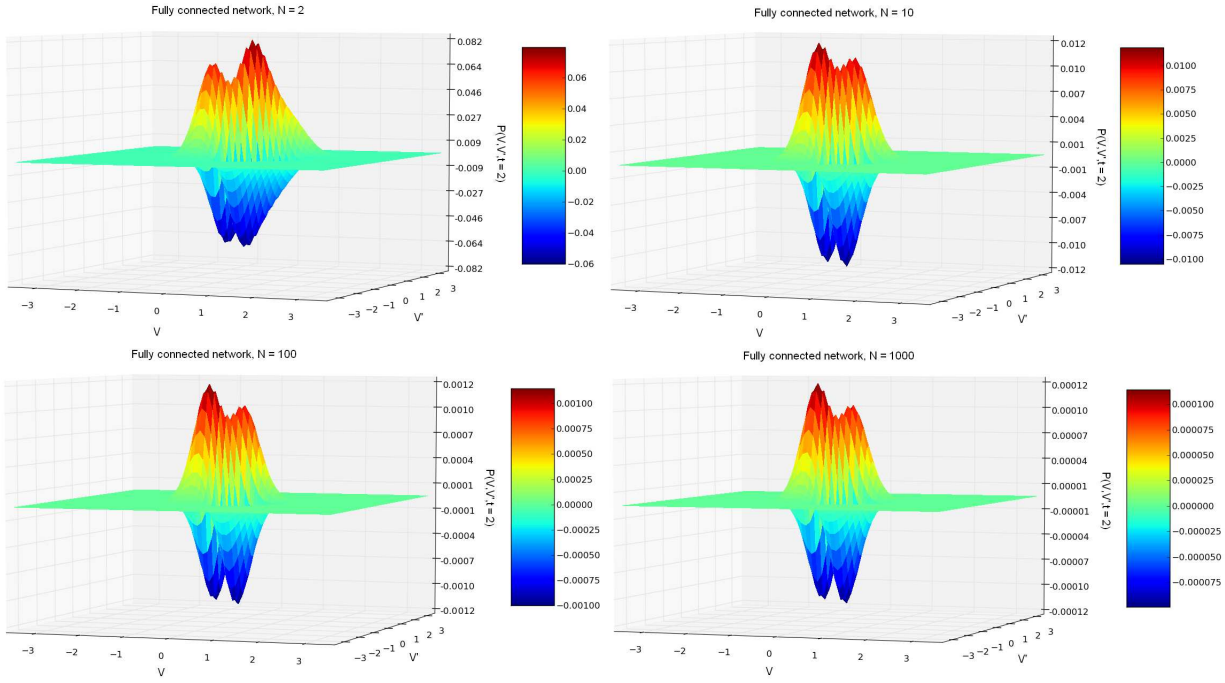

Figure 4.3: Front view of the function $P(V, V', t)$ obtained at $t = 2$ and for the values of the parameters reported in Table 4.1. This figure shows the front view of the pair correlation function in a fully connected network for different values of the network size, namely $N = 2$ (top-left), 10 (top-right), 100 (bottom-left) and 1000 (bottom-right). The presence of negative values, required in order to satisfy the constraint A.2, clearly shows that $P(V, V', t)$ cannot be interpreted as a probability density. Moreover, the maximum amplitude of the function, and therefore also the correlation between neurons, is approximately proportional to $\frac{1}{N}$. This result is proved analytically in Chapter 5 for weak sources of noise.

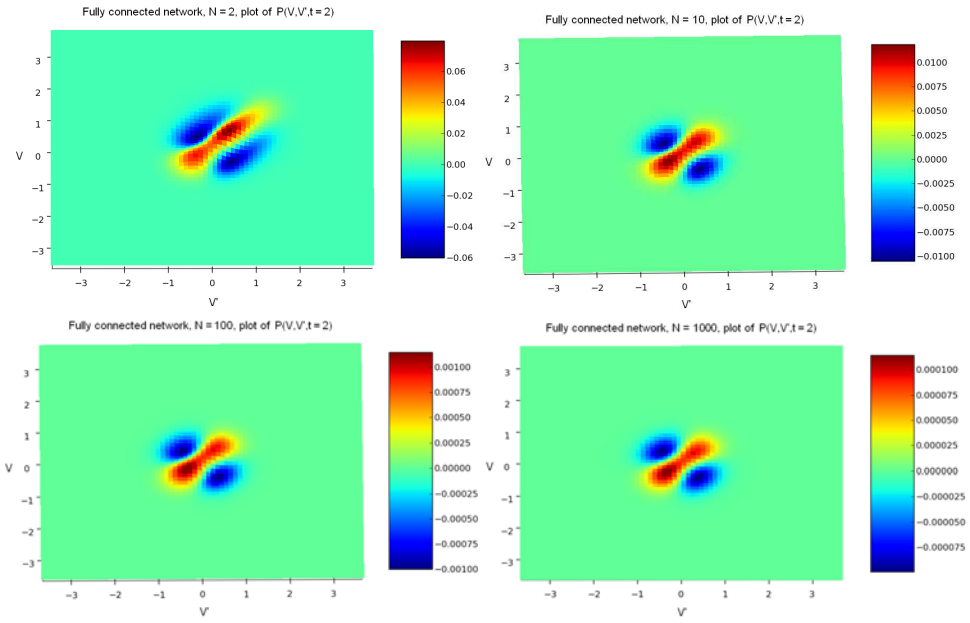

Figure 4.4: Top view of the function $P(V, V', t)$ obtained at $t = 2$ and for the values of the parameters reported in Table 4.1. This figure shows the top view of the pair correlation function in a fully connected network for different values of the network size, namely $N = 2$ (top-left), 10 (top-right), 100 (bottom-left) and 1000 (bottom-right). From it we can clearly see that $P(V, V', t) = P(V', V, t)$.

with N, with only the exception of small networks, and its amplitude seems to be roughly proportional to $\frac{1}{N}$, as we already said.

4.4 Numerical results for the Fisher information

In this section we show the numerical results for the Fisher information obtained from the following approximate expression:

$$\mathcal{I}(\theta,t) \approx$$

$$-\eta \int_{\mathbb{R}} p^{(1)}(V,t) \frac{\partial^2 log p^{(1)}(V,t)}{\partial\theta^2} dV$$

$$-\frac{\eta(\eta-1)}{2} \int_{\mathbb{R}^2} \left[p^{(1)}(V,t)\, p^{(1)}\left(V',t\right) + P\left(V,V',t\right)\right] \frac{\partial^2 d(V,V',t)}{\partial\theta^2} dV dV'$$

$$+\frac{\eta(\eta-1)}{4} \int_{\mathbb{R}^2} p^{(1)}(V,t)\, p^{(1)}\left(V',t\right) \frac{\partial^2 d^2(V,V',t)}{\partial\theta^2} dV dV' \tag{4.13}$$

$$+\frac{\eta(\eta-1)(\eta-2)}{2} \int_{\mathbb{R}^3} p^{(1)}(V,t)\, p^{(1)}\left(V',t\right) p^{(1)}\left(V'',t\right) \frac{\partial^2 \left[d(V,V',t)\, d(V,V'',t)\right]}{\partial\theta^2} dV dV' dV''$$

$$+\frac{\eta(\eta-1)(\eta-2)(\eta-3)}{8} \int_{\mathbb{R}^4} p^{(1)}(V,t)\, p^{(1)}\left(V',t\right) p^{(1)}\left(V'',t\right) p^{(1)}\left(V''',t\right) \frac{\partial^2 \left[d(V,V',t)\, d(V'',V''',t)\right]}{\partial\theta^2} dV dV' dV'' dV'''$$

where:

$$d(V_a,V_b,t) = \frac{P(V_a,V_b,t)}{p^{(1)}(V_a,t)\, p^{(1)}(V_b,t)}$$

The proof of this formula is rather complicated and can be found in the Appendix C (see formula C.9). Clearly, since we do not know the analytic expressions of $p^{(1)}(V,t)$ and $P(V,V',t)$, we cannot calculate exactly the integrals of 4.13. Moreover, the second order derivatives with respect to the parameters of the system are not known, therefore we have to use numerical techniques of integration. However, since these integrals are low-dimensional, there is no need to use Monte Carlo integration techniques, so we can simply use the standard numerical methods. In particular, here we use again the Newton-Cotes method of order 6, while the functions $p^{(1)}(V,t)$ and $P(V,V',t)$ are calculated numerically as explained in Section 4.3. For simplicity we use only the second order approximation of the Fisher information, namely formula 4.13. The second order derivatives with respect to

θ are calculated repeating three times the simulation of the PDE system and using formula 3.38 for $n = 1$, which gives:

$$\frac{d^2 f(x)}{dx^2} \approx \frac{1}{\Delta x^2} \left[f(x - \Delta x) - 2f(x) + f(x + \Delta x) \right] \tag{4.14}$$

In this thesis we are particularly interested in evaluating how well the neural network encodes the value of its external input current, therefore we choose $\theta = I$. Moreover the derivatives are calculated using $\Delta I = 10^{-4}$. It is important to observe that all the integrated quantities of this formula are well defined and converge quickly to zero when the membrane potentials are large, which implies that also the corresponding integrals are well defined. For example, the function:

$$\left[p^{(1)}(V,t)\, p^{(1)}(V',t) + P(V,V',t) \right] \frac{\partial^2 d(V,V',t)}{\partial I^2}$$

which appears in the second term of the right-hand side of 4.13, is shown in Figure 4.5.

Finally, Figure 4.6 shows the behavior of the Fisher information calculated at $t = 2$ as a function of N and for the values of the parameters shown in Table 4.1.

4.5 Partial conclusion

In this chapter we have developed an extension of the mean-field technique introduced in Chapter 3. With this approach, inspired by the physics of plasmas, we have described the finite size effects of the network and therefore its deviation from the situation of perfect independence of the neurons. All the results of Chapter 3 can be re-obtained with this new method if we let $N \to \infty$. We have also developed a perturbative expansion that allows us to calculate approximately the Fisher information of the system, obtaining that its encoding capability increases when the neurons are decorrelated. This seems to confirm the

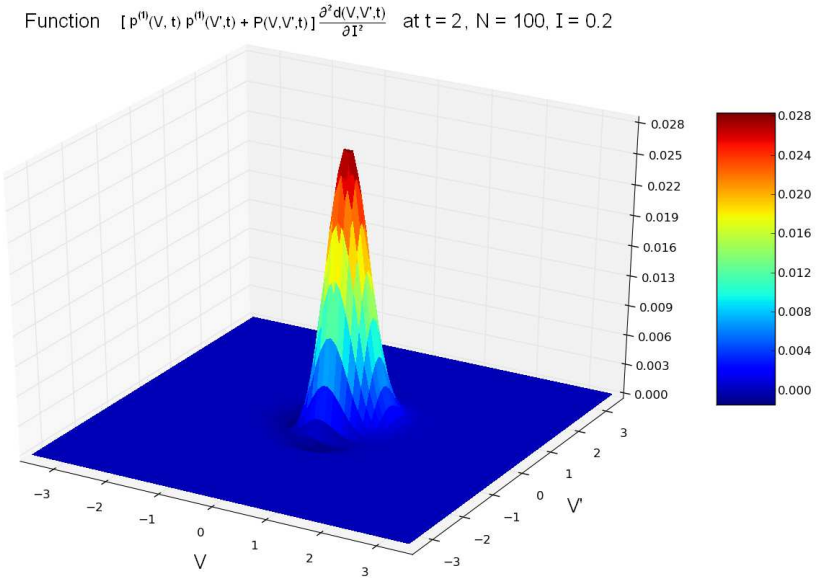

Figure 4.5: Behavior of $\left[p^{(1)}\left(V,t\right)p^{(1)}\left(V',t\right)+P\left(V,V',t\right)\right]\frac{\partial^2 d\left(V,V',t\right)}{\partial I^2}$ calculated for $t=2$, $N=100$ and for the values of the parameters reported in Table 4.1. Since this function converges quickly to zero for large values of the membrane potential, the second term of 4.13 is well defined. The same conclusion holds for all the other terms of the Fisher information, which are not shown here.

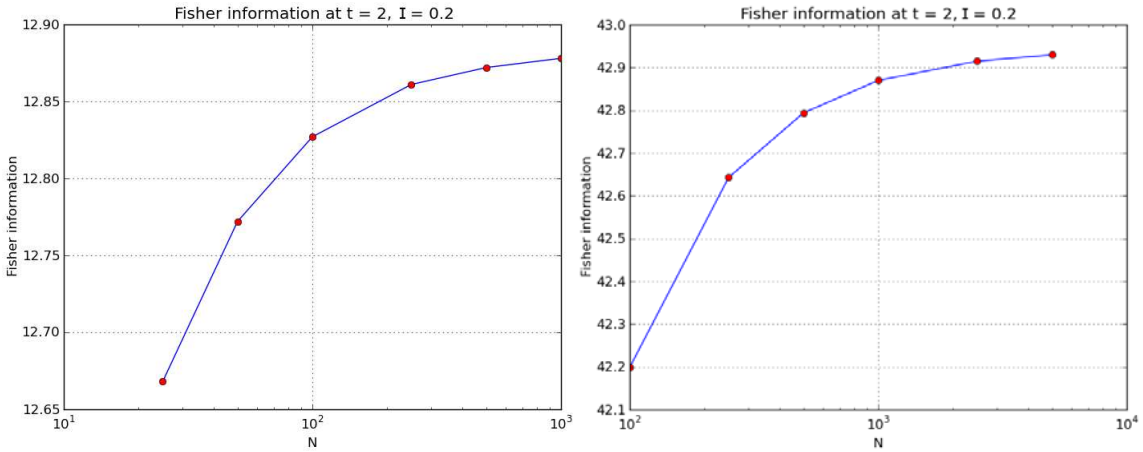

Figure 4.6: Fisher information of the subnetwork for $\eta=3$ and $N=[25,50,100,250,500,1000]$ (left-hand side), and for $\eta=10$ and $N=[100,250,500,1000,2500,5000]$ (right-hand side), obtained from 4.13 for $\Delta I=10^{-4}$, $t=2$ and for the values of the parameters shown in Table 4.1. In both cases the Fisher information of η neurons increases with N, reaching a saturation point when neurons get close to independence.

common belief that the network encodes better the external signals when the neurons are independent [178][181][182][183]. Notwithstanding, as we will see in the next chapters, the most interesting behavior of $\mathcal{I}(I,t)$ is obtained when the neurons are highly *correlated*. Unfortunately, with the Mayer's cluster expansion we cannot study the encoding capability of the system in the correlated case. This is due to the fact that the correlation of the neurons cannot be fixed arbitrarily, since it depends on the size of the network. Moreover the size must be high enough in order to use formula 4.12, so the neurons are highly decorrelated because propagation *to* chaos is very fast, as we have seen in Chapter 3. Therefore we cannot quantify the Fisher information keeping N fixed and changing the pair correlation. In order to do this, we need to introduce correlations between the Brownian motions that define the background noise, while in this chapter we have supposed that they are independent. So an extension of the model is required, and it will be developed in the next chapters. In other words, Figure 4.6 shows only one side of the coin, namely a small increase of the encoding capability of the network for highly decorrelated neurons. This result will be confirmed in Chapters 5 and 6 for the rate model with homogeneous synaptic weights, but there we will obtain much higher values of the Fisher information when the neurons are correlated and inhomogeneous. This proves that the belief of a better encoding capability for independent neurons is incorrect. The same result is obtained in Chapter 7 for the FitzHugh-Nagumo network. It is also important to observe that we have implemented numerically the Mayer's cluster expansion and the series expansion of the Fisher information only in the case of a fully connected network. Nevertheless in Chapter 6 we will prove analytically that the increase of $\mathcal{I}(I,t)$ for highly correlated and highly decorrelated neurons is a general phenomenon, that does not depend on the topology of the synaptic connections.

To conclude, we underline the main advantage of the Mayer's cluster expansion, namely the fact that it can be used to study the behavior of the network for arbitrarily high values of the noise, while the techniques developed in the next chapters work only for relatively weak sources of randomness.

Chapter 5

Perturbative analysis with strong synaptic weights

THIS chapter is devoted to the development of a perturbative technique which can be used to study the correlation structure of the rate model, for a generic number of neurons. This is somehow an extension of the previous analysis to the case of correlated Brownian motions in the background. In order to make the model more complete, we introduce also correlated initial conditions for the membrane potentials and correlated synaptic weights. In this approach we have to assume that the standard deviations of the three sources of randomness are relatively weak, since they are used as perturbative parameters, together with the existence of a stationary solution around which we perturb the solution. Moreover we have also to assume that the system is invariant under exchange of the neural indices, but the mean intensity of the synaptic weights can be arbitrarily high. This is an important point, since in Chapter 6 we will develop another perturbative approach with less hypotheses but with relatively weak mean synaptic weights. In Section 5.1 we describe the perturbative technique applied to the case of the rate model, while in Section 5.2 we use it in order to calculate the correlation structure of the network.

Instead, in Section 5.3 we apply this technique to special kinds of connectivity matrices, while in Section 5.4 we provide many numerical results that support the goodness of this approach. Moreover, in Section 5.5 we show the behavior of the correlation as a function of the intensity of the external input, and in Section 5.6 we prove in 3 different ways that in general the neurons do not become independent in this kind of network, preventing the use of the mean-field theory in the thermodynamic limit. To conclude, in Section 5.7 we use the perturbative method for calculating the Fisher information of the system in the case of approximately Gaussian behavior, obtaining that the network encodes better the value of the external input when the neurons are highly correlated.

5.1 Description of the model

Here we suppose again that the neural network is described by the rate equations that we have used in the Chapters 3 and 4, namely:

$$dV_i(t) = \left[-\frac{1}{\tau} V_i(t) + \sum_{j=0}^{N-1} J_{ij}(t) S(V_j(t)) + I_i(t) \right] dt + \sigma_1 dB_i(t) \tag{5.1}$$

with $i = 0, 1, ..., N-1$. We recall that the functions $B_i(t)$ are Brownian motions, which can be equivalently interpreted as a background noise for the membrane potentials $V_i(t)$ or as the stochastic component of the external input $I_i(t)$. In Chapters 3 and 4 the functions $B_i(t)$ were independent, while here we extend the model, supposing that in general the Brownian motions are correlated according to a covariance matrix Σ_1, whose components are:

$$[\Sigma_1]_{ij} = Cov\left(\frac{dB_i(t)}{dt}, \frac{dB_j(s)}{ds}\right) = C_{ij}^1 \delta(t-s) \tag{5.2}$$

$$C_{ij}^1 = \begin{cases} 1 & \text{if} \quad i = j \\ \\ C_1 & \text{if} \quad i \neq j \end{cases}$$

where $\delta(\cdot)$ is the Dirac delta function, while C_1 represents the correlation between two different Brownian motions (here the derivative of the Brownian motion is meant in the weak sense of distributions and is interpreted as white noise). In order to be a true covariance matrix, Σ_1 must be positive-semidefinite. Since it is symmetric, then it is positive-semidefinite if and only if its eigenvalues are non-negative. But Σ_1 is a circulant matrix, therefore its eigenvalues are $e_0 = 1 + C_1(N-1)$ and $e_i = 1 - C_1$, for $i = 1, 2, ..., N-1$. Therefore Σ_1 is positive-semidefinite if and only if $\frac{1}{1-N} \leq C_1 \leq 1$. We could increase the complexity of this correlation structure, since there is no technical difficulty in doing that, but we keep it simple for the sake of clarity.

Now we suppose also that the initial conditions are distributed according to the following multivariate normal probability density:

$$\overrightarrow{V}(0) \sim \mathcal{N}(\overrightarrow{\mu}, \Sigma_2) \tag{5.3}$$

where for simplicity:

$$\mu_i = \mu, \qquad i = 0, 1, ..., N-1 \tag{5.4}$$

$$\Sigma_2 = \sigma_2^2 \begin{bmatrix} 1 & C_2 & \cdots & C_2 \\ C_2 & 1 & \cdots & C_2 \\ \vdots & \vdots & \ddots & \vdots \\ C_2 & C_2 & \cdots & 1 \end{bmatrix} \tag{5.5}$$

Here σ_2 represents the initial standard deviation of each neuron, while C_2 is the initial correlation between pairs of neurons. As before, the matrix Σ_2 must be positive-semidefinite, and this is true if and only if $\frac{1}{1-N} \leq C_2 \leq 1$. Again, we could increase the complexity of this correlation structure, if desired.

About the synaptic connectivity matrix $J(t)$, we describe it using a different kind of model. In fact, in Chapter 3 we supposed that the synaptic weights fluctuated stochastically in time around a mean value. Here we suppose that their temporal evolution is deterministic, but the weights are distributed randomly over many repetitions of the network. This choice is dictated by the fact that adding stochastic fluctuations in the weights, according to 3.5, is somehow redundant, since we have already a similar kind of noise introduced by the Brownian motions in the background. Moreover, with this new kind of model for the synaptic weights we can match our results with those appeared in [184]. So, in detail, in this chapter the synaptic connectivity matrix $J(t)$ has random entries distributed according to the law:

$$J(t) \sim \mathcal{MN}\left(\overline{J} + \sigma_4 Z(t), \Omega_3, \Sigma_3\right) \tag{5.6}$$

This is the so called *matrix normal distribution* [185], namely the generalization of the multivariate normal distribution to the case of matrix-valued random variables. Here $\overline{J}$,

$Z(t)$, Ω_3 and Σ_3 are $N \times N$ deterministic matrices. In particular, $\overline{J} + \sigma_4 Z(t)$ represents the mean of $J(t)$, while Ω_3 and Σ_3 are its covariance matrices. We suppose that $\overline{J}$ has only two different kinds of entries, namely 0 (absence of connection) and Λ, where Λ is a free non-zero parameter. We also suppose that $Z(t)$ has general entries (with the obvious exception that $Z_{ij}(t) = 0$ if there is no connection from the j-th neuron to the i-th neuron, namely if $J_{ij}(t) = 0 \ \forall t$), therefore it is a source of inhomogeneity and time-variability for the connectivity matrix. We use for simplicity specific structures of the covariance matrices Ω_3 and Σ_3. Supposing that all the non-zero entries of $J(t)$ have the same standard deviation σ_3, it is possible to rewrite the matrix $J(t)$ in the following equivalent way:

$$J(t) = \overline{J} + \sigma_3 W + \sigma_4 Z(t) \tag{5.7}$$

$$W \sim \mathcal{MN}\left(0, \widetilde{\Omega}_3, \widetilde{\Sigma}_3\right) \tag{5.8}$$

where $\widetilde{\Omega}_3$ and $\widetilde{\Sigma}_3$ are normalized covariance matrices. Their explicit structure is not important, and the only thing that we need to know is that they are chosen in order to have:

$$Cov(W_{ij}, W_{kl}) = \begin{cases} 0 & \text{if} \quad (g(i,j) = 0) \vee (g(k,l) = 0) \\ 1 & \text{if} \quad (i = k) \wedge (j = l) \wedge (g(i,j) = 1) \\ C_3 & \text{otherwise} \end{cases} \tag{5.9}$$

where $g(x,y) = 0$ if there is no synaptic connection from the the y-th neuron to the x-th neuron (namely if $J_{xy}(t) = 0 \ \forall t$), and 1 otherwise, while C_3 is the correlation between two different and non-zero synaptic weights. We observe that the range of the possible values of C_3 in general depends on the topology of the connectivity matrix, and $W_{ij} = 0$ if

there is no connection from the j-th neuron to the i-th neuron (this is a consequence of the formulae 5.7 and 5.9). Again, as for the Brownian motions and the initial conditions, we could increase the complexity of this correlation structure, if desired.

Now we suppose that every neuron has *the same number of incoming connections*, that we call M. We observe that our assumptions imply that *the network is invariant under exchange of the neuronal indices*, which is the main hypothesis of this chapter. When M increases, each neuron receives a larger and larger input from the remainder of the network, therefore in order to fix this divergence we normalize the synaptic weight in the following way:

$$J\left(t\right) \to \frac{J\left(t\right)}{M}$$

as we did in Chapters 3 and 4. This normalization is intended to be used only when $M \neq 0$, because otherwise we obtain $J_{ij} = \frac{0}{0}$. For $M = 0$ the neurons have no incoming connections, therefore we have simply to set $J_{ij} = 0$.

To conclude, we also suppose that the external input current is deterministic (if we interpret $B_i\left(t\right)$ as the noise of the membrane potential) and given by:

$$\overrightarrow{I}\left(t\right) = \overrightarrow{\overline{I}} + \sigma_5 \overrightarrow{H}\left(t\right) \tag{5.10}$$

where the vector $\overrightarrow{\overline{I}}$ is time-independent and such that $\overline{I}_i = \overline{I}$, for $i = 0, 1, ..., N-1$. The vector $\overrightarrow{H}\left(t\right)$ has in general different and time-variable entries, so it is a source of inhomogeneity and time-variability. Now we define the following 2nd order perturbative expansion of the membrane potential:

$$V_i\left(t\right) \approx \mu + \sum_{m=1}^{5} \sigma_m Y_m^i\left(t\right) + \sum_{\substack{m,n=1 \\ m \leq n}}^{5} \sigma_m \sigma_n Y_{m,n}^i\left(t\right) \tag{5.11}$$

which will be used to obtain an approximate analytic solution of the system 5.1.

5.1.1 The system of equations

Now we put the perturbative expansion 5.11 and the expressions 5.7 and 5.10 for, respectively, the synaptic weights and the external input current inside the system 5.1. If all the parameters σ_m are small enough, we can expand the sigmoid function in a Taylor series around μ (see 5.4). In order to be rigorous, we have to determine the radius of convergence of the Taylor expansion of $S(V)$ for every value of V and to check if it is big enough compared to σ_m, because otherwise our technique cannot be applied. In fact, the various σ_m determine the order of magnitude of the fluctuations of V around μ, therefore it is important to check that V is inside the interval of convergence of the Taylor expansion of $S(V)$. To our knowledge, this calculation has been performed only for $V = 0$, so in Appendix D we show the general analysis, obtaining that in general the radius of convergence decreases with the slope parameter λ of the sigmoid function. So, supposing that λ is small enough, if we call:

$$\zeta_j = \sum_{m=1}^{5} \sigma_m Y_m^j(t) + \sum_{\substack{m,n=1 \\ m \leq n}}^{5} \sigma_m \sigma_n Y_{m,n}^j(t)$$

the Taylor expansion of the sigmoid function is:

$$
\begin{aligned}
S(\mu+\zeta_j) \approx & S(\mu)+S'(\mu)\zeta_j+\frac{1}{2}S''(\mu)\zeta_j^2 \\
\approx & S(\mu)+S'(\mu)\sum_{m=1}^{5}\sigma_m Y_m^j(t) \\
& +\sum_{\substack{m,n=1\\ m<n}}^{5}\sigma_m\sigma_n\left[S'(\mu)Y_{m,n}^j(t)+S''(\mu)Y_m^j(t)Y_n^j(t)\right] \\
& +\sum_{m=1}^{5}\sigma_m^2\left[S'(\mu)Y_{m,m}^j(t)+\frac{1}{2}S''(\mu)\left(Y_m^j(t)\right)^2\right]
\end{aligned}
$$

having neglected the terms with order higher than 2. Now we substitute this expansion of the sigmoid function inside the neural equation system and we equate the terms with the same σ coefficients, obtaining (here we report only the equations that we will actually use to compute the correlation structure in Section 5.2):

$$
\mu = \tau\left[\Lambda S(\mu)+\overline{I}\right] \tag{5.12}
$$

$$
dY_1^i(t) = \left[-\frac{1}{\tau}Y_1^i(t)+S'(\mu)\sum_{j=0}^{N-1}\overline{J}_{ij}Y_1^j(t)\right]dt+dB_i(t) \tag{5.13}
$$

$$
dY_2^i(t) = \left[-\frac{1}{\tau}Y_2^i(t)+S'(\mu)\sum_{j=0}^{N-1}\overline{J}_{ij}Y_2^j(t)\right]dt \tag{5.14}
$$

$$
dY_3^i(t) = \left[-\frac{1}{\tau}Y_3^i(t)+S'(\mu)\sum_{j=0}^{N-1}\overline{J}_{ij}Y_3^j(t)+S(\mu)\sum_{j=0}^{N-1}W_{ij}\right]dt \tag{5.15}
$$

$$dY_4^i(t) = \left[-\frac{1}{\tau} Y_4^i(t) + S'(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_4^j(t) + S(\mu) \sum_{j=0}^{N-1} Z_{ij}(t) \right] dt \tag{5.16}$$

$$dY_5^i(t) = \left[-\frac{1}{\tau} Y_5^i(t) + S'(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_5^j(t) + H_i(t) \right] dt \tag{5.17}$$

$$\vdots$$

$$dY_{1,4}^i(t) = \left[-\frac{1}{\tau} Y_{1,4}^i(t) + S'(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_{1,4}^j(t) + S'(\mu) \sum_{j=0}^{N-1} Z_{ij}(t) Y_1^j(t) + S''(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_1^j(t) Y_4^j(t) \right] dt \tag{5.18}$$

$$dY_{1,5}^i(t) = \left[-\frac{1}{\tau} Y_{1,5}^i(t) + S'(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_{1,5}^j(t) + S''(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_1^j(t) Y_5^j(t) \right] dt \tag{5.19}$$

$$\vdots$$

$$dY_{2,4}^i(t) = \left[-\frac{1}{\tau} Y_{2,4}^i(t) + S'(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_{2,4}^j(t) + S'(\mu) \sum_{j=0}^{N-1} Z_{ij}(t) Y_2^j(t) + S''(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_2^j(t) Y_4^j(t) \right] dt \tag{5.20}$$

$$dY_{2,5}^i(t) = \left[-\frac{1}{\tau} Y_{2,5}^i(t) + S'(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_{2,5}^j(t) + S''(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_2^j(t) Y_5^j(t) \right] dt \tag{5.21}$$

$$\vdots$$

$$dY_{3,4}^i(t) = \left[-\frac{1}{\tau} Y_{3,4}^i(t) + S'(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_{3,4}^j(t) + S'(\mu) \sum_{j=0}^{N-1} Z_{ij}(t) Y_3^j(t) + S'(\mu) \sum_{j=0}^{N-1} W_{ij} Y_4^j(t) + S''(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_3^j(t) Y_4^j(t) \right] dt \tag{5.22}$$

$$dY_{3,5}^i(t) = \left[-\frac{1}{\tau} Y_{3,5}^i(t) + S'(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_{3,5}^j(t) + S'(\mu) \sum_{j=0}^{N-1} W_{ij} Y_5^j(t) + S''(\mu) \sum_{j=0}^{N-1} \overline{J}_{ij} Y_3^j(t) Y_5^j(t) \right] dt \tag{5.23}$$

$$\vdots$$

Equation 5.12 is algebraic and non-linear, therefore must be solved numerically. 5.13 is the only stochastic differential equation of the set and can be solved analytically, since it is linear with constant coefficients. Equations 5.14 - 5.17 are ordinary, and can be solved in the same way as 5.13. To conclude, equations 5.18 - 5.23 determine the functions $Y^i_{m,n}(t)$, and depend on the terms $Y^i_m(t)$, which have been calculated at the previous step. Being linear and with constant coefficients, they can be integrated analytically as a function of the already known functions $Y^i_m(t)$.

5.1.2 The initial conditions

The perturbative expansion 5.11 at $t = 0$ gives:

$$V_i(0) \approx \mu + \sum_{m=1}^{5} \sigma_m Y^i_m(0) + \sum_{\substack{m,n=1 \\ m \leq n}}^{5} \sigma_m \sigma_n Y^i_{m,n}(0)$$

Moreover, according to 5.3, we have $V_i(0) \sim \mathcal{N}\left(\mu, \sigma_2^2\right) = \mu + \sigma_2 \mathcal{N}(0,1)$, so from the comparison it must be:

$$Y^i_2(0) \sim \mathcal{N}(0,1) \tag{5.24}$$

$$Y^i_m(0) = 0, \qquad m = 1,3,4,5 \tag{5.25}$$

$$Y^i_{m,n}(0) = 0, \qquad \forall (m,n) : \; m \leq n \tag{5.26}$$

So we have $V_i(0) = \mu + \sigma_2 Y^i_2(0)$ and therefore:

$$Cov\left(V_i\left(0\right),V_j\left(0\right)\right)=\sigma_2^2 Cov\left(Y_2^i\left(0\right),Y_2^j\left(0\right)\right)$$

But from 5.5 we also know that:

$$Cov\left(V_i\left(0\right),V_j\left(0\right)\right)=\begin{cases}\sigma_2^2 & \text{if} \quad i=j\\ \\ \sigma_2^2 C_2 & \text{if} \quad i\neq j\end{cases}$$

so from the comparison it must be that:

$$Cov\left(Y_2^i\left(0\right),Y_2^j\left(0\right)\right)=\begin{cases}1 & \text{if} \quad i=j\\ \\ C_2 & \text{if} \quad i\neq j\end{cases} \tag{5.27}$$

5.1.3 Solutions of the equations

As we said at the end of Section 5.1.1, the algebraic equation 5.12 is non-linear, therefore it cannot be solved exactly. However, the differential equations satisfied by all the functions $Y_m^i\left(t\right)$ and $Y_{m,n}^i\left(t\right)$ are linear with constant coefficients, therefore they can be solved analytically. In particular, the equations 5.13 - 5.17 can be solved directly. Instead the remaining equations are functions of the previous $Y_m^i\left(t\right)$, that we have already calculated. For example, according to 5.18, $Y_{1,4}^i\left(t\right)$ can be determined analytically as a function of $Y_1^i\left(t\right)$ and $Y_4^i\left(t\right)$, which are already known from the equations 5.13 and 5.16. Now we introduce the *fundamental matrix* $\Phi\left(t\right)$ such that:

$$\Phi(t) = e^{At} \tag{5.28}$$

$$A_{ij} = \begin{cases} -\frac{1}{\tau} & \text{if} \quad i = j \\ \\ \overline{J}_{ij} S'(\mu) & \text{if} \quad i \neq j \end{cases}$$

where:

$$\mathcal{J} = \overline{J} S'(\mu) \tag{5.29}$$

is the *effective connectivity* matrix of the network. Therefore the solutions of all the functions $Y_m^i(t)$ can be obtained straightforwardly as follows:

$$Y_1^i(t) = \sum_{j=0}^{N-1} \int_0^t \left[\Phi(t-s)\right]_{ij} dB_j(s) \tag{5.30}$$

$$Y_2^i(t) = \sum_{j=0}^{N-1} \Phi_{ij}(t) Y_2^j(0) \tag{5.31}$$

$$Y_3^i(t) = S(\mu) \sum_{j,k=0}^{N-1} W_{jk} \int_0^t \left[\Phi(t-s)\right]_{ij} ds \tag{5.32}$$

$$Y_4^i(t) = S(\mu) \sum_{j,k=0}^{N-1} \int_0^t \left[\Phi(t-s)\right]_{ij} Z_{jk}(s)\, ds \tag{5.33}$$

$$Y_5^i(t) = \sum_{j=0}^{N-1} \int_0^t \left[\Phi(t-s)\right]_{ij} H_j(s)\, ds \tag{5.34}$$

$$\vdots$$

$$\begin{aligned}Y_{1,4}^{i}(t) =& S'(\mu)\sum_{j,k,l=0}^{N-1}\int_0^t \left[\Phi(t-s)\right]_{ij}\left\{\int_0^s \left[\Phi(s-u)\right]_{kl} dB_l(u)\right\} Z_{jk}(s)\,ds \\ &+ S(\mu)S''(\mu)\sum_{j,k,l,m,n=0}^{N-1}\overline{J}_{jk}\int_0^t \left[\Phi(t-s)\right]_{ij}\left\{\int_0^s \left[\Phi(s-u)\right]_{kl} dB_l(u)\right\}\left\{\int_0^s \left[\Phi(s-u)\right]_{km} Z_{mn}(u)\,du\right\} ds\end{aligned} \tag{5.35}$$

$$Y_{1,5}^{i}(t) = S''(\mu)\sum_{j,k,l,m=0}^{N-1}\overline{J}_{jk}\int_0^t \left[\Phi(t-s)\right]_{ij}\left\{\int_0^s \left[\Phi(s-u)\right]_{kl} dB_l(u)\right\}\left\{\int_0^s \left[\Phi(s-u)\right]_{km} H_m(u)\,du\right\} ds \tag{5.36}$$

$$\vdots$$

$$\begin{aligned}Y_{2,4}^{i}(t) =& S'(\mu)\sum_{j,k,l=0}^{N-1} Y_2^l(0)\int_0^t \left[\Phi(t-s)\right]_{ij}\Phi_{kl}(s)\,Z_{jk}(s)\,ds \\ &+ S(\mu)S''(\mu)\sum_{j,k,l,m,n=0}^{N-1}\overline{J}_{jk}Y_2^l(0)\int_0^t \left[\Phi(t-s)\right]_{ij}\Phi_{kl}(s)\left\{\int_0^s \left[\Phi(s-u)\right]_{km} Z_{mn}(u)\,du\right\} ds\end{aligned} \tag{5.37}$$

$$Y_{2,5}^{i}(t) = S''(\mu)\sum_{j,k,l,m=0}^{N-1}\overline{J}_{jk}Y_2^l(0)\int_0^t \left[\Phi(t-s)\right]_{ij}\Phi_{kl}(s)\left\{\int_0^s \left[\Phi(s-u)\right]_{km} H_m(u)\,du\right\} ds \tag{5.38}$$

$$\vdots$$

$$\begin{aligned}Y_{3,4}^{i}(t) =& S(\mu)S'(\mu)\sum_{j,k,l,m=0}^{N-1} W_{lm}\int_0^t \left[\Phi(t-s)\right]_{ij}\left\{\int_0^s \left[\Phi(s-u)\right]_{kl} du\right\} Z_{jk}(s)\,ds \\ &+ S(\mu)S'(\mu)\sum_{j,k,l,m=0}^{N-1} W_{jk}\int_0^t \left[\Phi(t-s)\right]_{ij}\left\{\int_0^s \left[\Phi(s-u)\right]_{kl} Z_{lm}(u)\,du\right\} ds \\ &+ S^2(\mu)S''(\mu)\sum_{j,k,l,m,n,p=0}^{N-1}\overline{J}_{jk}W_{lm}\int_0^t \left[\Phi(t-s)\right]_{ij}\left\{\int_0^s \left[\Phi(s-u)\right]_{kl} du\right\}\left\{\int_0^s \left[\Phi(s-u)\right]_{kn} Z_{np}(u)\,du\right\} ds\end{aligned} \tag{5.39}$$

$$\begin{aligned}Y_{3,5}^{i}(t) =& S'(\mu)\sum_{j,k,l=0}^{N-1} W_{jk}\int_0^t \left[\Phi(t-s)\right]_{ij}\left\{\int_0^s \left[\Phi(s-u)\right]_{kl} H_l(u)\,du\right\} ds \\ &+ S(\mu)S''(\mu)\sum_{j,k,l,m,n=0}^{N-1}\overline{J}_{jk}W_{lm}\int_0^t \left[\Phi(t-s)\right]_{ij}\left\{\int_0^s \left[\Phi(s-u)\right]_{kl} du\right\}\left\{\int_0^s \left[\Phi(s-u)\right]_{kn} H_n(u)\,du\right\} ds\end{aligned} \tag{5.40}$$

$$\vdots$$

To conclude, we have performed a perturbative expansion around a stationary state μ because in this way the equations 5.13 - 5.23 have constant coefficients and therefore they can be solved exactly using the fundamental matrix 5.28. Had we performed the perturbative expansion around a non-stationary state, we would have obtained a system of differential equations with time-varying coefficients, whose general solution is not known. In this case, the best thing that we can try is to write the solution in terms of the *Magnus expansion* [186], but this introduces another approximation to the real solution of the neural network.

In this chapter we have also supposed that the system is invariant under exchange of the neural indices: for this reason we have used the same stationary solution μ, the same (unperturbed) input current $\overline{I}$ and the same number of incoming connections for all the neurons in the network. This invariance is required in order to ensure that the effective connectivity matrix $\mathcal{J}$ given by 5.29 has the same structure as the real and unperturbed connectivity matrix $\overline{J}$. In this way the fundamental matrix $\Phi(t)$ can be calculated using the properties of $\overline{J}$, as explained in Section 5.3. If the system is not invariant under exchange of the neural indices, $\mathcal{J}$ does not inherit the structure of $\overline{J}$, therefore the technique introduced in this chapter cannot be used anymore (see also the discussion at the end of Section 5.3.2). To conclude, it is important to observe that even if we have chosen structures of Σ_1, Σ_2, Σ_3 and Ω_3 that are invariant under exchange of the neural indices, their invariance is not required here: we have used it only to simplify the final formulae that we will obtain in Section 5.2. Therefore in principle inhomogeneous structures can be used for these covariance matrices.

5.2 Correlation structure of the network

In this section we want to calculate the correlation structure of the membrane potentials, according to the perturbative expansion 5.11. Since the covariance function is bilin-

ear, we have to compute it for all the possible combinations of the pairs $\left(Y_m^i(t), Y_n^j(t)\right)$, $\left(Y_m^i(t), Y_{n,p}^j(t)\right)$ and $\left(Y_{m,n}^i(t), Y_{p,q}^j(t)\right)$. However we do not have to consider the terms of order 4, like $\sigma_1^2\sigma_2^2 Cov\left(Y_{1,1}^i(t), Y_{2,2}^j(t)\right)$, because they are incomplete. In effect, in the perturbative expansion of $V_i(t)$, we did not consider the terms of order 3, like $\sigma_1^2\sigma_2 Y_{1,1,2}^i(t)$, that generate contributions of order 4 in the formula of the covariance. So the terms of order 4 cannot be considered in the expansion of the covariance, therefore the final formula is of order 3. For simplicity we suppose that the Brownian motions, the initial conditions and the uncertainty of the synaptic weights are 3 *independent* random processes (and indeed there is a priori no obvious reason to think that they are correlated), so all the cross terms like $\sigma_1\sigma_2 Cov\left(Y_1^i(t), Y_2^j(t)\right)$, $\sigma_1\sigma_2 Cov\left(Y_2^i(t), Y_1^j(t)\right)$, $\sigma_1^2\sigma_3 Cov\left(Y_{1,1}^i(t), Y_3^j(t)\right)$, ... are equal to zero (however, if desired, we could assume non-zero correlations between these 3 sources of randomness, since there is no technical difficulty in the calculations, only the problem to compute many non-zero cross terms). Let us show it with an example:

$$
\begin{aligned}
Cov\left(Y_1^i(t), Y_2^j(t)\right) =& Cov\left(\int_0^t \sum_{k=0}^{N-1} \left[\Phi(t-s)\right]_{ik} dB_k(s), \sum_{l=0}^{N-1} \Phi_{jl}(t) Y_2^l(0)\right) \\
=& \sum_{k,l=0}^{N-1} \Phi_{jl}(t) Cov\left(\int_0^t \left[\Phi(t-s)\right]_{ik} dB_k(s), Y_2^l(0)\right) \\
=& 0
\end{aligned}
$$

since $B_k(s)$ and $Y_2^l(0)$ are independent by assumption. Moreover, due to the Isserlis' theorem [187], we obtain also that all the terms in the covariance proportional to $\sigma_m^2\sigma_n$ with $m, n = 1, 2, 3$ are equal to zero, like $\sigma_1^2\sigma_2$ and σ_3^3. The same thing happens to all the terms proportional to $\sigma_m^2\sigma_n$, with $m = 4, 5$ and $n = 1, 2, 3$. This is due to the fact that, according to the Isserlis' theorem again, the mean of the product of any odd number of zero-mean normal processes is equal to zero. We show it with an example:

$$Cov\left(Y_2^i(t), Y_{2,2}^j(t)\right)$$

$$= Cov\left(\sum_{k=0}^{N-1} \Phi_{ik}(t) Y_2^k(0), \frac{1}{2} S''(\mu) \sum_{l,m,n,p=0}^{N-1} \overline{J}_{lm} Y_2^n(0) Y_2^p(0) \int_0^t \left[\Phi(t-s)\right]_{jl} \Phi_{mn}(s) \Phi_{mp}(s)\, ds\right)$$

$$= \frac{1}{2} S''(\mu) \sum_{k,l,m,n,p=0}^{N-1} \Phi_{ik}(t) \overline{J}_{lm} \left\{\int_0^t \left[\Phi(t-s)\right]_{jl} \Phi_{mn}(s) \Phi_{mp}(s)\, ds\right\} Cov\left(Y_2^k(0), Y_2^n(0) Y_2^p(0)\right)$$

$$= 0$$

because:

$$Cov\left(Y_2^k(0), Y_2^n(0) Y_2^p(0)\right) = \mathbb{E}\left[Y_2^k(0) Y_2^n(0) Y_2^p(0)\right] - \mathbb{E}\left[Y_2^k(0)\right] \mathbb{E}\left[Y_2^n(0) Y_2^p(0)\right] = 0$$

since $\mathbb{E}\left[Y_2^k(0) Y_2^n(0) Y_2^p(0)\right] = 0$ by the Isserlis' theorem and $\mathbb{E}\left[Y_2^k(0)\right] = 0$, because $Y_2^k(0) \sim \mathcal{N}(0,1)$ from 5.24. In the final formula of the covariance, also the terms proportional to $\sigma_m \sigma_n$ and $\sigma_m^2 \sigma_n$ with $m, n = 4, 5$ are zero, because the functions $Y_m^i(t)$ and $Y_{m,n}^i(t)$ are deterministic for $m, n = 4, 5$. In fact, for example, from the formulae 5.33 and 5.34 we can easily see that the functions $Y_4^i(t)$ and $Y_5^i(t)$ depend only on deterministic functions ($\Phi(t)$, $Z_{jk}(t)$ and $H_j(t)$), deterministic parameters (τ and all the parameters of $S(\cdot)$) and deterministic initial conditions ($Y_4^i(0) = Y_5^i(0) = 0$, from 5.25), and therefore they are deterministic as well. Also the terms proportional to $\sigma_m \sigma_n \sigma_p$ for $m = 4, 5$ and $n \neq p$ are zero, due to the independence of the sources of randomness or to the fact that $Y_m^i(t)$ is deterministic for $m = 4, 5$. In the same way the terms obtained from the covariance of $Y_m^i(t)$ for $m = 4, 5$ with $Y_{n,n}^i(t)$ for $n = 1, 2, 3$ are zero due to the fact that the first function is deterministic.

To conclude, the only non-zero terms in the final formula of the covariance are those proportional to σ_m^2 for $m = 1, 2, 3$, and those obtained from the covariance of $Y_{m,n}^i(t)$ with $Y_m^i(t)$, for $m = 1, 2, 3$ and $n = 4, 5$. So the final formula for the covariance is:

$$
\begin{aligned}
&Cov\left(V_i(t), V_j(t)\right) \\
&= \sigma_1^2 Cov\left(Y_1^i(t), Y_1^j(t)\right) + \sigma_2^2 Cov\left(Y_2^i(t), Y_2^j(t)\right) + \sigma_3^2 Cov\left(Y_3^i(t), Y_3^j(t)\right) \\
&+ \sigma_4 \Big\{ \sigma_1^2 \left[Cov\left(Y_1^i(t), Y_{1,4}^j(t)\right) + Cov\left(Y_{1,4}^i(t), Y_1^j(t)\right) \right] + \sigma_2^2 \left[Cov\left(Y_2^i(t), Y_{2,4}^j(t)\right) + Cov\left(Y_{2,4}^i(t), Y_2^j(t)\right) \right] \\
&+ \sigma_3^2 \left[Cov\left(Y_3^i(t), Y_{3,4}^j(t)\right) + Cov\left(Y_{3,4}^i(t), Y_3^j(t)\right) \right] \Big\} \\
&+ \sigma_5 \Big\{ \sigma_1^2 \left[Cov\left(Y_1^i(t), Y_{1,5}^j(t)\right) + Cov\left(Y_{1,5}^i(t), Y_1^j(t)\right) \right] + \sigma_2^2 \left[Cov\left(Y_2^i(t), Y_{2,5}^j(t)\right) + Cov\left(Y_{2,5}^i(t), Y_2^j(t)\right) \right] \\
&+ \sigma_3^2 \left[Cov\left(Y_3^i(t), Y_{3,5}^j(t)\right) + Cov\left(Y_{3,5}^i(t), Y_3^j(t)\right) \right] \Big\}
\end{aligned}
\tag{5.41}
$$

Even if the third order terms can be calculated exactly using the Isserlis' theorem (and even if in principle we can extend this perturbative expansion to any higher order), due to their complexity in this chapter we consider only the second order terms, that is equivalent to say that we truncate the perturbative expansion 5.11 of the membrane potential at the first order. After some algebra we obtain:

$$Cov\left(Y_1^i(t), Y_1^j(t)\right) = \sum_{k=0}^{N-1} \int_0^t \left[\Phi(t-s)\right]_{ik} \left[\Phi(t-s)\right]_{jk} ds + C_1 \sum_{\substack{k,l=0 \\ k\neq l}}^{N-1} \int_0^t \left[\Phi(t-s)\right]_{ik} \left[\Phi(t-s)\right]_{jl} ds \tag{5.42}$$

$$Cov\left(Y_2^i(t), Y_2^j(t)\right) = \sum_{k=0}^{N-1} \Phi_{ik}(t)\,\Phi_{jk}(t) + C_2 \sum_{\substack{k,l=0 \\ k\neq l}}^{N-1} \Phi_{ik}(t)\,\Phi_{jl}(t) \tag{5.43}$$

$$\begin{aligned} Cov\left(Y_3^i(t), Y_3^j(t)\right) =& \frac{S^2(\mu)}{M} \sum_{k=0}^{N-1} \left\{ \int_0^t \left[\Phi(t-s)\right]_{ik} ds \right\} \left\{ \int_0^t \left[\Phi(t-s)\right]_{jk} ds \right\} \\ &+ C_3 S^2(\mu) \left\{ \sum_{k,l=0}^{N-1} \left\{ \int_0^t \left[\Phi(t-s)\right]_{ik} ds \right\} \left\{ \int_0^t \left[\Phi(t-s)\right]_{jl} ds \right\} \right. \\ &\left. - \frac{1}{M} \sum_{k=0}^{N-1} \left\{ \int_0^t \left[\Phi(t-s)\right]_{ik} ds \right\} \left\{ \int_0^t \left[\Phi(t-s)\right]_{jk} ds \right\} \right\} \end{aligned} \tag{5.44}$$

So now the covariance $Cov\left(V_i(t), V_j(t)\right)$ is known for all the possible pairs (i,j), with $i,j = 0,1,...,N-1$, therefore we can determine the correlation structure of the network using the formula for the Pearson's correlation coefficient:

$$Corr_2\left(V_i(t), V_j(t)\right) = \frac{Cov\left(V_i(t), V_j(t)\right)}{\sqrt{Var\left(V_i(t)\right) Var\left(V_j(t)\right)}} \tag{5.45}$$

where:

$$Var\left(V_i(t)\right) = Cov\left(V_i(t), V_i(t)\right) \tag{5.46}$$

is the variance of the stochastic process $V_i(t)$. The subscript "2" means that this is a

correlation between a pair of neurons.

In order to determine the higher order correlations between triplets, quadruplets, quintuplets etc of neurons, we have to extend the Pearson's formula in the following way. The natural generalization of the covariance for n functions is:

$$\kappa_n\left(V_{i_0}(t), V_{i_1}(t), ..., V_{i_{n-1}}(t)\right) = \mathbb{E}\left[\prod_{j=0}^{n-1}\left(V_{i_j}(t) - \overline{V}_{i_j}(t)\right)\right] \tag{5.47}$$

This is known as the *joint cumulant* of the functions $V_{i_0}(t), V_{i_1}(t), ..., V_{i_{n-1}}(t)$. Unfortunately this is not enough, because as with the Pearson's correlation coefficient, we want to normalize the joint cumulant in order to find a function that is in the range $[-1, 1]$. To this purpose, we can observe that:

$$\left|\mathbb{E}\left[\prod_{j=0}^{n-1}\left(V_{i_j}(t) - \overline{V}_{i_j}(t)\right)\right]\right| \leq \mathbb{E}\left[\left|\prod_{j=0}^{n-1}\left(V_{i_j}(t) - \overline{V}_{i_j}(t)\right)\right|\right] \leq \left\{\prod_{j=0}^{n-1}\mathbb{E}\left[\left|V_{i_j}(t) - \overline{V}_{i_j}(t)\right|^n\right]\right\}^{\frac{1}{n}}$$

having used the fact that $|a+b| \leq |a|+|b|$ at the first step and a special case of the Hölder's inequality at the second. Therefore we have:

$$\left|\frac{\mathbb{E}\left[\prod_{j=0}^{n-1}\left(V_{i_j}(t) - \overline{V}_{i_j}(t)\right)\right]}{\sqrt[n]{\prod_{j=0}^{n-1}\mathbb{E}\left[\left|V_{i_j}(t) - \overline{V}_{i_j}(t)\right|^n\right]}}\right| \leq 1 \tag{5.48}$$

This means that the function:

$$Corr_n\left(V_{i_0}(t), V_{i_1}(t), ..., V_{i_{n-1}}(t)\right) \stackrel{def}{=} \frac{\mathbb{E}\left[\prod_{j=0}^{n-1}\left(V_{i_j}(t) - \overline{V}_{i_j}(t)\right)\right]}{\sqrt[n]{\prod_{j=0}^{n-1}\mathbb{E}\left[\left|V_{i_j}(t) - \overline{V}_{i_j}(t)\right|^n\right]}} \tag{5.49}$$

is in the range $[-1, 1]$, therefore it is a good formula to express higher order correlations. We can see that for $n = 2$ it gives the Pearson's formula, as it should be. Now, all these means $\mathbb{E}$ can be computed using the Isserlis' theorem as we did for the covariance, so in principle we can determine also the higher order correlation structure of the neural network. However, in practice, this gives rise to combinatorial problems with different levels of complexity when $V_{i_j}(t)$ does not have the same behavior for different values of i_j, namely if the deterministic matrix $\overline{J}_{ij} + \sigma_4 Z_{ij}(t)$ and the input vector $\overrightarrow{I}(t)$ do not have strong symmetries. Therefore, for simplicity, in the Appendix E we show only the fully connected case with the same synaptic weights and the same input current for all the neurons.

5.3 Calculation of the fundamental matrix

As we can see from the formulae 5.42, 5.43 and 5.44, the correlation structure is a function of the matrices $\Phi(t)$ and $\Phi(t)\,\Phi^T(t)$. Therefore we need to compute them for different kinds of connectivity matrices $\overline{J}$. In general this is not an easy task, but however in some special cases they can be obtained as discussed in Sections 5.3.1 and 5.3.2.

5.3.1 Block circulant matrices with circulant blocks

Given two positive integers R and S, with $1 \leq R, S \leq N$, we suppose that $\overline{J}$ is an $N \times N$ block circulant matrix (with $N = RS$) of the form:

$$\overline{J} = \frac{\Lambda}{M} \begin{bmatrix} b^{(0)} & b^{(1)} & \dots & b^{(R-1)} \\ b^{(R-1)} & b^{(0)} & \dots & b^{(R-2)} \\ \vdots & \vdots & \ddots & \vdots \\ b^{(1)} & b^{(2)} & \dots & b^{(0)} \end{bmatrix} \tag{5.50}$$

where $b^{(0)}, b^{(1)}, ..., b^{(R-1)}$ are $S \times S$ circulant matrices:

$$b^{(i)} = \begin{bmatrix} b_0^{(i)} & b_1^{(i)} & \cdots & b_{S-1}^{(i)} \\ b_{S-1}^{(i)} & b_0^{(i)} & \cdots & b_{S-2}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ b_1^{(i)} & b_2^{(i)} & \cdots & b_0^{(i)} \end{bmatrix} \tag{5.51}$$

All the entries $b_j^{(i)}$, for $i = 0, 1, ..., R-1$ and $j = 0, 1, ..., S-1$, can only be equal to 0 or 1, with only the exception of $b_0^{(0)}$ that must always be equal to 0 in order to avoid the self-connections. R can be interpreted as the number of neural populations, and S as the number of neurons per population. Due to this particular structure of the connectivity matrix, all the neurons have the same number of incoming synaptic connections M, as required. This analysis includes the special case when the matrix $\overline{J}$ is circulant (obtained for $R = 1$ or $S = 1$). As we said in Chapter 3, in the context of *Graph Theory*, a network whose adjacency matrix is circulant is called *circulant graph* (see Figure 3.1) and is usually represented by the notation $C_N(1, 2, ..., q)$. Moreover we have to recall that even if in Graph Theory the connections are often represented through undirected unweighted graphs, which means that the connectivity matrix is symmetric, in this section we do not assume in general that $\overline{J}$ is symmetric.

Now we want to calculate the matrices $\Phi(t)$ and $\Phi(t)\Phi^T(t)$ in terms of the eigenquantities of $\overline{J}$. The eigenvalues of $\overline{J}$ are the collection of the eigenvalues of the following matrices:

$$\widetilde{b}^{(i)} = \sum_{j=0}^{R-1} e^{\frac{2\pi}{R} ij\iota} b^{(j)} \tag{5.52}$$

where $\iota = \sqrt{-1}$. Since the matrices $\widetilde{b}^{(i)}$ are circulant, we can compute their eigenvalues $e_j^{(i)}$ as follows:

$$e_j^{(i)} = \sum_{k=0}^{S-1} e^{\frac{2\pi}{S} jk\iota} \left[\widetilde{b}^{(i)}\right]_{0k} = \sum_{k=0}^{S-1} \sum_{l=0}^{R-1} e^{2\pi\left(\frac{jk}{S} + \frac{il}{R}\right)\iota} b_k^{(l)} \tag{5.53}$$

Instead the matrix of the eigenvectors of $\overline{J}$ is:

$$\begin{aligned} Q =& F_R \otimes F_S \\ [F_K]_{ij} =& \ \tfrac{1}{\sqrt{K}} e^{\frac{2\pi}{K} ij\iota}, \qquad K = R, S, \qquad i, j = 0, 1, ..., K-1 \end{aligned} \tag{5.54}$$

where $\otimes$ is the Kronecker product. Now, for $k = 0, 1, ..., N-1$, we call a_k the eigenvalues of $A = -\frac{1}{\tau} Id_N + \overline{J} S'(\mu)$ (where Id_N is the $N \times N$ identity matrix) and e_k the eigenvalues of $\overline{J}$ (namely the collection of all the $e_j^{(i)}$, with $k = iS + j$), while we call $\overrightarrow{v}_k$ and $\overrightarrow{w}_k$ their respective eigenvectors. Therefore we have $a_k = -\frac{1}{\tau} + e_k S'(\mu)$ and $\overrightarrow{v}_k = \overrightarrow{w}_k$. Moreover, using also the fact that the matrix e^{At} can be diagonalized and is real, we can write:

$$\Phi(t) = e^{At} = QD(t) Q^*$$

$$\Phi(t) \Phi^T(t) = e^{At} \left(\left[e^{A(t)} \right]^T \right)^* = QD(t) Q^* QD^*(t) Q^* = QD(t) D^*(t) Q^*$$

where $*$ is the element-by-element complex conjugation, and $D(t) = diag\left(e^{a_0 t}, e^{a_1 t} ..., e^{a_{N-1} t}\right)$. Here we have used the fact that $D(t)$ and Q are symmetric matrices and also the identity:

$$Q^* Q = (F_R^* \otimes F_S^*)(F_R \otimes F_S) = (F_R^* F_R) \otimes (F_S^* F_S) = Id_{RS} = Id_N$$

due to the mixed-product property of the Kronecker product and to the elementary identity $F_K^* F_K = Id_K$. Now, since:

$$[F_R \otimes F_S]_{ij} = [F_R]_{mn} [F_S]_{pq} = \frac{1}{\sqrt{N}} e^{2\pi \left(\frac{mn}{R} + \frac{pq}{S} \right) \iota}$$

$$m = \left\lfloor \frac{i}{S} \right\rfloor, \qquad n = \left\lfloor \frac{j}{S} \right\rfloor, \qquad p = i - mS, \qquad q = j - nS$$

we conclude that:

$$\begin{aligned} \Phi_{ij}(t) &= \frac{1}{N} \sum_{k=0}^{N-1} e^{\left[-\frac{1}{\tau} + e_k S'(\mu)\right] t} f_{ijk} \\ \left[\Phi(t) \Phi^T(t)\right]_{ij} &= \frac{1}{N} \sum_{k=0}^{N-1} e^{2\left[-\frac{1}{\tau} + \Re(e_k) S'(\mu)\right] t} f_{ijk} \end{aligned} \tag{5.55}$$

where $\Re(e_k)$ represents the real part of e_k, while:

$$f_{ijk} = \left[F_R \otimes F_S\right]_{ik} \left[F_R \otimes F_S\right]^*_{kj} = e^{2\pi \left\{ \frac{1}{R} \left\lfloor \frac{k}{S} \right\rfloor \left(\left\lfloor \frac{i}{S} \right\rfloor - \left\lfloor \frac{j}{S} \right\rfloor \right) + \frac{k}{S}(i-j) \right\} \iota}$$

These formulae seem to give complex-valued functions, but due to the particular structure of the eigenvalues e_k and of the function f_{ijk}, their imaginary parts are equal to zero (see Appendix F). Therefore the covariance is a real function, as it should be.

Now we show an explicit example of this technique, namely the case when the blocks of the matrix $\overline{J}$ have the following symmetric circulant band structure:

$$b^{(i)} = \begin{bmatrix} 1-\delta_{i0} & 1 & \cdots & 1 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ 1 & 1-\delta_{i0} & \ddots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \ddots & 0 & \ddots & 1 \\ 1 & & \ddots & \ddots & \ddots & & \ddots & \ddots & & 0 \\ 0 & \ddots & & \ddots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & & \ddots & \ddots & \ddots & & \ddots & 0 \\ 0 & & & \ddots & & \ddots & \ddots & \ddots & & 1 \\ 1 & \ddots & 0 & & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & & \ddots & 1-\delta_{i0} & 1 \\ 1 & \cdots & 1 & 0 & \cdots & 0 & 1 & \cdots & 1 & 1-\delta_{i0} \end{bmatrix} \tag{5.56}$$

where, supposing for simplicity that $S \geq 3$, the first row of $b^{(i)}$ (excluding the term $\left[b^{(i)}\right]_{00}$, which is 0 for $i = 0$ and 1 for $i > 0$) can be written explicitly as:

$$\begin{cases} \left[b^{(i)}\right]_{0j} = & 1, \qquad (1 \leq j \leq \nu_i) \vee (\rho_i \leq j \leq S-1) \\ \\ \left[b^{(i)}\right]_{0j} = & 0, \qquad \nu_i < j < \rho_i \end{cases}$$

$$\rho_i = S - \nu_i + H\left(\nu_i - \left\lfloor \frac{S}{2} \right\rfloor + (-1)^S\right)$$

$$H(x) = \begin{cases} 0, & x \leq 0 \\ \\ 1, & x > 0 \end{cases}$$

with $1 \leq \nu_i \leq \left\lfloor \frac{S}{2} \right\rfloor$. Here we have to suppose that $S \geq 3$ because otherwise it is not possible to distinguish the diagonal band from the corner elements. Now, the bandwidth of $b^{(i)}$ is $2\nu_i + 1$, so this defines the integer parameters ν_i. Moreover, $2\nu_0 - H\left(\nu_0 - \left\lfloor \frac{S}{2} \right\rfloor + (-1)^S\right)$

represents the number of connections that every neuron in a given population receives from the neurons in the same population. Instead $2\nu_i + 1 - H\left(\nu_i - \lfloor \frac{S}{2} \rfloor + (-1)^S\right)$, for $i = 1, 2, ..., R-1$, is the number of connections that every neuron in the the k-th population receives from the neurons in the $(i+k)$-th mod R population, for $k = 0, 1, ..., R-1$. So the total number of incoming connections per neuron is $M = R-1+\sum_{i=0}^{R-1}\left[2\nu_i - H\left(\nu_i - \lfloor \frac{S}{2} \rfloor + (-1)^S\right)\right]$. It is important to observe that even if all the matrices $b^{(i)}$ are symmetric, the matrix $\overline{J}$ in general is not, since the number of connections in every block is different (the case of symmetric connectivity matrices is studied in Section 5.3.2). Now, using formula 5.53, we obtain that:

$$e_{mS+n} = \begin{cases} \frac{\Lambda}{M}\left[R - 1 + \sum_{k=0}^{R-1} f\left(n, \nu_k, S\right)\right], & m = 0, \forall n \\ \\ \frac{\Lambda}{M}\left[-1 + \sum_{k=0}^{R-1} e^{\frac{2\pi}{R}mk\iota} f\left(n, \nu_k, S\right)\right], & m \neq 0, \forall n \end{cases} \tag{5.57}$$

$$f\left(n, \nu_k, S\right) = \begin{cases} 2\nu_k - H\left(\nu_k - \lfloor \frac{S}{2} \rfloor + (-1)^S\right), & n = 0, \forall \nu_k \\ \\ -1, & n \neq 0,\ \nu_k = \lfloor \frac{S}{2} \rfloor \\ \\ \frac{sin\left(\frac{\pi n(2\nu_k+1)}{S}\right)}{sin\left(\frac{\pi n}{S}\right)} - 1, & n \neq 0,\ \nu_k < \lfloor \frac{S}{2} \rfloor \end{cases}$$

with $m = 0, 1, ..., R-1$ and $n = 0, 1, ..., S-1$.

Many different special cases can be studied. The simplest one is obtained for $\nu_0 = \nu_1 = ... = \nu_{R-1} \stackrel{def}{=} \nu$, and in this case formula 5.57 gives:

$$e_{mS+n} = \begin{cases} \frac{\Lambda}{M}\left[R - 1 + Rf\left(n, \nu, S\right)\right], & m = 0, \forall n \\ \\ -\frac{\Lambda}{M}, & m \neq 0, \forall n \end{cases} \tag{5.58}$$

with $M = R - 1 + R\left[2\nu - H\left(\nu - \left\lfloor \frac{S}{2} \right\rfloor + (-1)^S\right)\right]$. Therefore in this case the eigenvalues are real, as it must be, since with this special choice of the parameters the matrix $\overline{J}$ is symmetric. For $R = 1$ and $\nu < \left\lfloor \frac{N}{2} \right\rfloor$ we have $M = 2\nu$ and formula 5.58 gives the eigenvalues of the circulant network:

$$e_n = \begin{cases} \Lambda, & n = 0 \\ \\ \frac{\Lambda}{2\nu}\left[\frac{sin\left(\frac{\pi n(2\nu+1)}{N}\right)}{sin\left(\frac{\pi n}{N}\right)} - 1\right], & n \neq 0 \end{cases} \tag{5.59}$$

Instead for $\nu = \left\lfloor \frac{S}{2} \right\rfloor$ and $\forall R, S$ we have $M = N - 1$ and formula 5.58 gives the eigenvalues of the fully connected network:

$$e_n = \begin{cases} \Lambda, & n = 0 \\ \\ -\frac{\Lambda}{N-1}, & n \neq 0 \end{cases} \tag{5.60}$$

5.3.2 Symmetric matrices

Another case where the matrices $\Phi\left(t\right)$ and $\Phi\left(t\right)\Phi^T\left(t\right)$ can be computed easily is when we have a general symmetric matrix $\overline{J}$. Since its entries are real, it can be diagonalized by an orthogonal matrix Q (namely such that $Q^{-1} = Q^T$), therefore we have:

$$\overline{J} = Q\widetilde{D}Q^T$$

$$\widetilde{D} = diag\left(\widetilde{d}_1, \widetilde{d}_2, ..., \widetilde{d}_{N-1}\right)$$

So we obtain:

$$A = -\frac{1}{\tau}Id_N + \overline{J}S'(\mu) = Q\left[-\frac{1}{\tau}Id_N + \widetilde{D}S'(\mu)\right]Q^T$$

$$\Phi(t) = e^{At} = Qe^{\left[-\frac{1}{\tau}Id_N + \widetilde{D}S'(\mu)\right]t}Q^T = QD(t)Q^T$$

having defined the diagonal matrix $D(t)$ as follows:

$$D(t) = e^{\left[-\frac{1}{\tau}Id_N + \widetilde{D}S'(\mu)\right]t} = diag(d_1, d_2, ..., d_{N-1})$$

$$d_i = e^{\left[-\frac{1}{\tau} + \widetilde{d}_i S'(\mu)\right]t}$$

Moreover, also the matrix A is symmetric in this case, therefore:

$$\Phi(t)\Phi^T(t) = e^{2At} = QD^2(t)Q^T$$

so their components are:

$$\begin{aligned}\Phi_{ij}(t) &= \sum_{k=0}^{N-1} e^{D_k(t-s)} Q_{ik} Q_{jk} \\ \left[\Phi(t)\,\Phi^T(t)\right]_{ij} &= \sum_{k=0}^{N-1} e^{2D_k(t-s)} Q_{ik} Q_{jk}\end{aligned} \tag{5.61}$$

Again, now we need only the eigenquantities of $\overline{J}$, but it is not possible to find explicit expressions for a general symmetric connectivity matrix. Actually they can be calculated analytically only if $\overline{J}$ has some special kind of structure. However, since it is symmetric and all its non-zero entries have the same value $\frac{\Lambda}{M}$ (as we said in Section 5.1), it can be interpreted as the adjacency matrix of an undirected unweighted graph. Due to this correspondence, we can study the eigenquantities of $\overline{J}$ using the powerful techniques already developed in the context of Graph Theory for this kind of graphs. Lee and Yeh [188] have proved that it is possible to perform binary operations (in particular the Kronecker product $\otimes$ and the Cartesian product $\times$) on pairs of graphs G_1 and G_2, obtaining more complicated graphs, whose eigenvalues and eigenvectors can be calculated easily from those of the graphs G_1 and G_2. If e_G and $\overrightarrow{v}_G$ represent respectively the eigenvalues and eigenvectors of the graph G, then we obtain:

$$\begin{cases} e^{i,j}_{G_1 \otimes G_2} = & e^i_{G_1} e^j_{G_2} \\ \overrightarrow{v}^{\,i,j}_{G_1 \otimes G_2} = & \overrightarrow{v}^{\,i}_{G_1} \otimes \overrightarrow{v}^{\,j}_{G_2} \end{cases} \tag{5.62}$$

$$\begin{cases} e^{i,j}_{G_1 \times G_2} = & e^i_{G_1} + e^j_{G_2} \\ \overrightarrow{v}^{\,i,j}_{G_1 \times G_2} = & \overrightarrow{v}^{\,i}_{G_1} \otimes \overrightarrow{v}^{\,j}_{G_2} \end{cases} \tag{5.63}$$

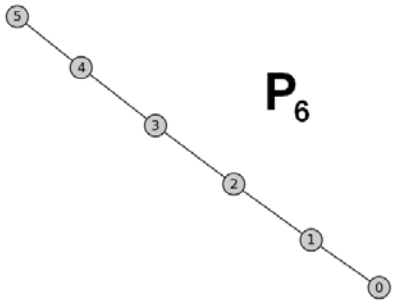

Figure 5.1: Example of the graph P_N, known as *path on N nodes*. Its connectivity matrix is tridiagonal without corner elements.

for $i, j = 0, 1, ..., N-1$. In particular we can choose G_1 and G_2 to be P_N and/or Cy_N, where P_N is the so called *path on N nodes* (see Figure 5.1), while Cy_N is the cycle graph (see Figure 3.1).

Their eigenquantities (in the case of unitary weights) are:

$$\begin{cases} e^i_{P_N} = & 2\cos\left[\frac{(i+1)\pi}{N+1}\right] \\ \\ \left[\overrightarrow{v}^i_{P_N}\right]_j = & \sin\left[\frac{(i+1)(j+1)\pi}{N+1}\right] \end{cases} \tag{5.64}$$

$$\begin{cases} e^i_{Cy_N} = & 2cos\left(\frac{2\pi i}{N}\right) \\ & \qquad\qquad , \iota = \sqrt{-1} \\ \left[\overrightarrow{v}^i_{Cy_N}\right]_j = & e^{\frac{2\pi ij}{N}\iota} \end{cases} \tag{5.65}$$

Combining them through the binary operations $\otimes$ and $\times$, we can create several classes of well-known graphs, like:

- Ladder: $L_n = P_n \times P_2$, with $2n = N$;
- Circular Ladder (also known as Annulus or Prism): $CL_n = Cy_n \times P_2$, with $2n = N$;

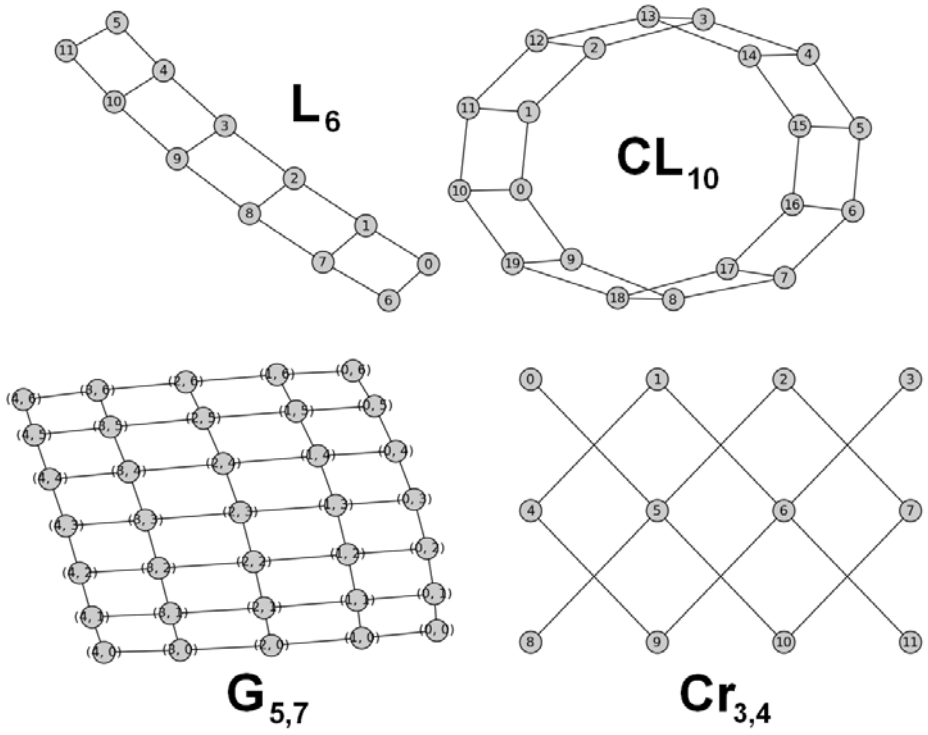

Figure 5.2: Some examples of graphs: the Ladder $L_n = P_n \times P_2$ (top left), the Circular Ladder $CL_n = Cy_n \times P_2$ (top right), the Grid $G_{m,n} = P_m \times P_n$ (bottom left) and the Cross $Cr_{m,n} = P_m \otimes P_n$ (bottom right).

- Grid: $G_{m,n} = P_m \times P_n$, with $mn = N$;
- Cylinder: $Cl_{m,n} = P_m \times Cy_n$, with $mn = N$;
- Torus: $T_{m,n} = Cy_m \times Cy_n$, with $mn = N$;
- Cross: $Cr_{m,n} = P_m \otimes P_n$, with $mn = N$;
- Hypercube: $H_n = \underbrace{P_2 \times P_2 \times ... \times P_2}_{n-times}$, with $2^n = N$;

and so on and so forth. Some of these examples are shown in the Figures 5.2 and 5.3. Even much more complicated graphs can be created in this way, like $Cr_{m,n} \otimes T_{p,q}$, or $G_{m,n} \otimes T_{p,q} \times H_r \otimes Cl_{x,y}$, and so on.

Using the mixed-product property of the Kronecker product, we obtain:

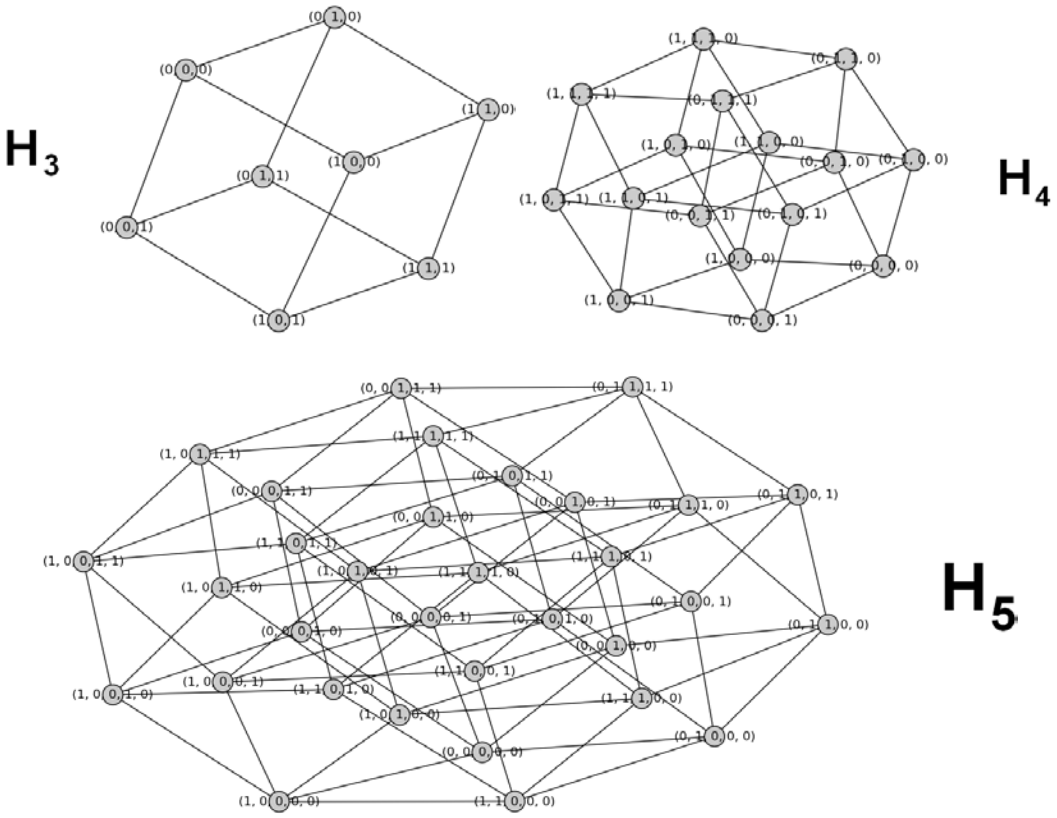

Figure 5.3: Three examples of the Hypercube graph H_n.

$$\overrightarrow{v}^{i,j}_{P_{N_1}\otimes P_{N_2}} \cdot \overrightarrow{v}^{k,l}_{P_{N_1}\otimes P_{N_2}} = \left(\overrightarrow{v}^{i}_{P_{N_1}} \cdot \overrightarrow{v}^{k}_{P_{N_1}}\right)\left(\overrightarrow{v}^{j}_{P_{N_2}} \cdot \overrightarrow{v}^{l}_{P_{N_2}}\right) = 0$$

if $i \neq k$ and/or $j \neq l$, since the eigenvectors of the path are orthogonal (of course the same is true for $\overrightarrow{v}^{i,j}_{P_{N_1}\times P_{N_2}} \cdot \overrightarrow{v}^{k,l}_{P_{N_1}\times P_{N_2}}$). Therefore the eigenvectors $\overrightarrow{v}^{i,j}_{P_{N_1}\otimes P_{N_2}}$ (or equivalently $\overrightarrow{v}^{i,j}_{P_{N_1}\times P_{N_2}}$) are orthogonal. Moreover $\overrightarrow{v}^{i,j}_{P_{N_1}\otimes P_{N_2}}$ have real entries, therefore they form an orthogonal matrix, that can be used directly to compute $\Phi(t)$ and $\Phi(t)\Phi^T(t)$ through formula 5.61.

For the eigenvectors $\overrightarrow{v}^{i,j}_{Cy_{N_1}\otimes Cy_{N_2}}$ the procedure is slightly more complicated, since the eigenvectors $\overrightarrow{v}^{i}_{Cy_N}$ have in general complex entries (with only the exception of the cases $i = 0$ and $i = \frac{N}{2}$ for N even) and therefore we cannot use them to form an orthogonal matrix. However, if $\overline{J}$ is the connectivity matrix corresponding to the graph Cy_N, we have:

$$\overline{J}\,\overrightarrow{v}^{i}_{Cy_N} = e^{i}_{Cy_N}\,\overrightarrow{v}^{i}_{Cy_N}$$

which implies:

$$\overline{J}^{*}\left(\overrightarrow{v}^{i}_{Cy_N}\right)^{*} = \left(e^{i}_{Cy_N}\right)^{*}\left(\overrightarrow{v}^{i}_{Cy_N}\right)^{*}$$

where $*$ is the element-by-element complex conjugation. Since $\overline{J}$ and $e^{i}_{Cy_N}$ are real, we obtain that $\overrightarrow{v}^{i}_{Cy_N}$ and $\left(\overrightarrow{v}^{i}_{Cy_N}\right)^{*}$ are both eigenvectors of $\overline{J}$, corresponding to the same eigenvalue $e^{i}_{Cy_N}$. Therefore, if for all the complex eigenvectors $\overrightarrow{v}^{i}_{Cy_N}$ we define the new vectors:

$$\begin{aligned}\overrightarrow{V}^{i}_{Cy_N} =& \frac{1}{2}\left(\overrightarrow{v}^{i}_{Cy_N} + \left[\overrightarrow{v}^{i}_{Cy_N}\right]^{*}\right)\\ \overrightarrow{W}^{i}_{Cy_N} =& \frac{1}{2\iota}\left(\overrightarrow{v}^{i}_{Cy_N} - \left[\overrightarrow{v}^{i}_{Cy_N}\right]^{*}\right)\end{aligned} \tag{5.66}$$

we conclude that they are eigenvectors of $\overline{J}$ with eigenvalue $e^{i}_{Cy_N}$. Now, it is easy to see that $\overrightarrow{V}^{i}_{Cy_N} \cdot \overrightarrow{W}^{j}_{Cy_N} = 0 \; \forall i,j$. Moreover $\overrightarrow{V}^{i}_{Cy_N}$ and $\overrightarrow{W}^{i}_{Cy_N}$ are orthogonal also to $\overrightarrow{V}^{0}_{Cy_N}$ and $\overrightarrow{V}^{\frac{N}{2}}_{Cy_N}$ in the case of N even, and their entries are real. Therefore, if we use this set of real eigenvectors with the rules 5.62 or 5.63, we obtain a set of eigenvectors for $Cy_{N_1} \otimes Cy_{N_2}$ or $Cy_{N_1} \times Cy_{N_2}$ which are orthogonal and real (the proof is similar to the case $P_{N_1} \otimes P_{N_2}$ seen before). So they can be used to form an orthogonal matrix, through which we can compute $\Phi(t)$ and $\Phi(t)\,\Phi^{T}(t)$, according to formula 5.61.

To conclude, taking for example the cases we have written previously, it is important to observe that only the graphs CL_n, $T_{m,n}$ and H_n can be considered in our analysis. In effect these three graphs have the same number of incoming connections per neuron, a feature that is not shown by the ladder, the grid etc, due to their boundaries. The latter graphs can be studied using this approach only in the thermodynamic limit $N \to \infty$. In fact only in this case the number of incoming connections per neuron is the same for all the neurons in the network, because when $N \to \infty$ the system "loses" its boundaries, since they are pushed to infinity. For example, in the graph L_n all the neurons have 3 incoming connections (see neurons $1-4$ and $7-10$ in the case of the graph L_6 shown in Figure 5.2), with only the exception of those at the boundaries, which have only 2 connections (see neurons 0, 5, 6 and 11 in Figure 5.2). When $N \to \infty$, if we start to travel on the ladder from its center toward the boundaries, we will never reach them, since they are at infinity, therefore during the trip we meet only neurons with the same number (namely 3) of incoming connections. Therefore in the thermodynamic limit all the neurons of the graphs with boundaries behave in the same way. This means that we have obtained the

invariance of the system under exchange of the neural indices, which is what we need in order to apply the perturbative approach introduced in this chapter.

5.4 Numerical comparison

The Figures 5.4 - 5.10 show the numerical comparison obtained with the first-order perturbative expansion. For simplicity in this case we have chosen $\sigma_4 = \sigma_5 = 0$, since according to 5.41 these two parameters affect the covariance only at a higher order. These figures report both the results obtained from the exact network equations 5.1 (blue lines) and from the first-order perturbative expansion (red lines), the latter being generated with the equations 5.12 - 5.15. Moreover we have shown the comparison with the analytic results for the variance, covariance and correlation generated by the formulae 5.41 - 5.46 (green lines). Instead the Figure 5.11 shows the results for the second-order perturbative expansion, obtained from the equations 5.12 - 5.23. For the sake of brevity, here we have reported only the results for the correlation, but we have not shown the comparison with its analytic formula (green lines), due to the complexity of the higher order terms of the variance and covariance. In this case we have used $\sigma_4 = \sigma_5 = 1$, $Z(t) = e^{-t\overline{J}}$ and $\overrightarrow{H}(t) = sin(2\pi t)\overrightarrow{1}$, where $\overrightarrow{1}$ is the vector whose entries are all ones. For all these simulations we have used the parameters reported in Table 5.1, while the statistics have been calculated with $10,000$ Monte Carlo simulations. Moreover the equations 5.1 and 5.12 - 5.23 have been solved numerically using the Euler-Maruyama scheme, while the time-integrals involved in the formulae 5.42, 5.43 and 5.44 have been calculated with the trapezoidal rule. The integration time step is $\Delta t = 0.1$. The covariance and correlation have always been calculated between the 0th and the 1st neuron, while the potentials and the variances have been reported only for the former. The general conclusion is that for small enough values of the parameters σ_1-σ_5 there is a very good agreement between the real network and the first-order perturbative expansion and that for higher values of these parameters the second-order expansion should be used. The match depends on the

Neuron	Input	Synaptic Weights	Sigmoid Function
$\tau = 1$	$\overline{I} = 0$	$\Lambda = 1$	$T_{MAX} = 1$
$C_2 = 0.4$	$C_1 = 0.3$	$C_3 = 0.5$	$\lambda = 1$
			$V_T = 0$

Table 5.1: Parameters used for all the numerical simulations of the Figures 5.4 - 5.13.

dynamics of the neurons, on the synaptic connectivity and on the network size, and in the case of the variance, covariance and correlation, it also depends on the number of Monte Carlo simulations used to evaluate the statistics.

5.5 Correlation as a function of the input

From the formulae 5.42, 5.43, 5.44 and 5.55 the effect of the non-linearity introduced by the sigmoid function $S(V)$ is evident. Through its slope, it generates an effective connectivity matrix $\overline{J}S'(\mu)$, which can be interpreted as the real connectivity matrix of the system if it were linear. Now, the stationary solution μ depends on the external input current $\overline{I}$ through the formula 5.12, therefore the effective synaptic strength and the correlation structure depend on $\overline{I}$ as well. In particular, it is interesting to observe that if $\left|\overline{I}\right|$ is very large, then $|\mu|$ is also very large, therefore $S'(\mu)$ and the entries of the effective connectivity matrix are small. In other words, the neurons become (effectively) disconnected. An important consequence of this phenomenon is that, for $C_1 = C_2 = C_3 = 0$ and for large values of $\left|\overline{I}\right|$, the neurons become independent, even if the size of the network is finite. This intuition is confirmed numerically in Figure 5.12, which has been obtained for the graph Cy_5 (which is made of 10 neurons) simulated with the exact equations 5.1, for $\overline{I} = -5, 0, 5$ and $50,000$ Monte Carlo simulations. The sources of randomness have intensities $\sigma_1 = \sigma_2 = \sigma_3 = 0.1$, and moreover $\sigma_4 = \sigma_5 = 0$, while all the remaining parameters are those of Table 5.1. As usual, the numerical scheme is the Euler-Maruyama one, with integration time step $\Delta t = 0.1$.

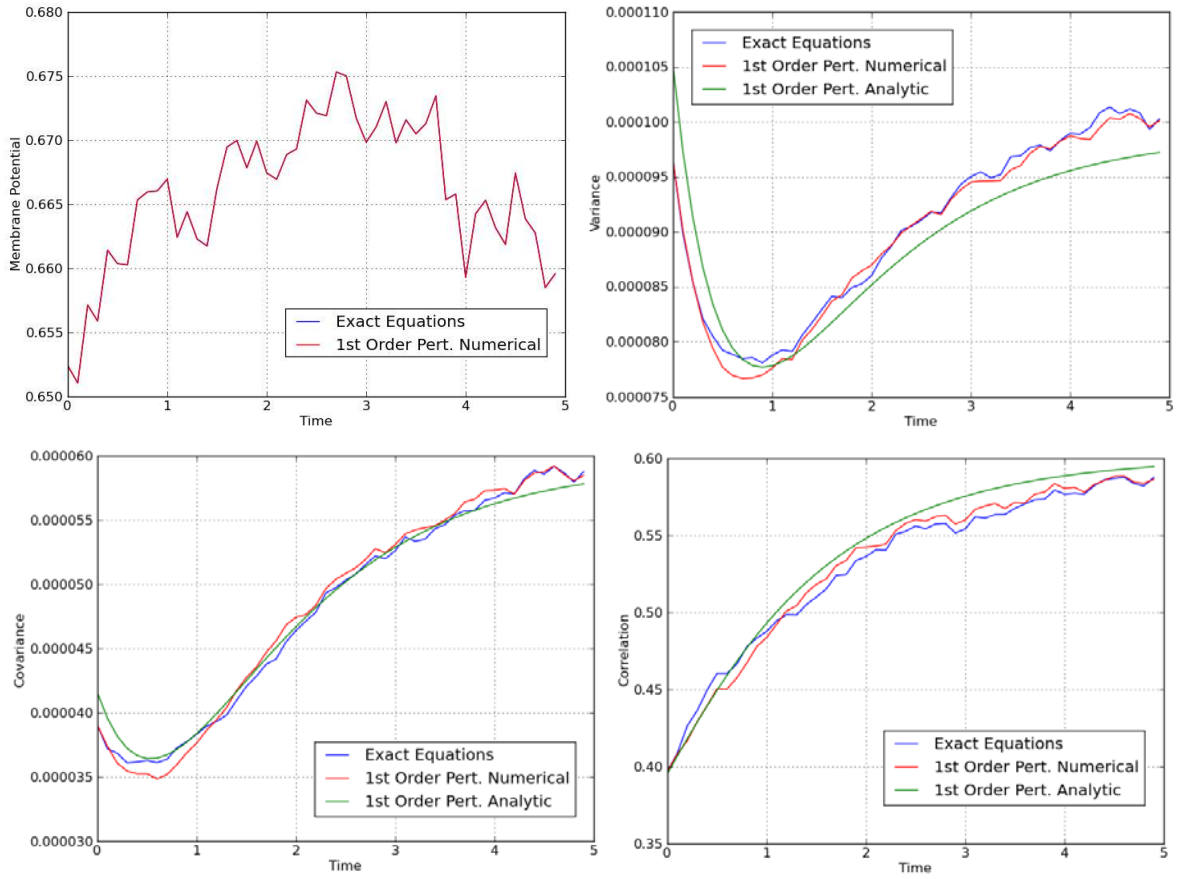

Figure 5.4: First-order perturbative expansion ($\sigma_4 = \sigma_5 = 0$) for a network with connectivity matrix CL_{10} (namely $N = 20$). These results have been obtained for the values of the parameters reported in Table 5.1, for $\sigma_1 = \sigma_2 = \sigma_3 = 0.01$ and with the statistics evaluated through $10,000$ Monte Carlo simulations. Since σ_1, σ_2 and σ_3 are small, in the picture of the membrane potentials $V_i(t)$ (top-left) there is a perfect agreement between the result obtained from the exact network equations 5.1 (blue line) and that obtained from the first-order perturbative expansion, namely from the equations 5.12 - 5.15 (red line). Instead the comparison between the variances (top-right), covariances (bottom-left) and correlations (bottom-right) is less good because small values of σ_1, σ_2 and σ_3 determine small values of the variance and covariance, therefore a higher number of Monte Carlo simulations is required in order to improve the match. The green line represents the analytic result obtained for the first-order perturbative expansion for an infinite number of Monte Carlo simulations (formulae 5.41 - 5.46), therefore it is the limit curve reached by the red line when the number of simulations is increased indefinitely. The blue and red lines have been obtained numerically by solving the corresponding equations with the Euler-Maruyama scheme, while the integrals with respect to time involved in the formulae for the evaluation of the green line have been calculated with the trapezoidal rule. In all the cases the integration time step is $\Delta t = 0.1$.

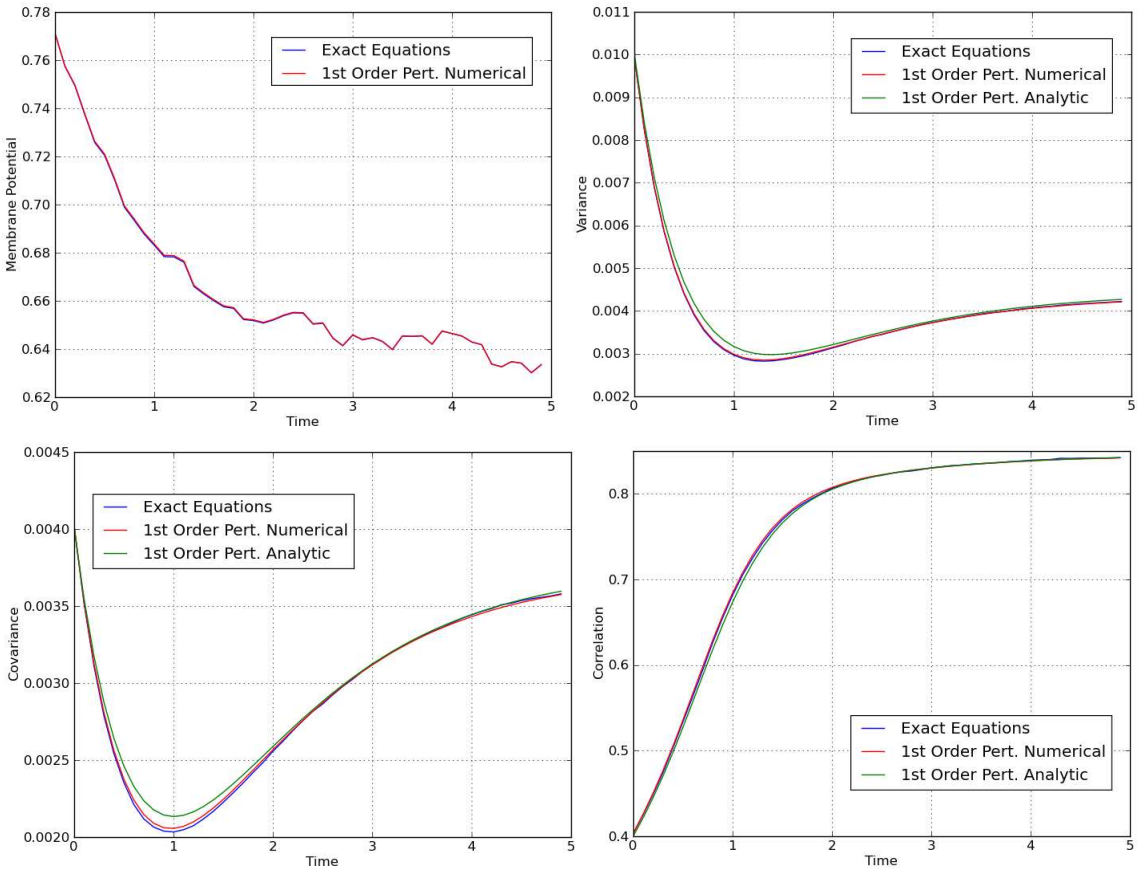

Figure 5.5: First-order perturbative expansion for a network with connectivity matrix CL_{10}. These results have been obtained for the values of the parameters reported in Table 5.1, for $\sigma_1 = 0.01$, $\sigma_2 = \sigma_3 = 0.1$ and with the statistics evaluated through $10,000$ Monte Carlo simulations. The parameter σ_1 is small because high values determine large fluctuations of the variance and covariance (see Figure 5.8), so in that case a higher number of Monte Carlo simulations is required in order to obtain a good match.

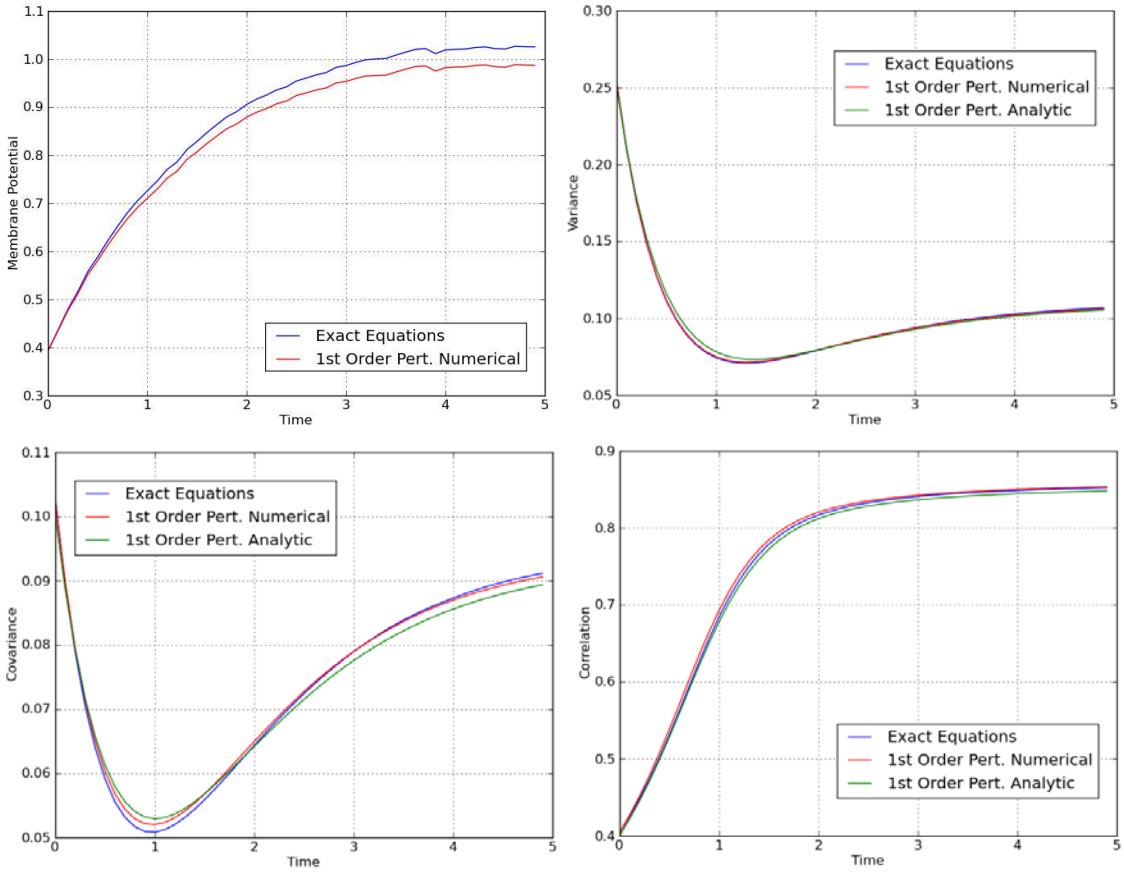

Figure 5.6: First-order perturbative expansion for a network with connectivity matrix CL_{10}. These results have been obtained for the values of the parameters reported in Table 5.1, for $\sigma_1 = 0.01$, $\sigma_2 = \sigma_3 = 0.5$ and with the statistics evaluated through $10,000$ Monte Carlo simulations.

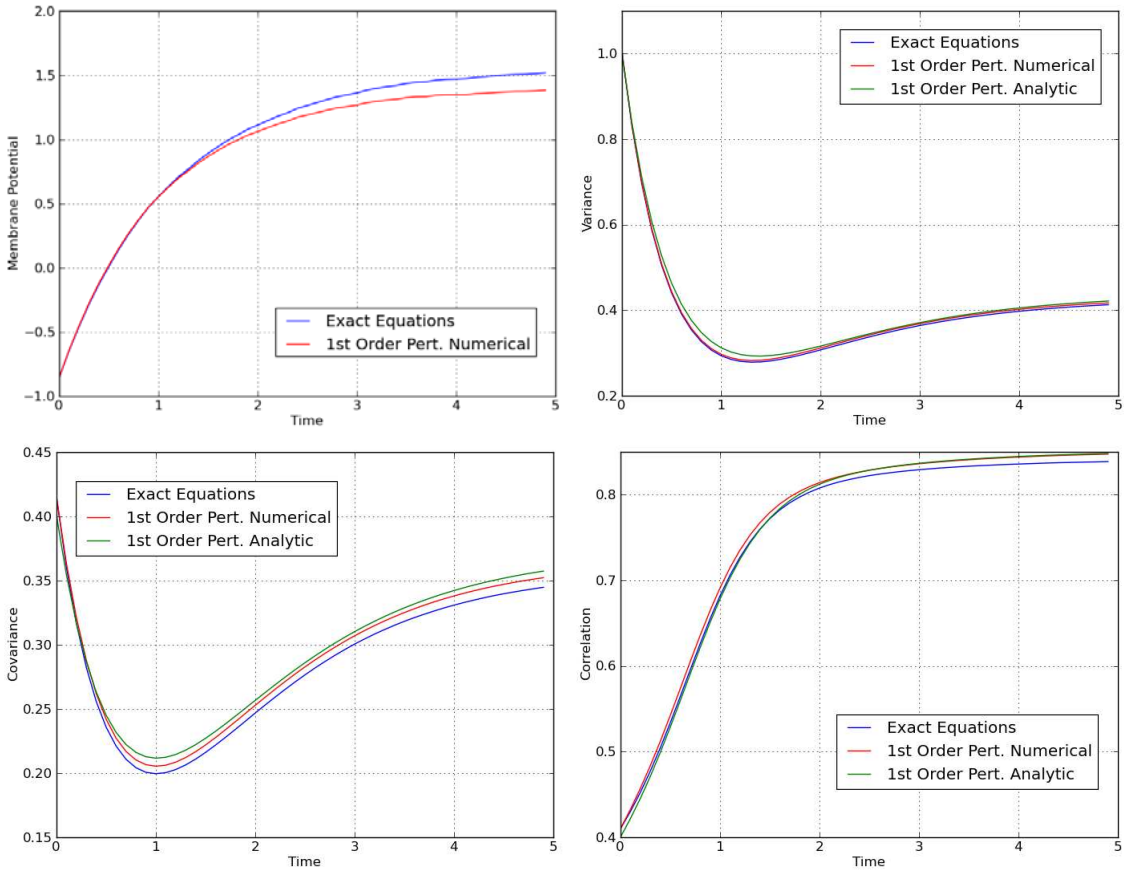

Figure 5.7: First-order perturbative expansion for a network with connectivity matrix CL_{10}. These results have been obtained for the values of the parameters reported in Table 5.1, for $\sigma_1 = 0.01$, $\sigma_2 = \sigma_3 = 1$ and with the statistics evaluated through $10,000$ Monte Carlo simulations. The match between the exact behavior and the first-order perturbative expansion is still reasonably good, even if σ_2 and σ_3 are large.

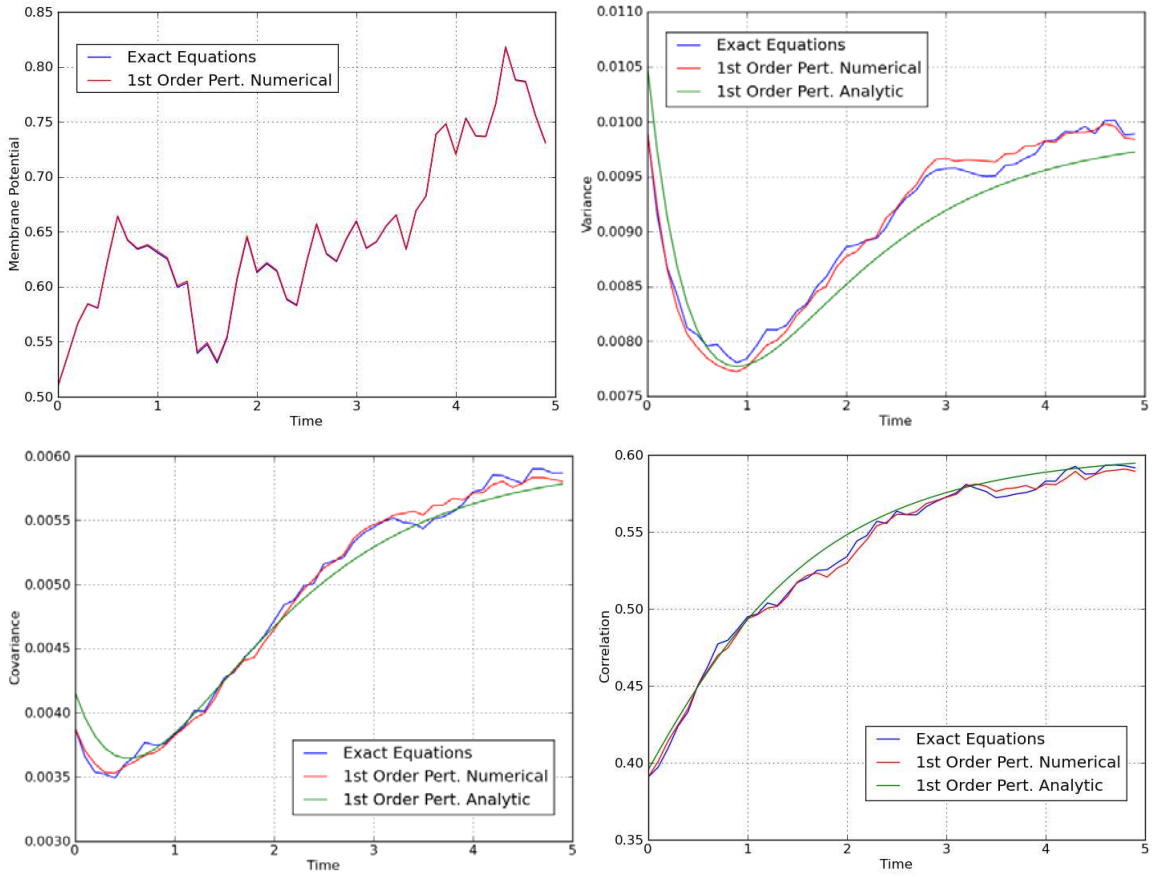

Figure 5.8: First-order perturbative expansion for a network with connectivity matrix CL_{10}. These results have been obtained for the values of the parameters reported in Table 5.1, for $\sigma_1 = \sigma_2 = \sigma_3 = 0.1$ and with the statistics evaluated through $10,000$ Monte Carlo simulations. The match is not as good as in the previous figures because large values of σ_1 determine large fluctuations of the variance and covariance. In other terms, the variance (over many repetitions of groups made up of $10,000$ Monte Carlo simulations each) of the variance and covariance is large if σ_1 is big.

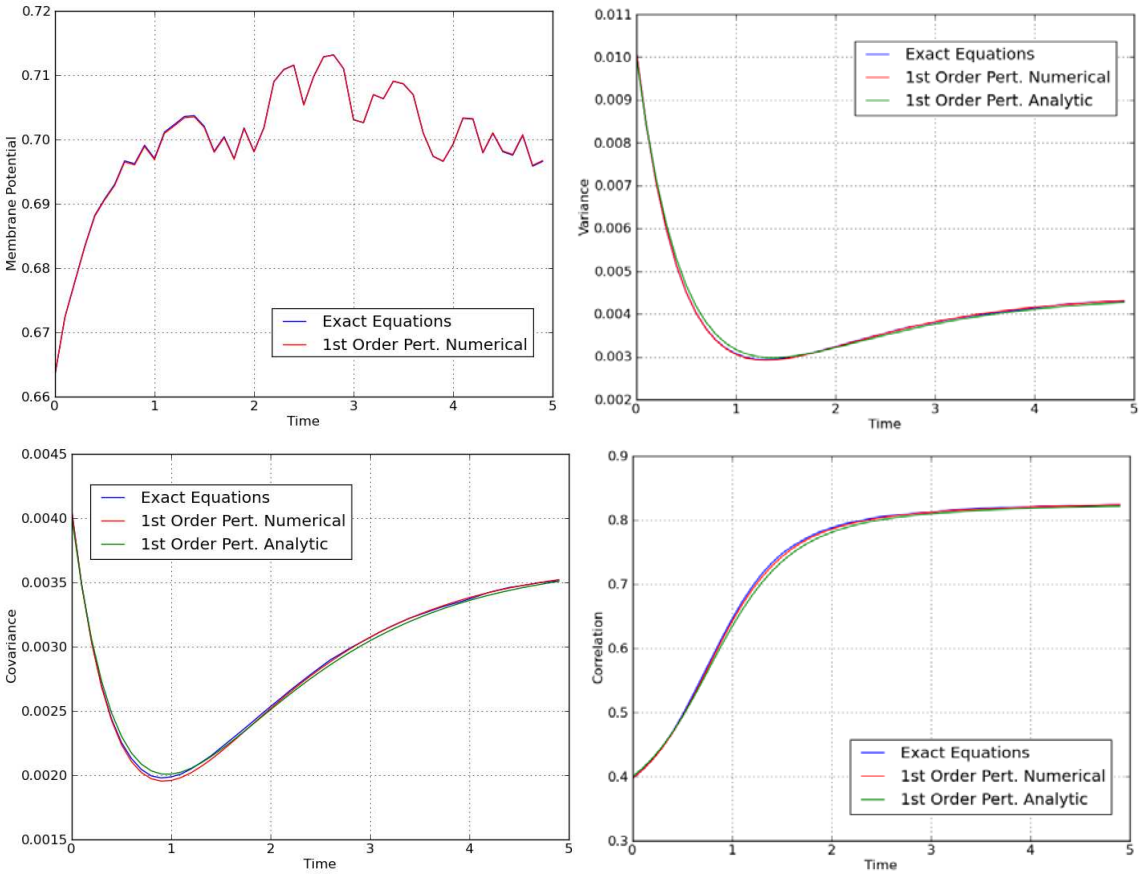

Figure 5.9: First-order perturbative expansion for a network with connectivity matrix H_3 (namely $N = 8$). These results have been obtained for the values of the parameters reported in Table 5.1, for $\sigma_1 = 0.01$, $\sigma_2 = \sigma_3 = 0.1$ and with the statistics evaluated through $10,000$ Monte Carlo simulations.

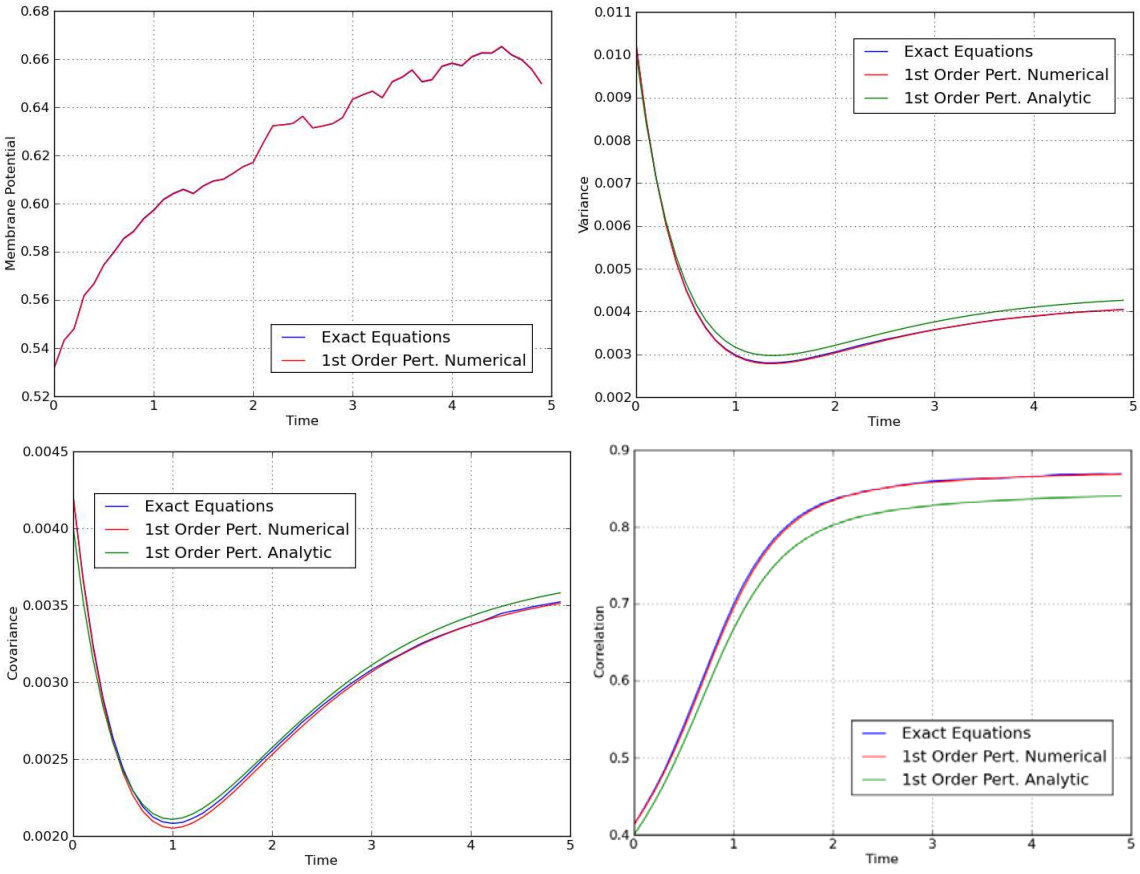

Figure 5.10: First-order perturbative expansion for a network with connectivity matrix $C_{10}\,(1, 2, 0, ..., 0)$. These results have been obtained for the values of the parameters reported in Table 5.1, for $\sigma_1 = 0.01$, $\sigma_2 = \sigma_3 = 0.1$ and with the statistics evaluated through $10,000$ Monte Carlo simulations. This figure clearly shows that the goodness of the match between the curves depends on the connectivity matrix of the network, for a fixed number of Monte Carlo simulations.

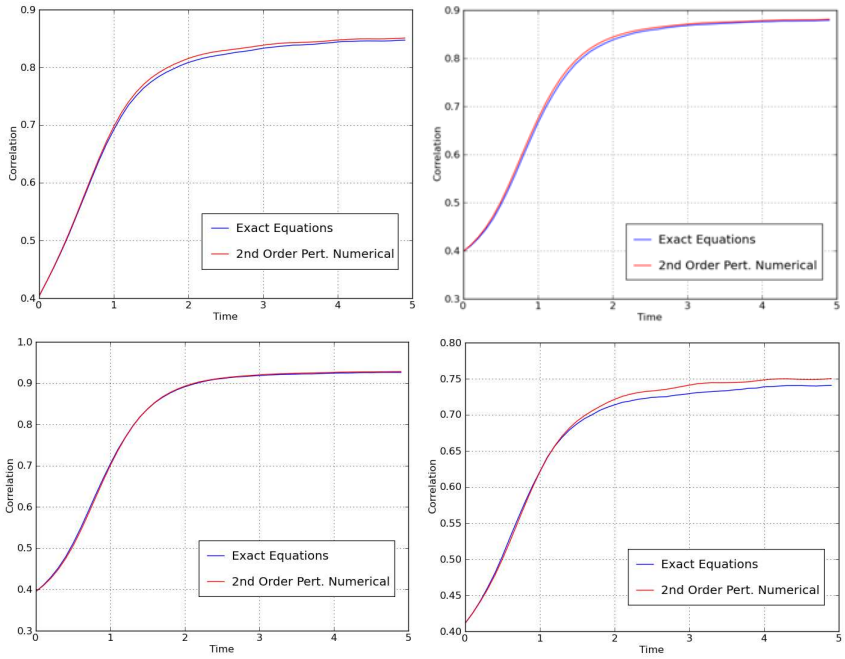

Figure 5.11: Correlation function obtained with the second-order perturbative expansion for a network with connectivity matrix CL_{10} (top-left), H_5 (top-right), K_{10} (bottom-left) and Cy_{15} (bottom-right). These results have been obtained for the values of the parameters reported in Table 5.1, for $\sigma_1 = 0.01$, $\sigma_2 = \sigma_3 = 0.1$, $\sigma_4 = \sigma_5 = 1$, $Z(t) = e^{-t}\overline{J}$ and $\overrightarrow{H}(t) = sin(2\pi t)\ \overrightarrow{1}$ and with the statistics evaluated through $10,000$ Monte Carlo simulations. The match is good even if σ_4 and σ_5 are large.

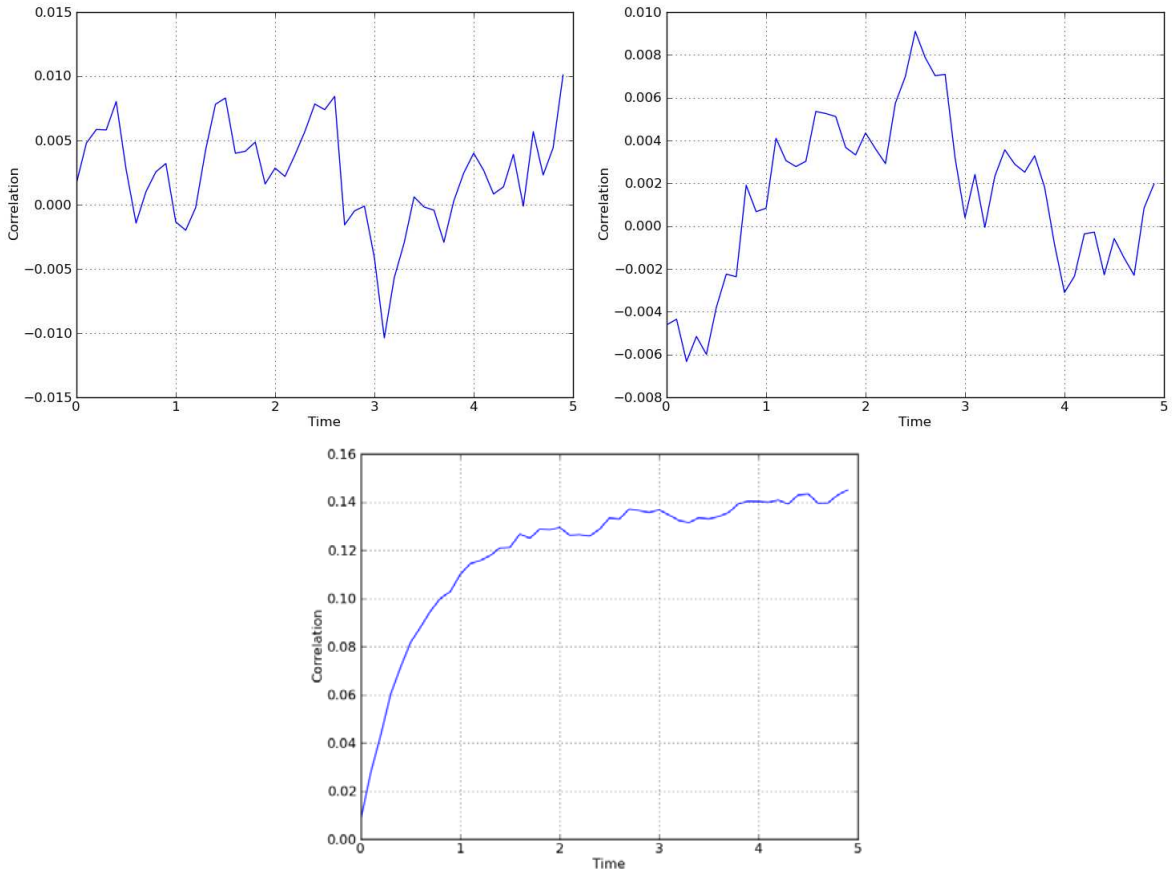

Figure 5.12: Correlation function obtained for the graph Cy_5 and $\overline{I} = -5$ (top-left), 5 (top-right) and 0 (bottom). These results have been obtained from the exact equations 5.1, numerically solved using the Euler-Maruyama scheme with integration time step $\Delta t = 0.1$ and with $50,000$ Monte Carlo simulations. The parameters used are $C_1 = C_2 = C_3 = 0$, $\sigma_1 = \sigma_2 = \sigma_3 = 0.1$ and $\sigma_4 = \sigma_5 = 0$, while all the remaining parameters are those of Table 5.1. From this figure it is possible to see that the correlation between pairs of neurons strongly decreases for high values of $|\overline{I}|$, confirming its relation with the effective connectivity matrix $\overline{J}S'(\mu)$.

5.6 Failure of the mean-field theory

In this section we show three different reasons that invalidate the use of the mean-field theory for the mathematical analysis of a neural network. A neural network is generally described by a large set of stochastic differential equations, that makes it hard to understand the underlying behavior of the system. However, if the neurons become independent, their dynamics can be described with the mean-field theory using a highly reduced set of equations, that are much simpler to analyze (see Chapter 3). For this reason the mean-field theory is a powerful tool that can be used to understand the network. One of the mechanisms through which the independence of the neurons can be obtained is the phenomenon known as *propagation of chaos* [51][63][64][65]. Propagation of chaos refers to the fact that, if we choose independent initial conditions for the membrane potentials at $t = 0$ (which may be called *initial chaos*), then the neurons are always perfectly independent $\forall t > 0$. Therefore the term *propagation* refers to the "transfer" of the independence of the membrane potentials from $t = 0$ to $t > 0$. Under simplified assumptions about the nature of the network (namely that the other sources of randomness in the system, in our case the Brownian motions and the synaptic weights, are independent), propagation of chaos does occur in the so called *thermodynamic limit* of the system, namely when the number of neurons in the system grows to infinity. However in Sections 5.6.1, 5.6.2 and 5.6.3 we show that for a system with correlated Brownian motions, initial conditions and synaptic weights, with a general connectivity matrix or with an arbitrarily large (but still finite) size, the correlation between pairs of neurons can be high. Therefore in general the neurons cannot be independent, invalidating the use of the mean-field theory.

5.6.1 Independence does not occur for $N \to \infty$ if C_1, C_2 or C_3 are not equal to zero

Let us consider the case when at least one of the parameters C_1, C_2 and C_3 (defined by 5.2, 5.5 and 5.9) is not equal to zero. For example we analyze the term proportional to C_1 in the formula 5.42, for a fully connected network. Using the technique developed in Section 5.3.1, it is easy to prove that this term for $i \neq j$ is:

$$C_1 \sum_{\substack{k,l=0 \\ k \neq l}}^{N-1} \int_0^t \left[\Phi\left(t-s\right)\right]_{ik} \left[\Phi\left(t-s\right)\right]_{jl} ds$$

$$= \frac{C_1}{2} \left\{ \left(1 - \frac{1}{N}\right) \frac{1}{\frac{1}{\tau} - \Lambda S'\left(\mu\right)} \left[1 - e^{-2\left(\frac{1}{\tau} - \Lambda S'(\mu)\right)t}\right] + \frac{1}{N} \frac{1}{\frac{1}{\tau} + \frac{\Lambda S'(\mu)}{N-1}} \left[1 - e^{-2\left(\frac{1}{\tau} + \frac{\Lambda S'(\mu)}{N-1}\right)t}\right] \right\}$$

while for $i = j$ it is:

$$C_1 \sum_{\substack{k,l=0 \\ k \neq l}}^{N-1} \int_0^t \left[\Phi\left(t-s\right)\right]_{ik} \left[\Phi\left(t-s\right)\right]_{il} ds$$

$$= \frac{C_1}{2} \left(1 - \frac{1}{N}\right) \left\{ \frac{1}{\frac{1}{\tau} - \Lambda S'\left(\mu\right)} \left[1 - e^{-2\left(\frac{1}{\tau} - \Lambda S'(\mu)\right)t}\right] - \frac{1}{\frac{1}{\tau} + \frac{\Lambda S'(\mu)}{N-1}} \left[1 - e^{-2\left(\frac{1}{\tau} + \frac{\Lambda S'(\mu)}{N-1}\right)t}\right] \right\}$$

So the covariance (and therefore also the correlation) does not go to zero for $N \to \infty$, or in other words the neurons are not independent, even in the thermodynamic limit.

The reader can easily check that the same result holds for the terms of the covariance proportional to C_2 and C_3.

5.6.2 Propagation of chaos does not occur for a general connectivity matrix

We study propagation of chaos as a function of the number of connections in the circulant network. To this purpose, we have to set $C_2 = 0$ (initial chaos) and also $C_1 = C_3 = 0$, because otherwise the neurons cannot be independent, as explained in Section 5.6.1. Using the formulae 5.42, 5.43, 5.44 and 5.55 we obtain that in this case the covariance is:

$$Cov\left(V_i\left(t\right), V_j\left(t\right)\right) = \sigma_1^2 \int_0^t \left[\Phi\left(t-s\right)\Phi^T\left(t-s\right)\right]_{ij} ds + \sigma_2^2 \left[\Phi\left(t\right)\Phi^T\left(t\right)\right]_{ij}$$

$$+ \sigma_3^2 \frac{S^2\left(\mu\right)}{M} \sum_{k=0}^{N-1} \left[\int_0^t \Phi_{ik}\left(t-s\right) ds\right] \left[\int_0^t \Phi_{jk}\left(t-s\right) ds\right] \tag{5.67}$$

where:

$$\int_0^t \left[\Phi\left(t-s\right)\Phi^T\left(t-s\right)\right]_{ij} ds = \frac{1}{2N} \sum_{k=0}^{N-1} \frac{cos\left[\frac{2\pi}{N} k\left(i-j\right)\right]}{-\frac{1}{\tau} + e_k S'\left(\mu\right)} \left\{1 - e^{2\left[-\frac{1}{\tau} + S'(\mu) e_k\right] t}\right\}$$

$$\left[\Phi\left(t\right)\Phi^T\left(t\right)\right]_{ij} = \frac{1}{N} \sum_{k=0}^{N-1} e^{2\left[-\frac{1}{\tau} + S'(\mu) e_k\right] t} cos\left[\frac{2\pi}{N} k\left(i-j\right)\right]$$

$$\begin{aligned}
&\sum_{k=0}^{N-1} \left[\int_0^t \Phi_{ik}\left(t-s\right) ds\right] \left[\int_0^t \Phi_{jk}\left(t-s\right) ds\right] \\
&= \frac{1}{N^2} \sum_{l,m=0}^{N-1} e^{\frac{2\pi}{N} l i_\iota} e^{\frac{2\pi}{N} m j_\iota} \left[\sum_{k=0}^{N-1} e^{-\frac{2\pi}{N}(l+m) k_\iota}\right] \left\{\frac{1 - e^{\left[-\frac{1}{\tau} + S'(\mu) e_l\right] t}}{-\frac{1}{\tau} + e_l S'\left(\mu\right)}\right\} \left\{\frac{1 - e^{\left[-\frac{1}{\tau} + S'(\mu) e_m\right] t}}{-\frac{1}{\tau} + e_m S'\left(\mu\right)}\right\} \\
&= \frac{1}{N} \sum_{l=0}^{N-1} \left\{\frac{1 - e^{\left[-\frac{1}{\tau} + S'(\mu) e_l\right] t}}{-\frac{1}{\tau} + e_l S'\left(\mu\right)}\right\}^2 cos\left[\frac{2\pi}{N} l\left(i-j\right)\right]
\end{aligned}$$

while the eigenvalues e_k are given by formula 5.59 or by formula 5.60. Now, for $N \to \infty$

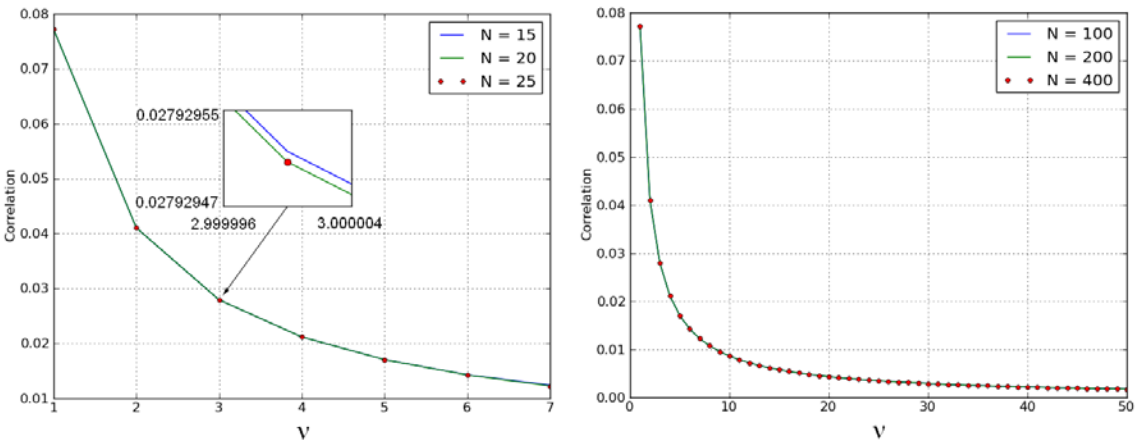

Figure 5.13: Propagation of chaos for $t = 1$ as a function of $\nu = \frac{M}{2}$, in the case of a circulant connectivity matrix. This result has been obtained for $C_1 = C_2 = C_3 = 0$ (while all the remaining parameters are those of Table 5.1), using the analytic formula 5.67 (normalized with the variance).

the right-hand side of formula 5.67 converges to a non-zero function (see Figure 5.13), therefore for every finite value of ν (which is the number of incoming connections per neuron divided by 2) propagation of chaos does not occur.

Moreover correlation decreases with ν, therefore propagation of chaos occurs in the circulant network only in the thermodynamic limit $N \to \infty$ and if ν is an increasing function of N, namely if $\lim_{N\to\infty} \nu = \infty$ (compare with 3.8). For example, in the fully connected network $\nu = \lfloor \frac{N}{2} \rfloor$, so it explains why in this case correlation goes to zero in the thermodynamic limit. Instead in a network described by a cycle graph, perfect decorrelation is never possible, also for $N \to \infty$, since $\nu = 1$. In other words, having infinitely many neurons is not a sufficient condition for getting propagation of chaos, because also infinite connections per neuron are required. This result had already been obtained numerically in Chapter 3 (see Figure 3.8).

5.6.3 Stochastic synchronization

In this section we show that for every finite and arbitrarily large number of neurons N in the network, it is possible to choose special values of the parameters of the system such

that, at some finite and arbitrarily large time instant $\overline{t}$, the correlation between pairs of neurons is (approximately) 1. In general $\overline{t}$ increases with N.

The general theory

We show that even when $C_1 = C_2 = C_3 = 0$, if the matrix $A = -\frac{1}{\tau} Id_N + \overline{J} S'(\mu)$ has an eigenvalue of multiplicity 1 with non-negative real part, while all the other eigenvalues have negative real parts, then correlation goes to 1 for $t \to +\infty$, for every finite N. In other terms, the stochastic components of the membrane potentials become perfectly synchronized. From now on we refer to this phenomenon as *stochastic synchronization*. To prove this, we suppose that A has an eigenvalue $\overline{a}$ with non-negative real part and with a generic multiplicity $m > 0$, while all the other eigenvalues have negative real parts. Now we recall that $e^{At} = Qe^{Dt}Q^{-1}$, where D is the diagonal matrix of the eigenvalues of A, and Q is the matrix of its eigenvectors. So for $t \to +\infty$ we have:

$$e^{Dt} \prec diag(0, 0, ..., 0, \underbrace{e^{\overline{a}t}, e^{\overline{a}t}, ..., e^{\overline{a}t}}_{m-\text{times}}, 0, 0, ..., 0)$$

where $\prec$ means *dominated by*, because all the eigenvalues have negative real part but $\overline{a}$. If the $\overline{a}$s are the r-th, $(r+1)$-th, ..., $(r+m-1)$-th eigenvalues of A and if we call $Q^{-1} = B$ in order to simplify the notation, we obtain:

$$Qe^{Dt}B \prec e^{\overline{a}t}\begin{bmatrix} 0 & 0 & \cdots & 0 & Q_{0,r} & Q_{0,r+1} & \cdots & Q_{0,r+m-1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & Q_{1,r} & Q_{1,r+1} & \cdots & Q_{1,r+m-1} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & Q_{N-1,r} & Q_{N-1,r+1} & \cdots & Q_{N-1,r+m-1} & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$\times \begin{bmatrix} B_{0,0} & B_{0,1} & ... & B_{0,N-1} \\ B_{1,0} & B_{1,1} & ... & B_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ B_{N-1,0} & B_{N-1,1} & ... & B_{N-1,N-1} \end{bmatrix}$$

and therefore:

$$e^{At} = Qe^{Dt}B \prec e^{\overline{a}t}E$$

$$E_{pq} = \sum_{k=0}^{m-1} Q_{p,r+k}B_{r+k,q}$$

This means that:

$$\begin{aligned} Cov\left(V_i\left(t\right), V_j\left(t\right)\right) =& \sigma_1^2 \sum_{k=0}^{N-1} \int_0^t \left[e^{A(t-s)}\right]_{ik} \left[e^{A(t-s)}\right]_{jk} ds + \sigma_2^2 \sum_{k=0}^{N-1} \left[e^{At}\right]_{ik} \left[e^{At}\right]_{jk} \\ &+ \sigma_3^2 \frac{S^2\left(\mu\right)}{M} \sum_{k=0}^{N-1} \left\{\int_0^t \left[e^{A(t-s)}\right]_{ik} ds\right\} \left\{\int_0^t \left[e^{A(t-s)}\right]_{jk} ds\right\} \\ &\prec \left[\frac{\sigma_1^2}{2\overline{a}} + \sigma_2^2 + \frac{\sigma_3^2}{\overline{a}^2}\frac{S^2\left(\mu\right)}{M}\right] e^{2\overline{a}t} \sum_{k=0}^{N-1} E_{ik}E_{jk} \end{aligned} \tag{5.68}$$

so the variance is:

$$Var\left(V_i\left(t\right)\right) = Cov\left(V_i\left(t\right), V_i\left(t\right)\right) \prec \left[\frac{\sigma_1^2}{2\overline{a}} + \sigma_2^2 + \frac{\sigma_3^2}{\overline{a}^2}\frac{S^2\left(\mu\right)}{M}\right] e^{2\overline{a}t} \sum_{k=0}^{N-1} \left(E_{ik}\right)^2 \tag{5.69}$$

Therefore the correlation is:

$$Corr\left(V_i\left(t\right), V_j\left(t\right)\right) = \frac{Cov\left(V_i\left(t\right), V_j\left(t\right)\right)}{\sqrt{Var\left(V_i\left(t\right)\right) Var\left(V_j\left(t\right)\right)}} \rightarrow \frac{\sum_{k=0}^{N-1} E_{ik}E_{jk}}{\sqrt{\left[\sum_{k=0}^{N-1} \left(E_{ik}\right)^2\right]\left[\sum_{k=0}^{N-1} \left(E_{jk}\right)^2\right]}} \tag{5.70}$$

when $t \rightarrow +\infty$. Now, in the special case $m = 1$ we obtain:

$$E_{pq} = Q_{pr}B_{rq}$$

$$\sum_{k=0}^{N-1} E_{ik}E_{jk} = Q_{ir}Q_{jr} \sum_{k=0}^{N-1} \left(B_{rk}\right)^2$$

$$\sum_{k=0}^{N-1} \left(E_{ik}\right)^2 = \left(Q_{ir}\right)^2 \sum_{k=0}^{N-1} \left(B_{rk}\right)^2$$

so we conclude that $Corr\left(V_i\left(t\right), V_j\left(t\right)\right) \rightarrow 1$ when $t \rightarrow +\infty$. This proves that if $C_1 = C_2 = C_3 = 0$ and the matrix A has an eigenvalue of multiplicity 1 with non-negative real part while all the other eigenvalues have negative real parts, then propagation of chaos does not occur. For continuity, for every finite N we have that $Corr\left(V_i\left(t\right), V_j\left(t\right)\right) \rightarrow 1$ also for $\mathcal{R}\left(\overline{a}\right) \rightarrow 0^-$ (where $\mathcal{R}$ means *the real part of*), i.e. correlation is very big also when the system is stable but close to the instability region $\mathcal{R}\left(\overline{a}\right) > 0$. It is also interesting to observe that, due to the Perron-Frobenius theorem [189], if $\Lambda > 0$ and if $\overline{J}$ is an irreducible matrix (namely if its corresponding directed graph is *strongly connected*, which means that it is possible to reach each vertex in the graph from any other vertex, by moving on

the edges according to their connectivity directions), then it has a unique largest positive eigenvalue, which can be used to generate stochastic synchronization. We conclude that in general propagation of chaos does not always occur, even if $C_1 = C_2 = C_3 = 0$, therefore this invalidates the use of the mean-field theory, at least in this special case.

The example of the fully connected network

We show how to set the parameters of the system such that the phenomenon of stochastic synchronization does occur. For simplicity we assume a fully connected network. In this case, from formula 5.60, we know that the matrix A has eigenvalues:

$$a_0 = -\tfrac{1}{\tau} + \Lambda S'(\mu), \qquad a_1 = -\tfrac{1}{\tau} - \tfrac{\Lambda S'(\mu)}{N-1} \tag{5.71}$$

The multiplicity of a_0 and a_1 is respectively 1 and $N-1$, therefore in order to obtain the stochastic synchronization, according to Section 5.6.3, we have to set $a_0 \geq 0$. Let us consider the case $a_0 = 0$, namely $\Lambda S'(\mu) = \frac{1}{\tau}$. Now, since:

$$S'(\mu) = \lambda \left[S(\mu) - \frac{S^2(\mu)}{T_{MAX}} \right] \tag{5.72}$$

we obtain the algebraic equation:

$$\Lambda\lambda \left[S(\mu) - \frac{S^2(\mu)}{T_{MAX}} \right] = \frac{1}{\tau}$$

whose solutions are:

$$S(\mu_{1,2}) = T_{MAX} \frac{1 \pm \sqrt{1 - \frac{4}{\tau\Lambda\lambda T_{MAX}}}}{2} \tag{5.73}$$

where $\mu_{1,2}$ are two possible stationary solutions of the membrane potential. Moreover, from equation 5.12 we know that:

$$\mu_{1,2} = \tau \left[\Lambda S \left(\mu_{1,2} \right) + \overline{I} \right] \tag{5.74}$$

Putting together the formulae 5.73 and 5.74 we obtain:

$$\mu_{1,2} = \tau \left(\Lambda T_{MAX} \frac{1 \pm \sqrt{1 - \frac{4}{\tau \Lambda \lambda T_{MAX}}}}{2} + \overline{I} \right) \tag{5.75}$$

Replace this value of $\mu_{1,2}$ in 5.74 to obtain the final result:

$$T_{MAX} \frac{1 \pm \sqrt{1 - \frac{4}{\tau \Lambda \lambda T_{MAX}}}}{2} = S \left(\tau \left(\Lambda T_{MAX} \frac{1 \pm \sqrt{1 - \frac{4}{\tau \Lambda \lambda T_{MAX}}}}{2} + \overline{I} \right) \right) \tag{5.76}$$

This non-linear algebraic equation is the constraint that must be satisfied by all the parameters of the system in order to have correlation equal to 1 in the limit $t \to +\infty$. An example of solution of this equation is $\lambda = T_{MAX} = 1$, $V_T = 0$, $\Lambda = -2\overline{I}$ and $\tau = -\frac{2}{\overline{I}}$, $\forall \overline{I} < 0$. In this case $\mu_{1,2} = 0$ and it can be used as initial condition in order to ensure the stationarity of the system. In Figure 5.14 we show the phenomenon of stochastic synchronization in the case of a fully connected network, for the values of the parameters reported in Table 5.2, which satisfy the constraint 5.76. As we can see, correlation goes to 1 more and more slowly if we increase the number of neurons N in the network. It reaches the value 1 asymptotically with an inverse exponential-like behavior, with a time constant that increases with the size of the network. For $N \to \infty$ the time constant diverges, therefore for every finite time the system has correlation 0. This proves that in the thermodynamic limit there is still propagation of chaos, provided that $C_1 = C_2 = C_3 = 0$. This is in perfect agreement with the result on propagation of chaos proved in [51][64][65] for independent Brownian motions, initial conditions and synaptic weights.

Neuron	Input	Synaptic Weights	Sigmoid Function
$\tau = 0.1$	$\overline{I} = -20$	$\Lambda = 40$	$T_{MAX} = 1$
$C_2 = 0$	$C_1 = 0$	$C_3 = 0$	$\lambda = 1$
			$V_T = 0$

Table 5.2: Parameters used for the numerical simulations of Figure 5.14.

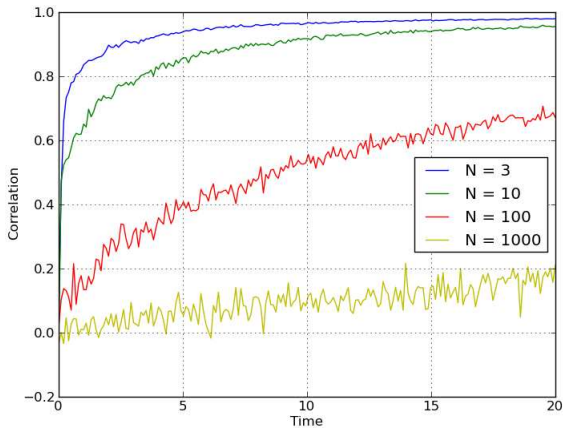

Figure 5.14: Stochastic synchronization in a fully connected network. Correlation gets closer and closer to 1 with a speed that depends on the number of neurons N in the system. These results have been obtained with the exact non-linear equations 5.1 and with $1,000$ Monte Carlo simulations. The parameters are those of Table 5.2, which are chosen in order to satisfy the constraint 5.76 for $\Lambda = -2\overline{I}$ and $\tau = -\frac{2}{\overline{I}}$. The value of the external current is purposely large ($\overline{I} = -20$) because it causes a faster convergence of the correlation to the value 1.

5.7 Fisher information

At the first perturbative order, the probability density of the membrane potentials is a multivariate normal distribution, therefore using formula 2.6 with $\theta = \overline{I}$ (therefore we have to set $\sigma_5 = 0$, in order to have a unique input current for all the neurons) and evaluating analytically the integrals of this definition, we obtain that the Fisher information is given by the well-known formula:

$$\mathcal{I}\left(\overline{I}, t\right) = \frac{\partial \overrightarrow{m}^T(t)}{\partial \overline{I}} \Sigma^{-1}(t) \frac{\partial \overrightarrow{m}(t)}{\partial \overline{I}} + \frac{1}{2} Tr\left(\frac{\partial \Sigma(t)}{\partial \overline{I}} \Sigma^{-1}(t) \frac{\partial \Sigma(t)}{\partial \overline{I}} \Sigma^{-1}(t)\right) \tag{5.77}$$

where $\overrightarrow{m}(t)$ is the vector of the means of the membrane potentials, while $\Sigma(t)$ is their covariance matrix, namely:

$$m_i(t) = \mathbb{E}\left[V_i(t)\right] \tag{5.78}$$

$$\Sigma_{ij}(t) = Cov\left(V_i(t), V_j(t)\right) \tag{5.79}$$

for $i, j = 0, 1, ..., N-1$.

In order to match the results of the Fisher information of this chapter with those of Chapter 6, we modify slightly our model: we assume that the entries of the synaptic connectivity matrix J are generic, deterministic and constant in time. For this reason we set $\sigma_3 = \sigma_4 = 0$, obtaining that $J = \overline{J}$, where $\overline{J}$ is a generic matrix with time-constant entries. Since $\mathbb{E}\left[Y_1^i(t)\right] = \mathbb{E}\left[Y_2^i(t)\right] = 0$, from 5.11 we obtain that:

$$m_i(t) = \mu_i \qquad \forall t$$

where μ_i is given by:

$$\mu_i = \tau \left[\sum_{j=0}^{N-1} J_{ij} S\left(\mu_j\right) + \overline{I} \right] \tag{5.80}$$

which is the extension of 5.12 to the case of a generic connectivity matrix J. This system of equations is solved with the Newton's method, using the inverse of its Jacobian matrix. Defining the function:

$$f_i\left(\overrightarrow{\mu}, \overline{I}\right) = \mu_i - \tau \left[\sum_{j=0}^{N-1} J_{ij} S\left(\mu_j\right) + \overline{I} \right], \qquad i = 0, 1, ..., N-1 \tag{5.81}$$

we obtain that the Jacobian matrix $\frac{\partial \overrightarrow{f}\left(\overrightarrow{\mu}, \overline{I}\right)}{\partial \overrightarrow{\mu}}$ has the following components:

$$\frac{\partial f_i\left(\overrightarrow{\mu}, \overline{I}\right)}{\partial \mu_j} = \begin{cases} 1 & \text{if} \quad i = j \\ \\ -\tau J_{ij} S'\left(\mu_j\right) & \text{if} \quad i \neq j \end{cases}$$

Since the dependence on the initial conditions vanishes quickly, we also assume for simplicity that $\sigma_2 = 0$. Therefore the covariance matrix of the system has the simple form:

$$\Sigma_{ij}\left(t\right) = \sigma_1^2 \left[\sum_{k=0}^{N-1} \int_0^t \Phi_{ik}\left(t-s\right) \Phi_{jk}\left(t-s\right) ds + C_1 \sum_{\substack{k,l=0 \\ k \neq l}}^{N-1} \int_0^t \Phi_{ik}\left(t-s\right) \Phi_{jl}\left(t-s\right) ds \right]$$

$$\Phi\left(t\right) = e^{At}$$

$$A_{ij} = \begin{cases} -\frac{1}{\tau} & \text{if} \quad i = j \\ & = -\frac{1}{\tau}\frac{\partial f_i\left(\overrightarrow{\mu}, \overline{I}\right)}{\partial \mu_j} \\ J_{ij} S'\left(\mu_j\right) & \text{if} \quad i \neq j \end{cases}$$

The derivatives of $\overrightarrow{m}\left(t\right)$ with respect to $\overline{I}$, that appear in 5.77, can be calculated as follows. From formulae 5.80 and 5.81, we obtain that $f_i\left(\overrightarrow{\mu}, \overline{I}\right) = 0$, and differentiating this relation we obtain:

$$df_i\left(\overrightarrow{\mu}, \overline{I}\right) = \sum_{j=0}^{N-1} \frac{\partial f_i(\overrightarrow{\mu}, \overline{I})}{\partial \mu_j} d\mu_j + \frac{\partial f_i(\overrightarrow{\mu}, \overline{I})}{\partial \overline{I}} d\overline{I} = 0, \qquad i = 0, 1, ..., N-1$$

Now, if we divide this formula by $d\overline{I}$, we obtain a system of N equations with N unknowns, namely the terms $\frac{d\mu_j}{d\overline{I}}$. Inverting the system, we obtain:

$$\frac{d\overrightarrow{\mu}}{d\overline{I}} = -\left[\frac{\partial \overrightarrow{f}\left(\overrightarrow{\mu}, \overline{I}\right)}{\partial \overrightarrow{\mu}}\right]^{-1} \frac{\partial \overrightarrow{f}\left(\overrightarrow{\mu}, \overline{I}\right)}{\partial \overline{I}}$$

where in our case:

$$\frac{\partial \overrightarrow{f}\left(\overrightarrow{\mu}, \overline{I}\right)}{\partial \overline{I}} = -\tau \overrightarrow{1}$$

This way of calculating the derivatives with respect to $\overline{I}$ is numerically convenient, because the inverse of the Jacobian matrix is known, since it has already been calculated in order to solve the system 5.80. This trick cannot be used to determine the derivatives $\frac{d\Sigma}{d\overline{I}}$ which appear in the second term of 5.77, therefore we have calculated them using formula 3.38. However, in all our simulations the second term of the Fisher information has always been much smaller than the first term, therefore it can be safely neglected. This is due to the fact that the covariance matrix Σ depends weakly on $\overline{I}$, as we will see in Chapter

Neuron	Input	Synaptic Weights	Sigmoid Function
$\tau = 1$	$\overline{I} = 0.5$	$\alpha = 1$	$T_{MAX} = 1$
	$\sigma_1 = 0.01$	$\beta = [0, 1, 20]$	$\lambda = 1$
			$V_T = 0$

Table 5.3: Parameters used for the numerical simulations of Figure 5.15.

6. Therefore now we have an algorithm to calculate numerically the Fisher information of the neural network.

In particular, we suppose that in a single trial the synaptic weights are independent and identically distributed as $J_{ij} \sim \mathcal{N}\left(\frac{\alpha}{N-1}, \left(\frac{\beta}{N-1}\right)^2\right) \; \forall\,(i,j) \,:\, i \neq j$, where α and β are two generic parameters. The simulations have been performed for the parameters of Table 5.3 and are reported in Figure 5.15, which shows that the Fisher information depends strongly on the inhomogeneities of the matrix J, namely on the parameter β. These inhomogeneities can be chosen arbitrarily high, since β is not a perturbative parameter. It is important to observe that these results have not been obtained by averaging the Fisher information over several realizations of the synaptic weights, because we are supposing that the J_{ij}s are deterministic, as we said previously. Clearly they are generated by the random distribution $\mathcal{N}\left(\frac{\alpha}{N-1}, \left(\frac{\beta}{N-1}\right)^2\right)$, but only once, with the purpose of creating quickly a sequence of "frozen" (and therefore deterministic) numbers that are used in the numerical calculation of $\mathcal{I}\left(\overline{I}, t\right)$.

To conclude, the main message is that the Fisher information can be higher for values of C_1 close to 1 (i.e. for highly correlated neurons) than for values of C_1 close to 0 (i.e. decorrelated neurons), depending on the inhomogeneities of the synaptic connectivity matrix.

An intuitive explanation of this phenomenon is not available, however in Chapter 6 we will provide an analytic formula that explains the emergence of the peaks of the Fisher information in the case of weak synaptic weights. We can also see that for a homogeneous network the Fisher information has a peak only for $C_1 \to 0$. This corresponds to the result found in Chapter 4 (see Figure 4.6), since there we analyzed the behavior of the neural

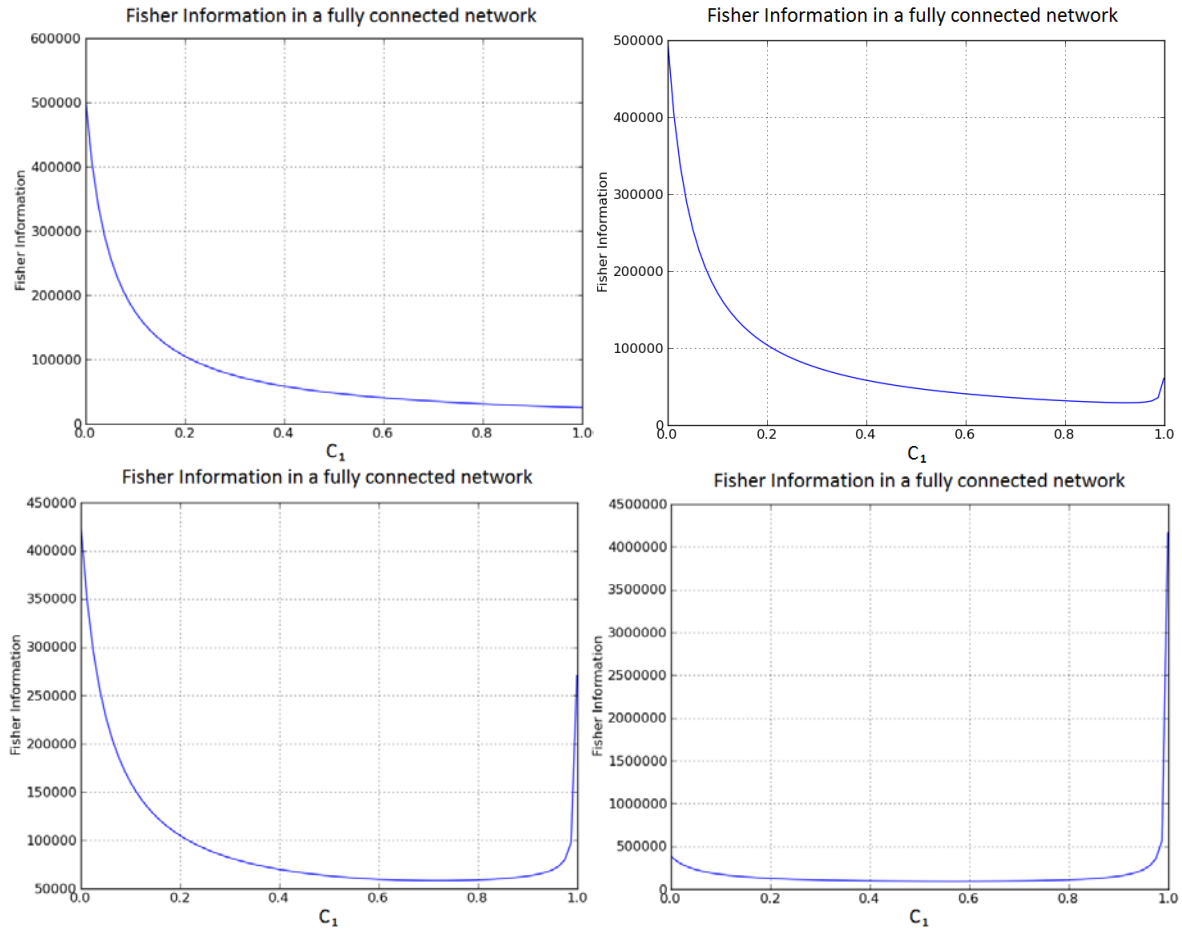

Figure 5.15: Four possible examples of Fisher information, obtained from the perturbative expansion with strong synaptic weights in a fully connected network with $N = 20$, $t = 2$, and for the parameters reported in Table 5.3. In detail, the figures on the top has been obtained for $\beta = 0$ (left-hand side) and $\beta = 1$ (right-hand side), while those at the bottom are two examples obtained for $\beta = 20$. The result depends on the values of the synaptic weights from trial to trial, showing higher encoding efficiency for $C_1 \to 0$ and/or $C_1 \to 1$. The non-linear system of algebraic equations 5.80 have been solved with the Newton's method, iterated times, while the integrals with respect to the time have been calculated with the trapezoidal rule, with integration time step $\Delta t = 0.1$.

network for constant and homogeneous synaptic weights.

5.8 Partial conclusion

In this chapter we have introduced a new technique, that extends the results found in Chapter 4. It is based on a perturbative expansion of the membrane potentials in terms of the sources of randomness of the system, namely the Brownian motions in the background, the initial conditions and the distribution of the synaptic weights. The expansion is performed around the stationary states of a finite network of rate neurons, assuming that the system is invariant under exchange of the neural indices. The variations of the membrane potentials in time and the inhomogeneities of the synaptic weights and of the external input currents are also introduced perturbatively. If the sources of randomness have sufficiently small variances, the expansion can be truncated at the first order. So this technique provides a way to study analytically the finite size effects of the network for many topologies of the connectivity matrix. It is important to observe that this approach works also for small networks, unlike the method used in Chapter 4, where we had to study large systems in order to avoid the higher order correlations.

The numerical comparison of this expansion with the real stochastic differential equations of the network provides a good match for relatively weak sources of randomness, especially at the second perturbative order. Moreover, by calculating the correlation structure of the system, it proves in three different ways that in general the mean-field theory cannot be used to describe the neural network, even in the thermodynamic limit. This is due to the impossibility to obtain the independence of the neurons when:

- the sources of randomness are correlated;
- the number of incoming synaptic connections per neuron is not high enough;

- even for decorrelated sources of randomness, special values of the parameters are chosen.

The second case had already been proved numerically in Chapter 3 with the cycle network, while the third case corresponds to the phenomenon that we have called stochastic synchronization. It consists in a perfect correlation between the neurons, generated by carefully tuned values of the parameters of the network. A perfect correlation means that the random fluctuations of the membrane potentials are exactly synchronized, so this explains our choice of the name for this phenomenon.

To conclude, we have used this perturbative expansion to calculate the Fisher information of the network. This analysis shows that the encoding capability of the system increases when the neurons are highly correlated or decorrelated. In particular, the increase of $\mathcal{I}\left(\overline{I},t\right)$ for decorrelated neurons had already been proved in Chapter 4 with the Mayer's cluster expansion. However the perturbative expansion has also revealed the second side of the coin, namely that the Fisher information could be much higher for correlated neurons, depending on the degree of inhomogeneity of the synaptic connections. This result will be proved analytically in the next chapter.

Chapter 6

Perturbative analysis with weak synaptic weights

THIS chapter is devoted to the development of another kind of perturbative analysis, similar to that considered in Chapter 5. However, now the perturbative parameters are not only the standard deviations of the Brownian motions, of the initial conditions and of the synaptic weights, but also the mean strength of the weights themselves. For this reason, this perturbative analysis provides a good agreement with the exact neural equations only in the case of relatively weak synaptic weights. The power of this new technique is in the fact that it does not require stationary states and the invariance of the system under exchange of the neural indices, as in Chapter 5. Moreover, it lets us analyze the case when also the topology of the network is random. In Section 6.1 we describe this approach, applied to the case of the rate model, while in Sections 6.2 and 6.3 we use it in order to calculate the correlation structure of the system. In Section 6.4 we apply this idea to the case of a fractal connectivity matrix with small-world properties, and we conclude in Section 6.5 by showing numerical evidence that supports this perturbative technique.

6.1 Description of the model

In this chapter we suppose again that the neural network is described by the rate equations that we have used in the Chapters 3, 4 and 5, namely:

$$dV_i(t) = \left[-\frac{1}{\tau} V_i(t) + \sum_{j=0}^{N-1} J_{ij}(t) S(V_j(t)) + I_i(t) \right] dt + \sigma_1 dB_i(t) \tag{6.1}$$

with $i = 0, 1, ..., N-1$. Randomness is present in the system through three different variables, the Brownian motions, the initial conditions and the strength of the synaptic weights, which are treated perturbatively. Their distributions are still supposed to be normal, because this will let us calculate analytically the correlation structure of the network using the Isserlis' theorem. We also introduce a fourth *non-perturbative* source of randomness, namely the *topology* of the synaptic connections. This means that not only the intensities of the synaptic weights are considered as random, but also the existence or not of a connection between two given neurons is not certain anymore. For the first three variables, we use the same covariance structures as in Chapter 5. For the Brownian motions it is given by the matrix Σ_1, whose entries are:

$$\begin{aligned} \left[\Sigma_1\right]_{ij} =& Cov\left(\frac{dB_i(t)}{dt}, \frac{dB_j(s)}{ds}\right) = C_{ij}^1 \delta(t-s) \\ C_{ij}^1 =& \begin{cases} 1 & \text{if} \quad i = j \\ \\ C_1 & \text{if} \quad i \neq j \end{cases} \end{aligned} \tag{6.2}$$

where C_1 is a free parameter that represents the correlation between two Brownian motions, while $\delta(\cdot)$ is the Dirac delta function. The matrix Σ_1 is a genuine covariance matrix

only if it is positive-semidefinite, namely if $\frac{1}{1-N} \leq C_1 \leq 1$.

The initial conditions are defined in terms of the following multivariate normal process:

$$\overrightarrow{V}(0) \sim \mathcal{N}(\overrightarrow{\mu}, \Sigma_2) \tag{6.3}$$

where:

$$\Sigma_2 = \sigma_2^2 \begin{bmatrix} 1 & C_2 & \cdots & C_2 \\ C_2 & 1 & \cdots & C_2 \\ \vdots & \vdots & \ddots & \vdots \\ C_2 & C_2 & \cdots & 1 \end{bmatrix} \tag{6.4}$$

We remind the reader that σ_2 represents the standard deviation of the initial conditions, while C_2 is their correlation. Again we have to choose $\frac{1}{1-N} \leq C_2 \leq 1$.

In this chapter we consider networks with random topologies, which means that the fact to have or not a connection between two given neurons is a (known) random variable: if in one realization of the network there is a connection from the j-th neuron to the i-th neuron, in another realization this connection could be missing. Therefore we suppose that the synaptic weights are given by the following formulae:

$$J_{ij}(t) = \frac{1}{M_i}\left[\sigma_4 \overline{J}_{ij}(t) + \sigma_3 W_{ij}\right] \tag{6.5}$$

$$\overline{J}_{ij}(t) = \widehat{\overline{J}}_{ij}(t) \circ T \tag{6.6}$$

$$W = \widehat{W} \circ T \tag{6.7}$$

$$\widehat{W} \sim \mathcal{MN}(0, \Omega_3, \Sigma_3) \tag{6.8}$$

$$M_i = \sum_{j=0}^{N-1} T_{ij} \tag{6.9}$$

where σ_3 and σ_4 are two perturbative parameters that represent (after the division by M_i), respectively, the standard deviation and the mean strength of the synaptic connections. M_i is the number (in general random) of incoming connections to the i-th neuron, and is used to prevent the explosion of the term $\sum_{j=0}^{N-1} J_{ij}(t) S(V_j(t))$ in equation 6.1 when M_i grows arbitrarily large. The symbol "$\circ$" represents the *Hadamard product*, therefore $C = A \circ B$ means that $C_{ij} = A_{ij}B_{ij}$, $\forall i, j$. T is a generic binary random matrix which represents the topology of the synaptic connections. More explicitly, we have $T_{ij} = 0$ if there is no connection from the j-th to the i-th neuron (namely if $J_{ij}(t) = 0 \ \forall t$), while $T_{ij} = 1$ if this connection is present. Below we show an example of connectivity matrix and its corresponding topology:

$$\overline{J}(t) = \begin{bmatrix} 0 & 0 & 2cos(t) & 3.6 \\ sin(5t) & 0 & 10 & 0 \\ 1 & \pi & 0 & arctan(7t) \\ 0 & (1+t)^{-5} & e^{-3t} & 0 \end{bmatrix}, \qquad T = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

The matrix $\hat{\overline{J}}(t)$ is completely deterministic, while the matrix $\widehat{W}$ is random only in the amplitudes of the synaptic weights (which follow a matrix normal distribution), but not in the topology. The covariance matrices Ω_3 and Σ_3 of $\widehat{W}$ are chosen in order to have:

$$Cov\left(\widehat{W}_{ij}, \widehat{W}_{kl}\right) = \begin{cases} 1 & \text{if} \quad (i=k) \wedge (j=l) \\ \\ C_3 & \text{otherwise} \end{cases} \tag{6.10}$$

The free parameter C_3 represents the correlation between two different and non-zero

synaptic weights, and the range of its plausible values depends on the topology of the connections, which is supposed to be completely generic. Moreover we assume that $\widehat{W}$ and T are independent.

To finish, we suppose that also the Brownian motions, the initial conditions, the amplitudes of the synaptic weights and the topology are independent from each other, therefore their reciprocal covariances are equal to zero:

$$Cov\left(B_i\left(t\right), V_j\left(0\right)\right) = Cov\left(B_i\left(t\right), \widehat{W}_{jk}\right) = Cov\left(B_i\left(t\right), T_{jk}\right)$$

$$= Cov\left(V_i\left(0\right), \widehat{W}_{jk}\right) = Cov\left(V_i\left(0\right), T_{jk}\right) = 0, \qquad \forall i,j,k \tag{6.11}$$

In principle, as we said in Chapter 5, the inner and mutual covariance structure of $B_i\left(t\right)$, $V_i\left(0\right)$ and $\widehat{W}_{ij}$ can be arbitrarily chosen. However in the current chapter we use only the simple structure defined by formulae 6.2, 6.4, 6.10 and 6.11, because this will generate simple analytic results for the correlation structure of the membrane potentials. We now are ready to introduce a perturbative expansion of $V_i\left(t\right)$ in terms of the parameters σ:

$$V_i\left(t\right) \approx Y_0^i\left(t\right) + \sum_{m=1}^{4} \sigma_m Y_m^i\left(t\right) + \sum_{\substack{m,n=1 \\ m \leq n}}^{4} \sigma_m \sigma_n Y_{m,n}^i\left(t\right) \tag{6.12}$$

where the functions $Y_m^i\left(t\right)$ and $Y_{m,n}^i\left(t\right)$ are to be determined through equation 6.1. In principle this expansion can be extended to any perturbative order, but in this chapter we truncate it at the second because the complexity of the results becomes quickly intractable.

6.1.1 The system of equations

In order to evaluate the functions $Y_m^i(t)$ and $Y_{m,n}^i(t)$, we have to replace the expansion 6.12 inside the equation 6.1, and to identify the coefficients of the same monomials in σ. Before doing this, we need the expansion of the sigmoid function in terms of σ. Therefore, defining, as in Chapter 5:

$$\zeta_j = \sum_{m=1}^{4} \sigma_m Y_m^j(t) + \sum_{\substack{m,n=1 \\ m \leq n}}^{4} \sigma_m \sigma_n Y_{m,n}^j(t)$$

the Taylor expansion of the sigmoid function is:

$$\begin{aligned} S(\mu + \zeta_j) \approx & S(\mu) + S'(\mu)\zeta_j + \frac{1}{2}S''(\mu)\zeta_j^2 \\ \approx & S(\mu) + S'(\mu)\sum_{m=1}^{4} \sigma_m Y_m^j(t) \\ & + \sum_{\substack{m,n=1 \\ m<n}}^{4} \sigma_m \sigma_n \left[S'(\mu) Y_{m,n}^j(t) + S''(\mu) Y_m^j(t) Y_n^j(t)\right] \\ & + \sum_{m=1}^{4} \sigma_m^2 \left[S'(\mu) Y_{m,m}^j(t) + \frac{1}{2}S''(\mu)\left(Y_m^j(t)\right)^2\right] \end{aligned}$$

having neglected the terms with order higher than 2. As in Chapter 5, we remind that this expansion can be used provided that its radius of convergence is large enough and that the rigorous analysis of the radius of convergence can be found in the Appendix D, if the activation function $S(\cdot)$ is given by 3.3. Now, if we replace this expansion and 6.12 inside the equation 6.1, comparing the coefficients of the same monomials in σ we obtain the following equations:

$$dY_0^i(t) = \left[-\frac{1}{\tau}Y_0^i(t) + I_i(t)\right]dt \tag{6.13}$$

$$dY_1^i(t) = -\frac{1}{\tau}Y_1^i(t)\,dt + dB_i(t) \tag{6.14}$$

$$dY_2^i(t) = -\frac{1}{\tau}Y_2^i(t)\,dt \tag{6.15}$$

$$dY_3^i(t) = \left[-\frac{1}{\tau}Y_3^i(t) + \frac{1}{M_i}\sum_{j=0}^{N-1} W_{ij}S\left(Y_0^j(t)\right)\right]dt \tag{6.16}$$

$$dY_4^i(t) = \left[-\frac{1}{\tau}Y_4^i(t) + \frac{1}{M_i}\sum_{j=0}^{N-1} \overline{J}_{ij}(t)\,S\left(Y_0^j(t)\right)\right]dt \tag{6.17}$$

$$\vdots$$

$$dY_{1,4}^i(t) = \left[-\frac{1}{\tau}Y_{1,4}^i(t) + \frac{1}{M_i}\sum_{j=0}^{N-1} \overline{J}_{ij}(t)\,S'\left(Y_0^j(t)\right)Y_1^j(t)\right]dt \tag{6.18}$$

$$dY_{2,4}^i(t) = \left[-\frac{1}{\tau}Y_{2,4}^i(t) + \frac{1}{M_i}\sum_{j=0}^{N-1} \overline{J}_{ij}(t)\,S'\left(Y_0^j(t)\right)Y_2^j(t)\right]dt \tag{6.19}$$

$$dY_{3,4}^i(t) = \left[-\frac{1}{\tau}Y_{3,4}^i(t) + \frac{1}{M_i}\sum_{j=0}^{N-1} \overline{J}_{ij}(t)\,S'\left(Y_0^j(t)\right)Y_3^j(t) + \frac{1}{M_i}\sum_{j=0}^{N-1} W_{ij}S'\left(Y_0^j(t)\right)Y_4^j(t)\right]dt \tag{6.20}$$

$$dY_{4,4}^i(t) = \left[-\frac{1}{\tau}Y_{4,4}^i(t) + \frac{1}{M_i}\sum_{j=0}^{N-1} \overline{J}_{ij}(t)\,S'\left(Y_0^j(t)\right)Y_4^j(t)\right]dt \tag{6.21}$$

$$\vdots$$

We have only written the equations that will be used in Section 6.2. The others do not

influence the perturbative expansions of the variance and covariance truncated at the 3rd perturbative order, therefore they are not shown here.

6.1.2 The initial conditions

The perturbative expansion 6.12 at $t = 0$ gives:

$$V_i(0) \approx Y_0^i(0) + \sum_{m=1}^{4} \sigma_m Y_m^i(0) + \sum_{\substack{m,n=1 \\ m \leq n}}^{4} \sigma_m \sigma_n Y_{m,n}^i(0)$$

From 6.3 we have $V_i(0) \sim \mathcal{N}\left(\mu_i, \sigma_2^2\right) = \mu_i + \sigma_2 \mathcal{N}(0,1)$, so comparing the two expressions we obtain:

$$Y_0^i(0) = \mu_i \tag{6.22}$$

$$Y_2^i(0) \sim \mathcal{N}(0,1) \tag{6.23}$$

$$Y_m^i(0) = 0, \qquad m = 1,3,4 \tag{6.24}$$

$$Y_{m,n}^i(0) = 0, \qquad \forall (m,n) : \; m \leq n \tag{6.25}$$

Therefore we can write the initial conditions as $V_i(0) = \mu_i + \sigma_2 Y_2^i(0)$, from which we obtain:

$$Cov\left(V_i(0), V_j(0)\right) = \sigma_2^2 Cov\left(Y_2^i(0), Y_2^j(0)\right)$$

Since from 6.4 we also know that:

$$Cov\left(V_i\left(0\right),V_j\left(0\right)\right)=\begin{cases}\sigma_2^2 & \text{if} \quad i=j \\ \\ \sigma_2^2 C_2 & \text{if} \quad i\neq j\end{cases}$$

from the comparison of these two expressions of the covariance matrix of $V_i\left(0\right)$ we obtain:

$$Cov\left(Y_2^i\left(0\right),Y_2^j\left(0\right)\right)=\begin{cases}1 & \text{if} \quad i=j \\ \\ C_2 & \text{if} \quad i\neq j\end{cases} \tag{6.26}$$

6.1.3 Solutions of the equations

Since equations 6.13 - 6.21 are linear, they can be solved analytically, giving the following solutions:

$$Y_0^i\left(t\right)=e^{-\frac{t}{\tau}}\left[\mu_i+\int_0^t e^{\frac{s}{\tau}}I_i\left(s\right)ds\right] \tag{6.27}$$

$$Y_1^i\left(t\right)=e^{-\frac{t}{\tau}}\int_0^t e^{\frac{s}{\tau}}dB_i\left(s\right) \tag{6.28}$$

$$Y_2^i\left(t\right)=e^{-\frac{t}{\tau}}Y_2^i\left(0\right) \tag{6.29}$$

$$Y_3^i\left(t\right)=\frac{e^{-\frac{t}{\tau}}}{M_i}\sum_{j=0}^{N-1}W_{ij}\int_0^t e^{\frac{s}{\tau}}S\left(Y_0^j\left(s\right)\right)ds \tag{6.30}$$

$$Y_4^i\left(t\right)=\frac{e^{-\frac{t}{\tau}}}{M_i}\sum_{j=0}^{N-1}\int_0^t e^{\frac{s}{\tau}}\overline{J}_{ij}\left(s\right)S\left(Y_0^j\left(s\right)\right)ds \tag{6.31}$$

$$\vdots$$

$$Y_{1,4}^{i}\left(t\right)=\frac{e^{-\frac{t}{\tau}}}{M_i}\sum_{j=0}^{N-1}\int_0^t\overline{J}_{ij}\left(s\right)S'\left(Y_0^j\left(s\right)\right)\left[\int_0^s e^{\frac{u}{\tau}}dB_j\left(u\right)\right]ds \tag{6.32}$$

$$Y_{2,4}^{i}\left(t\right)=\frac{e^{-\frac{t}{\tau}}}{M_i}\sum_{j=0}^{N-1}Y_2^j\left(0\right)\int_0^t\overline{J}_{ij}\left(s\right)S'\left(Y_0^j\left(s\right)\right)ds \tag{6.33}$$

$$\begin{aligned}Y_{3,4}^{i}\left(t\right)=&\frac{e^{-\frac{t}{\tau}}}{M_i}\left\{\sum_{j,k=0}^{N-1}\frac{W_{jk}}{M_j}\int_0^t\overline{J}_{ij}\left(s\right)S'\left(Y_0^j\left(s\right)\right)\left[\int_0^s e^{\frac{u}{\tau}}S\left(Y_0^k\left(u\right)\right)du\right]ds\right.\\&\left.+\sum_{j,k=0}^{N-1}\frac{W_{ij}}{M_j}\int_0^t S'\left(Y_0^j\left(s\right)\right)\left[\int_0^s e^{\frac{u}{\tau}}\overline{J}_{jk}\left(u\right)S\left(Y_0^k\left(u\right)\right)du\right]ds\right\}\end{aligned} \tag{6.34}$$

$$Y_{4,4}^{i}\left(t\right)=\frac{e^{-\frac{t}{\tau}}}{M_i}\sum_{j,k=0}^{N-1}\frac{1}{M_j}\int_0^t\overline{J}_{ij}\left(s\right)S'\left(Y_0^j\left(s\right)\right)\left[\int_0^s e^{\frac{u}{\tau}}\overline{J}_{jk}\left(u\right)S\left(Y_0^k\left(u\right)\right)du\right]ds \tag{6.35}$$

$$\vdots$$

Now we can use these results to calculate the correlation structure of the membrane potentials.

6.2 Correlation structure of the network

In this section we analyze the general case of random topologies, and we consider the networks with deterministic connections as a special case. From the perturbative expansion 6.12 with all the functions $Y_m^i\left(t\right)$ and $Y_{m,n}^i\left(t\right)$ evaluated as shown in Section 6.1.3, in order to calculate the covariance matrix of the membrane potentials we need to determine all the pair covariances between all the possible combinations of these functions. This is a consequence of the bilinearity property of the covariance operator. However, using the

Isserlis' theorem and the relations 6.11, it is easy to see that many of these terms are equal to zero. Moreover we have also to remove the 4th order terms in the expression of the covariance, like $\sigma_1^2\sigma_3^2 Cov\left(Y_{1,3}^i(t), Y_{1,3}^j(t)\right)$, since they are not complete. This is due to the fact that there are also 4th order terms like $\sigma_1^2\sigma_3^2 Cov\left(Y_1^i(t), Y_{1,1,3}^j(t)\right)$. These terms are due to 3rd order functions, like $Y_{1,1,3}^j(t)$ in this case, in the perturbative expansion 6.12, which have not been taken into account since we have truncated the expansion of the membrane potential at the 2nd order. Therefore the expansion of the covariance must be truncated at the 3rd order. So, to conclude, we obtain the following result:

$$
\begin{aligned}
&Cov\left(V_i(t), V_j(t)\right) \\
&= \sigma_1^2 Cov\left(Y_1^i(t), Y_1^j(t)\right) + \sigma_2^2 Cov\left(Y_2^i(t), Y_2^j(t)\right) \\
&+ \sigma_3^2 Cov\left(Y_3^i(t), Y_3^j(t)\right) + \sigma_4^2 Cov\left(Y_4^i(t), Y_4^j(t)\right) \\
&+ \sigma_4\left\{\sigma_1^2\left[Cov\left(Y_1^i(t), Y_{1,4}^j(t)\right) + Cov\left(Y_{1,4}^i(t), Y_1^j(t)\right)\right] + \sigma_2^2\left[Cov\left(Y_2^i(t), Y_{2,4}^j(t)\right) + Cov\left(Y_{2,4}^i(t), Y_2^j(t)\right)\right]\right. \\
&\left.+\sigma_3^2\left[Cov\left(Y_3^i(t), Y_{3,4}^j(t)\right) + Cov\left(Y_{3,4}^i(t), Y_3^j(t)\right)\right] + \sigma_4^2\left[Cov\left(Y_4^i(t), Y_{4,4}^j(t)\right) + Cov\left(Y_{4,4}^i(t), Y_4^j(t)\right)\right]\right\}
\end{aligned}
\tag{6.36}
$$

where, due to formulae 6.28, 6.29 and 6.30, for $i \neq j$ we obtain:

$$Cov\left(Y_1^i(t), Y_1^j(t)\right) = \frac{\tau C_1}{2}\left(1 - e^{-\frac{2t}{\tau}}\right) \tag{6.37}$$

$$Cov\left(Y_2^i(t), Y_2^j(t)\right) = C_2 e^{-\frac{2t}{\tau}} \tag{6.38}$$

$$Cov\left(Y_3^i(t), Y_3^j(t)\right) = C_3 e^{-\frac{2t}{\tau}} \sum_{k,l=0}^{N-1} \left[\int_0^t e^{\frac{s}{\tau}} S\left(Y_0^k(s)\right) ds\right] \left[\int_0^t e^{\frac{s}{\tau}} S\left(Y_0^l(s)\right) ds\right] \mathbb{E}\left[\frac{T_{ik}T_{jl}}{M_i M_j}\right] \tag{6.39}$$

and for $i = j$:

$$Var\left(Y_1^i(t)\right) = \frac{\tau}{2}\left(1 - e^{-\frac{2t}{\tau}}\right) \tag{6.40}$$

$$Var\left(Y_2^i(t)\right) = e^{-\frac{2t}{\tau}} \tag{6.41}$$

$$\begin{aligned} Var\left(Y_3^i(t)\right) = e^{-\frac{2t}{\tau}} \Bigg\{ & \sum_{k=0}^{N-1} \left[\int_0^t e^{\frac{s}{\tau}} S\left(Y_0^k(s)\right) ds\right]^2 \mathbb{E}\left[\left(\frac{T_{ik}}{M_i}\right)^2\right] \\ & + C_3 \sum_{\substack{k,l \\ k \neq l}} \left[\int_0^t e^{\frac{s}{\tau}} S\left(Y_0^k(s)\right) ds\right] \left[\int_0^t e^{\frac{s}{\tau}} S\left(Y_0^l(s)\right) ds\right] \mathbb{E}\left[\frac{T_{ik}T_{il}}{M_i^2}\right] \Bigg\} \end{aligned} \tag{6.42}$$

Because of formulae 6.28 - 6.35, for all i, j we obtain:

$$Cov\left(Y_4^i(t), Y_4^j(t)\right) = e^{-\frac{2t}{\tau}} \sum_{k,l=0}^{N-1} \left[\int_0^t e^{\frac{s}{\tau}} \hat{J}_{ik}(s) S\left(Y_0^k(s)\right) ds\right] \left[\int_0^t e^{\frac{s}{\tau}} \hat{J}_{jl}(s) S\left(Y_0^l(s)\right) ds\right] Cov\left(\frac{T_{ik}}{M_i}, \frac{T_{jl}}{M_j}\right) \tag{6.43}$$

$$\begin{aligned} &Cov\left(Y_1^i(t), Y_{1,4}^j(t)\right) = \\ &= \frac{\tau}{2} e^{-\frac{2t}{\tau}} \left\{ \mathbb{E}\left[\frac{T_{ji}}{M_j}\right] \left[\int_0^t \widehat{\overline{J}}_{ji}(s) S'\left(Y_0^i(s)\right) \left(e^{\frac{2s}{\tau}} - 1\right) ds\right] + C_1 \sum_{\substack{k=0 \\ k\neq i}}^{N-1} \mathbb{E}\left[\frac{T_{jk}}{M_j}\right] \left[\int_0^t \widehat{\overline{J}}_{jk}(s) S'\left(Y_0^k(s)\right) \left(e^{\frac{2s}{\tau}} - 1\right) ds\right] \right\} \end{aligned} \tag{6.44}$$

$$Cov\left(Y_2^i(t), Y_{2,4}^j(t)\right) = e^{-\frac{2t}{\tau}} \left\{ \mathbb{E}\left[\frac{T_{ji}}{M_j}\right] \left[\int_0^t \widehat{\overline{J}}_{ji}(s) S\left(Y_0^i(s)\right) ds\right] + C_2 \sum_{\substack{k=0 \\ k\neq i}}^{N-1} \mathbb{E}\left[\frac{T_{jk}}{M_j}\right] \left[\int_0^t \widehat{\overline{J}}_{jk}(s) S\left(Y_0^k(s)\right) ds\right] \right\} \tag{6.45}$$

$$\begin{aligned} &Cov\left(Y_3^i(t), Y_{3,4}^j(t)\right) \\ &= e^{-\frac{2t}{\tau}} \left\{ \sum_{k=0}^{N-1} \mathbb{E}\left[\left(\frac{T_{ik}}{M_i}\right)^2 \frac{T_{ji}}{M_j}\right] \left[\int_0^t e^{\frac{s}{\tau}} S\left(Y_0^k(s)\right) ds\right] \int_0^t \widehat{\overline{J}}_{ji}(s) S'\left(Y_0^i(s)\right) \left[\int_0^s e^{\frac{u}{\tau}} S\left(Y_0^k(u)\right) du\right] ds \right. \\ &+ C_3 \sum_{k,l=0}^{N-1} \mathbb{E}\left[\frac{T_{ik}T_{il}T_{ji}}{M_i^2 M_j}\right] \left[\int_0^t e^{\frac{s}{\tau}} S\left(Y_0^k(s)\right) ds\right] \int_0^t \widehat{\overline{J}}_{ji}(s) S'\left(Y_0^i(s)\right) \left[\int_0^s e^{\frac{u}{\tau}} S\left(Y_0^l(u)\right) du\right] ds \\ &+ C_3 \sum_{k,l=0}^{N-1} \mathbb{E}\left[\frac{T_{ik}T_{lk}T_{jl}}{M_i M_j M_l}\right] \left[\int_0^t e^{\frac{s}{\tau}} S\left(Y_0^k(s)\right) ds\right] \int_0^t \widehat{\overline{J}}_{jl}(s) S'\left(Y_0^l(s)\right) \left[\int_0^s e^{\frac{u}{\tau}} S\left(Y_0^k(u)\right) du\right] ds \\ &\left. + C_3 \sum_{k,l,m=0}^{N-1} \mathbb{E}\left[\frac{T_{ik}T_{lm}T_{jl}}{M_i M_j M_l}\right] \left[\int_0^t e^{\frac{s}{\tau}} S\left(Y_0^k(s)\right) ds\right] \int_0^t \widehat{\overline{J}}_{jl}(s) S'\left(Y_0^l(s)\right) \left[\int_0^s e^{\frac{u}{\tau}} S\left(Y_0^m(u)\right) du\right] ds \right\} \end{aligned} \tag{6.46}$$

$$\begin{aligned} &Cov\left(Y_4^i(t), Y_{4,4}^j(t)\right) \\ &= e^{-\frac{2t}{\tau}} \sum_{k,l,m=0}^{N-1} Cov\left(\frac{T_{ik}}{M_i}, \frac{T_{jl}T_{lm}}{M_j M_l}\right) \left[\int_0^t e^{\frac{s}{\tau}} \widehat{\overline{J}}_{ik}(s) S\left(Y_0^k(s)\right) ds\right] \left[\int_0^t \widehat{\overline{J}}_{jl}(s) S'\left(Y_0^l(s)\right) \left[\int_0^s e^{\frac{u}{\tau}} \widehat{\overline{J}}_{lm}(s) S\left(Y_0^m(u)\right) du\right] ds\right] \end{aligned} \tag{6.47}$$

Formula 6.44 is obtained using the following identity (which is a consequence of the mutual independence of the random variables):

$$Cov\left(B_i(t), B_j(t)\frac{\overline{J}_{kl}(t)}{M_k}\right) = \mathbb{E}\left[B_i(t)B_j(t)\widehat{\overline{J}}_{kl}(t)\frac{T_{kl}}{M_k}\right] - \mathbb{E}\left[B_i(t)\right]\mathbb{E}\left[B_j(t)\widehat{\overline{J}}_{kl}(t)\frac{T_{kl}}{M_k}\right]$$

$$= \widehat{\overline{J}}_{kl}(t)\left(\mathbb{E}\left[B_i(t)B_j(t)\right]\mathbb{E}\left[\frac{T_{kl}}{M_k}\right] - \mathbb{E}\left[B_i(t)\right]\mathbb{E}\left[B_j(t)\right]\mathbb{E}\left[\frac{T_{kl}}{M_k}\right]\right)$$

$$= \widehat{\overline{J}}_{kl}(t)\,\mathbb{E}\left[\frac{T_{kl}}{M_k}\right]Cov\left(B_i(t), B_j(t)\right)$$

A similar relation can be found for the initial conditions $\overrightarrow{V}(0)$ and the topology T:

$$Cov\left(V_i(0), V_j(0)\frac{\overline{J}_{kl}(t)}{M_k}\right) = \widehat{\overline{J}}_{kl}(t)\,\mathbb{E}\left[\frac{T_{kl}}{M_k}\right]Cov\left(V_i(0), V_j(0)\right)$$

from which we have obtained formula 6.45. Instead, in order to obtain formula 6.46, we have used the following result:

$$Cov\left(\frac{W_{ij}}{M_i}, \frac{W_{kl}}{M_k}\frac{\overline{J}_{mn}(t)}{M_m}\right) = Cov\left(\widehat{W}_{ij}\frac{T_{ij}}{M_i}, \widehat{W}_{kl}\frac{T_{kl}}{M_k}\widehat{\overline{J}}_{mn}(t)\frac{T_{mn}}{M_m}\right)$$

$$= \widehat{\overline{J}}_{mn}(t)\left(\mathbb{E}\left[\widehat{W}_{ij}\frac{T_{ij}}{M_i}\widehat{W}_{kl}\frac{T_{kl}}{M_k}\frac{T_{mn}}{M_m}\right] - \mathbb{E}\left[\widehat{W}_{ij}\frac{T_{ij}}{M_i}\right]\mathbb{E}\left[\widehat{W}_{kl}\frac{T_{kl}}{M_k}\frac{T_{mn}}{M_m}\right]\right)$$

$$= \widehat{\overline{J}}_{mn}(t)\left(\mathbb{E}\left[\widehat{W}_{ij}\widehat{W}_{kl}\right]\mathbb{E}\left[\frac{T_{ij}T_{kl}T_{mn}}{M_iM_kM_m}\right] - \mathbb{E}\left[\widehat{W}_{ij}\right]\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]\mathbb{E}\left[\widehat{W}_{kl}\right]\mathbb{E}\left[\frac{T_{kl}T_{mn}}{M_kM_m}\right]\right)$$

$$= \widehat{\overline{J}}_{mn}(t)\,Cov\left(\widehat{W}_{ij}, \widehat{W}_{kl}\right)\mathbb{E}\left[\frac{T_{ij}T_{kl}T_{mn}}{M_iM_kM_m}\right]$$

which is a consequence of the independence between $\widehat{W}$ and T. In the same way it is possible to prove that:

$$Cov\left(\frac{W_{ij}}{M_i},\frac{\overline{J}_{kl}\left(t\right)}{M_k}\frac{\overline{J}_{mn}\left(t\right)}{M_m}\right)=\widehat{\overline{J}}_{kl}\left(t\right)\widehat{\overline{J}}_{mn}\left(t\right)\mathbb{E}\left[\widehat{W}_{ij}\right]\left(\mathbb{E}\left[\frac{T_{ij}T_{kl}T_{mn}}{M_iM_kM_m}\right]-\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]\mathbb{E}\left[\frac{T_{kl}T_{mn}}{M_kM_m}\right]\right)=0$$

so for this reason the term $Cov\left(Y_3^i\left(t\right),Y_{4,4}^i\left(t\right)\right)$ does not appear in formula 6.36.

Once the covariance matrix of the membrane potentials has been determined, we can evaluate their correlation structure using the Pearson's correlation coefficient. The only quantities that remain unspecified are $\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]$, $\mathbb{E}\left[\frac{T_{ik}T_{jl}}{M_iM_j}\right]$ and $\mathbb{E}\left[\frac{T_{ik}T_{lm}T_{jl}}{M_iM_jM_l}\right]$, that depend on the distribution of the matrix T. This can be accomplished by a multidimensional Taylor expansion. For example, for $\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]$ we Taylor-expand the function:

$$f:\ \ \left(T_{i0},...,T_{i,N-1}\right)\ \ \rightarrow\ \ \frac{T_{ij}}{M_i} \tag{6.48}$$

at the point $\left(\mathbb{E}\left[T_{i0}\right],...,\mathbb{E}\left[T_{i,N-1}\right]\right)$ to obtain:

$$\begin{aligned}\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]=&\mathbb{E}\left[\frac{T_{ij}}{\sum_{k=0}^{N-1}T_{ik}}\right]\\&=\sum_{n_0=0}^{\infty}\sum_{n_1=0}^{\infty}\cdots\sum_{n_{N-1}=0}^{\infty}\frac{\mathbb{E}\left[\left(T_{i0}-\mathbb{E}\left[T_{i0}\right]\right)^{n_0}\cdots\left(T_{i,N-1}-\mathbb{E}\left[T_{i,N-1}\right]\right)^{n_{N-1}}\right]}{n_0!\cdots n_{N-1}!}\left(\frac{\partial^{n_0+...+n_{N-1}}f}{\partial T_{i0}^{n_0}\cdots\partial T_{i,N-1}^{n_{N-1}}}\right)\left(\mathbb{E}\left[T_{i0}\right],\cdots,\mathbb{E}\left[T_{i,N-1}\right]\right)\end{aligned} \tag{6.49}$$

In detail, we have up to the third order:

$$\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]\approx\frac{\mathbb{E}\left[T_{ij}\right]}{\sum_{k=0}^{N-1}\mathbb{E}\left[T_{ik}\right]}+\frac{1}{2}\sum_{k,l=0}^{N-1}Cov\left(T_{ik},T_{il}\right)\left(\frac{\partial^2f}{\partial T_{ik}\partial T_{il}}\right)\left(\mathbb{E}\left[T_{i0}\right],\cdots,\mathbb{E}\left[T_{i,N-1}\right]\right) \tag{6.50}$$

where:

$$\left(\frac{\partial^2 f}{\partial T_{ik}\partial T_{il}}\right)\left(\mathbb{E}\left[T_{i0}\right],\cdots,\mathbb{E}\left[T_{i,N-1}\right]\right)$$

$$=\begin{cases}\dfrac{2\mathbb{E}[T_{ij}]}{\left(\sum\limits_{m=0}^{N-1}\mathbb{E}[T_{im}]\right)^3} & \text{if} \quad k,l\neq j \\ \dfrac{2\mathbb{E}[T_{ij}]-\sum\limits_{m=0}^{N-1}\mathbb{E}[T_{im}]}{\left(\sum\limits_{m=0}^{N-1}\mathbb{E}[T_{im}]\right)^3} & \text{if} \quad ((k\neq j)\wedge(l=j))\vee((k=j)\wedge(l\neq j)) \\ -\dfrac{2\sum\limits_{\substack{m=0\\ m\neq j}}^{N-1}\mathbb{E}[T_{im}]}{\left(\sum\limits_{m=0}^{N-1}\mathbb{E}[T_{im}]\right)^3} & \text{if} \quad k,l=j\end{cases}$$

The function 6.48 is analytic everywhere, but when $M_i=0$. However, we remind that the case $M_i=0$ is not contemplated by formula 6.5, since it gives $J_{ij}=\frac{0}{0}$, therefore we simply set $J_{ij}=0$ since there are no incoming connections to the i-th neuron. For this reason we have always to consider $M_i\neq 0$, so the multidimensional Taylor series of $f\left(T_{i0},...,T_{i,N-1}\right)$ has a finite radius of convergence and it does converge to $\frac{T_{ij}}{\sum\limits_{k=0}^{N-1}T_{ik}}$ everywhere.

After this analysis, the conclusion is that we can calculate $\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]$ once we know the quantities $\mathbb{E}\left[T_{ik}\right]$, $\mathbb{E}\left[T_{ik}T_{il}\right]$ etc. The same reasoning can be applied to $\mathbb{E}\left[\frac{T_{ik}T_{jl}}{M_iM_j}\right]$ and $\mathbb{E}\left[\frac{T_{ik}T_{lm}T_{jl}}{M_iM_jM_l}\right]$. In Section 6.4 we show how to determine these quantities for the fractal connectivity matrix introduced by Sporns in [190]. These results can also be used for networks with deterministic topologies, but we have to set $\mathbb{E}\left[\prod\frac{T}{M}\right]=\prod\frac{T}{M}$ in formulae 6.39, 6.42, 6.44, 6.45,

6.46, and we have to set to zero the covariance functions of $\frac{T}{M}$ in formulae 6.43 and 6.47 (so that $Cov\left(Y_4^i(t), Y_4^j(t)\right) = Cov\left(Y_4^i(t), Y_{4,4}^j(t)\right) = 0$).

6.3 A problem with the initial conditions

Before we start to analyze a concrete example of connectivity matrix, we have to show a problem with the initial conditions. In fact, if we choose $\sigma_2, \sigma_4 \neq 0$, $\sigma_1, \sigma_3 = 0$ and $C_2 = 0$, at least in the case of a deterministic topology the correlation function that we have calculated perturbatively is not necessarily in the range $[-1, 1]$ as required. This can be seen from formulae 6.36 - 6.47, which for these values of the parameters and a deterministic T, give:

$$Corr\left(V_i(t), V_j(t)\right) = \frac{\sigma_2^2 Cov\left(Y_2^i(t), Y_2^j(t)\right) + \sigma_4\sigma_2^2\left[Cov\left(Y_2^i(t), Y_{2,4}^j(t)\right) + Cov\left(Y_{2,4}^i(t), Y_2^j(t)\right)\right]}{\sigma_2^2 Var\left(Y_2^i(t)\right) + \sigma_4\sigma_2^2\left[Cov\left(Y_2^i(t), Y_{2,4}^i(t)\right) + Cov\left(Y_{2,4}^i(t), Y_2^i(t)\right)\right]}$$

$$= \sigma_4\left\{\frac{T_{ji}}{M_j}\left[\int_0^t \widehat{\overline{J}}_{ji}(s)\, S\left(Y_0^i(s)\right) ds\right] + \frac{T_{ij}}{M_i}\left[\int_0^t \widehat{\overline{J}}_{ij}(s)\, S\left(Y_0^j(s)\right) ds\right]\right\} \tag{6.51}$$

where for simplicity we have also supposed that all the neurons behave in the same way, so that $Var\left(V_i(t)\right) = Var\left(V_j(t)\right)$. Therefore, if $\widehat{\overline{J}}_{ij}(t)$, $\widehat{\overline{J}}_{ji}(t)$, $S\left(Y_0^i(t)\right)$ and $S\left(Y_0^j(t)\right)$ are for example constant in time, from formula 6.51 we obtain that $Corr\left(V_i(t), V_j(t)\right)$ increases linearly with time, therefore at some point it will be outside the range $[-1, 1]$. This can be seen also from Figure 6.1 (left-hand side), which has been obtained from the numerical simulation of the equations 6.13 - 6.21 (the details of the numerical scheme will be provided in Section 6.5) for the values of the parameters reported in Table 6.1.

This problem does not happen when $\sigma_1, \sigma_4 \neq 0$ and $\sigma_2, \sigma_3 = 0$, or when $\sigma_3, \sigma_4 \neq 0$ and $\sigma_1, \sigma_2 = 0$, or when $\sigma_4 \neq 0$ and $\sigma_1, \sigma_2, \sigma_3 = 0$, therefore it is only related to the initial

Neuron	Input	Synaptic Weights	Sigmoid Function
$\tau = 1$	$I_i = 0$	$\widehat{\overline{J}}_{ij} = 3$	$T_{MAX} = 1$
$\sigma_2 = 0.1$	$\sigma_1 = 0$	$\sigma_3 = 0$	$\lambda = 1$
		$\sigma_4 = 0.1$	$V_T = 0$
$C_2 = 0$	$C_1 = 0$	$C_3 = 0$	
$\mu = 0$			

Table 6.1: Values of the parameters used to generate Figure 6.1.

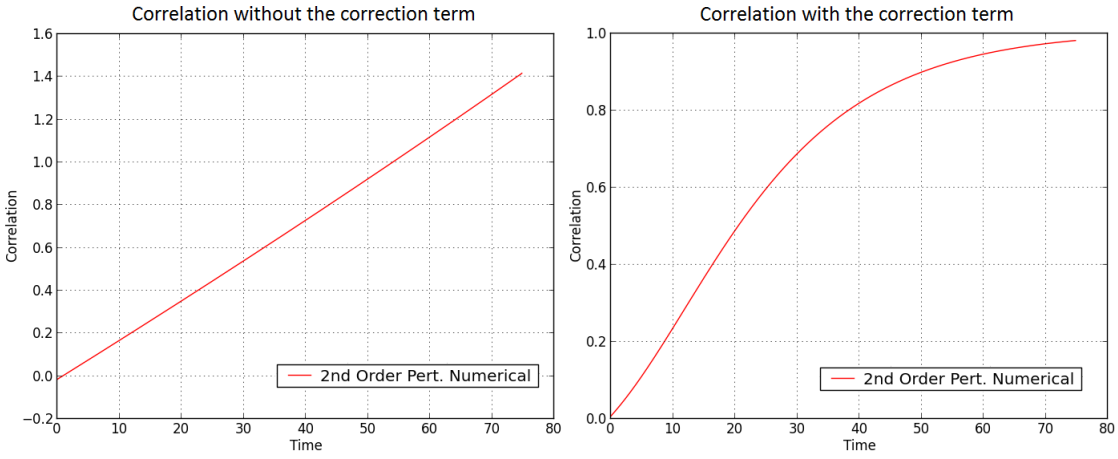

Figure 6.1: Correlation obtained from formula 6.36 using the numerical simulation of formulae 6.13 - 6.21 (left-hand side), and the same function obtained from formula 6.53 (right-hand side). The values of the parameters are shown in Table 6.1, while the topology of the network is K_{10}. In the first figure the correlation does not stay in the range $[-1, 1]$ for all time, and the problem is corrected in the second figure, see text.

conditions. It is of course due to our approximation. In fact, if we want to calculate the variance and covariance between two perturbative expansions of the form $F_i(t) = F_0^i(t) + \epsilon F_1^i(t) + \epsilon^2 F_2^i(t)$, where $F_0^i(t)$ is deterministic, we obtain:

$$Var(F_i(t)) = \epsilon^2 Var\left(F_1^i(t)\right) + 2\epsilon^3 Cov\left(F_1^i(t), F_2^i(t)\right) + \epsilon^4 Var\left(F_2^i(t)\right)$$

$$Cov(F_i(t), F_j(t)) = \epsilon^2 Cov\left(F_1^i(t), F_1^j(t)\right) + \epsilon^3 Cov\left(F_1^i(t), F_2^j(t)\right) + \epsilon^3 Cov\left(F_2^i(t), F_1^j(t)\right)$$

$$+ \epsilon^4 Cov\left(F_2^i(t), F_2^j(t)\right)$$

Due to the Cauchy-Schwarz inequality, we always have:

$$[Cov(F_i(t), F_j(t))]^2 \leq Var(F_i(t)) Var(F_j(t))$$

namely $|Corr(F_i(t), F_j(t))| \leq 1$. However, if we neglect the terms proportional to ϵ^4, as we did in Section 6.2, this inequality is not guaranteed to hold anymore. Therefore even if the approximations of the variance and covariance are good, the correlation could be completely wrong. This is the origin of the problem we have mentioned before. Moreover, it happens only when we deal with the initial conditions and not with the other random variables, because only for $\sigma_2, \sigma_4 \neq 0$ and $\sigma_1, \sigma_3 = 0$ do we have 4th order terms and the variance and covariance converge to zero for $t \to +\infty$, giving rise to an undefined correlation of the form $\frac{0}{0}$.

The solution is to keep the 4th order terms generated by the initial conditions in the formula of the variance and covariance. Now, for $\sigma_2, \sigma_4 \neq 0$ and $\sigma_1, \sigma_3 = 0$ we have:

$$V_i(t) = Y_0^i(t) + \sigma_2 Y_2^i(t) + \sigma_4 Y_4^i(t) + \sigma_2\sigma_4 Y_{2,4}^i(t) + \sigma_4^2 Y_{4,4}^i(t)$$

since it can be easily proved that $Y_{2,2}^{i}\left(t\right)=0\,\forall t$. Therefore in this case the exact covariance function is:

$$
\begin{aligned}
&Cov\left(V_i\left(t\right),V_j\left(t\right)\right)\\
&=\sigma_2^2 Cov\left(Y_2^i\left(t\right),Y_2^j\left(t\right)\right)+\sigma_4^2 Cov\left(Y_4^i\left(t\right),Y_4^j\left(t\right)\right)\\
&+\sigma_4\sigma_2^2\left[Cov\left(Y_2^i\left(t\right),Y_{2,4}^j\left(t\right)\right)+Cov\left(Y_{2,4}^i\left(t\right),Y_2^j\left(t\right)\right)\right]+\sigma_4^3\left[Cov\left(Y_4^i\left(t\right),Y_{4,4}^j\left(t\right)\right)+Cov\left(Y_{4,4}^i\left(t\right),Y_4^j\left(t\right)\right)\right]\\
&+\sigma_2^2\sigma_4^2 Cov\left(Y_{2,4}^i\left(t\right),Y_{2,4}^j\left(t\right)\right)+\sigma_4^4 Cov\left(Y_{4,4}^i\left(t\right),Y_{4,4}^j\left(t\right)\right)
\end{aligned}
\tag{6.52}
$$

The 4th order term $\sigma_2\sigma_4^3 Cov\left(Y_{2,4}^i\left(t\right),Y_{4,4}^j\left(t\right)\right)$ has not been taken into account because it is proportional to $Cov\left(Y_2^k\left(0\right)\frac{T_{ik}}{M_i},\frac{T_{jk}}{M_j}\frac{T_{kl}}{M_k}\right)$, which is equal to zero, as proved below:

$$
\begin{aligned}
Cov\left(Y_2^k\left(0\right)\frac{T_{ik}}{M_i},\frac{T_{jk}}{M_j}\frac{T_{kl}}{M_k}\right)=&\mathbb{E}\left[Y_2^k\left(0\right)\frac{T_{ik}}{M_i}\frac{T_{jk}}{M_j}\frac{T_{kl}}{M_k}\right]-\mathbb{E}\left[Y_2^k\left(0\right)\frac{T_{ik}}{M_i}\right]\mathbb{E}\left[\frac{T_{jk}}{M_j}\frac{T_{kl}}{M_k}\right]\\
=&\mathbb{E}\left[Y_2^k\left(0\right)\right]\left(\mathbb{E}\left[\frac{T_{ik}}{M_i}\frac{T_{jk}}{M_j}\frac{T_{kl}}{M_k}\right]-\mathbb{E}\left[\frac{T_{ik}}{M_i}\right]\mathbb{E}\left[\frac{T_{jk}}{M_j}\frac{T_{kl}}{M_k}\right]\right)\\
=&0
\end{aligned}
$$

We can simplify 6.52 further by noticing that for $\sigma_4\neq 0$ and $\sigma_1,\sigma_2,\sigma_3=0$ the problem of the correlation does not appear anymore if we calculate it using the truncated covariance function 6.36. Since for these values of the perturbative parameters the covariance 6.52 becomes simply:

$$Cov\left(V_i\left(t\right),V_j\left(t\right)\right)=\sigma_4^2Cov\left(Y_4^i\left(t\right),Y_4^j\left(t\right)\right)$$

$$+\sigma_4^3\left[Cov\left(Y_4^i\left(t\right),Y_{4,4}^j\left(t\right)\right)+Cov\left(Y_{4,4}^i\left(t\right),Y_4^j\left(t\right)\right)\right]+\sigma_4^4Cov\left(Y_{4,4}^i\left(t\right),Y_{4,4}^j\left(t\right)\right)$$

which differs from formula 6.36 (calculated for $\sigma_4 \neq 0$ and $\sigma_1,\sigma_2,\sigma_3=0$) only in the 4th order term $\sigma_4^4Cov\left(Y_{4,4}^i\left(t\right),Y_{4,4}^j\left(t\right)\right)$, this means that there is no need to add this term in order to correct the perturbative expansion. Therefore we see from 6.52 that the only term which is required to alleviate the problem of the correlation is $\sigma_2^2\sigma_4^2Cov\left(Y_{2,4}^i\left(t\right),Y_{2,4}^j\left(t\right)\right)$. To conclude, the final formula for the covariance that we have to use is:

$$Cov\left(V_i\left(t\right),V_j\left(t\right)\right)$$

$$=\sigma_1^2Cov\left(Y_1^i\left(t\right),Y_1^j\left(t\right)\right)+\sigma_2^2Cov\left(Y_2^i\left(t\right),Y_2^j\left(t\right)\right)$$

$$+\sigma_3^2Cov\left(Y_3^i\left(t\right),Y_3^j\left(t\right)\right)+\sigma_4^2Cov\left(Y_4^i\left(t\right),Y_4^j\left(t\right)\right)$$

$$+\sigma_4\left\{\sigma_1^2\left[Cov\left(Y_1^i\left(t\right),Y_{1,4}^j\left(t\right)\right)+Cov\left(Y_{1,4}^i\left(t\right),Y_1^j\left(t\right)\right)\right]+\sigma_2^2\left[Cov\left(Y_2^i\left(t\right),Y_{2,4}^j\left(t\right)\right)+Cov\left(Y_{2,4}^i\left(t\right),Y_2^j\left(t\right)\right)\right]\right.$$

$$\left.+\sigma_3^2\left[Cov\left(Y_3^i\left(t\right),Y_{3,4}^j\left(t\right)\right)+Cov\left(Y_{3,4}^i\left(t\right),Y_3^j\left(t\right)\right)\right]+\sigma_4^2\left[Cov\left(Y_4^i\left(t\right),Y_{4,4}^j\left(t\right)\right)+Cov\left(Y_{4,4}^i\left(t\right),Y_4^j\left(t\right)\right)\right]\right\}$$

$$+\sigma_2^2\sigma_4^2Cov\left(Y_{2,4}^i\left(t\right),Y_{2,4}^j\left(t\right)\right) \tag{6.53}$$

where:

$$Cov\left(Y_{2,4}^{i}\left(t\right),Y_{2,4}^{j}\left(t\right)\right)$$

$$=e^{-\frac{2t}{\tau}}\left\{\sum_{k=0}^{N-1}\mathbb{E}\left[\frac{T_{ik}T_{jk}}{M_{i}M_{j}}\right]\left[\int_{0}^{t}\widehat{\overline{J}}_{ik}\left(s\right)S'\left(Y_{0}^{k}\left(s\right)\right)ds\right]\left[\int_{0}^{t}\widehat{\overline{J}}_{jk}\left(s\right)S'\left(Y_{0}^{k}\left(s\right)\right)ds\right]\right.$$

$$\left.+C_{2}\sum_{\substack{k,l=0\\k\neq l}}^{N-1}\mathbb{E}\left[\frac{T_{ik}T_{jl}}{M_{i}M_{j}}\right]\left[\int_{0}^{t}\widehat{\overline{J}}_{ik}\left(s\right)S'\left(Y_{0}^{k}\left(s\right)\right)ds\right]\left[\int_{0}^{t}\widehat{\overline{J}}_{jl}\left(s\right)S'\left(Y_{0}^{l}\left(s\right)\right)ds\right]\right\} \tag{6.54}$$

We remind the reader that if he/she is interested only in the calculation of the variance and covariance, the term $Cov\left(Y_{2,4}^{i}\left(t\right),Y_{2,4}^{j}\left(t\right)\right)$ is not important, but it must be used if he/she needs to evaluate the correlation function. Indeed, using formula 6.53, the problem of the correlation is corrected, as it can be seen from Figure 6.1 (right-hand side).

6.4 Fractal connectivity matrix

As we reported in Section 2.1, the brain is often characterized by a small-world topology. A famous algorithm that generates networks with this property has been introduced by Watts and Strogatz [47]. Even if in principle it is possible to calculate analytically the covariance structure of the neurons over the random topology generated by this algorithm, in practice it is not a simple task, because the exact evaluation of $\mathbb{E}\left[T_{ij}\right]$, $\mathbb{E}\left[T_{ik}T_{jl}\right]$, $\mathbb{E}\left[T_{il}T_{jm}T_{kn}\right]$ etc, which is required for example by formula 6.49, can be accomplished through a complicated combinatorial analysis. Moreover this algorithm does not mimic the nested structure of the connectivity matrix of the brain. In fact, Watts and Strogatz tried to replicate only two features of the brain, namely its path length (which represents the shortest distance between two vertices in terms of the number of edges) and its

clustering coefficient (which, for a given vertex, quantifies the connectivity degree of its neighbourhood, i.e. of the vertices directly connected to it), without taking into account its nested structure. A more tractable algorithm, which reproduces more biologically realistic connections, has been introduced by Sporns in [190]. Since the connectome of the brain has a nested structure, Sporns suggested to describe it using a fractal connectivity matrix. One of the cases he studied is what he called the *fractal pattern* (frc). It is obtained by choosing two integer numbers, μ and η (Sporns called them m and n, but we prefer to use different symbols to avoid confusion with the vector and matrix indices) with $\mu \leq \eta$, and a real non-negative number E. The total number of neurons in the network is $N = 2^\eta$, and the different levels of the fractal structure are described by a parameter $\kappa = 0, 1, ..., \eta - \mu$ (Sporns called it k). As shown in Figure 6.2, we start with an elementary block of 2^μ neurons, which forms the level 0 of the fractal structure ($\kappa = 0$). Within this block the neurons are fully connected and without self-connections. Then we duplicate this block. The connection density between the two elementary blocks is the number of actual connections between them divided by the total number of possible connections. So we connect them with a connection density E^{-1} (here $\kappa = 1$, namely we are at the level 1). This means that the number of connections between the two blocks in one direction is the integer part of $4^\mu E^{-1}$. We emphasize the fact that these connections are randomly chosen. The resulting network is then "duplicated", namely we produce another pair of groups with 2^μ fully interconnected neurons in each one, and interconnected between them with a connection density E^{-1} (the connections are chosen randomly again, so this is not an identical copy). Then we connect the two "copies" with a connection density E^{-2} ($\kappa = 2$), and so on and so forth. The process is repeated iteratively until we reach the level $\kappa = \eta - \mu$. It is also important to observe that these connections are directed, therefore the connectivity matrix is generally not symmetric. Two examples are shown in Figure 6.3.

According to [190], the parameter E determines the path length, the clustering coefficient and the complexity of the network. The latter was first introduced in [191], and quantifies the extent to which a system is both functionally segregated and functionally integrated.

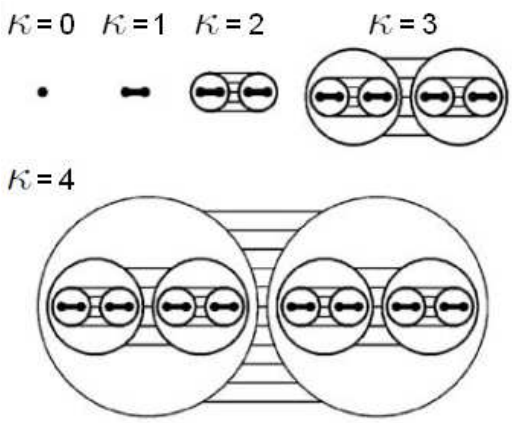

Figure 6.2: Sporns' algorithm for the fractal connectivity matrix. At the level $\kappa = 0$ a single dot represents a group of 2^{μ} fully connected neurons. At $\kappa = 1$ we duplicate this elementary block, obtaining two groups of 2^{μ} neurons which are linked together with a connection density E^{-1}. This structure is generated again at the level $\kappa = 2$, and connected to the previous one with a connection density E^{-2}, and so on. This figure has been taken and adapted from [190].

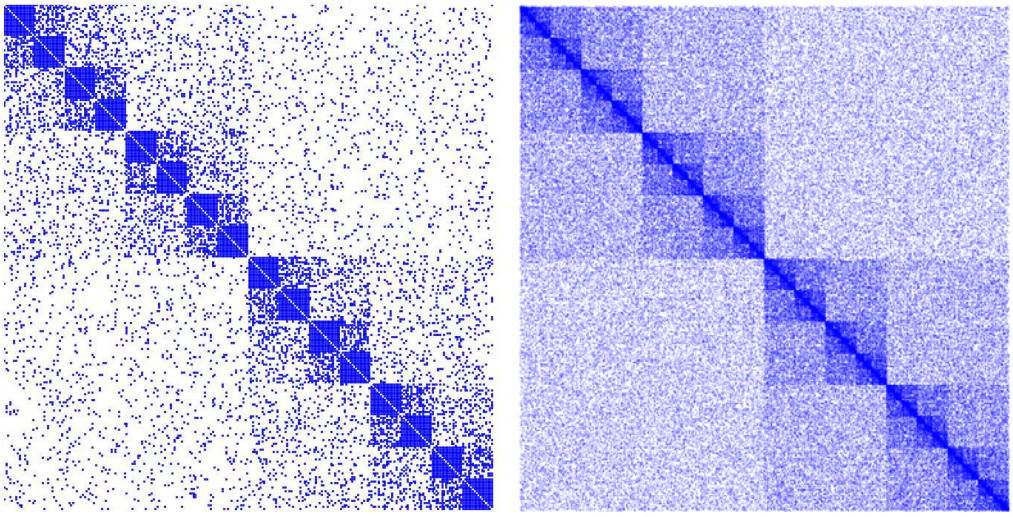

Figure 6.3: Two examples of fractal matrix obtained with the Sporns' algorithm, for $\eta = 8$, $\mu = 4$ and $E = 2.0$ (left-hand side) and for $\eta = 11$, $\mu = 2$ and $E = 1.5$ (right-hand side). A blue dot corresponds to a 1 in the topology matrix, while the absence of the dot corresponds to a 0. The figure on the right-hand side has been resized in order to the have the same spatial extension as the figure on the left-hand side. For this reason it does not clearly show the diagonal white line corresponding to $J_{ii}(t) = 0$, namely to the absence of self-connections.

This means that both the degree of independence of the blocks and their level of cooperation are taken into account by a single quantity, the complexity of the network, which for the fractal topology is maximum when $E \approx 2$.

Now we have to determine the quantities $\mathbb{E}[T_{ij}]$, $\mathbb{E}[T_{ik}T_{jl}]$, $\mathbb{E}[T_{il}T_{jm}T_{kn}]$ etc. Therefore we need to analyze the algorithm that generates the fractal connectivity matrix. If all the connections are at the level $\kappa = 0$, where the neurons are always fully connected, then we trivially have:

$$\begin{aligned}
\mathbb{E}[T_{ij}] &= 1 - \delta_{ij} \\
\mathbb{E}[T_{ik}T_{jl}] &= (1 - \delta_{ik})(1 - \delta_{jl}) \\
\mathbb{E}[T_{il}T_{jm}T_{kn}] &= (1 - \delta_{il})(1 - \delta_{jm})(1 - \delta_{kn}) \\
&\vdots
\end{aligned}$$

because in this case the entries of the topology are deterministic. Moreover, if we have an entry of the topology matrix, for example T_{ik}, at the level $\kappa = 0$, and another entry, for example T_{jl}, at a different level, we obtain $\mathbb{E}[T_{ik}T_{jl}] = (1 - \delta_{ik})\,\mathbb{E}[T_{jl}]$, and so on and so forth.

We next compute these statistical quantities when the connections are not at the level $\kappa = 0$. At a given level $\kappa > 1$, the total number of possible connections (in one direction) is $\alpha_\kappa = 4^{\mu+\kappa-1}$, among which the algorithm has to choose randomly $\beta_\kappa = \lfloor E^{-\kappa}\alpha_\kappa \rfloor$ connections.

At the level κ the probability that T_{ij} is chosen at some time after β_κ steps, regardless the step at which it has been actually chosen, is:

$$p\left(T_{ij}=1\right)=\frac{\beta_\kappa}{\alpha_\kappa}$$

since we can draw uniformly among α_κ possible connections, therefore:

$$\mathbb{E}\left[T_{ij}\right]=0\times p\left(T_{ij}=0\right)+1\times p\left(T_{ij}=1\right)=\frac{\beta_\kappa}{\alpha_\kappa}$$

Now we want to evaluate $\mathbb{E}\left[T_{ij}T_{kl}\right]$. If, in the picture of the connectivity matrix, T_{ij} and T_{kl} are in two different squares, then clearly they are not correlated, therefore in that case we have $\mathbb{E}\left[T_{ij}T_{kl}\right]=\mathbb{E}\left[T_{ij}\right]\mathbb{E}\left[T_{kl}\right]=\frac{\beta_{\kappa_1}}{\alpha_{\kappa_1}}\frac{\beta_{\kappa_2}}{\alpha_{\kappa_2}}$. If instead they are in the same square, we have:

$$\mathbb{E}\left[T_{ij}T_{kl}\right]=\frac{\beta_\kappa\left(\beta_\kappa-1\right)}{\alpha_\kappa\left(\alpha_\kappa-1\right)}$$

since they are selected sequentially and independently from each other. In general, for n entries of the topology in the same square, with $n\leq\beta_\kappa$, we obtain:

$$\mathbb{E}\left[T_{i_0j_0}T_{i_1j_1}...T_{i_{n-1}j_{n-1}}\right]=\frac{\beta_\kappa\left(\beta_\kappa-1\right)...\left(\beta_\kappa-n+1\right)}{\alpha_\kappa\left(\alpha_\kappa-1\right)...\left(\alpha_\kappa-n+1\right)}=\frac{\beta_\kappa!\left(\alpha_\kappa-n\right)!}{\alpha_\kappa!\left(\beta_\kappa-n\right)!}$$

thereby the problem of determining the correlation structure of the neural network with the fractal connectivity matrix is solved.

6.5 Numerical experiments

As we did in Chapter 5, we want to show that this perturbative expansion provides a good match with the exact equations of the network. For this reason in Figures 6.4 and 6.5 we have shown the comparison between the membrane potential, variance, covariance and

correlation of pairs of neurons for two kinds of connectivity matrices (fully connected and cycle graphs), obtained from the simulation of equations 6.1 (blue line), of equations 6.13 - 6.17 (red line) and from formulae 6.37 - 6.43, 6.53 and 6.54 (green line). Therefore, as in Chapter 5, we have obtained these figures without considering the second order terms in the perturbative expansion of $V_i(t)$. In other words, we have omitted the third order terms 6.44 - 6.47 in the variance and covariance, due to the difficulty of implementing them numerically. Instead in Figures 6.6 and 6.7 we have shown the comparison between equations 6.1 (blue line) and equations 6.13 - 6.21 (red line), therefore considering also the higher order terms, because the numerical calculation of the variance and covariance through the simulation of equations 6.18 - 6.21 is much easier than the implementation of the terms 6.44 - 6.47.

For the networks with random topology, the analytic formulae of the variance, covariance and correlation are rather complex to implement. In fact usually the approximation of order 0 of the quantities $\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]$, $\mathbb{E}\left[\frac{T_{ik}T_{jl}}{M_iM_j}\right]$ and $\mathbb{E}\left[\frac{T_{ik}T_{lm}T_{jl}}{M_iM_jM_l}\right]$ is not precise enough, forcing us to add the higher order corrections. For example, for a network with independent random connections with $p(T_{ij}=1)=p\ \forall i,j:\ i\neq j$, the approximation of order 0 of $\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]$ is:

$$\mathbb{E}\left[\frac{T_{ij}}{M_i}\right] \approx \frac{\mathbb{E}[T_{ij}]}{\sum_{k=0}^{N-1}\mathbb{E}[T_{ik}]} = \frac{p}{(N-1)p} = \frac{1}{N-1}$$

which does not depend on p and therefore does not contain information about the randomness of the topology. This means that in general this approximation is a too poor description of the random topology, and therefore the higher order corrections must be included. Unfortunately, according to 6.50, the approximations of order 1 are always equal to zero, therefore we have to extend the approximation up to the 2nd order. In other terms, we have to compute the second order derivatives in the multidimensional Taylor expansions of $\mathbb{E}\left[\frac{T_{ij}}{M_i}\right]$, $\mathbb{E}\left[\frac{T_{ik}T_{jl}}{M_iM_j}\right]$ and $\mathbb{E}\left[\frac{T_{ik}T_{lm}T_{jl}}{M_iM_jM_l}\right]$. This is a feasible but complex task, and it is particularly hard for the fractal connectivity matrix, since it depends on the blocks the synaptic

connections belong to. For this reason we have opted for showing only the comparison between the numerical simulations of the stochastic differential equations (red and blue lines), without using the analytic formulae. Figures 6.8 - 6.14 show these results for a network with independent random connections and for the Sporns' fractal matrix. The differential equations have been solved numerically using the Euler-Maruyama scheme, while the integrals with respect to time have been calculated using the trapezoidal rule, in both cases with an integration time step $\Delta t = 0.1$. All the statistics have been evaluated with $10,000$ Monte Carlo simulations (where we have independently generated repetitions of the four sources of randomness of the system), while the remaining parameters are reported in Table 6.2. The covariance and correlation have always been calculated between the 0th and the 1st neuron. The only exceptions are in Figures 6.12, 6.13 and 6.14, where the comparison is between the 0th and the 8th neuron. Instead the membrane potentials and the variances have always been reported only for the the 0th neuron. In general we have obtained a better agreement with the exact equations when we use also the second order corrections of the membrane potential.

It is important to observe that a detailed analysis of the error introduced by the perturbative expansion as a function of the approximation order, the values of all the parameters of the system and the infinitely many connectivity matrices is missing and is beyond the purpose of this thesis.

6.6 Fisher information

We consider again only the first order terms in the perturbative expansion of the membrane potentials, because this allows us to use formulae 5.77, 5.78 and 5.79 for the analytic evaluation of the Fisher information. In general it is not easy to calculate the inverse $\Sigma^{-1}(t)$ of the covariance matrix. This can be done straightforwardly only in the case when $\sigma_3 = 0$, because otherwise the term 6.39 determines a complicated structure of $\Sigma(t)$, which

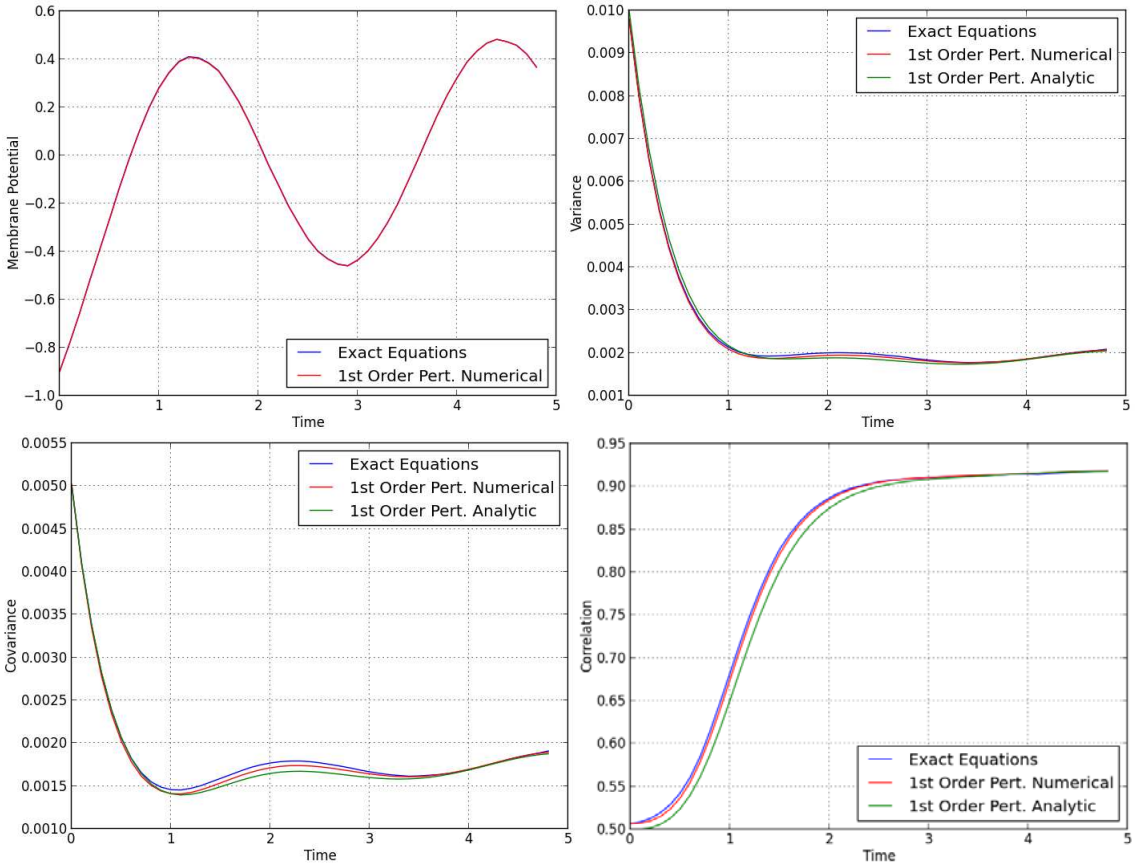

Figure 6.4: Comparison of the variance, covariance and correlation obtained from the simulation of equations 6.1 (blue line), of equations 6.13 - 6.17 (red line) and from formulae 6.37 - 6.43, 6.53 and 6.54 (green line). Therefore the perturbative expansion of the membrane potential has been truncated at the first order, while those of the variance and covariance at the second order. We have compared the 0th and the 1st neuron, using the Euler-Maruyama scheme (blue and red lines) and the trapezoidal rule (green line) with integration time step $\Delta t = 0.1$. The statistics have been evaluated with $10,000$ Monte Carlo simulations, for the values of the parameters reported in Table 6.2. The topology is K_{10} (see Chapter 3) and therefore deterministic.

Neuron	Input
$\tau = 1$	$I_i(t) = \begin{cases} sin(2t), & i = 0 \div \frac{N}{2} - 1 \\ 0.2 + e^{-t}, & i = \frac{N}{2} \div N - 1 \end{cases}$
$C_2 = 0.5$	$C_1 = 0.4$
$\sigma_2 = 0.1$	$\sigma_1 = 0.01$
$\mu_i = \begin{cases} -1, & i = 0 \div \frac{N}{2} - 1 \\ 0.5, & i = \frac{N}{2} \div N - 1 \end{cases}$	

Synaptic Weights	Sigmoid Function
$\widehat{J}_{ij}(t) = \begin{cases} (1+t^2)^{-1}, & i,j = 0 \div \frac{N}{2} - 1 \\ \cos(t), & i = 0 \div \frac{N}{2} - 1, \quad j = \frac{N}{2} \div N - 1 \\ \sqrt{1 + \frac{2}{\pi}\arctan(t)}, & i = \frac{N}{2} \div N - 1, \quad j = 0 \div \frac{N}{2} - 1 \\ e^{-t}\sin(t), & i,j = \frac{N}{2} \div N - 1 \end{cases}$	$T_{MAX} = 1$
$C_3 = 0.6$	$\lambda = 1$
$\sigma_3 = 0.1$	$V_T = 0$

Table 6.2: Parameters used to generate Figures 6.4 - 6.14

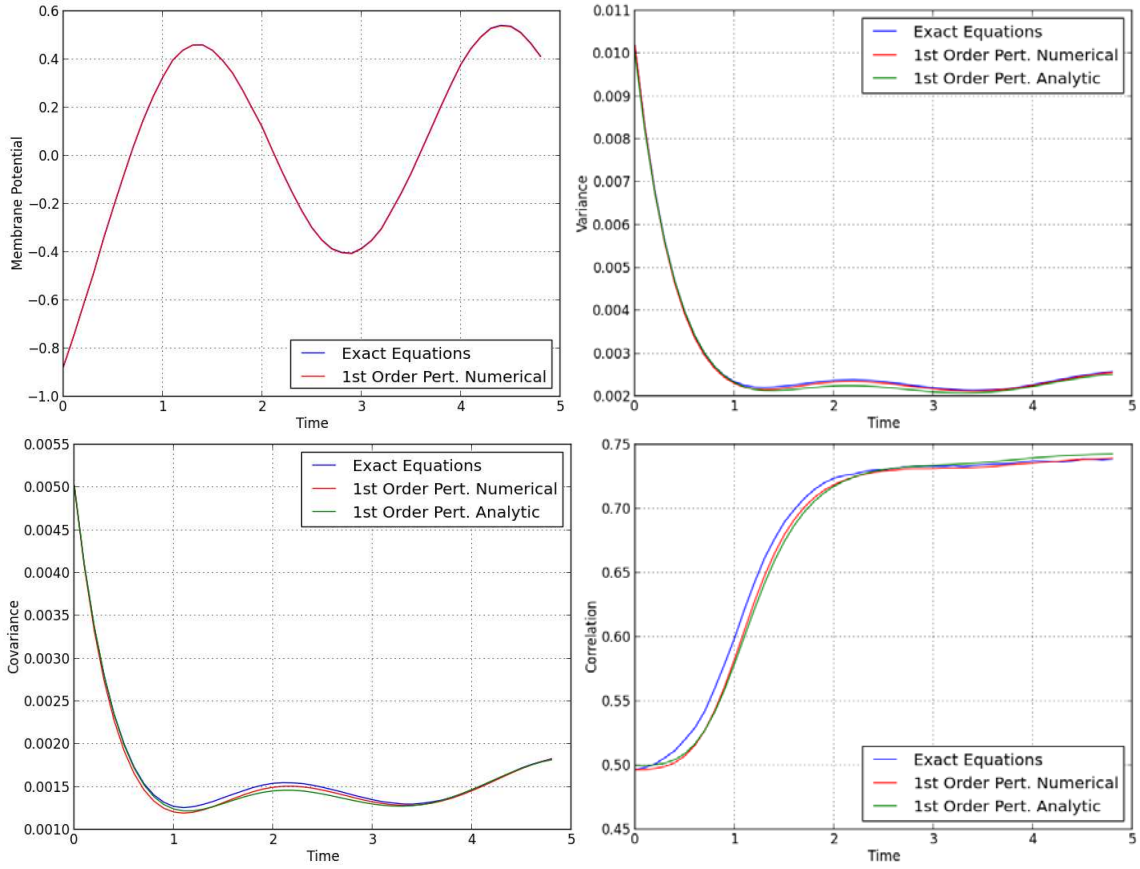

Figure 6.5: Comparison of the variance, covariance and correlation obtained for the deterministic topology Cy_{10} (see Chapter 3), for the values of the parameters reported in Table 6.2.

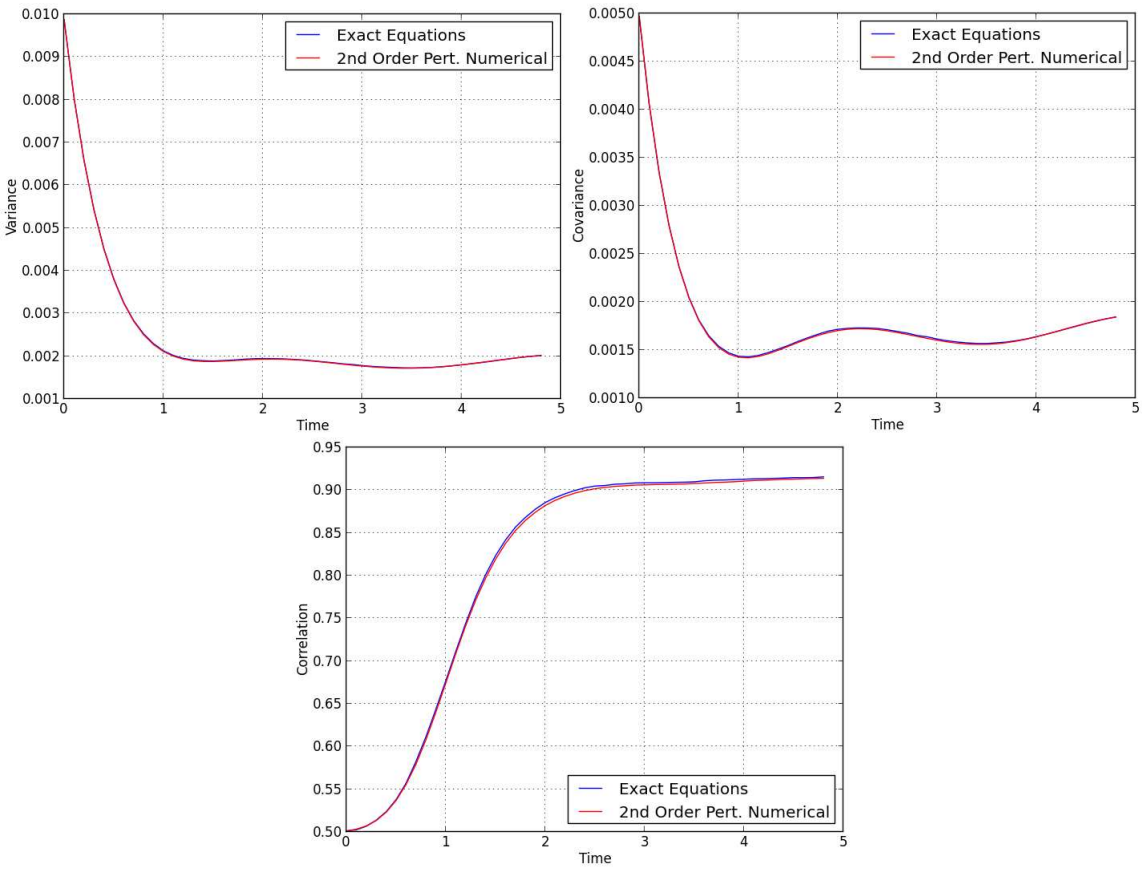

Figure 6.6: Comparison of the variance, covariance and correlation obtained for the deterministic topology K_{10}, for the values of the parameters reported in Table 6.2, but considering also the second order corrections of the membrane potential. Clearly the match has been improved by the addition of these terms, as the reader can easily check from the comparison with Figure 6.4.

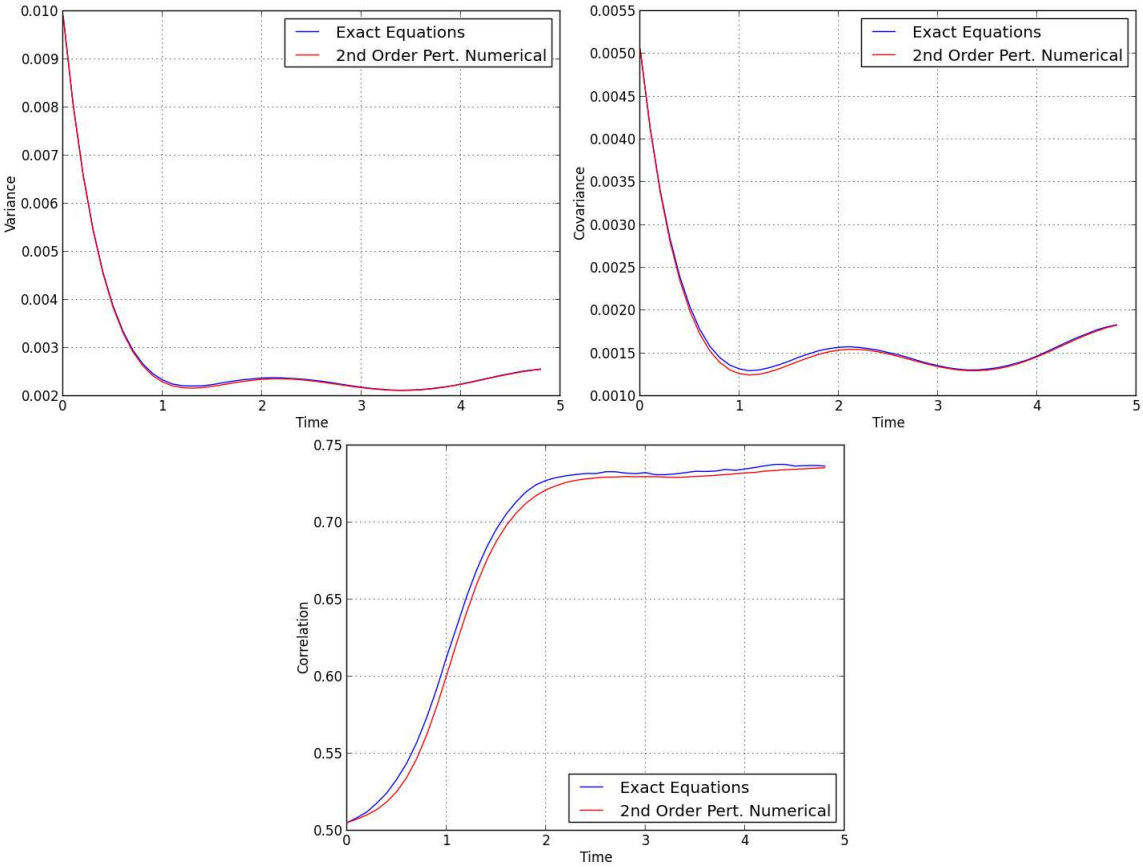

Figure 6.7: Comparison of the variance, covariance and correlation obtained for the deterministic topology Cy_{10}, for the values of the parameters reported in Table 6.2, but considering also the second order corrections of the membrane potential. This time the improvement of the match is not evident, if compared with Figure 6.5, which proves that the goodness of the perturbative expansion depends also on the topology of the network. It is important to observe that the second order corrections are generally small, therefore their magnitude could be of the same order of the numerical error introduced by the finite number of Monte Carlo simulations.

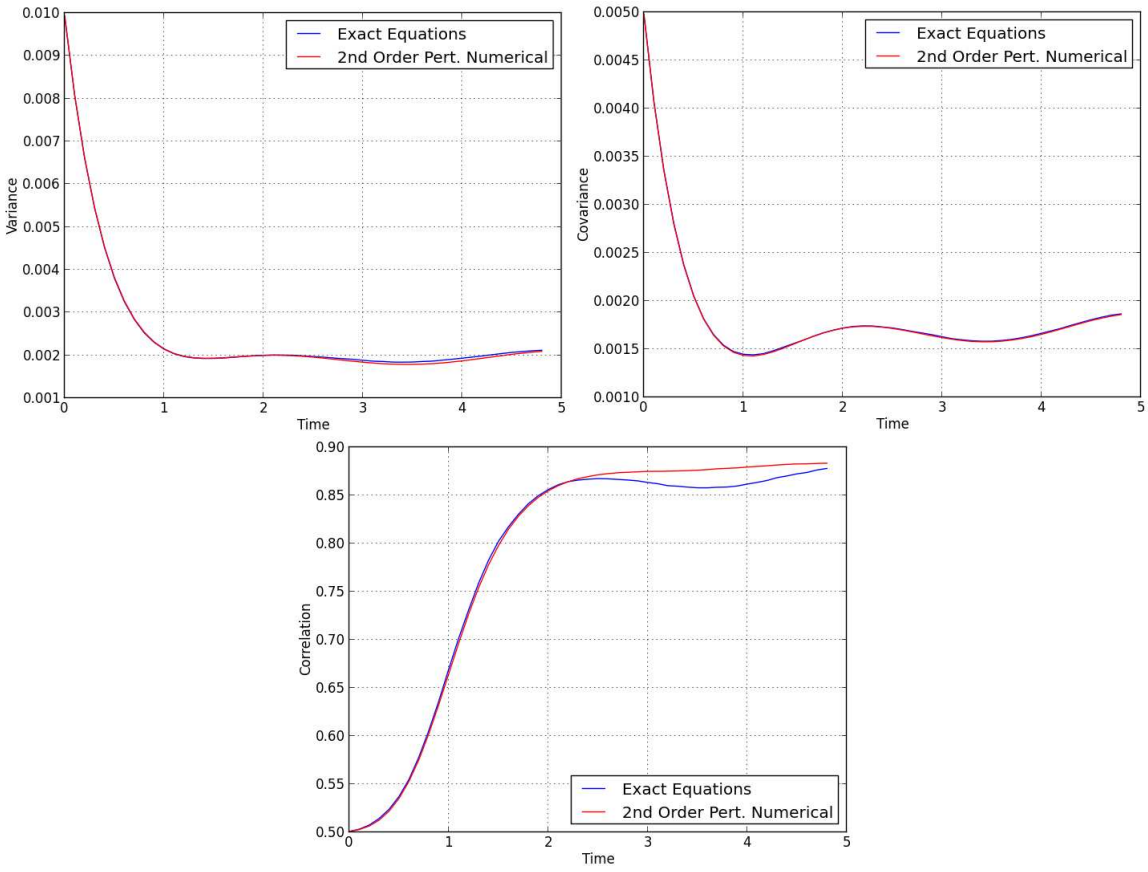

Figure 6.8: Comparison of the variance, covariance and correlation obtained for a random topology, for the values of the parameters reported in Table 6.2, considering also the second order corrections of the membrane potential. In detail, here we have assumed that each pair of neurons is connected independently from the others and with probability $p = 0.7$. Even if the match of the variance and covariance is quantitatively very good, the approximation of the correlation is not satisfying for $t > 2$. This is due to the fact that the ratio of small quantities (in this case the variance and covariance) is very sensitive to small errors in the numerator and denominator. Nevertheless the second order expansion provides a satisfying result, because the variance and covariance are in very good agreement with the exact neural equations. It is important to observe that the discrepancy is also due to the finite number of Monte Carlo simulations, which should be increased especially for small values of the variance and covariance.

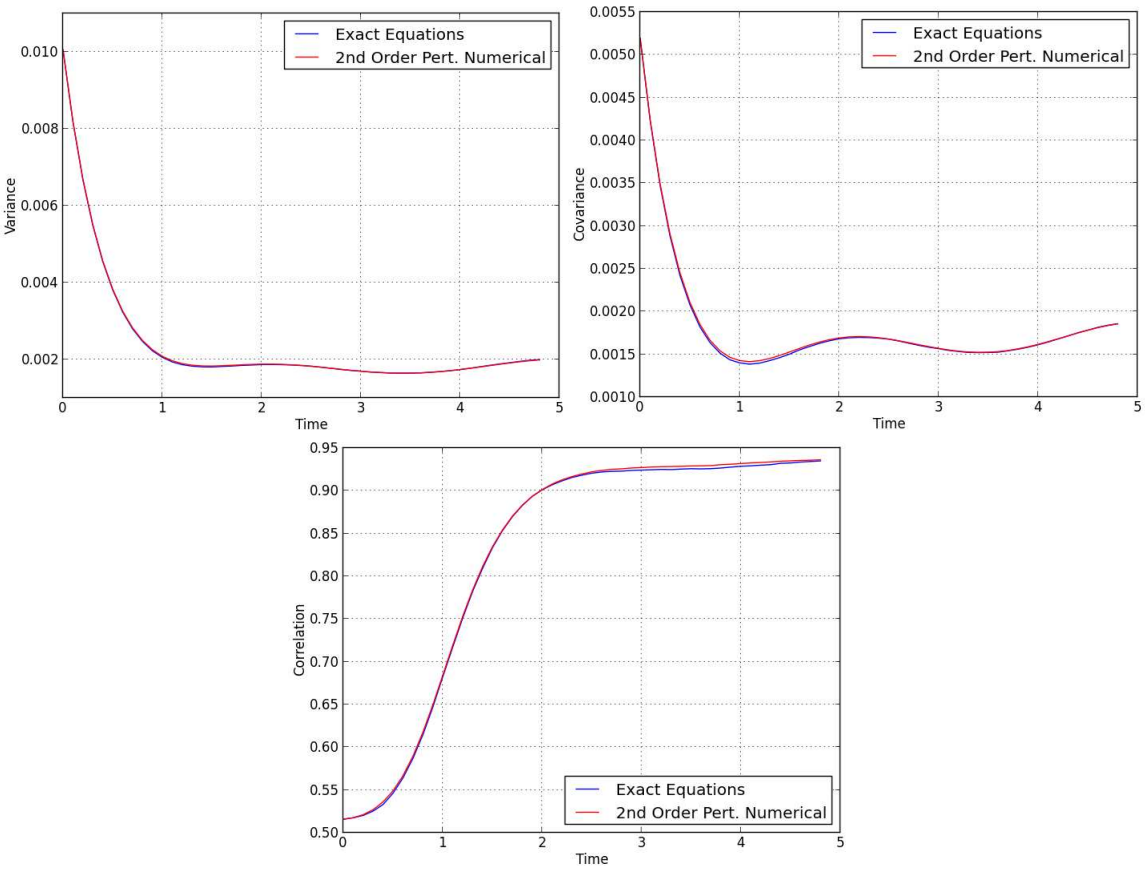

Figure 6.9: Comparison of the variance, covariance and correlation obtained for the Sporns' topology, for the values of the parameters reported in Table 6.2, considering also the second order corrections of the membrane potential. In this example we have set $\eta = 4$ ($N = 16$), $\mu = 2$ and $E = 1.1$, therefore the network is almost fully connected. The two neurons are in the same block, therefore they are connected at the level $\kappa = 0$.

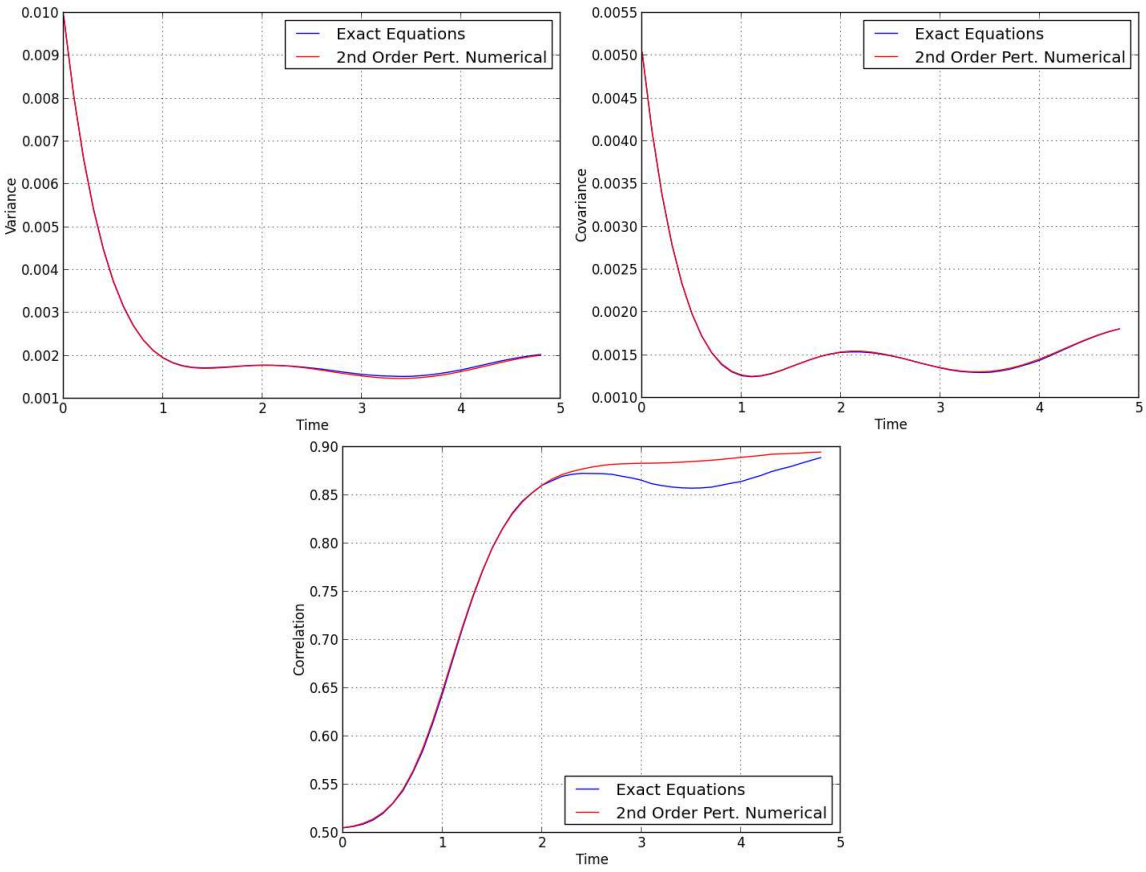

Figure 6.10: As in the Figure 6.9, but with $E = 2$. This, according to [190], is approximately the point of maximum complexity of the network, see text.

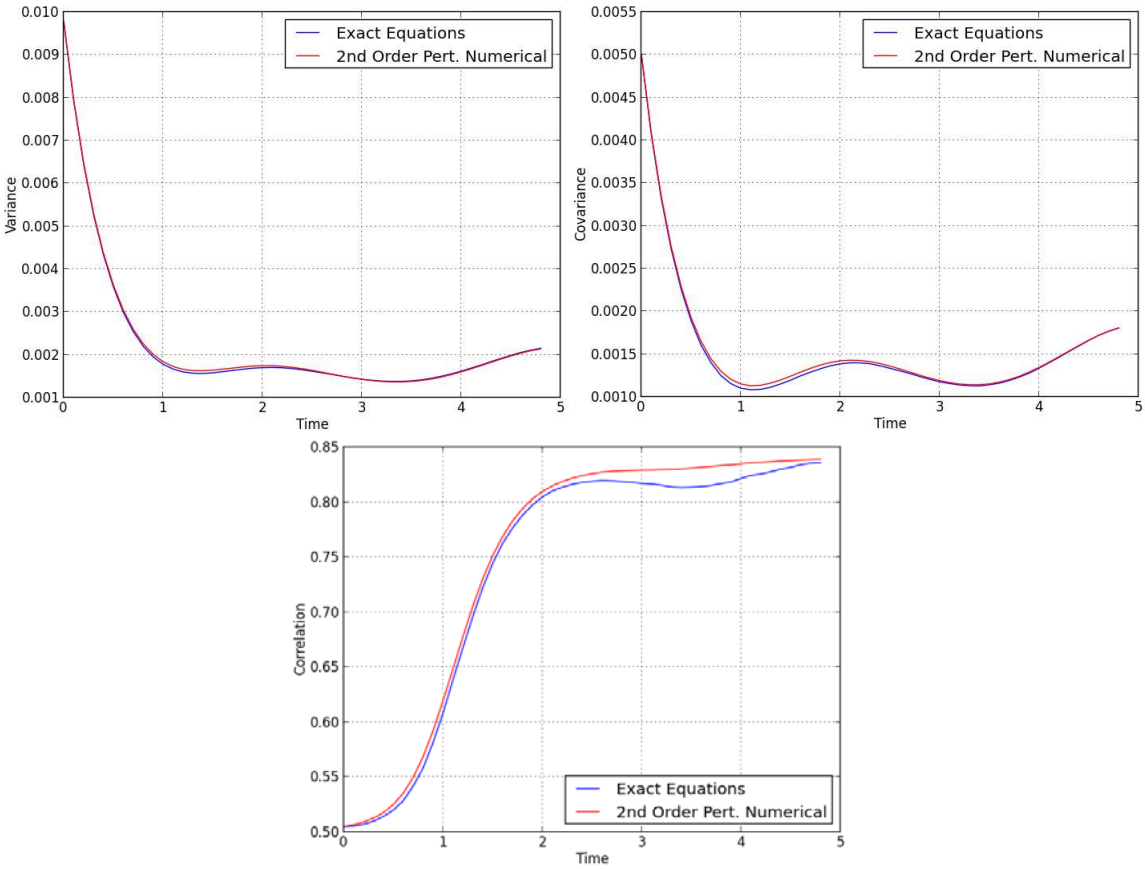

Figure 6.11: As in the Figure 6.9, but with $E = 5$. In this case the blocks are almost completely disconnected. From the comparison with Figures 6.9 and 6.10, the reader can easily check that the increase of the parameter E determines the reduction of the correlation at large t, as a consequence of the diminution of the number of connections.

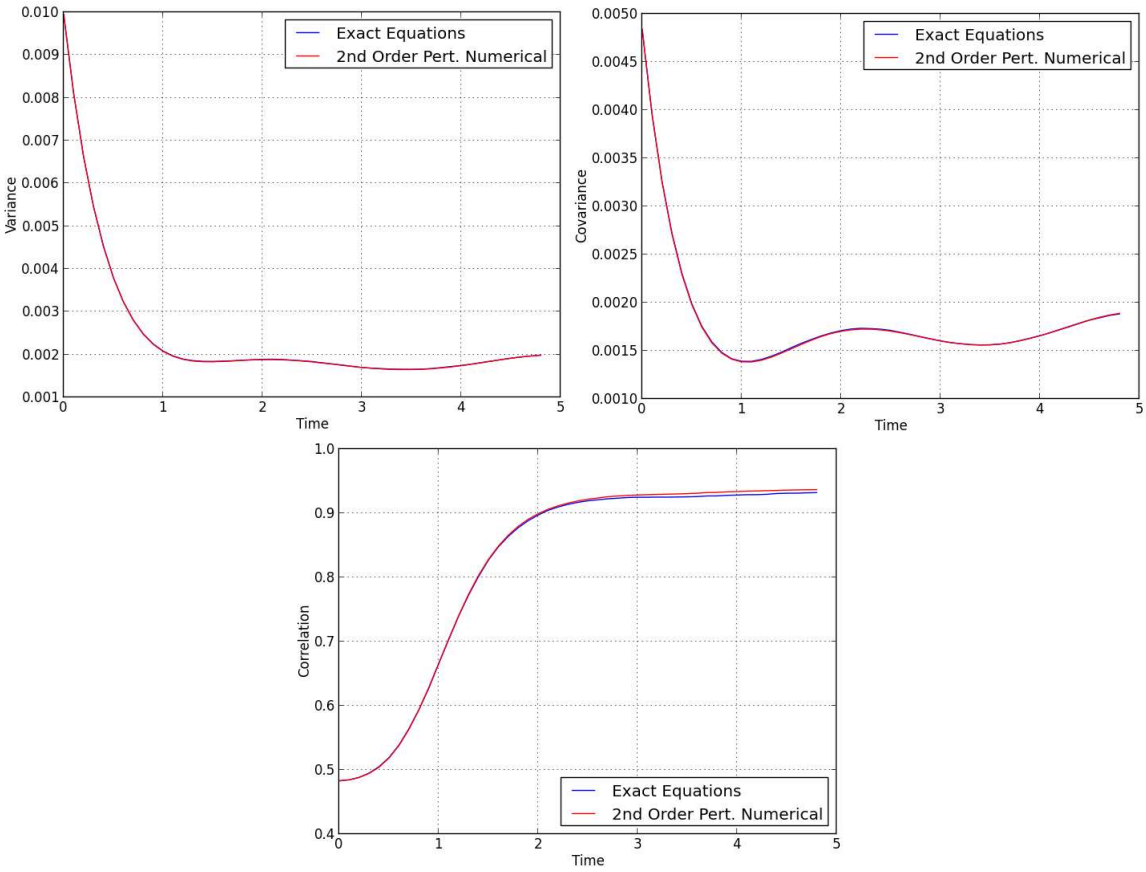

Figure 6.12: Comparison of the variance, covariance and correlation obtained for the Sporns' topology, for the values of the parameters reported in Table 6.2, considering also the second order corrections of the membrane potential. In this example we have set $\eta = 4$ ($N = 16$), $\mu = 2$ and $E = 1.1$, as in Figure 6.9, but now the neurons are in two different blocks, and they are connected at the level $\kappa = 2$.

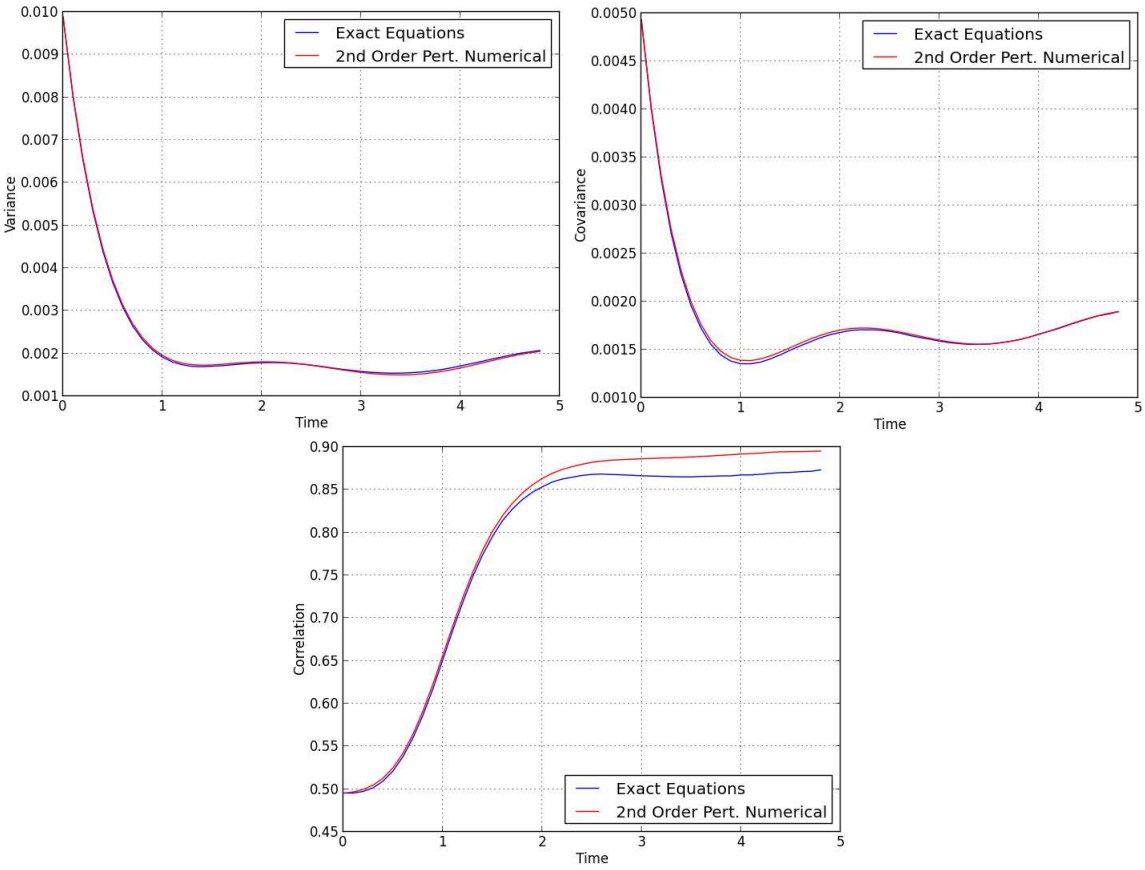

Figure 6.13: As in the Figure 6.12, but with $E = 2$.

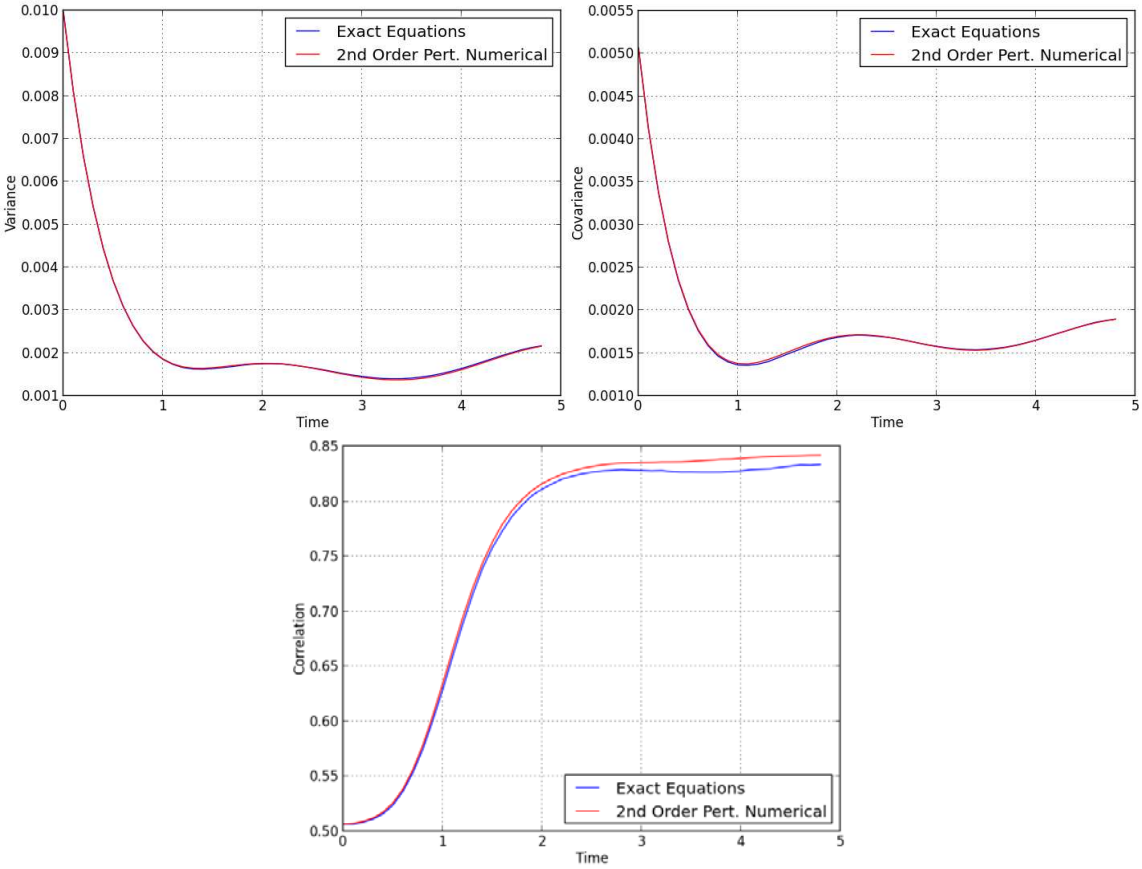

Figure 6.14: As in the Figure 6.12, but with $E = 5$. Again, the increase of the parameter E determines the reduction of the correlation for large t. It is important to observe that the difference between the two cases with the neurons in the same block or in two different blocks is very small. This is due to the fact that the values of the parameters C_1, C_2 and C_3 are relatively high (see Table 6.2), therefore they strongly determine the behavior of the correlation, for every topology. When these parameters are set to zero, a richer behavior of the correlation emerges. This analysis is not shown in the thesis, because the purpose of this work is to develop mathematical tools that allow us to understand a neural network, not the analysis of the consequences of the formulae.

in general cannot be inverted easily. Moreover, for simplicity, we consider again the case when the synaptic weights are constant in time and the initial conditions are deterministic as well (i.e. $\sigma_2 = 0$). In order to obtain concise formulae, we consider only the stationary case, namely the limit for $t \to +\infty$. Therefore we have:

$$m_i = \mathbb{E}\left[V_i\left(t\right)\right] = \overline{I}\tau + \frac{1}{\tau} S\left(\overline{I}\tau\right) \sum_{j=0}^{N-1} J_{ij} \tag{6.55}$$

$$\Sigma_{ij} = \begin{cases} \sigma_1^2 \frac{\tau}{2} & \text{if} \quad i = j \\ \\ \sigma_1^2 \frac{\tau C_1}{2} & \text{if} \quad i \neq j \end{cases} \tag{6.56}$$

$$\left[\Sigma^{-1}\right]_{ij} = \begin{cases} \frac{2}{\sigma_1^2 \tau} \frac{1 - C_1 + C_1(N-1)}{(1-C_1)[1+C_1(N-1)]} & \text{if} \quad i = j \\ \\ -\frac{2}{\sigma_1^2 \tau} \frac{C_1}{(1-C_1)[1+C_1(N-1)]} & \text{if} \quad i \neq j \end{cases} \tag{6.57}$$

having used formulae 6.12, 6.31 and 6.37. This is formally equivalent to the additive noise case studied in [158] (the difference consists in the fact that Abbott and Dayan performed the analysis in terms of the firing rates, while here we are using the membrane potentials), with the advantage that now we have an expression for the Fisher information as a function of the dynamics of the network (when t is finite), of its parameters and of its connectivity matrix, which are not taken into account in [158]. According to Abbott and Dayan, the first term of the Fisher information is:

$$\mathcal{I}_1\left(\overline{I}\right) = \frac{\partial \overrightarrow{m}^T}{\partial \overline{I}} \Sigma^{-1} \frac{\partial \overrightarrow{m}}{\partial \overline{I}} = \frac{2}{\sigma_1^2 \tau} \left[\frac{C_1 N^2 \left[F_1\left(\overline{I}\right) - F_2\left(\overline{I}\right)\right]}{\left(1 - C_1\right)\left[1 + C_1\left(N-1\right)\right]} + \frac{N F_1\left(\overline{I}\right)}{1 + C_1\left(N-1\right)} \right] \tag{6.58}$$

where:

$$F_1\left(\overline{I}\right) = \frac{1}{N}\sum_{i=0}^{N-1}\left(\tau + a_i\right)^2$$

$$F_2\left(\overline{I}\right) = \left[\frac{1}{N}\sum_{i=0}^{N-1}\left(\tau + a_i\right)\right]^2$$

and:

$$a_i = S'\left(\overline{I}\tau\right)\sum_{j=0}^{N-1} J_{ij}$$

Instead the second term of the Fisher information is:

$$\mathcal{I}_2\left(\overline{I}\right) = \frac{1}{2}Tr\left(\frac{\partial\Sigma}{\partial\overline{I}}\Sigma^{-1}\frac{\partial\Sigma}{\partial\overline{I}}\Sigma^{-1}\right) = 0 \tag{6.59}$$

since Σ does not depend on $\overline{I}$. This is compatible with the result found in Chapter 5, namely that the trace term of the Fisher information is negligible. So formula 6.58 explains the origin of the two peaks of the Fisher information we found in Section 5.7. The right-hand side of 6.58 has two sub-terms, the first proportional to $F_1\left(\overline{I}\right) - F_2\left(\overline{I}\right)$ and the second proportional to $F_1\left(\overline{I}\right)$. It is easy to see that when $F_1\left(\overline{I}\right) - F_2\left(\overline{I}\right) \approx 0$ the second term dominates for small values of the correlation, and actually it increases for $C_1 \to 0$. However also a small difference between $F_1\left(\overline{I}\right)$ and $F_2\left(\overline{I}\right)$ generates an explosion of the Fisher information for $C_1 \to 1$, because the denominator goes to zero but not the numerator, so this explains the formation of the two peaks obtained in Section 5.7. Some

Neuron	Input	Synaptic Weights	Sigmoid Function
$\tau = 1$	$\overline{I} = 0.5$	$\alpha = 0.1$	$T_{MAX} = 1$
	$\sigma_1 = 0.01$	$\beta = [0, 0.2]$	$\lambda = 1$
			$V_T = 0$

Table 6.3: Parameters used for the numerical simulations of Figure 6.15.

numerical examples are shown in Figure 6.15, where in a single trial we have used again independent and identically distributed synaptic weights given by $J_{ij} \sim \mathcal{N}\left(\frac{\alpha}{N-1}, \left(\frac{\beta}{N-1}\right)^2\right) \forall\,(i,j) : i \neq j$. The values of the parameters are shown in Table 6.3.

However it must be noted that formula 6.58 gives a fairly poor approximation of the Fisher information, since it has been obtained from formula 6.56, which, we know, is not precise. In fact, if we set $C_1 = 0$, formula 6.56 predicts independent neurons, which is not true for synaptic weights of generic strength. Nevertheless, this approximation of the Fisher information predicts the two peaks we have found in Section 5.7, so it can be used to predict qualitative results. Moreover it clearly shows that the formation of the two peaks does not depend on the topology of the synaptic connections, since formula (6.58) is true for a generic connectivity matrix J.

6.7 Partial conclusion

In this chapter we have introduced another perturbative expansion, that somehow extends the results of Chapter 5. While in Chapter 5 we expanded the membrane potentials around the stationary solutions of the network, which was supposed to be invariant under the exchange of the neural indices, here the expansion is performed with respect to the network without connections. In this way the correlation structure of the system can be calculated for arbitrary (and in general time-varying) states and for any connectivity matrix. The disadvantage, compared to the technique of Chapter 5, is in the fact that this new

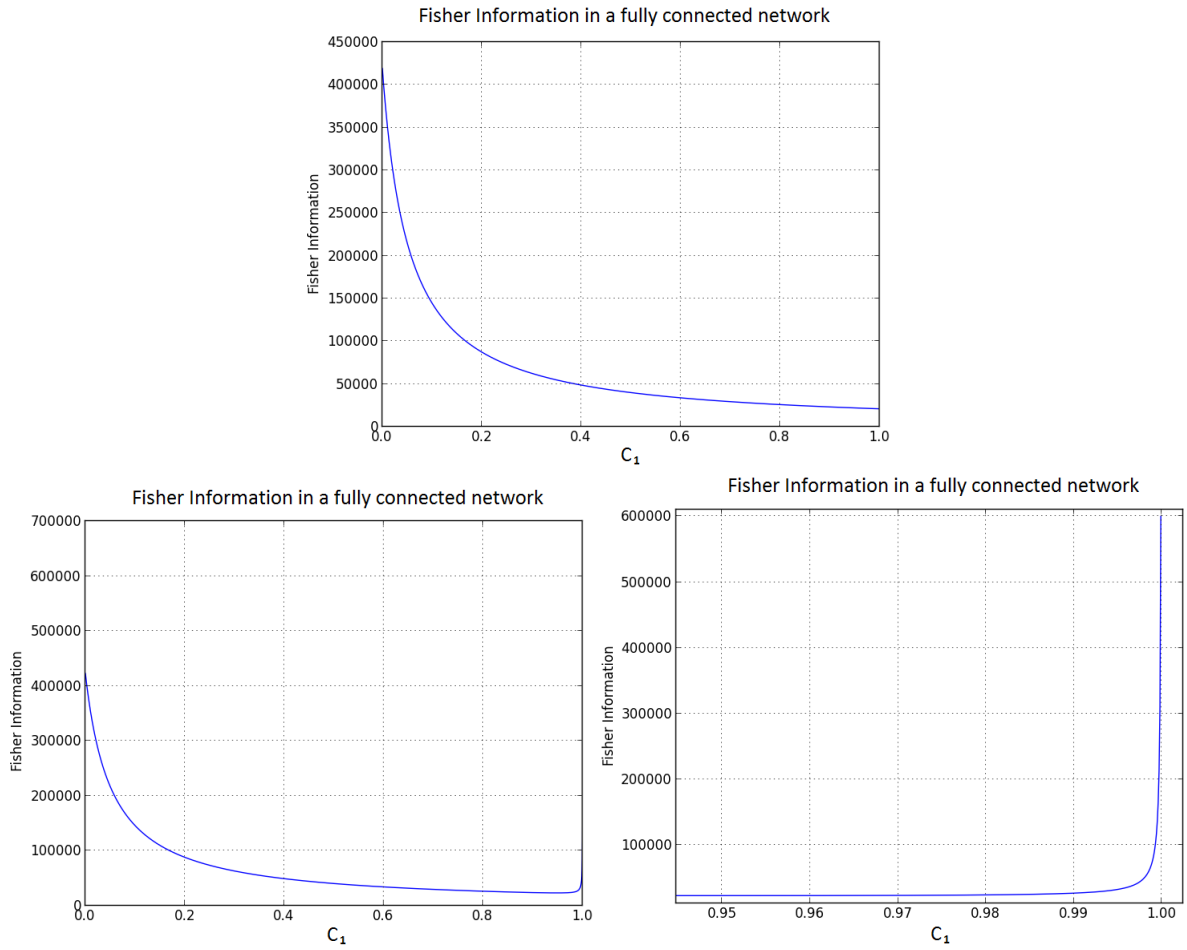

Figure 6.15: Two possible examples of Fisher information, obtained from the perturbative expansion with weak synaptic weights in a fully connected network with $N = 20$ and for the parameters reported in Table 6.3. The result depends on the values of the synaptic weights from trial to trial, showing higher encoding efficiency for $C_1 \to 0$ and/or $C_1 \to 1$, as in Section 5.7. In detail, the figure on the top has been obtained for $\beta = 0$, $F_1 = 1.0475$ and $F_1 - F_2 = 0$, while the two figures at the bottom have been obtained for $\beta = 0.2$, $F_1 = 1.0566$ and $F_1 - F_2 = 1.447 \times 10^{-4}$. The figure on the bottom-right side represents the zoom of the figure on the bottom-left side for $C_1 \to 1$, from which it is possible to see the explosion of the Fisher information for highly correlated membrane potentials.

expansion can be used only for relatively weak synaptic weights. However now the connections can have a random topology, so we have used the new method to study biologically realistic cases. An important example, characterized by biological relevance and mathematical tractability, is the connectivity matrix introduced by Sporns [190]. This is a first attempt to describe the main features of the human connectome, namely its small-world properties and the nested structure (see Chapters 1 and 2). Even if we have completely calculated the correlation matrix generated by this connectome, obtaining a good match with the numerical simulation of the exact neural equations, we have not analyzed in detail its implications for the network. Being very complex, this analysis is reserved for future work.

To conclude, we have used the perturbative expansion to calculate the Fisher information of the network, obtaining an approximate analytic expression that predicts two peaks of $\mathcal{I}\left(\overline{I}\right)$ for highly correlated and highly decorrelated neurons with inhomogeneous synaptic weights. This confirms the result found in Chapter 5, and clearly shows that the formation of the two peaks is a generic phenomenon, that does not depend on the topology of the connections.

Chapter 7

Numerical calculation of the Fisher information

THIS chapter introduces an algorithm to calculate the Fisher information of a neural network, which is not based on any analytic approximation of the probability density. Therefore this technique is purely numerical, and in principle can be applied to every kind of neural model. The method is based on the Expectation-Maximization algorithm with Gaussian mixtures, and on the Monte Carlo integration with importance sampling, as shown in Section 7.1. In Section 7.2 we show the results provided by the algorithm for the case of the FitzHugh-Nagumo network with correlated Brownian motions. In particular, we confirm the fact that the Fisher information is higher when the membrane potentials are highly correlated than when they are independent.

7.1 Description of the algorithm

In the context of *Statistical Estimation Theory*, many efficient algorithms have been invented for the numerical evaluation of the Fisher information using a Monte Carlo ap-

proach (see for example [192][193][194][195][196]). Nevertheless, for the sake of simplicity and clarity, in this chapter we only introduce an elementary algorithm, that allows us to confirm qualitatively the results found in Chapters 5 and 6 for a different kind of neural model. In other words, this algorithm is particularly useful when the probability density of the system is strongly non-Gaussian, therefore we apply it to the case of the FitzHugh-Nagumo network. In fact, the techniques for the calculation of the Fisher information introduced in Chapters 5 and 6 can be applied only to the rate model, since its probability density is approximately Gaussian. Instead for a spiking network the probability density in general is highly non-Gaussian, therefore a new technique must be used. We remind the reader that, according to 2.6, the Fisher information is defined as:

$$\mathcal{I}(\theta,t)=\int_{\mathbb{R}^{kN}}\left(\frac{\partial}{\partial\theta}\log p\left(\overrightarrow{x},\theta,t\right)\right)^{2}p\left(\overrightarrow{x},\theta,t\right)d\overrightarrow{x} \tag{7.1}$$

Therefore we need to evaluate the joint probability density of the solutions of the equations describing the network, and this can be achieved by repeating the simulation of the network many times. In other words, we could take a finite portion of the phase space, divide it in many bins (namely construct a grid) and build a histogram that approximates the probability density using a Monte Carlo method. However, unfortunately, this can be done only in principle. In fact, suppose for example that we would like to build the histogram of a network made up of 100 neurons, and suppose also that every dimension in the phase space is divided in 10 bins. In this case, the number of subdivisions of the whole phase space is 10^{100}. If now we would like to increase the precision of the grid, or to increase the size of the network, this number becomes quickly prohibitive in terms of memory consumption. Therefore we need to find a more efficient way for evaluating the probability density of the network. As we said, the basic idea presented in this chapter is to use a simplified version of the techniques developed in the context of Statistical Estimation Theory. In particular we will use an *Expectation-Maximization* (EM) algorithm [197] with Gaussian mixtures in order to evaluate the joint probability density of the network.

Supposing that the system is made up of N neurons and that it is has been simulated a certain number of times (that we call MCS, which stands for *Monte Carlo Simulation*) with a Monte Carlo method, then we have generated a collection of $N \times MCS$ samples. Using these data, the EM algorithm can estimate the joint probability density of the system as a mixture (namely a sum) of Gaussian multivariate distributions. This result is achieved without the subdivision of the phase space into bins, therefore it is not affected by the problem of the previous method. Therefore using this technique we can study also large neural networks without memory problems. The use of Gaussian mixtures is particularly useful for systems whose probability density is strongly non-Gaussian. In particular, from Section 3.2.2 we know that the probability density of the FitzHugh-Nagumo network deviates considerably from the Gaussian distribution when the neurons are spiking. Therefore it is an optimal candidate for testing this approach. More in detail, we suppose that the probability density of the network could be approximated as a weighted sum of multivariate Gaussian distributions. The weights of the sum, the mean vectors and the covariance matrices of the Gaussian distributions are not known and must be evaluated through the samples that come from the Monte Carlo simulations of the network. Moreover, every single Gaussian distribution has hyper-volume 1, therefore the sum of the weights must be equal to 1, for the normalization condition of the probability densities. In this section we suppose that the Brownian motions $B_i^V(t)$, $B_i^J(t)$ and $B_i^y(t)$ have all correlation C, and Figures 7.1 and 7.2 show two examples of this approximation, respectively when most of the neurons are spiking or are at rest. In these results the EM algorithm has been used with a mixture of 6 Gaussian distributions, a percentage of the log likelihood difference between 2 iterations equal to 10^{-3} and with a maximum of 200 iterations allowed.

Now, the joint probability density has to be used for the evaluation of the Fisher information. However, this distribution is useful also for another reason. According to formula 7.1, the Fisher information is obtained after the calculation of a high-dimensional integral. In the case of the FitzHugh-Nagumo network this integral has $3N$ dimensions, and when N is large it cannot be calculated with the standard numerical methods. However it is

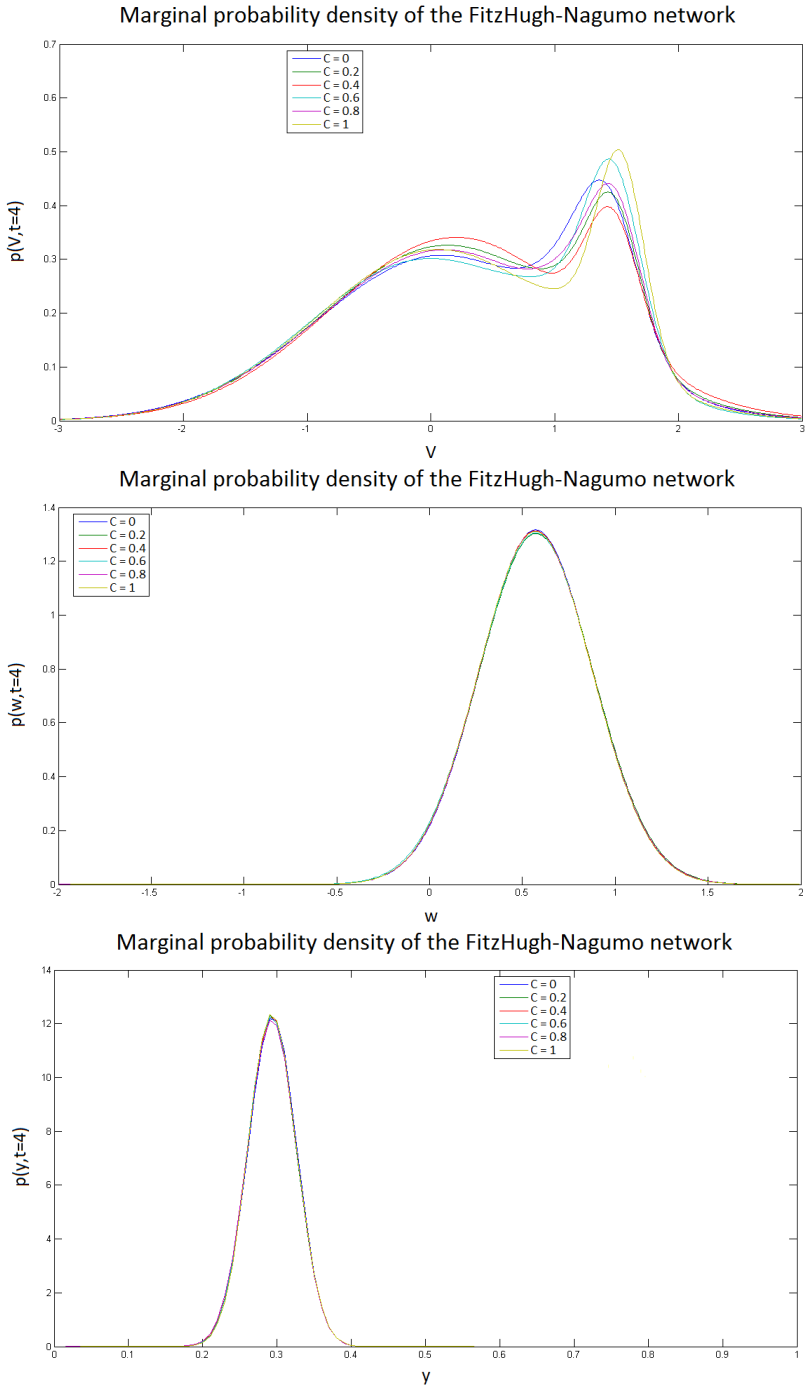

Figure 7.1: Marginal probability densities $p(V,t)$, $p(w,t)$ and $p(y,t)$ (respectively top, middle and bottom of the figure) of the FitzHugh-Nagumo network, according to the EM algorithm. These results have been obtained for $N = 5$, $t = 4$ and $10,000$ Monte Carlo simulations of the network. Moreover the values of the parameters are those of Table 7.1, while the EM algorithm has been used with a mixture of 6 Gaussian distributions. The formation of the two peaks of the probability density explained in Figures 3.9, 3.10 and 3.11 is particularly evident from the behavior of $p(V,t)$, and represents the spiking activity of the neurons on the limit cycle.

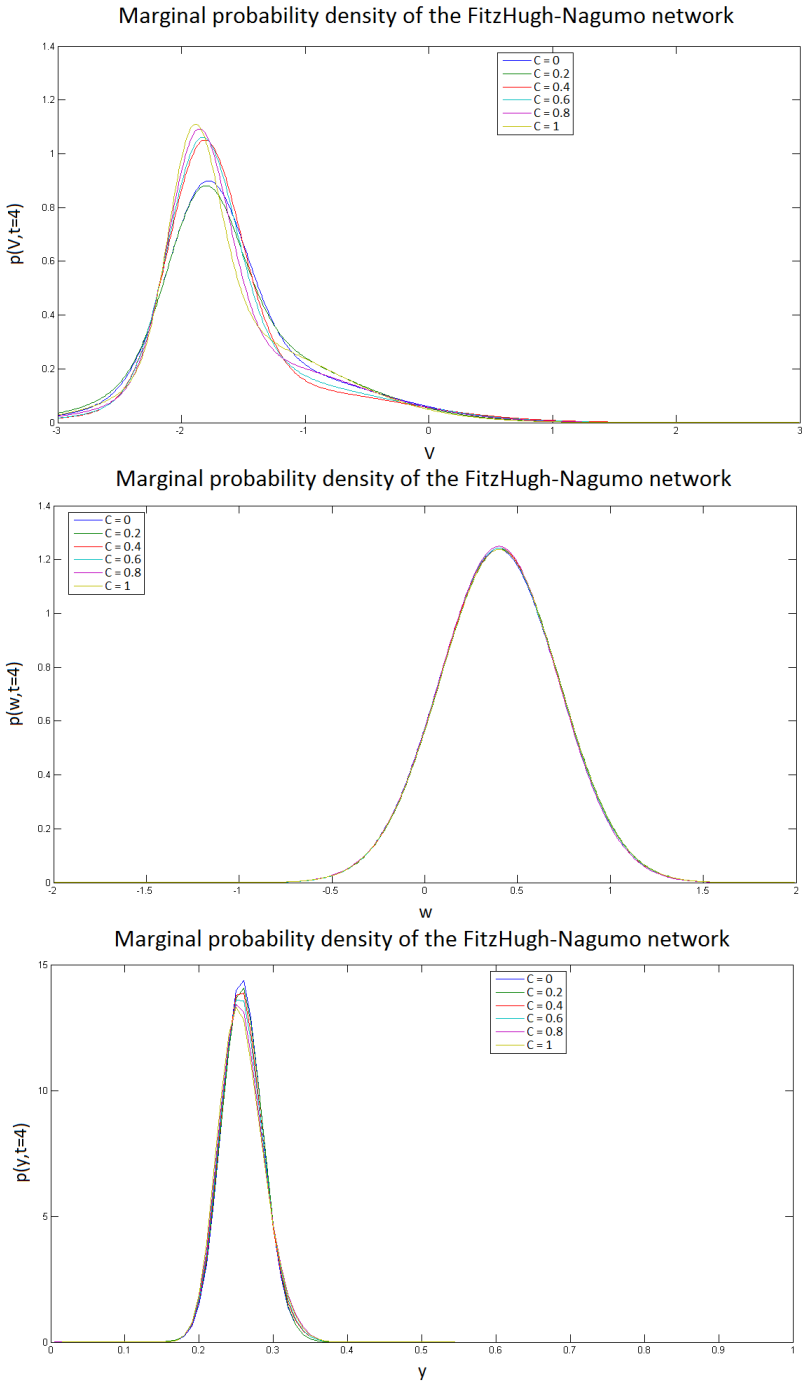

Figure 7.2: Marginal probability densities $p\left(V,t\right)$, $p\left(w,t\right)$ and $p\left(y,t\right)$ (respectively top, middle and bottom of the figure) of the FitzHugh-Nagumo network, according to the EM algorithm. These results have been obtained for $N=10$, $t=4$ and $10,000$ Monte Carlo simulations of the network. Moreover the values of the parameters are those of Table 7.1, with only the exception of the external input current and the background noise, which have been set to $I=-0.8$ and $\sigma_1=0.45$, while the EM algorithm has been used with a mixture of 6 Gaussian distributions. The single peak of the probability density is explained in Figure 3.12 and corresponds to a rest state, where most of the neurons are not spiking.

Initial conditions	FitzHugh-Nagumo	Synaptic weights	Synapses	Other
$\mu^V = 0$	$a = 0.7$	$\Lambda = 1$	$V_{rev} = 1$	$\Delta t = 0.1$
$\mu^w = 0.5$	$b = 0.8$	$\sigma_3 = 0.2$	$\alpha = 1$	$C = [0, 0.2, 0.4, 0.6, 0.8, 1]$
$\mu^y = 0.3$	$c = 0.08$		$\beta = 1$	
$\sigma_2^V = 0.4$	$I = 0.4$		$T_{MAX} = 1$	
$\sigma_2^w = 0.4$	$\sigma_1 = 0$		$\lambda = 0.2$	
$\sigma_2^y = 0.05$			$V_T = 2$	
			$\Gamma = 0.1$	
			$\Upsilon = 0.5$	

Table 7.1: Values of the parameters of equation 3.34 used to obtain Figures 7.1 and 7.4. These are the same parameters of Table 3.2, with only the exception of the correlation between the Brownian motions.

well known that an efficient technique for highly dimensional integrals is the Monte Carlo approach. Supposing that we want to integrate a multidimensional function $f(\overrightarrow{x})$, we can do it using the following trick:

$$\int_{\Omega} f(\overrightarrow{x})\, d\overrightarrow{x} = Vol(\Omega) \int_{\Omega} f(\overrightarrow{x}) \frac{1}{Vol(\Omega)} d\overrightarrow{x} = Vol(\Omega)\, \mathbb{E}_B[f(\overrightarrow{x})]$$

where Ω is the hyper-dimensional set with hyper-volume $Vol(\Omega)$ on which we want to integrate the function $f(\overrightarrow{x})$, while $\mathbb{E}_B[\cdot]$ is the mean evaluated with respect to the hyper-dimensional box probability density $B(\overrightarrow{x})$, defined as follows:

$$B(\overrightarrow{x}) = \begin{cases} 0 & \text{if} \quad \overrightarrow{x} \notin \Omega \\ \\ \frac{1}{Vol(\Omega)} & \text{if} \quad \overrightarrow{x} \in \Omega \end{cases}$$

Now we can evaluate this mean using infinitely many samples generated by the distribution $B(\overrightarrow{x})$, obtaining finally:

$$\int_{\Omega} f\left(\overrightarrow{x}\right) d\overrightarrow{x} = Vol\left(\Omega\right) \lim_{MCI \to \infty} \left[\frac{1}{MCI} \sum_{j=0}^{MCI-1} f\left(\overrightarrow{x}^{j}\right) \right]$$

where $\overrightarrow{x}^{j}$ is the j-th sample generated by $B\left(\overrightarrow{x}\right)$, while MCI (which stands for *Monte Carlo Integration*, and that must not be confused with MCS) is the total number of samples. Therefore in practice the integral can be approximated through the calculation of the mean of the function $f\left(\overrightarrow{x}\right)$, evaluated with many different samples $\overrightarrow{x}^{j}$, generated randomly by the hyper-dimensional box probability density. Due to the law of large numbers, the error of this integration scheme is always $O\left(MCI^{-\frac{1}{2}}\right)$, regardless of the dimensionality D of the function $f\left(\overrightarrow{x}\right)$. Instead the standard integration schemes suffer from the curse of dimensionality. For example, the error of the trapeziodal rule is $O\left(MCI^{-\frac{2}{D}}\right)$, while for the Simpson's rule it is $O\left(MCI^{-\frac{4}{D}}\right)$, which clearly shows that for $D \gg 1$ the error of the standard integration schemes could be unacceptably large. Instead the Monte Carlo method does not depend on D, therefore it is the ideal candidate for our purpose.

However, if the function $f\left(\overrightarrow{x}\right)$ has non-zero values only in a very limited subset of Ω, many random points $\overrightarrow{x}^{j}$ generated by $B\left(\overrightarrow{x}\right)$ will give $f\left(\overrightarrow{x}\right) = 0$ and therefore they will not contribute to the evaluation of the integral. In other terms if the sampling probability density is not chosen accurately, the points it generates will be wasted. Therefore, in order to use the Monte Carlo integration in an efficient way, we have to choose a sampling distribution $P\left(\overrightarrow{x}\right)$ which generates points only in the range where the function $f\left(\overrightarrow{x}\right)$ is not negligible. This is called *importance sampling* [198], and is formally described as follows:

$$\int_{\Omega} f\left(\overrightarrow{x}\right) d\overrightarrow{x} = \int_{\Omega} \frac{f\left(\overrightarrow{x}\right)}{P\left(\overrightarrow{x}\right)} P\left(\overrightarrow{x}\right) d\overrightarrow{X} = \mathbb{E}_{P}\left[\frac{f\left(\overrightarrow{x}\right)}{P\left(\overrightarrow{x}\right)}\right] = \lim_{MCI \to \infty} \left[\frac{1}{MCI} \sum_{j=0}^{MCI-1} \frac{f\left(\overrightarrow{x}^{j}\right)}{P\left(\overrightarrow{x}^{j}\right)} \right]$$

where now the samples $\overrightarrow{x}^{j}$ are randomly generated using the probability density $P\left(\overrightarrow{x}\right)$. This is the method that we will use to evaluate numerically the Fisher information of the

neural network. From formula 7.1 and since we have $\theta = I$, we can see that in our case the function to integrate is:

$$f(\overrightarrow{x}) = \left(\frac{\partial}{\partial I}\log p(\overrightarrow{x}, I, t)\right)^2 p(\overrightarrow{x}, I, t) \tag{7.2}$$

and that the integration set is $\Omega = \mathbb{R}^{3N}$ (therefore $D = 3N$). Now we have only to determine a good sampling distribution $P(\overrightarrow{x})$. It is obvious that the areas of the phase space where the probability density $p(\overrightarrow{x}, I, t)$ is flat do not contribute to the Fisher information. Moreover the probability density of the FitzHugh-Nagumo network is flat only when it is close to zero. This means that the areas of the phase space that contribute to the Fisher information are only those where $p(\overrightarrow{x}, I, t)$ is significantly larger than zero. Therefore the most natural sampling distribution in this case is $P(\overrightarrow{x}) = p(\overrightarrow{x}, I, t)$, since this distribution generates the random points mainly in correspondence of its peaks, where the function 7.2 contributes more to the Fisher information. However we already know the distribution $p(\overrightarrow{x}, I, t)$, since we have calculated it through the EM algorithm, so we can use it to generate many samples for the Monte Carlo integration of the function 7.2. Actually this last step is not necessarily required, since the samples obtained from the simulation of the network (the same samples that we used to determine $p(\overrightarrow{x}, I, t)$ through the EM algorithm) are distributed according to $p(\overrightarrow{x}, I, t)$ and therefore they could be used to integrate the function 7.2. However in this case, if we need more samples for improving the precision of the Monte Carlo integration, we have to increase the number of simulations of the whole network, in order to generate more samples. Unfortunately, if the network is large this could be a very slow procedure. Therefore the best way to integrate the function 7.2 is to generate many new samples using the EM approximation of $p(\overrightarrow{x}, I, t)$. In general, to produce samples distributed according to a given probability density is a complicated task. However in our case $p(\overrightarrow{x}, I, t)$ is a mixture of Gaussian distributions, so it can be done easily. In fact, given a mixture of m variables $\overrightarrow{y}_i \sim \mathcal{N}(\overrightarrow{\mu}_i, \Sigma_i)$ with weights w_i, for $i = 0, 1, ..., m-1$, and such that $\sum_{i=0}^{m-1} w_i = 1$, it is possible to generate a sample s_j from

the mixture distribution through the following two steps:

- first, choosing randomly an index $\bar{i}$ in the set $\{0, 1, ..., m-1\}$, where the i-th element of this set has a probability of being chosen equal to w_i;
- second, generating the sample $\overrightarrow{s}_j$ from the multivariate normal distribution with mean $\overrightarrow{\mu}_{\bar{i}}$ and covariance $\Sigma_{\bar{i}}$.

Finally, we still have to evaluate the derivative $\frac{\partial}{\partial I} \log p\left(\overrightarrow{x}, I, t\right)$ required in the calculation of the Fisher information. Since the execution of this algorithm on a regular laptop is quite slow, we want to avoid the use of formula 3.38, since it requires at least 3 repetitions (for $n = 1$) of the simulation, for slightly different values of I. Therefore we have opted for a simple forward difference scheme of first order, namely:

$$f\left(\overrightarrow{x}\right) = \left[\frac{1}{\Delta I}\left(\log p\left(\overrightarrow{x}, I + \Delta I, t\right) - \log p\left(\overrightarrow{x}, I, t\right)\right)\right]^2 p\left(\overrightarrow{x}, I, t\right) \tag{7.3}$$

since it allows us to evaluate $f\left(\overrightarrow{x}\right)$ using only 2 repetitions of the simulation. Therefore, to recapitulate, the whole algorithm is based on two different Monte Carlo methods. The first is the repetition, for MCS times, of the network simulation in order to generate the samples used for the evaluation of $p\left(\overrightarrow{x}, I, t\right)$ through the EM algorithm. Instead the second is the generation of MCI new samples from the approximation of $p\left(\overrightarrow{x}, I, t\right)$, that are used to integrate the function 7.3. A schematic representation of the process is shown in Figure 7.3.

7.2 Numerical results

Figure 7.4 shows the Fisher information calculated by this algorithm for different values of the correlation between the Brownian motions.

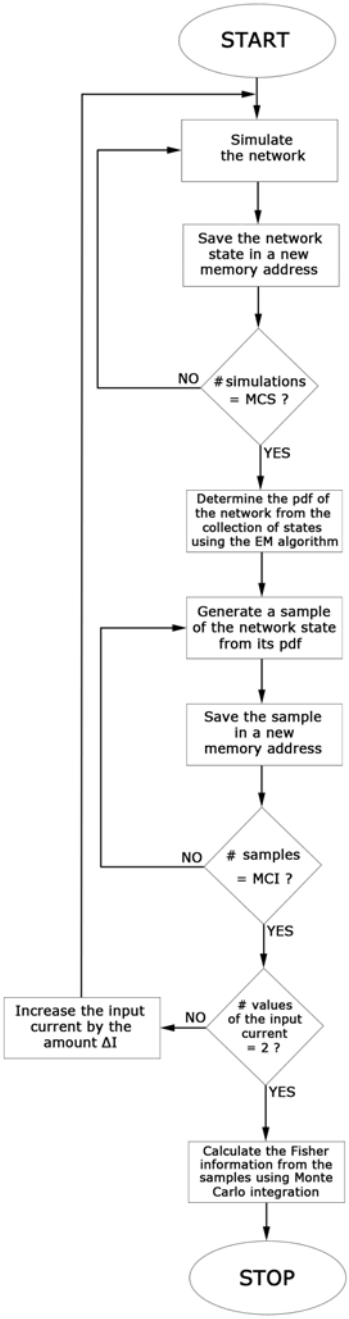

Figure 7.3: Flow chart of the algorithm used for the numerical evaluation of the Fisher information. MCS and MCI are, respectively, the total number of Monte Carlo simulations of the network equations and the total number of samples used for the Monte Carlo integration with importance sampling.

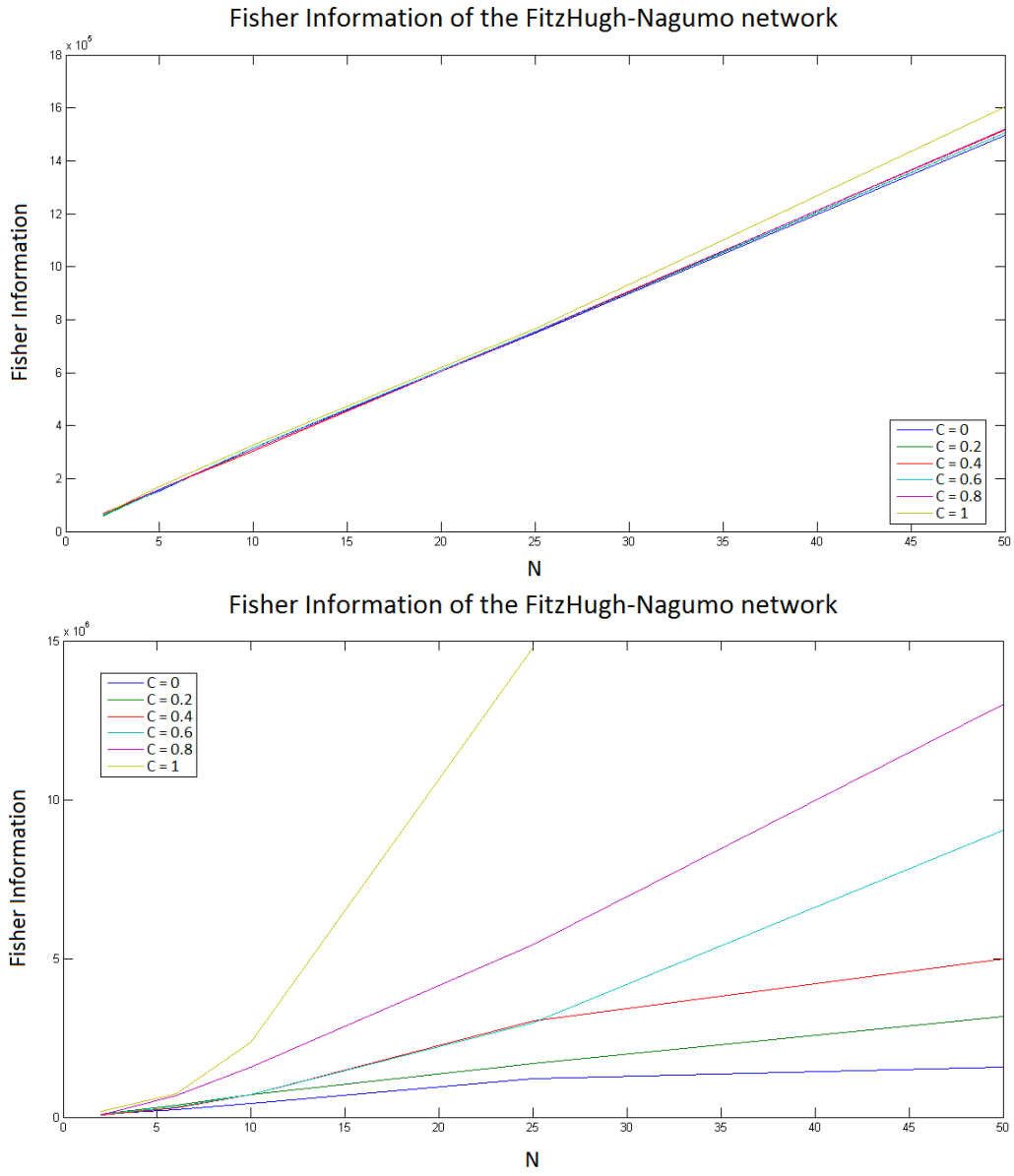

Figure 7.4: Interpolation of the Fisher information for spiking (top) and non-spiking (bottom) neurons, obtained for the FitzHugh-Nagumo network with $N = 2, 6, 10, 25, 50$ and for $t = 4$. In these simulations we have used the parameters of Table 7.1 for the figure on the top, while for the figure at the bottom we have used the same parameters with only the exception of the external input and the intensity of the background noise, which have been set respectively to $I = -0.8$ and $\sigma_1 = 0.45$. The partial derivative with respect to the input has been calculated with $\Delta I = 0.01$. To conclude, the EM algorithm has been used with $10,000$ repetitions of the network, while the Monte Carlo integration has been performed with $100,000$ samples. **NOTE:** for the figure at the bottom, the Fisher information has not been reported for $C = 1$ and $N = 50$ because the algorithm does not provide a plausible result, due to the extremely fast divergence of the curve.

The error analysis of this numerical method has not been performed, therefore its precision has not been quantified yet. This means that the figures shown in this chapter must be taken as indicative, namely they are only qualitative results. Nevertheless they clearly show again, as in Chapters 5 and 6, that the Fisher information is larger when the correlation between the Brownian motions is high. Furthermore, now the analysis has been split into two parts, namely the cases when the neurons are spiking and when they are at rest, showing that the Fisher information is always larger in the case of strong correlations. These results also prove that qualitatively the behavior of $\mathcal{I}(\theta, t)$ depends on the state of the network and is characterized by a higher variability when the neurons are not spiking.

7.3 Partial conclusion

In this chapter we have developed an algorithm that calculates the Fisher information of any neural network in a purely numerical way, without analytic approximations. This approach is based on statistical techniques, namely the Expectation-Maximization algorithm with Gaussian mixtures and the Monte Carlo integration with importance sampling. This numerical method in principle works also for systems with highly non-Gaussian probability distributions, therefore we applied it to the case of the FitzHugh-Nagumo network. Unlike the rate model, in a spiking network the neurons can be essentially in two different states, namely the rest or the spiking state. We analyzed both cases with the algorithm, obtaining that the Fisher information is always higher for correlated membrane potentials, especially if the neurons are in the rest state. This confirms the results obtained in Chapters 5 and 6, and highlights the strong dependence of the Fisher information on the state of the network. This relation is less evident in the case of the rate model, since its probability density is usually close to a Gaussian distribution. It is also important to observe that the increase of the Fisher information for correlated FitzHugh-Nagumo neurons is probably due to the presence of two different sources of inhomogeneity. The first is the inhomogeneity of the synaptic weights, generated by σ_3 (compare with Chapters 5 and

6), while the second is that generated by the bifurcations of the system. In detail, if the neurons are in a spiking state, the bifurcation is described by the split of the trajectories in the phase space (see Figure 3.11). Instead, if the neurons are in a rest state, the bifurcation is due to the non-negligible probability of generating a spike through their random fluctuations.

To conclude, we remind that the results of this algorithm must be considered as qualitative, because they are based on statistical techniques, whose error analysis has not been performed yet.

Chapter 8

Conclusion

In this thesis we have introduced different techniques that describe the dynamics and the statistics of the neural networks, with a special interest for their information encoding capabilities.

In Chapter 1 we have discussed the structure of the brain, in particular of the cerebral cortex, which is responsible for our higher cognitive functions. In particular, the main feature of the brain that has emerged from this analysis is its nested structure, namely its subdivision into areas containing other smaller areas with specific functions. The extension of these regions goes from the macroscopic scale of $10^6 \div 10^9$ neurons to the microscopic scale of single neurons, passing through an intermediate mesoscopic scale of $10^1 \div 10^5$ neurons. This property of the brain reflects its working principles, therefore must be taken into account in the description of large neural networks. In Chapter 6 we have shown how this can be accomplished from the mathematical point of view, using the Sporns' fractal topology.

In Chapter 2 we have explained that, according to the Theory of Complexity, it may be more relevant to describe correctly the interactions between the neurons, than to use extremely realistic models for the single cells. In fact, as we said, the functions of the brain

are mainly due to its connectivity matrix, therefore we have underlined the importance of developing a mathematical theory which is able to describe the behavior of the network for different kinds of synaptic connectivity. In this chapter we have clarified the differences between computation and information processing, explaining that they are both performed by the brain. Since in this thesis we have considered stochastic networks, which are characterized by a probability density, and since we have determined this density in different ways, it is natural to apply these techniques for studying the information processing capabilities of the system. In particular, here we have considered information encoding, which is quantified by the Fisher information of the neural network.

In Chapter 3 we have started with the mean-field theory, showing that, if the correlations between neurons are negligible, it is possible to reduce the full system of the network equations to a smaller set, which describes a single neuron with a mean-field interaction. This is known as the *McKean-Vlasov equation* of the system, which is a stochastic differential equation with an implicit interaction term that depends on the marginal probability density of a single neuron. Since the marginal density is not known a priori, the mean-field equation cannot be solved directly. For this reason it is convenient to transform it in the corresponding Fokker-Planck equation (FPE). The FPE is a partial integro-differential equation (PIDE) whose unknown is the marginal probability density of a single neuron. The equation has a non-linear integral term, inherited from the implicit interaction of the McKean-Vlasov equation, therefore in general it cannot be solved analytically. Once the solution is known, usually numerically or semi-analytically, it can be used to calculate the joint probability density of a sub-system of the network. In fact, due to independence, the joint probability density of the sub-system can be reconstructed through the product of the marginal densities. Therefore this represents a very compact and convenient way to describe the activity of the network. In this thesis we have derived the FPE in two different intuitive ways and we have solved it for two kinds of network, made up of rate or FitzHugh-Nagumo neurons. In the latter case, the solution can be obtained only numerically, but this has proved to be a difficult task, due to the stiffness of the equations and

to the quite high dimensionality of the phase space of a single neuron. This has led us to evaluate the solution using powerful calculators like GPUs [64][176]. Therefore the simulation of systems with an even higher dimensionality, like the Hodgkin-Huxley neuron, represents a real challenge. Instead, in the case of the rate model, not only the FPE is easy to solve numerically due to its low dimensionality, but it can also be converted into a system of $2P$ coupled ordinary differential equations (ODEs), in the case of P neural populations. This system is much more convenient to solve than a PIDE, and describes the evolution of the mean and the variance of the stochastic process. This is a sufficient description for the probability density of the single neuron, because in the case of the rate model, if the initial conditions are Gaussian, then the process is always Gaussian at every time instant. In this thesis we have shown an alternative derivation, compared to that appeared in [51], of the ODE system of the mean and the variance and of the probability density of the process, based on the solution of the FPE.

In Chapter 4 we have extended the previous mean-field analysis in order to describe finite-size effects. In fact the mean-field theory works properly only in the thermodynamic limit, namely when the number of neurons in the system is ideally infinite. Therefore an extension of the theory is required in order to quantify the behavior of a finite number of interacting neurons. The idea expressed in Chapter 4 is to use the same mathematical formalism already developed in physics of plasma, which is known as the *Mayer's cluster expansion* (MCE). In fact, a network with independent Brownian motions and a sufficient number of connections, and a plasma, share similar statistical properties, namely small values of correlation. However, if the amount of correlation is not small enough to be neglected, an extension of the mean-field theory is required. The MCE is a formal decomposition of the global interaction between a finite number of particles (in our case neurons) in the contributions generated by the interplay of pairs, triplets, quadruplets etc. In this way, the negligible part of the interaction can be removed, allowing us to write a convenient system of PIDEs for the probability density of the finite size system. Usually only the pairwise interactions are considered, while higher order correlations are neglected. This

idea works properly if we study the behavior of a small sub-system and if the total number of neurons in the whole network is high enough. Otherwise higher order correlations could give a relevant contribution to the joint probability density of the sub-system. In this chapter we have also solved numerically the PIDE system obtained from the MCE, and we have used the probability density to evaluate the Fisher information of the network. Compared to the mean-field case, the Fisher information is now much more difficult to calculate, due to the presence of correlation. In fact, since the neurons are not independent anymore, the Fisher information of the whole system is not the sum of the Fisher information of the single neurons. Therefore we have solved the problem by expanding the Fisher information up to the third order, using the small pairwise correlations as perturbative parameters. The simulations clearly show that the Fisher information increases when the neurons are more and more independent, a result that we have confirmed in Chapters 5 and 6.

In Chapters 5 and 6 we have emphasized the role of the connectivity matrix and for the first time in the thesis we have used three different sources of external correlation, that we have introduced in the Brownian motions of the background noise, in the initial conditions and in the distribution of the synaptic weights. These networks have been studied using two complementary perturbative expansions, that for simplicity have been applied only to rate neurons. The first technique can be used only around stationary solutions and if the network is invariant under exchange of the neural indices. Instead the second expansion can be applied only if the synaptic weights are weak enough, but works for general connectivity matrices with random topologies. The formulae have been obtained for a generic number of neurons and describe also some non-linear effects of the network. In the first approach the inhomogeneities of the neural behavior can be added perturbatively. Moreover it does not require weak synaptic weights and, first of all, has been used to study block circulant connectivity matrices with circulant blocks and with a generic number of incoming connections per neuron, that we have called M. In the special case when the connectivity matrix is circulant, we have proved that, if the three external sources of correlation are set to zero, correlation decreases with M, and not with the number of neurons, as previously

thought [64]. Then, using graph theory, we have studied special symmetric connectivity matrices, generated by the Cartesian or the Kronecker product of circulant graphs. We have also proved that if the intensities of the three external sources of correlation, namely those of the Brownian motions, the initial conditions and the synaptic weights, are not negligible, then the neurons cannot become independent. Moreover, even if these external correlations are set to zero, for special values of the parameters of the system the neurons can be perfectly correlated. We have called this phenomenon *stochastic synchronization*, and it can occur for any finite number of neurons. The time instant at which the neurons become perfectly correlated is finite and increases with the network size. Therefore in the mean-field limit this phenomenon occurs only at $t \to +\infty$, or in other words it never happens, confirming the phenomenon of propagation of chaos discussed in [64] for independent initial conditions. In the case of networks with weak synaptic weights, the second kind of perturbative expansion can be used. This approach can be applied to any kind of connectivity matrix, also with random topology, therefore we have used it to describe the behavior of the network in the case of biologically realistic matrices. In particular we have chosen the fractal connectivity introduced in [190], which roughly describes the nested structure of the brain discussed in Chapter 1. All the results of Chapters 5 and 6 have been verified also numerically. Finally, we have also used these perturbative expansions truncated at the first order for evaluating the Fisher information of the network. Since at the first order the probability density is Gaussian, there is a simple formula that quantifies the Fisher information in terms of the mean-vector and the covariance matrix of the membrane potentials. The two perturbative methods have provided the same result, namely that the Fisher information increases for highly independent neurons, as in Chapter 4, but also that for networks with inhomogeneities the Fisher information diverges when the neurons are highly correlated, contrarily to common belief [178][181][182][183].

Finally, in Chapter 7 we have developed an algorithm for the numerical evaluation of the Fisher information of the FitzHugh-Nagumo network. Due to the exponential explosion of the amount of data that are necessary for the construction of a histogram, we have

used another way to approximate the probability density of the whole network. Particularly convenient is the use of an Expectation-Maximization algorithm, that approximates the probability density with a mixture of multivariate Gaussian distributions. Then the the integral that defines the Fisher information has been evaluated with a Monte Carlo method and importance sampling. This numerical technique is affected by three different kinds of error:

- the finite number of Monte Carlo simulations that are used to determine the joint probability density of the network;
- the approximation of the derivative with respect to the external input;
- the finite number of Monte Carlo repetitions that are used to evaluate the multidimensional integral.

We have not analyzed in detail the magnitude of these three sources of error, therefore the results shown in Chapter 7 should be taken as qualitative indications of the behavior of the Fisher information. Nevertheless, the algorithm clearly shows that for both spiking and non-spiking FitzHugh-Nagumo neurons, the Fisher information is higher for correlated membrane potentials. In particular, for non-spiking neurons the function quickly diverges for sufficiently large networks. This confirms the results found in Chapters 5 and 6.

Probably, the incorrect belief of higher encoding capabilities for independent neurons is due to the comparison with the case of the Shannon information. In fact, in the case of transmission through a communication channel, usually the signal is coded in such a way that its redundancy is increased, in order to protect it against noise. In fact even if the code has been damaged by the random fluctuations of the communication channel, its information content was repeated so many times (due to the high redundancy) that the original message can still be reconstructed. So, if we have many neurons with essentially the same membrane potentials (namely highly correlated), their redundancy is very high,

and therefore the Shannon information content is very low. However this is true only for the Shannon information, and not for the Fisher information. This could explain why the belief that independent neurons are required for high encoding capabilities is so common.

To conclude, this thesis contributes to the comprehension of the mechanisms that underlie the dynamics and the statistics of stochastic neural networks, for many different kinds of connectivity matrices. Moreover, it provides an analytic relation between the anatomical connectivity and the functional connectivity of the system, a problem that is currently intensively investigated [54][55][56]. This work opens also the way to the quantification of the information processing capability of the neural networks. In particular, we have studied information encoding, through the evaluation of the Fisher information of the system as a function of the correlation between the neurons in the network. Other information quantities can be calculated in the same way, like the Shannon differential entropy, the mutual information, the transfer entropy [199][200][201], and so on and so forth. This allows us to quantify the information storage, transfer and modification, following the same ideas already developed in the field of automata theory [119][120][121][122]. A kind of information that we are not yet able to quantify is the so called *semantic information* [127][128][129][130]. A full comprehension of this rather elusive concept is necessary in order to understand the brain in all its complexity. Instead, at the other end of the spectrum, we have the Shannon's information theory, that provides a syntactic notion of information. In other words, Shannon's theory quantifies the amount of information conveyed in a message as a function of its syntax and probability density, regardless of its semantic content. Therefore Shannon information and semantic information can be seen as two complementary aspects of the brain. In this thesis we have laid the foundations for studying only the former side of the coin. Moreover, it must be noted that the networks described in this work are not necessarily performing a specific task, since their function is determined by the connectivity matrix, which has been chosen arbitrarily. In a real neural network, the strength of the connections is modeled by synaptic plasticity and learning, in response to the stimuli coming from the environment. However, the perturbative techniques in-

troduced in Chapters 5 and 6 can be extended in order to include this phenomenon. In particular, the second perturbative expansion can be also applied to the case of spiking neurons, as now we are going to explain. The first expansion requires the existence of a stationary state, around which we perturb the solution of the neural equations. Clearly, in the conductance-based models, like those of FitzHugh-Nagumo or Hodgkin-Huxley, when the neurons are spiking their membrane potentials are not in a stationary state, therefore the first perturbative approach cannot be used. Instead, in the case of the second expansion, the perturbation is performed around the case without synaptic connections (i.e. $J(t) = 0$), therefore the existence of a stationary state is not required. This means that the second perturbative approach can be also used for the conductance-based models, as we stated, therefore now we are in a much better position for understanding in detail the working principles of extremely realistic neural networks.

Conclusion générale (en Français)

Dans cette thèse nous avons présenté une variété de techniques mathématiques, statistiques et numériques permettant l'analyse de réseaux neuronaux stochastiques pour un large spectre de matrices de connectivité. Toutes les techniques utilisées, à savoir la théorie de champ moyen, l'expansion de groupe de Mayer, deux développements perturbatifs distincts et les différentes méthodes numériques de calcul de l'information de Fisher du réseau, ont fourni des résultats compatibles. Ainsi nous avons une vision beaucoup plus claire des principes sous-jacents qui régissent la relation entre la matrice de connectivité synaptique d'un réseau neuronal et sa structure de corrélation, qui détermine à son tour la capacité de codage du système.

La contribution de cette thèse est double. D'une part, elle fournit une extension de la théorie de champ moyen, largement utilisée dans la description des réseaux neuronaux, dans le cas où cette dernière ne peut pas être appliquée du fait de la non-indépendance des neurones. Cela inclue les réseaux de taille finie, ceux comprenant des sources aléatoires corrélées, ou encore ceux à densité faible de connexion et avec à valeurs spéciales de paramètre impliquant une corrélation forte (ceci est le phénomène que nous avons appelé la *synchronisation stochastique*). D'autre part cette thèse propose le calcul de l'information de Fisher pour les modèles de neurones "rate" et "spiking", avec différents ordres d'approximation, montrant que la capacité de codage du système est plus haute pour des neurones fortement corrélés. Ceci est contraire à l'intuition et aux hypothèses faites

communément et joue un rôle important dans le cadre du traitement de l'information exécuté par le cerveau.

Les techniques perturbatives développées aux Chapitres 5 et 6 peuvent également être utilisées pour calculer toutes les quantités d'informations définies par la théorie de l'information de Shannon, tels l'information mutuelle, le transfert d'entropie, l'entropie différentielle, ou encore le stockage et la modification de l'information. En raison des contraintes de temps cette analyse n'a pas été effectué et est reportée à une future étude. En outre ces techniques peuvent être appliquées à toutes sortes d'équations neurales même dans le cas de la plasticité synaptique et de l'apprentissage. Leurs inconvénients est qu'on ne peut les utiliser que dans le cas de sources aléatoires relativement faibles. Si l'incertitude est élevée l'ordre d'approximation de la perturbation doit être augmenté. Malheureusement la complexité induite par cette élévation augmente fortement, devenant rapidement insoluble. À l'inverse l'expansion de groupe de Mayer développée au Chapitre 4 peut être appliquée pour toute source aléatoire mais ne fournit pas de résultat analytique et restera un complément de l'approche perturbative.

En conclusion, cette thèse propose des techniques efficaces pouvant améliorer notre compréhension du cerveau en terme de corrélation neuronale et de traitement de l'information. Ces techniques ont été appliquées avec succès à diverses topologies de connexions synaptiques, des plus basiques aux plus réalistes. Cependant dans tous nos modèles le poids des connexions est fixe avec des fluctuations stochastiques ou évolue selon une loi d'évolution quelconque indépendamment des principes biologiques. Ainsi la prochaine étape est l'extension de cette théorie aux réseaux de neurones dont les connexions évoluent selon un apprentissage comme observé expérimentalement. Cela permettra d'étudier des réseaux acquérant des fonctions et des comportements spécifiques du fait de l'interaction avec l'environnement et donc d'améliorer la plausibilité du modèle.

Appendix A

Self-consistency constraints of the Mayer's cluster expansion

In order to have a self-consistent Mayer's cluster expansion, the following constraints must be satisfied:

$$\begin{aligned}
\int_{\mathbb{R}} p^{(1)}(V,t)\,dV &= 1 \\
\int_{\mathbb{R}} p^{(2)}(V,V',t)\,dV' &= p^{(1)}(V,t) \\
\int_{\mathbb{R}} p^{(3)}(V,V',V'',t)\,dV'' &= p^{(2)}(V,V',t) \\
&\vdots
\end{aligned} \tag{A.1}$$

Now, the first condition of A.1 is assumed to be true, since it can be easily satisfied by mul-

tiplying eventually the function $p^{(1)}(V,t)$ by a normalization constant. Moreover, according to (4.1):

$$\int_{\mathbb{R}} dV' p^{(2)}(V,V',t) = p^{(1)}(V,t) \int_{\mathbb{R}} p^{(1)}(V',t)\, dV' + \int_{\mathbb{R}} P(V,V',t)\, dV'$$

$$= p^{(1)}(V,t) + \int_{\mathbb{R}} P(V,V',t)\, dV'$$

having used the normalization condition on $p^{(1)}(V,t)$. Therefore the second constraint of A.1 is satisfied if and only if:

$$\int_{\mathbb{R}} P(V,V',t)\, dV' = 0 \qquad \forall V \tag{A.2}$$

Clearly the same constraint is satisfied if we integrate $P(V,V',t)$ with respect to V. This is the only condition that must be satisfied by the pair correlation function P.

To conclude, we have also:

$$\int_{\mathbb{R}} p^{(3)}(V,V',V'',t)\, dV''$$

$$= \int_{\mathbb{R}} \Big[p^{(1)}(V,t)\, p^{(1)}(V',t)\, p^{(1)}(V'',t) + p^{(1)}(V,t)\, P(V',V'',t)$$

$$+ p^{(1)}(V',t)\, P(V,V'',t) + p^{(1)}(V'',t)\, P(V,V',t) + T(V,V',V'',t) \Big]\, dV''$$

$$= p^{(1)}(V,t)\, p^{(1)}(V',t) \int_{\mathbb{R}} p^{(1)}(V'',t)\, dV'' + p^{(1)}(V,t) \int_{\mathbb{R}} P(V',V'',t)\, dV''$$

$$+ p^{(1)}(V', t) \int_{\mathbb{R}} P(V, V'', t)\, dV'' + P(V, V', t) \int_{\mathbb{R}} p^{(1)}(V'', t)\, dV'' + \int_{\mathbb{R}} T(V, V', V'', t)\, dV''$$

$$= p^{(1)}(V, t)\, p^{(1)}(V', t) + P(V, V', t) + \int_{\mathbb{R}} T(V, V', V'', t)\, dV''$$

Therefore the second condition is satisfied if and only if:

$$\int_{\mathbb{R}} T(V, V', V'', t)\, dV'' = 0 \qquad \forall V, V' \tag{A.3}$$

and so on and so forth.

Appendix B

Spike count correlation

Usually in experiments the researchers measure the spike-count of a single neuron in a given time range $[t; t + \Delta t]$, that is formally defined as:

$$N_i\left(t; t + \Delta t\right) = \int_t^{t+\Delta t} \rho_i\left(\tau\right) d\tau$$

where:

$$\rho_i\left(t\right) = \sum_{k=0}^{n_i - 1} \delta\left(t - t_{k,i}\right), \qquad t \leq t_{k,i} \leq t + \Delta t \qquad \forall k$$

is the neural response function of the i-th neuron, i.e. the sum of all the spikes (assumed to be Dirac delta functions and detected at the time instants $t_{k,i}$) produced by that neuron in the time range $[t; t + \Delta t]$ of a single Monte Carlo simulation (in this case we have of course $N_i\left(t; t + \Delta t\right) = n_i$). Instead in the literature the instantaneous spike-count rate is defined as:

$$scr_i\left(t\right) = \lim_{\Delta t \to 0} \frac{N_i\left(t; t + \Delta t\right)}{\Delta t} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} \rho_i\left(\tau\right) d\tau$$

and from it we can define the instantaneous mean firing rate as the mean $\mathbb{E}[\cdot]$ of the instantaneous spike-count rate over many Monte Carlo simulations:

$$R_i(t) = \mathbb{E}[scr_i(t)] = \lim_{\Delta t \to 0} \frac{\mathbb{E}[N_i(t; t+\Delta t)]}{\Delta t} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} \mathbb{E}[\rho_i(\tau)]\, d\tau = \lim_{\Delta t \to 0} R_i(t; t+\Delta t) \quad \text{(B.1)}$$

where:

$$R_i(t; t+\Delta t) = \frac{\mathbb{E}[N_i(t; t+\Delta t)]}{\Delta t} = \frac{1}{\Delta t} \int_t^{t+\Delta t} \mathbb{E}[\rho_i(\tau)]\, d\tau$$

is the average mean firing rate in the time range $[t; t+\Delta t]$. Here the adjective "average" means "average over time", i.e. in the time range $[t; t+\Delta t]$ (this is inspired by physics, where for example the average speed over a given space is equal to the length of the space divided by the time spent to travel on it), while "mean" is referred to the mean over many Monte Carlo simulations. Instead in the rate model we can quantify the instantaneous mean firing rate, defined as:

$$R_i(t) = \int_{\mathbb{R}} S(V')\, p_i(V', t)\, dV'$$

where $p_i(\cdot, t)$ is the marginal probability density function of the i-th neuron at time t (see 3.17), and $S(\cdot)$ is the activation function. This is equivalent to the mean of the function $S(\cdot)$ over infinitely many Monte Carlo simulations, therefore according to formula B.1 we can identify $S(V(t))$ with $scr(t)$. Now, the correlation between the spike-counts of 2 neurons i and j is defined as:

$$Corr(N_i(t; t+\Delta t), N_j(t; t+\Delta t)) = \frac{Cov(N_i(t; t+\Delta t), N_j(t; t+\Delta t))}{\sqrt{Var(N_i(t; t+\Delta t))\, Var(N_j(t; t+\Delta t))}}$$

where:

$$Cov\left(N_i\left(t;t+\Delta t\right),N_j\left(t;t+\Delta t\right)\right)=\mathbb{E}\left[N_i\left(t;t+\Delta t\right)N_j\left(t;t+\Delta t\right)\right]-\mathbb{E}\left[N_i\left(t;t+\Delta t\right)\right]\mathbb{E}\left[N_j\left(t;t+\Delta t\right)\right]$$

$$Var\left(N_i\left(t;t+\Delta t\right)\right)=Cov\left(N_i\left(t;t+\Delta t\right),N_i\left(t;t+\Delta t\right)\right)$$

Now we observe that:

$$\lim_{\Delta t\to 0}\frac{1}{\left(\Delta t\right)^2}\mathbb{E}\left[N_i\left(t;t+\Delta t\right)N_j\left(t;t+\Delta t\right)\right]=\lim_{\Delta t\to 0}\mathbb{E}\left[\frac{N_i\left(t;t+\Delta t\right)}{\Delta t}\frac{N_j\left(t;t+\Delta t\right)}{\Delta t}\right]=\mathbb{E}\left[scr_i\left(t\right)scr_j\left(t\right)\right]$$

and remembering that $scr\left(t\right)=S\left(V\left(t\right)\right)$, we can write:

$$\mathbb{E}\left[scr_i\left(t\right)scr_j\left(t\right)\right]=\mathbb{E}\left[S\left(V_i\left(t\right)\right)S\left(V_j\left(t\right)\right)\right]=\int_{\mathbb{R}^2}S\left(V'\right)S\left(V''\right)p_{i,j}\left(V',V'',t\right)dV'dV''$$

where $p_{i,j}\left(\cdot,\cdot,t\right)$ is the marginal probability density function of the i-th and j-th neurons at time t. Finally, supposing that $\int_{\mathbb{R}^2}S\left(V'\right)S\left(V''\right)p_{i,j}\left(V',V'',t\right)dV'dV''$ does not change too much in the time range Δt, we can write:

$$\mathbb{E}\left[N_i\left(t;t+\Delta t\right)N_j\left(t;t+\Delta t\right)\right]\approx\left[\int_{\mathbb{R}^2}S\left(V'\right)S\left(V''\right)p_{i,j}\left(V',V'',t\right)dV'dV''\right]\left(\Delta t\right)^2$$

In a similar way we can prove also that:

$$\mathbb{E}\left[N_i\left(t;t+\Delta t\right)\right]\approx\left[\int_{\mathbb{R}}S\left(V'\right)p_i\left(V',t\right)dV'\right]\Delta t$$

So we can conclude that the spike-count correlation usually evaluated by experimentalists can be calculated with the rate model using the following formula:

$$Corr\left(N_i\left(t;t+\Delta t\right),N_j\left(t;t+\Delta t\right)\right)\approx\frac{f_{i,j}\left(t\right)}{\sqrt{f_i\left(t\right)f_j\left(t\right)}}$$

where:

$$f_{i,j}(t) = \left[\int_{\mathbb{R}^2} S(V') S(V'') p_{i,j}(V', V'', t)\, dV' dV''\right] - \left[\int_{\mathbb{R}} S(V') p_i(V', t)\, dV'\right]\left[\int_{\mathbb{R}} S(V') p_j(V', t)\, dV'\right]$$

$$f_i(t) = f_{i,i}(t)$$

Now we have to compute, according to the rate model, the different terms involved in this formula. Let us show the procedure for the term $\int_{\mathbb{R}^2} S(V') S(V'') p_{i,j}(V', V'', t)\, dV' dV''$. For relatively weak stochastic fluctuations we can approximately expand the function $S(V') S(V'')$ in a $2D$ Taylor series up to the second order with respect to the point:

$$(m_i(t), m_j(t)) = \int_{\mathbb{R}^2} (V', V'') p_{i,j}(V', V'', t)\, dV' dV'' = (\mathbb{E}[V_i(t)], \mathbb{E}[V_j(t)])$$

obtaining:

$$\begin{aligned} S(V') S(V'') \approx & S(m_i(t)) S(m_j(t)) \\ & + S(m_i(t)) S^{(1)}(m_j(t)) (V'' - m_j(t)) + S(m_j(t)) S^{(1)}(m_i(t)) (V' - m_i(t)) \\ & + S^{(1)}(m_i(t)) S^{(1)}(m_j(t)) (V' - m_i(t)) (V'' - m_j(t)) \\ & + \frac{1}{2} S(m_i(t)) S^{(2)}(m_j(t)) (V'' - m_j(t))^2 + \frac{1}{2} S(m_j(t)) S^{(2)}(m_i(t)) (V' - m_i(t))^2 \end{aligned}$$

since under our hypothesis the fluctuations $(V_i - m_i(t), V_j - m_j(t))$ are small. Therefore, using the Isserlis' theorem, we obtain:

$$\int_{\mathbb{R}^2} S\left(V'\right) S\left(V''\right) p_{i,j}\left(V', V'', t\right) dV' dV'' \approx S\left(m_i\left(t\right)\right) S\left(m_j\left(t\right)\right) + S^{(1)}\left(m_i\left(t\right)\right) S^{(1)}\left(m_j\left(t\right)\right) \Sigma_{ij}\left(t\right)$$

$$+ \frac{1}{2} S\left(m_i\left(t\right)\right) S^{(2)}\left(m_j\left(t\right)\right) \Sigma_{jj}\left(t\right) + \frac{1}{2} S\left(m_j\left(t\right)\right) S^{(2)}\left(m_i\left(t\right)\right) \Sigma_{ii}\left(t\right)$$

where:

$$\Sigma_{ij}\left(t\right) = Cov\left(V_i\left(t\right), V_j\left(t\right)\right)$$

In the same way we can also prove that:

$$\int_{\mathbb{R}} S\left(V'\right) p_i\left(V', t\right) dV' \approx S\left(m_i\left(t\right)\right) + \frac{1}{2} S^{(2)}\left(m_i\left(t\right)\right) \Sigma_{ii}\left(t\right)$$

Finally, putting all the results together and neglecting the higher order terms proportional to $Var\left(V_i\left(t\right)\right) Var\left(V_j\left(t\right)\right)$, we obtain:

$$Corr\left(N_i\left(t; t + \Delta t\right), N_j\left(t; t + \Delta t\right)\right) \approx Corr\left(V_i\left(t\right), V_j\left(t\right)\right)$$

Therefore for relatively small stochastic fluctuations the correlation of the spike-counts is approximately equal to the correlation of the membrane potentials, which can be calculated from the neural equations of the rate model using the methods developed in this thesis. So the values of the parameters of the system with respect to which we calculate the Fisher information are encoded by both the spike-counts and the membrane potentials, but it is better to reason in terms of the latter, since it is more natural. In fact our neural network is described for example by 3.1, namely by differential equations whose unknowns are the membrane potentials, not the spike-counts. Due to this approximate

equality of the correlation structures we can conclude that the spike-counts and the membrane potentials encode the value of the external input qualitatively in the same way, as it must be. In fact, there is an interplay between these two variables since, roughly speaking, the pre-synaptic membrane potentials determine the values of the firing rates (and therefore of the spike-counts), while the firing rates determine the values of the post-synaptic membrane potentials. However this does not mean that the Fisher information calculated over the membrane potentials is the same of that calculated over the spike-counts. In fact the Fisher information is related not only to the covariance matrix of the process, but also to its mean vector, which is not the same for our two variables:

$$\mathbb{E}\left[N_i\left(t;t+\Delta t\right)\right] \approx \left[\int_{\mathbb{R}} S\left(V'\right) p_i\left(V',t\right) dV'\right] \Delta t \approx \left[S\left(m_i\left(t\right)\right) + \frac{1}{2} S^{(2)}\left(m_i\left(t\right)\right) \Sigma_{ii}\left(t\right)\right] \Delta t$$

From this formula we see that there is no simple relation between the Fisher information calculated over the spike-counts and that evaluated over the membrane potentials.

Appendix C

Perturbative expansion of the Fisher information

In this appendix we consider for simplicity only the case of a fully connected network. According to the discussion at the end of Section 4.2, formula 4.12 can be used only for a subset of $\eta \ll N$ neurons. Moreover, it can be rewritten in the following way:

$$p\left(\overrightarrow{V},t\right)=\prod_{i=0}^{\eta-1}p^{(1)}\left(V_i,t\right)+\widetilde{P}\left(\overrightarrow{V},t\right)=\left[\prod_{i=0}^{\eta-1}p^{(1)}\left(V_i,t\right)\right]\left[1+D\left(\overrightarrow{V},t\right)\right] \tag{C.1}$$

where:

$$\widetilde{P}\left(\overrightarrow{V},t\right)=\sum_{\substack{a,b\\a<b}}\left\{\left[\prod_{c\neq a,b}p^{(1)}\left(V_c,t\right)\right]P\left(V_a,V_b,t\right)\right\}$$

$$D\left(\overrightarrow{V},t\right)=\frac{\widetilde{P}\left(\overrightarrow{V},t\right)}{\prod_{i=0}^{\eta-1}p^{(1)}\left(V_i,t\right)}$$

Clearly $D\left(\overrightarrow{V},t\right) \ll 1$ because $P \ll p^{(1)}p^{(1)}$, as we numerically proved in Section (4.3). Now observe that $D\left(\overrightarrow{V},t\right)$ can be rewritten as:

$$D\left(\overrightarrow{V},t\right) = \sum_{\substack{a,b \\ a<b}} d\left(V_a, V_b, t\right)$$

where:

$$d\left(V_a, V_b, t\right) = \frac{P\left(V_a, V_b, t\right)}{p^{(1)}\left(V_a, t\right) p^{(1)}\left(V_b, t\right)}$$

According to 2.6, the Fisher information of the subset of η neurons with respect to a parameter θ is defined as:

$$\mathcal{I}\left(\theta, t\right) = \int_{\mathbb{R}^\eta} \left(\frac{\partial \log p\left(\overrightarrow{V},\theta,t\right)}{\partial\theta} \right)^2 p\left(\overrightarrow{V},\theta,t\right) d\overrightarrow{V} \tag{C.2}$$

where $p\left(\overrightarrow{V},\theta,t\right)$ is given by C.1. Clearly, plugging C.1 into C.2 does not provide a useful formula for calculating the Fisher information of the system. In the case of independent neurons, in Section 3.3 we have seen that $\mathcal{I}\left(\theta,t\right)$ can be calculated easily in terms of a single one-dimensional integral, which can be evaluated easily with standard numerical techniques. Instead, in the case studied in Chapter 4, this trick cannot be used, since in general the neurons are correlated. However, the main idea is still to find a way to express the Fisher information of a finite network in terms of low-dimensional integrals, in order to avoid the use of Monte Carlo integration techniques (which instead are used in Chapter 7). To this purpose, we have to write formula C.2 in an another form. Supposing that the following regularity conditions are satisfied:

- the support of $p\left(\overrightarrow{V},\theta,t\right)$ does not depend on θ,

- $p\left(\overrightarrow{V},\theta,t\right)$ is twice differentiable,
- differentiation under integral sign can be applied,

it is well known that the Fisher information can be equivalently rewritten as:

$$\mathcal{I}(\theta,t)=-\int_{\mathbb{R}^{\eta}}\frac{\partial^{2}logp\left(\overrightarrow{V},\theta,t\right)}{\partial\theta^{2}}p\left(\overrightarrow{V},\theta,t\right)d\overrightarrow{V} \tag{C.3}$$

We use formula C.3 in order to write a perturbative expansion of $\mathcal{I}(\theta,t)$. Together with C.1 it can be rewritten as:

$$\begin{aligned}\mathcal{I}(\theta,t)=&-\int_{\mathbb{R}^{\eta}}\left[\prod_{i=0}^{\eta-1}p^{(1)}(V_i,t)+\widetilde{P}\left(\overrightarrow{V},t\right)\right]\frac{\partial^{2}log\left\{\left[\prod_{i=0}^{\eta-1}p^{(1)}(V_i,t)\right]\left[1+D\left(\overrightarrow{V},t\right)\right]\right\}}{\partial\theta^{2}}d\overrightarrow{V}\\ =&-\int_{\mathbb{R}^{\eta}}\left[\prod_{i=0}^{\eta-1}p^{(1)}(V_i,t)\right]\frac{\partial^{2}log\left[\prod_{i=0}^{\eta-1}p^{(1)}(V_i,t)\right]}{\partial\theta^{2}}d\overrightarrow{V}-\int_{\mathbb{R}^{\eta}}\left[\prod_{i=0}^{\eta-1}p^{(1)}(V_i,t)\right]\frac{\partial^{2}log\left[1+D\left(\overrightarrow{V},t\right)\right]}{\partial\theta^{2}}d\overrightarrow{V}\\ &-\int_{\mathbb{R}^{\eta}}\widetilde{P}\left(\overrightarrow{V},t\right)\frac{\partial^{2}log\left[\prod_{i=0}^{\eta-1}p^{(1)}(V_i,t)\right]}{\partial\theta^{2}}d\overrightarrow{V}-\int_{\mathbb{R}^{\eta}}\widetilde{P}\left(\overrightarrow{V},t\right)\frac{\partial^{2}log\left[1+D\left(\overrightarrow{V},t\right)\right]}{\partial\theta^{2}}d\overrightarrow{V}\end{aligned} \tag{C.4}$$

Now we show how to calculate the four terms of (C.4).

C.1 First term

This term is the easiest to calculate, and it is simply given by:

$$-\int_{\mathbb{R}^{\eta}}\left[\prod_{i=0}^{\eta-1}p^{(1)}(V_i,t)\right]\frac{\partial^{2}log\left[\prod_{i=0}^{\eta-1}p^{(1)}(V_i,t)\right]}{\partial\theta^{2}}d\overrightarrow{V}=-\eta\int_{\mathbb{R}}p^{(1)}(V,t)\frac{\partial^{2}logp^{(1)}(V,t)}{\partial\theta^{2}}dV$$

C.2 First order approximation of the second term

This term cannot be calculated exactly, so we have to use the Taylor expansion of the logarithmic function. In particular, now we consider only the expansion at the first order, namely

$$log\left[1 + D\left(\overrightarrow{V}, t\right)\right] \approx D\left(\overrightarrow{V}, t\right) \tag{C.5}$$

The second and third orders will be taken into account in Sections C.6 and C.7 respectively. So from C.5 the second term of C.4 can be approximated as follows:

$$-\int_{\mathbb{R}^\eta}\left[\prod_{i=0}^{\eta-1} p^{(1)}\left(V_i, t\right)\right]\frac{\partial^2 log\left[1 + D\left(\overrightarrow{V}, t\right)\right]}{\partial\theta^2}d\overrightarrow{V}$$

$$\approx -\int_{\mathbb{R}^\eta}\left[\prod_{i=0}^{\eta-1} p^{(1)}\left(V_i, t\right)\right]\frac{\partial^2 D\left(\overrightarrow{V}, t\right)}{\partial\theta^2}d\overrightarrow{V}$$

$$= -\sum_{\substack{a,b \\ a<b}}\int_{\mathbb{R}^\eta}\left[\prod_{i=0}^{\eta-1} p^{(1)}\left(V_i, t\right)\right]\frac{\partial^2 d\left(V_a, V_b, t\right)}{\partial\theta^2}d\overrightarrow{V}$$

$$= -\sum_{\substack{a,b \\ a<b}}\left[\prod_{i\neq a,b}\int_{\mathbb{R}} p^{(1)}\left(V_i, t\right)dV_i\right]\left[\int_{\mathbb{R}^2} p^{(1)}\left(V_a, t\right)p^{(1)}\left(V_b, t\right)\frac{\partial^2 d\left(V_a, V_b, t\right)}{\partial\theta^2}dV_a dV_b\right]$$

$$= -\sum_{\substack{a,b \\ a<b}}\left[\int_{\mathbb{R}^2} p^{(1)}\left(V_a, t\right)p^{(1)}\left(V_b, t\right)\frac{\partial^2 d\left(V_a, V_b, t\right)}{\partial\theta^2}dV_a dV_b\right]$$

$$= -\frac{\eta\left(\eta-1\right)}{2}\int_{\mathbb{R}^2} p^{(1)}\left(V,t\right)p^{(1)}\left(V',t\right)\frac{\partial^2 d\left(V,V',t\right)}{\partial\theta^2}dVdV'$$

having used the normalization condition of the function $p^{(1)}$ and the fact that the function $d\left(\cdot,\cdot,t\right)$ does not depend on the neural indices in the case of a fully connected network.

C.3 Third term

This term can be calculated exactly as follows:

$$-\int_{\mathbb{R}^\eta}\widetilde{P}\left(\overrightarrow{V},t\right)\frac{\partial^2 log\left[\prod_{i=0}^{\eta-1}p^{(1)}\left(V_i,t\right)\right]}{\partial\theta^2}d\overrightarrow{V}$$

$$= -\sum_{i=0}^{\eta-1}\int_{\mathbb{R}^\eta}\widetilde{P}\left(\overrightarrow{V},t\right)\frac{\partial^2 logp^{(1)}\left(V_i,t\right)}{\partial\theta^2}d\overrightarrow{V}$$

$$= -\sum_{\substack{a,b\\a<b}}\sum_{i=0}^{\eta-1}\int_{\mathbb{R}^\eta}\left[\prod_{c\neq a,b}p^{(1)}\left(V_c,t\right)\right]P\left(V_a,V_b,t\right)\frac{\partial^2 logp^{(1)}\left(V_i,t\right)}{\partial\theta^2}d\overrightarrow{V}$$

$$= -\sum_{\substack{a,b\\a<b}}\sum_{i\neq a,b}\left[\prod_{c\neq a,b,i}\int_{\mathbb{R}}p^{(1)}\left(V_c,t\right)dV_c\right]\left[\int_{\mathbb{R}^2}P\left(V_a,V_b,t\right)dV_adV_b\right]\left[\int_{\mathbb{R}}p^{(1)}\left(V_i,t\right)\frac{\partial^2 logp^{(1)}\left(V_i,t\right)}{\partial\theta^2}dV_i\right]$$

$$-\sum_{\substack{a,b\\a<b}}\left[\prod_{c\neq a,b}\int_{\mathbb{R}}p^{(1)}\left(V_c,t\right)dV_c\right]\left\{\int_{\mathbb{R}}\left[\frac{\partial^2 logp^{(1)}\left(V_a,t\right)}{\partial\theta^2}\int_{\mathbb{R}}P\left(V_a,V_b,t\right)dV_b\right]dV_a\right\}$$

$$-\sum_{\substack{a,b\\a<b}}\left[\prod_{c\neq a,b}\int_{\mathbb{R}}p^{(1)}\left(V_c,t\right)dV_c\right]\left\{\int_{\mathbb{R}}\left[\frac{\partial^2 logp^{(1)}\left(V_b,t\right)}{\partial\theta^2}\int_{\mathbb{R}}P\left(V_a,V_b,t\right)dV_a\right]dV_b\right\}$$

$$=0$$

having used the constraint A.2 on the function P.

C.4 Fourth term

Using again the approximation C.5, we obtain:

$$-\int_{\mathbb{R}^\eta}\widetilde{P}\left(\overrightarrow{V},t\right)\frac{\partial^2 log\left[1+D\left(\overrightarrow{V},t\right)\right]}{\partial\theta^2}d\overrightarrow{V}\approx-\int_{\mathbb{R}^\eta}\widetilde{P}\left(\overrightarrow{V},t\right)\frac{\partial^2 D\left(\overrightarrow{V},t\right)}{\partial\theta^2}d\overrightarrow{V}$$

$$=-\sum_{\substack{a,b\\a<b}}\sum_{\substack{a',b'\\a'<b'}}\int_{\mathbb{R}^\eta}\left[\prod_{c\neq a,b}p^{(1)}\left(V_c,t\right)\right]P\left(V_a,V_b,t\right)\frac{\partial^2 d\left(V_{a'},V_{b'},t\right)}{\partial\theta^2}d\overrightarrow{V}$$

Now, we can see that many terms are equal to zero due to the constraint A.2. The only non-zero terms are those for which $(a=a')\wedge(b=b')$, therefore we obtain:

$$-\sum_{\substack{a,b\\a<b}}\sum_{\substack{a',b'\\a'<b'}}\int_{\mathbb{R}^\eta}\left[\prod_{c\neq a,b}p^{(1)}\left(V_c,t\right)\right]P\left(V_a,V_b,t\right)\frac{\partial^2 d\left(V_{a'},V_{b'},t\right)}{\partial\theta^2}d\overrightarrow{V}$$

$$= -\sum_{\substack{a,b \\ a<b}} \int_{\mathbb{R}^\eta} \left[\prod_{c \neq a,b} p^{(1)}\left(V_c, t\right) \right] P\left(V_a, V_b, t\right) \frac{\partial^2 d\left(V_a, V_b, t\right)}{\partial \theta^2} d\overrightarrow{V}$$

$$= -\sum_{\substack{a,b \\ a<b}} \left[\prod_{c \neq a,b} \int_{\mathbb{R}} p^{(1)}\left(V_c, t\right) dV_c \right] \int_{\mathbb{R}^2} P\left(V_a, V_b, t\right) \frac{\partial^2 d\left(V_a, V_b, t\right)}{\partial \theta^2} dV_a dV_b$$

$$= -\frac{\eta\left(\eta - 1\right)}{2} \int_{\mathbb{R}^2} P\left(V, V', t\right) \frac{\partial^2 d\left(V, V', t\right)}{\partial \theta^2} dV dV'$$

However this term is of second order since it contains the product $P\left(V, V', t\right) \frac{\partial^2 d(V,V',t)}{\partial \theta^2}$, therefore it does not appear in the first order approximation of the Fisher information. So we will use it in Section C.6, for the calculation of the Fisher information at the second order.

C.5 Putting everything together: First order approximation of the Fisher information

Collecting all these results, the final formula for the Fisher information at the first order of approximation is:

$$\begin{aligned} \mathcal{I}\left(\theta, t\right) \approx & -\eta \int_{\mathbb{R}} p^{(1)}\left(V, t\right) \frac{\partial^2 log p^{(1)}\left(V, t\right)}{\partial \theta^2} dV \\ & - \frac{\eta\left(\eta - 1\right)}{2} \int_{\mathbb{R}^2} p^{(1)}\left(V, t\right) p^{(1)}\left(V', t\right) \frac{\partial^2 d\left(V, V', t\right)}{\partial \theta^2} dV dV' \end{aligned} \tag{C.6}$$

Hence, we have decomposed the high-dimensional integral in formula C.3 into terms of low-dimensional integrals, which can be calculated with standard integration techniques. In the next sections we extend this result to higher order approximations.

C.6 Second order approximation of the Fisher information

We consider again the second term of the Fisher information that we have already calculated in Section C.2, and include the second order term in the Taylor expansion of the logarithm:

$$log\left[1+D\left(\overrightarrow{V},t\right)\right]\approx D\left(\overrightarrow{V},t\right)-\frac{1}{2}D^{2}\left(\overrightarrow{V},t\right) \tag{C.7}$$

Using the same methods already shown in the previous sections, we obtain that the contribution to the Fisher information generated by $-\frac{1}{2}D^{2}\left(\overrightarrow{V},t\right)$ is:

$$\frac{1}{2}\int_{\mathbb{R}^{\eta}}\left[\prod_{i=0}^{\eta-1}p^{(1)}\left(V_{i},t\right)\right]\frac{\partial^{2}D^{2}\left(\overrightarrow{V},t\right)}{\partial\theta^{2}}d\overrightarrow{V}$$

$$=\frac{1}{2}\sum_{\substack{a,b\\a<b}}\sum_{\substack{a',b'\\a'<b'}}\int_{\mathbb{R}^{\eta}}\left[\prod_{i=0}^{\eta-1}p^{(1)}\left(V_{i},t\right)\right]\frac{\partial^{2}\left[d\left(V_{a},V_{b},t\right)d\left(V_{a'},V_{b'},t\right)\right]}{\partial\theta^{2}}d\overrightarrow{V}$$

$$=\frac{1}{2}\left\{\frac{\eta\left(\eta-1\right)}{2}\int_{\mathbb{R}^{2}}p^{(1)}\left(V,t\right)p^{(1)}\left(V',t\right)\frac{\partial^{2}d^{2}\left(V,V',t\right)}{\partial\theta^{2}}dVdV'\right. \tag{C.8}$$

$$+\eta\left(\eta-1\right)\left(\eta-2\right)\int_{\mathbb{R}^{3}}p^{(1)}\left(V,t\right)p^{(1)}\left(V',t\right)p^{(1)}\left(V'',t\right)\frac{\partial^{2}\left[d\left(V,V',t\right)d\left(V,V'',t\right)\right]}{\partial\theta^{2}}dVdV'dV''$$

$$\left.+\frac{\eta\left(\eta-1\right)\left(\eta-2\right)\left(\eta-3\right)}{4}\int_{\mathbb{R}^{4}}p^{(1)}\left(V,t\right)p^{(1)}\left(V',t\right)p^{(1)}\left(V'',t\right)p^{(1)}\left(V''',t\right)\frac{\partial^{2}\left[d\left(V,V',t\right)d\left(V'',V''',t\right)\right]}{\partial\theta^{2}}dVdV'dV''dV'''\right\}$$

It is important to observe that the second order term in C.7 has generated three integrals, where the terms inside the derivative are of the form:

- $d^2\left(V, V', t\right)$
- $d\left(V, V', t\right) d\left(V, V'', t\right)$
- $d\left(V, V', t\right) d\left(V'', V''', t\right)$

These are all the possible different terms that can be generated by $d\left(V_a, V_b, t\right) d\left(V_{a'}, V_{b'}, t\right)$, with $(a < b) \wedge (a' < b')$. Moreover the sum of their corresponding coefficients is:

$$\frac{\eta\left(\eta-1\right)}{2} + \eta\left(\eta-1\right)\left(\eta-2\right) + \frac{\eta\left(\eta-1\right)\left(\eta-2\right)\left(\eta-3\right)}{4} = \left[\frac{\eta\left(\eta-1\right)}{2}\right]^2$$

as it must be, since the double sum:

$$\sum_{\substack{a,b \\ a<b}} \sum_{\substack{a',b' \\ a'<b'}}$$

which appears in C.8 generates exactly $\left[\frac{\eta(\eta-1)}{2}\right]^2$ terms. To conclude, if we put together all the results collected up to now, including the second order term found in Section C.4, we obtain that the Fisher information at the second order is:

$$\mathcal{I}(\theta,t)$$

$$\approx -\eta\int_{\mathbb{R}} p^{(1)}(V,t)\frac{\partial^2 logp^{(1)}(V,t)}{\partial\theta^2}dV$$

$$-\frac{\eta(\eta-1)}{2}\int_{\mathbb{R}^2}\left[p^{(1)}(V,t)\,p^{(1)}\left(V',t\right)+P\left(V,V',t\right)\right]\frac{\partial^2 d(V,V',t)}{\partial\theta^2}dVdV'$$

$$+\frac{\eta(\eta-1)}{4}\int_{\mathbb{R}^2} p^{(1)}(V,t)\,p^{(1)}\left(V',t\right)\frac{\partial^2 d^2(V,V',t)}{\partial\theta^2}dVdV' \tag{C.9}$$

$$+\frac{\eta(\eta-1)(\eta-2)}{2}\int_{\mathbb{R}^3} p^{(1)}(V,t)\,p^{(1)}\left(V',t\right)p^{(1)}\left(V'',t\right)\frac{\partial^2\left[d(V,V',t)\,d(V,V'',t)\right]}{\partial\theta^2}dVdV'dV''$$

$$+\frac{\eta(\eta-1)(\eta-2)(\eta-3)}{8}\int_{\mathbb{R}^4} p^{(1)}(V,t)\,p^{(1)}\left(V',t\right)p^{(1)}\left(V'',t\right)p^{(1)}\left(V''',t\right)\frac{\partial^2\left[d(V,V',t)\,d(V'',V''',t)\right]}{\partial\theta^2}dVdV'dV''dV'''$$

This is the formula that we use in Section 4.4 for the calculation of the Fisher information.

C.7 Third order approximation of the Fisher information

At the third perturbative order the Taylor expansion of the logarithmic function is:

$$log\left[1+D\left(\overrightarrow{V},t\right)\right]\approx D\left(\overrightarrow{V},t\right)-\frac{1}{2}D^2\left(\overrightarrow{V},t\right)+\frac{1}{3}D^3\left(\overrightarrow{V},t\right)$$

$D^3\left(\overrightarrow{V},t\right)$ can be expressed as:

$$D^3\left(\overrightarrow{V},t\right)=\sum_{\substack{a,b\\a<b}}\sum_{\substack{a',b'\\a'<b'}}\sum_{\substack{a'',b''\\a''<b''}} d(V_a,V_b,t)\,d(V_{a'},V_{b'},t)\,d(V_{a''},V_{b''},t)$$

therefore we have seven different terms of the form:

- $d^3\left(V, V', t\right)$
- $d^2\left(V, V', t\right) d\left(V', V'', t\right)$
- $d^2\left(V, V', t\right) d\left(V'', V''', t\right)$
- $d\left(V, V', t\right) d\left(V', V'', t\right) d\left(V, V'', t\right)$
- $d\left(V, V', t\right) d\left(V', V'', t\right) d\left(V, V''', t\right)$
- $d\left(V, V', t\right) d\left(V', V'', t\right) d\left(V''', V'''', t\right)$
- $d\left(V, V', t\right) d\left(V'', V''', t\right) d\left(V'''', V''''', t\right)$

whose corresponding coefficients are:

- $\frac{\eta(\eta-1)}{2}$
- $2\eta\left(\eta-1\right)\left(\eta-2\right)$
- $\frac{\eta(\eta-1)(\eta-2)(3\eta-5)}{4}$
- $\frac{\eta(\eta-1)(\eta-2)(\eta-3)}{2}$
- $\frac{\eta(\eta-1)(\eta-2)(5\eta-11)}{4}$
- $\eta\left(\eta-1\right)\left(\eta-2\right)^2\left(\eta-3\right)$
- $\frac{\eta(\eta-1)(\eta-2)^2(\eta-3)^2}{8}$

These coefficients can be obtained with a complicated combinatorial calculus, which is not explained here, since it goes beyond the purpose of this thesis. However the reader can easily check that their sum is equal to $\left[\frac{\eta(\eta-1)}{2}\right]^3$, as it must be. To conclude, in order to obtain the Fisher information at the third order, we have to calculate the third order contribution generated by the fourth term of C.4, namely:

$$\frac{1}{2}\int_{\mathbb{R}^\eta}\widetilde{P}\left(\overrightarrow{V},t\right)\frac{\partial^2 D^2\left(\overrightarrow{V},t\right)}{\partial\theta^2}d\overrightarrow{V}$$

$$=\frac{1}{2}\sum_{\substack{a,b\\a<b}}\sum_{\substack{a',b'\\a'<b'}}\sum_{\substack{a'',b''\\a''<b''}}\int_{\mathbb{R}^\eta}\left[\prod_{c\neq a,b}p^{(1)}\left(V_c,t\right)\right]P\left(V_a,V_b,t\right)\frac{\partial^2\left[d\left(V_{a'},V_{b'},t\right)d\left(V_{a''},V_{b''},t\right)\right]}{\partial\theta^2}d\overrightarrow{V}$$

$$=\frac{1}{2}\left\{\sum_{\substack{a,b\\a<b}}\sum_{\substack{a'',b''\\a''<b''}}\int_{\mathbb{R}^\eta}\left[\prod_{c\neq a,b}p^{(1)}\left(V_c,t\right)\right]P\left(V_a,V_b,t\right)\frac{\partial^2\left[d\left(V_a,V_b,t\right)d\left(V_{a''},V_{b''},t\right)\right]}{\partial\theta^2}d\overrightarrow{V}\right.$$

$$\left.+\sum_{\substack{a,b\\a<b}}\sum_{\substack{a',b'\\a'<b'}}\int_{\mathbb{R}^\eta}\left[\prod_{c\neq a,b}p^{(1)}\left(V_c,t\right)\right]P\left(V_a,V_b,t\right)\frac{\partial^2\left[d\left(V_{a'},V_{b'},t\right)d\left(V_a,V_b,t\right)\right]}{\partial\theta^2}d\overrightarrow{V}\right\}$$

since the only non-zero terms are those such that $(a=a')\wedge(b=b')$ or $(a=a'')\wedge(b=b'')$. The two double sums are equivalent, therefore finally we obtain:

$$\sum_{\substack{a,b\\a<b}}\sum_{\substack{a',b'\\a'<b'}}\int_{\mathbb{R}^\eta}\left[\prod_{c\neq a,b}p^{(1)}\left(V_c,t\right)\right]P\left(V_a,V_b,t\right)\frac{\partial^2\left[d\left(V_a,V_b,t\right)d\left(V_{a'},V_{b'},t\right)\right]}{\partial\theta^2}d\overrightarrow{V}$$

$$=\frac{\eta\left(\eta-1\right)}{2}\int_{\mathbb{R}^2}P\left(V,V',t\right)\frac{\partial^2\left[d^2\left(V,V',t\right)\right]}{\partial\theta^2}dVdV'$$

$$+\eta\left(\eta-1\right)\left(\eta-2\right)\int_{\mathbb{R}^3}P\left(V,V',t\right)p^{(1)}\left(V'',t\right)\frac{\partial^2\left[d\left(V,V',t\right)d\left(V,V'',t\right)\right]}{\partial\theta^2}dVdV'dV''$$

$$+\frac{\eta\left(\eta-1\right)\left(\eta-2\right)\left(\eta-3\right)}{4}\int_{\mathbb{R}^4}P\left(V,V',t\right)p^{(1)}\left(V'',t\right)p^{(1)}\left(V''',t\right)\frac{\partial^2\left[d\left(V,V',t\right)d\left(V'',V''',t\right)\right]}{\partial\theta^2}dVdV'dV''dV'''$$

Collecting all the terms, we can write the final expression for the Fisher information at

the third order of approximation. Being very long, we do not show this formula here. The interested reader can work it out by himself.

To conclude, the general structure that emerges from this analysis is clear. We have proved that the Fisher information for correlated neurons can be expanded in terms of integrals with increasing dimension and polynomial coefficients in η with increasing order. Moreover, as we have already said, according to our numerical simulations the function $p^{(1)}(V,t)$ depends weakly on N, while $P(V,V',t)$ depends approximately on $\frac{1}{N}$ (in Chapter 5 we derive the explicit formulae in the case of weak noise). Therefore essentially we have expressed the Fisher information as a series of terms which depend on η and N. For $N \to \infty$, these terms converge to zero, therefore we obtain:

$$\mathcal{I}(\theta,t) = -\eta \int_{\mathbb{R}} p^{(1)}(V,t) \frac{\partial^2 log p^{(1)}(V,t)}{\partial \theta^2} dV$$

as it must be in the thermodynamic limit (compare with 3.43).

Appendix D

Radius of convergence of the sigmoid and arctangent functions

In this section we compute numerically the radius of convergence of two examples of the activation function $S(\cdot)$. For simplicity we consider only the case with $T_{MAX} = 1$ and $V_T = 0$, but this analysis can be extended easily to the most general case.

D.1 The sigmoid function

According to [202], the n-th order derivative of the sigmoid function:

$$S(x) = \frac{1}{1 + e^{-\lambda x}}$$

is:

$$S^{(n)}(x) = \lambda^n \sum_{k=1}^{n} (-1)^{k-1} A(n, k-1) [S(x)]^k [1 - S(x)]^{n+1-k}$$

where $A(n,k)$ are the so called *Eulerian numbers* [203]. Now we can rewrite this expression in the following way:

$$\begin{aligned} S^{(n)}(x) =& \lambda^n S(x)[1-S(x)]^n \sum_{k=1}^{n} (-1)^{k-1} A(n,k-1)[S(x)]^{k-1}[1-S(x)]^{-(k-1)} \\ =& \lambda^n S(x)[1-S(x)]^n \sum_{k=0}^{n-1} (-1)^{k} A(n,k)[S(x)]^{k}[1-S(x)]^{-k} \\ =& \lambda^n S(x)[1-S(x)]^n \sum_{k=0}^{n-1} A(n,k)\left(-e^{-\lambda x}\right)^{-k} \end{aligned}$$

Now from [204], we know that:

$$Li_{-n}(x) = \frac{x^n}{(1-x)^{n+1}} \sum_{k=0}^{n-1} A(n,k)\, x^{-k}, \qquad n>0,\ |x|<1 \tag{D.1}$$

where $Li_{-n}(\cdot)$ represents the so called *polylogarithm* (with negative order). Here we have omitted the n-th term of the sum since $A(n,n)=0\ \forall n>0$. So we can write:

$$S^{(n)}(x) = \lambda^n S(x)[1-S(x)]^n \frac{\left(1+e^{-\lambda x}\right)^{n+1}}{\left(-e^{-\lambda x}\right)^{n}} Li_{-n}\left(-e^{-\lambda x}\right) = (-\lambda)^n Li_{-n}\left(-e^{-\lambda x}\right) \tag{D.2}$$

This result is true only for $\left|-e^{-\lambda x}\right|<1$, i.e. only for $x>0$. Instead, for $x<0$, we can use the relation $S(-x)=1-S(x)$, from which we deduce that:

- $S^{(n)}(-x) = (-1)^{n-1} S^{(n)}(x),\ \forall n>0$;
- $S(-x)$ has the same radius of convergence of $S(x)$.

So formula D.2 can be used to express $S^{(n)}(x)\ \forall x\neq 0$. Instead for $x=0$ it gives $Li_{-n}(-1)$, that is defined by an analytic continuation of the polylogarithm function. In this way

we can determine $S^{(n)}(0)$. Another way is to use the following property of the Eulerian numbers:

$$\sum_{k=1}^{n}(-1)^{k-1}A(n,k-1)=2^{n+1}\left(2^{n+1}-1\right)\frac{B_{n+1}}{n+1} \tag{D.3}$$

where B_n are the so called *Bernoulli numbers* [205], from which we obtain:

$$S^{(n)}(0)=\frac{\lambda^n}{2^{n+1}}\sum_{k=1}^{n}(-1)^{k-1}A(n,k-1)=\lambda^n\left(2^{n+1}-1\right)\frac{B_{n+1}}{n+1} \tag{D.4}$$

Now we can compute the radius of convergence $R(x_0)$ of the Taylor series:

$$S(x)=\sum_{n=0}^{+\infty}\frac{S^{(n)}(x_0)}{n!}(x-x_0)^n$$

using the Cauchy root test:

$$R(x_0)=\frac{1}{\underset{n\to+\infty}{lim\,sup}\sqrt[n]{\left|\frac{S^{(n)}(x_0)}{n!}\right|}}$$

For $x_0=0$ we obtain:

$$R(0)=\frac{1}{\underset{n\to+\infty}{lim\,sup}\sqrt[n]{\left|\frac{\lambda^n\left(2^{n+1}-1\right)\frac{B_{n+1}}{n+1}}{n!}\right|}}=\frac{\pi}{\lambda}$$

This can be proved after the substitution $n\to 2n-1$ (which is motivated by the fact that $B_{2n+1}=0\;\forall n>0$), using the following asymptotic expansion of the Bernoulli numbers:

$$B_{2n}\sim(-1)^{n-1}4\sqrt{\pi n}\left(\tfrac{n}{\pi e}\right)^{2n},\qquad n\to+\infty$$

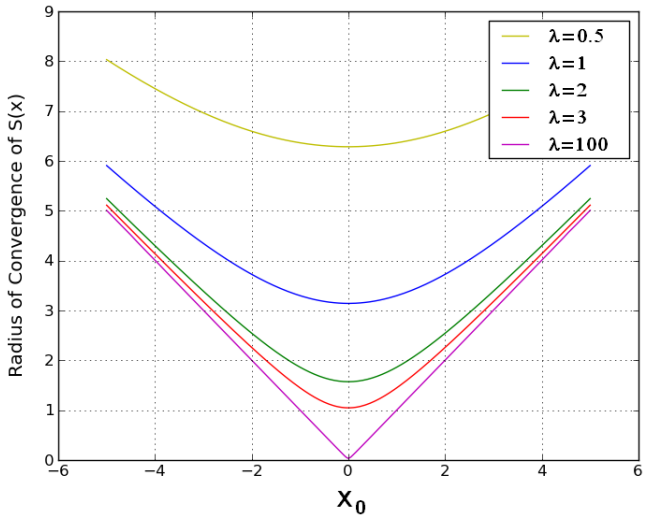

Figure D.1: Radius of convergence R of the Taylor series of the sigmoid function, in terms of the point x_0 about which the expansion is performed. $R(x_0)$ has been computed numerically, for many values of the parameter λ, which determines the slope of the sigmoid function. For large x_0 the radius of converge increases linearly since the sigmoid function is asymptotically flat. Instead for $\lambda \to +\infty$ we obtain $R(0) \to 0$, because in that limit the sigmoid function $S(x)$ becomes a Heaviside step function with a discontinuity in $x = 0$.

and the Stirling approximation of $(2n-1)!$. We are not aware of any asymptotic expansion of $Li_{-n}\left(-e^{-x_0}\right)$ for $n \to +\infty$ and $x_0 \neq 0$, so we have to compute the radius of convergence numerically $\forall x_0 \neq 0$.

Figure D.1 shows the result for different values of λ. From it we can see that the radius of convergence of the Taylor series of $S(x)$ around the point $x = x_0$ increases with x_0. This is reasonable, since the function $S(x)$ becomes flat when x is large. Moreover for large λ it converges to $R(x_0) = |x_0|$ and therefore it is equal to zero only for $x_0 = 0$, as it must be. In fact, for $\lambda \to +\infty$ the function $S(x)$ converges to the Heaviside step function, which has a vertical jump at $x = 0$.

D.2 The arctangent function

Now we calculate the radius of convergence of the arctangent function. According to [206], the n-th order derivative of this function is:

$$arctan^{(n)}(\lambda x) = \lambda^n \frac{(-1)^{n-1}(n-1)!}{\left[1+(\lambda x)^2\right]^{\frac{n}{2}}} \sin\left[n \arcsin\left(\frac{1}{\sqrt{1+(\lambda x)^2}}\right)\right]$$

So from the root test we obtain:

$$R(x_0) = \frac{\sqrt{1+(\lambda x_0)^2}}{\lambda \underset{n\to+\infty}{lim\ sup} \frac{\sqrt[n]{\left|\sin\left[n \arcsin\left(\frac{1}{\sqrt{1+(\lambda x_0)^2}}\right)\right]\right|}}{\sqrt[n]{n}}}$$

Now, since:

$$\underset{n\to+\infty}{lim} \sqrt[n]{\left|\sin\left(n \arcsin\left(\frac{1}{\sqrt{1+(\lambda x_0)^2}}\right)\right)\right|} = 1$$

due to the fact that:

$$\left|\sin\left[n \arcsin\left(\frac{1}{\sqrt{1+(\lambda x_0)^2}}\right)\right]\right| \in [0,1]$$

and moreover $\underset{n\to+\infty}{lim} \sqrt[n]{n} = 1$, we obtain finally:

$$R(x_0) = \frac{1}{\lambda}\sqrt{1+(\lambda x_0)^2}$$

Therefore the radius of convergence increases with x_0, as it must be. Moreover in the limit $\lambda \to +\infty$ it gives $R(x_0) = |x_0|$, as with the sigmoid function.

Appendix E

Higher order correlations for a fully connected neural network

Here we show how it is possible to use the perturbative expansion to calculate the higher order correlations between the neurons. For simplicity, we consider only the simplest case, namely a fully connected network, even if this analysis could be extended to more complicated connectivity matrices. Moreover we want to avoid long expressions for the joint cumulants, therefore we consider only the expansion of the membrane potential at the first perturbative order. In principle this calculation can be performed at any perturbative order, but starting from the second order (namely from the third order terms in the covariance) the functions $Z(t)$ and $\overrightarrow{H}(t)$ in general introduce inhomogeneities in the covariance structure of the network, therefore the higher order correlations should be calculated using combinatorial techniques applied to the Isserlis' theorem. In this section we avoid the issue and we focus only on the first order perturbations. In this case the probability density of every $V_i(t)$ is normal and this is true also for the quantities $V_i(t) - \overline{V}_i(t)$, which have all zero mean and the same variance, that we call $Var(V(t))$. Since they have zero mean we can use the Isserlis' theorem, that for a fully connected network gives simply:

$$\mathbb{E}\left[\prod_{j=0}^{n-1}\left(V_{i_j}(t)-\overline{V}_{i_j}(t)\right)\right]=\begin{cases}0, & n \text{ odd}\\ \\ \frac{n!}{2^{\frac{n}{2}}\left(\frac{n}{2}\right)!}\left[Cov\left(V_i(t),V_j(t)\right)\right]^{\frac{n}{2}}, & n \text{ even}\end{cases} \tag{E.1}$$

because in this case all the pairs of neurons are equivalent, since they are all-to-all connected (instead, if the network is not fully connected, the connected pairs give a different contribution with the Isserlis' theorem compared to the disconnected pairs). Moreover, the central absolute moments of a normal distribution are:

$$\mathbb{E}\left[\left|V_{i_j}(t)-\overline{V}_{i_j}(t)\right|^n\right]=\frac{2^{\frac{n}{2}}\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}}\left[Var\left(V(t)\right)\right]^{\frac{n}{2}} \tag{E.2}$$

and if n is even we have $\Gamma\left(\frac{n+1}{2}\right)=\frac{n!}{2^n\left(\frac{n}{2}\right)!}\sqrt{\pi}$. Therefore putting everything together we obtain:

$$Corr_n\left(V_{i_0}(t),V_{i_1}(t),...,V_{i_{n-1}}(t)\right)=\begin{cases}0, & n \text{ odd}\\ \\ \left[Corr_2\left(V_i(t),V_j(t)\right)\right]^{\frac{n}{2}}, & n \text{ even}\end{cases} \tag{E.3}$$

From this result it is interesting to observe that if there is a perfect stochastic synchronization between pairs of neurons, then it is "propagated" to all the higher order correlations with even order, namely $Corr_2\left(V_i(t),V_j(t)\right)=1$ implies $Corr_n\left(V_{i_0}(t),V_{i_1}(t),...,V_{i_{n-1}}(t)\right)=1, \forall n$ even. It is also curious to observe that all the odd order correlations are always equal to zero, even if the size of the network is finite.

Appendix F

Proof that formula 5.55 gives real functions

In this appendix we want to prove that the quantities $\Phi_{ij}(t)$ and $\left[\Phi(t)\Phi^T(t)\right]_{ij}$, given by formula 5.55, are real functions. The proof can be divided into four cases, namely when R and S are both even, both odd, R even and S odd, or vice versa. Here we analyze only the first case, while the others can be proved in a similar way.

So, if R and S are both even, the function $\Phi_{ij}(t)$, according to 5.55, can be equivalently rewritten as:

$$\Phi_{ij}(t) = \sum_{x=0}^{R-1}\sum_{y=0}^{S-1} e^{\left[-\frac{1}{\tau}+e_{xS+y}S'(\mu)\right]t} f_{i,j,xS+y} \tag{F.1}$$

where now the subscripts are separated by commas, in order to avoid confusion. Defining:

$$g_{x,y}^{ij} = e^{\left[-\frac{1}{\tau}+e_{xS+y}S'(\mu)\right]t} f_{i,j,xS+y}$$

formula F.1 can be rewritten in the following symmetric way, with respect to R and S:

$$\begin{aligned}\Phi_{ij}\left(t\right) =& g_{0,0}^{ij} + g_{\frac{R}{2},0}^{ij} + g_{0,\frac{S}{2}}^{ij} + g_{\frac{R}{2},\frac{S}{2}}^{ij} \\ &+ \sum_{x=1}^{\frac{R}{2}-1}\left[g_{x,0}^{ij} + g_{R-x,0}^{ij}\right] + \sum_{y=1}^{\frac{S}{2}-1}\left[g_{0,y}^{ij} + g_{0,S-y}^{ij}\right] + \sum_{y=1}^{\frac{S}{2}-1}\left[g_{\frac{R}{2},y}^{ij} + g_{\frac{R}{2},S-y}^{ij}\right] + \sum_{x=1}^{\frac{R}{2}-1}\left[g_{x,\frac{S}{2}}^{ij} + g_{R-x,\frac{S}{2}}^{ij}\right] \\ &+ \sum_{x=1}^{\frac{R}{2}-1}\sum_{y=1}^{\frac{S}{2}-1}\left[g_{x,y}^{ij} + g_{R-x,S-y}^{ij}\right] + \sum_{x=1}^{\frac{R}{2}-1}\sum_{y=1}^{\frac{S}{2}-1}\left[g_{x,S-y}^{ij} + g_{R-x,y}^{ij}\right] \end{aligned} \tag{F.2}$$

The quantities $g_{0,0}^{ij}$, $g_{\frac{R}{2},0}^{ij}$, $g_{0,\frac{S}{2}}^{ij}$ and $g_{\frac{R}{2},\frac{S}{2}}^{ij}$ are real numbers. Moreover, since:

$$f_{i,j,xS+y} = e^{2\pi\left\{\frac{x}{R}\left(\left\lfloor\frac{i}{S}\right\rfloor - \left\lfloor\frac{j}{S}\right\rfloor\right) + \frac{y}{S}(i-j)\right\}\iota} = f_{i,j,(R-x)S+(S-y)}^{*}$$

$$f_{i,j,xS+(S-y)} = e^{2\pi\left\{\frac{x}{R}\left(\left\lfloor\frac{i}{S}\right\rfloor - \left\lfloor\frac{j}{S}\right\rfloor\right) - \frac{y}{S}(i-j)\right\}\iota} = f_{i,j,(R-x)S+y}^{*}$$

for $0 \leq y < S$ and $0 \leq x < R$, and also, according to 5.53:

$$e_{xS+y} = \sum_{k=0}^{S-1}\sum_{l=0}^{R-1} e^{2\pi\left(\frac{yk}{S} + \frac{xl}{R}\right)\iota} b_k^{(l)} = e_{(R-x)S+(S-y)}^{*}$$

$$e_{xS+(S-y)} = \sum_{k=0}^{S-1}\sum_{l=0}^{R-1} e^{2\pi\left(-\frac{yk}{S} + \frac{xl}{R}\right)\iota} b_k^{(l)} = e_{(R-x)S+y}^{*}$$

we conclude that:

$$g^{ij}_{x,y} = \left(g^{ij}_{R-x,S-y}\right)^*$$

$$g^{ij}_{x,S-y} = \left(g^{ij}_{R-x,y}\right)^*$$

For this reason, all the quantities in the square parenthesis in formula F.2 are real, and therefore also $\Phi_{ij}(t)$. A similar proof can be obtained for $\left[\Phi(t)\Phi^T(t)\right]_{ij}$, and in the cases when only one of R and S, or both, are odd.

Bibliography

[1] N. Burger. Limitless. Relativity Media, Virgin Produced, 2011.

[2] FACETS-ITN: Fast Analog Computing with Emergent Transient States - Initial Training Network. facets-itn.eu.

[3] G. van Rossum. Python. www.python.org.

[4] T. Oliphant. NumPy. www.numpy.org.

[5] E. Jones, T. Oliphant, and P. Peterson. SciPy. www.scipy.org.

[6] Los Alamos National Laboratory. NetworkX. networkx.github.io.

[7] J. Hunter. matplotlib. matplotlib.org.

[8] MathWorks Incorporation. MATLAB. Natick, Massachusetts, United States, www.mathworks.com.

[9] P. Tsui. Fast EM_GM Algorithm. PAMI Research Group, Department of Electrical and Computer Engineering, University of Waterloo, 2006. www.mathworks.com/matlabcentral/fileexchange/9659-fast-emgm.

[10] P. D. MacLean and V. A. Kral. *A Triune Concept of the Brain and Behaviour*. Clarence M. Hincks Memorial Lectures. Ontario Mental Health Foundation, 1973.

[11] P. D. MacLean. *The Triune Brain in Evolution: Role in Paleocerebral Functions*. Springer, 1990.

[12] V. B. Mountcastle. Modality and Topographic Properties of Single Neurons of Cat's Somatic Sensory Cortex. *Journal of Neurophysiology*, 20:408–434, 1957.

[13] F. Grimbert. *Mesoscopic Models of Cortical Structures*. PhD thesis, 2007.

[14] W. Weaver. Science and Complexity. *American Scientist*, 36(4):536–544, 1948.

[15] Y. Bar-Yam. *Dynamics of Complex Systems*. Addison-Wesley Studies in Nonlinearity. Westview Press, 1997.

[16] H. Haken. *Information and Self-Organization: A Macroscopic Approach to Complex Systems*. Springer Complexity. Springer London, Limited, 2006.

[17] C. G. Langton. Computation at the Edge of Chaos: Phase Transitions and Emergent Computation. *Physica D: Nonlinear Phenomena*, 42(1-3):12–37, June 1990.

[18] M. Mitchell, P. T. Hraber, and J. P. Crutchfield. Revisiting the Edge of Chaos: Evolving Cellular Automata to Perform Computations. *Complex Systems*, 7:89–130, 1993.

[19] M. M. Waldrop. *Complexity: The Emerging Science at the Edge of Order and Chaos*. Touchstone Book. Simon & Schuster, 1993.

[20] R. Lewin. *Complexity: Life at the Edge of Chaos*. University Of Chicago Press, 1st edition, February 2000.

[21] C. W. Reynolds. Flocks, Herds and Schools: A Distributed Behavioral Model. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '87, pages 25–34, New York, NY, USA, 1987. ACM.

[22] G. Nicolis and I. Prigogine. *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations*. A Wiley-Interscience Publication. Wiley, 1977.

[23] Frégnac Y., Rudolph M., Davison A. P., and Destexhe A. Complexity in Neuronal Networks. In François Képès, editor, *Biological Networks*. 2007.

[24] S.-I. Amari. Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields. *Biological Cybernetics*, 27:77–87, 1977.

[25] S.-I. Amari. Field Theory of Self-Organizing Neural Nets. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(5):741–748, 1983.

[26] S. Coombes. Waves, Bumps, and Patterns in Neural Field Theories. *Biological Cybernetics*, 93(2):91–108, August 2005.

[27] C. B. Curtis, Z. H. Victor, S. Yoshiko, and G. Kirsty. Synaptic Plasticity in a Cerebellum-Like Structure Depends on Temporal Order. *Nature*, 387(6630):278–281, 1997.

[28] G. Q. Bi and M. M. Poo. Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type. *Journal of Neuroscience*, 18(24):10464–10472, 1998.

[29] G. Q. Bi and M. M. Poo. Synaptic Modification by Correlated Activity: Hebb's Postulate Revisited. *Annual Review of Neuroscience*, 24:139–166, January 2001.

[30] L. F. Abbott and S. B. Nelson. Synaptic Plasticity : Taming the Beast. *Nature Neuroscience Supplement*, 3:1178–1183, November 2000.

[31] J. G. White, E. Southgate, J. N. Thomson, and S. Brenner. The Structure of the Nervous System of the Nematode Caenorhabditis Elegans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 314(1165):1–340, 1986.

[32] L. R. Varshney, B. L. Chen, E. Paniagua, D. H. Hall, and D. B. Chklovskii. Structural Properties of the Caenorhabditis Elegans Neuronal Network. *PLoS Computational Biology*, 7(2):e1001066, February 2011.

[33] Mouse Connectome Project. `www.mouseconnectome.org`.

[34] K. L. Briggman, M. Helmstaedter, and W. Denk. Wiring Specificity in the Direction-Selectivity Circuit of the Retina. *Nature*, 471(7337):183–188, March 2011.

[35] D. D. Bock, W.-C. A. Lee, A. M. Kerlin, M. L. Andermann, G. Hood, A. W. Wetzel, S. Yurgenson, E. R. Soucy, H. S. Kim, and R. C. Reid. Network Anatomy and In Vivo Physiology of Visual Cortical Neurons. *Nature*, 471(7337):177–182, March 2011.

[36] G. A. P. C. Burns and M. P. Young. Analysis of the Connectional Organization of Neural Systems Associated with the Hippocampus in Rats. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 355(1393):55–70, 2000.

[37] O. Schmitt, P. Eipert, K. Philipp, R. Kettlitz, G. Fuellen, and A. Wree. The Intrinsic Connectome of the Rat Amygdala. *Front Neural Circuits*, 6:81, 2012.

[38] J. W. Scannell, C. Blakemore, and M. P. Young. Analysis of Connectivity in the Cat Cerebral Cortex. *Journal of Neuroscience*, 15:1463, 1995.

[39] J. W. Scannell, G. A. P. C. Burns, C. C. Hilgetag, M. A. O'Neil, and M. P. Young. The Connectional Organization of the Cortico-Thalamic System of the Cat. *Cerebral Cortex*, 9(3):277–299, April 1999.

[40] Daniel J. F. and David C. V. E. Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex*, pages 1–47, 1991.

[41] O. Sporns, G. Tononi, and R. Kötter. The Human Connectome: A Structural Description of the Human Brain. *PLoS Computational Biology*, 1(4):e42, September 2005.

[42] O. Sporns. The Human Connectome: A Complex Network. *Annals of the New York Academy of Sciences*, 1224(1):109–125, 2011.

[43] Human Connectome Project. `www.humanconnectomeproject.org`.

[44] C. Stam and J. Reijneveld. Graph Theoretical Analysis of Complex Networks in the Brain. *Nonlinear Biomedical Physics*, 1(1):3+, 2007.

[45] E. Bullmore and O. Sporns. Complex Brain Networks: Graph Theoretical Analysis of Structural and Functional Systems. *Nature Reviews. Neuroscience*, 10(3):186–198, March 2009.

[46] D. S. Bassett and E. T. Bullmore. Human Brain Networks in Health and Disease. *Current Opinion in Neurology*, 22:340–347, 2009.

[47] D. J. Watts and S. H. Strogatz. Collective Dynamics of 'Small-World' Networks. *Nature*, 393(6684):409–10, 1998.

[48] S.-I. Amari. Characteristics of Random Nets of Analog Neuron-Like Elements. *IEEE Transactions on Systems, Man, and Cybernetics*, 2(5):643–657, 1972.

[49] H. R. Wilson and J. D. Cowan. Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophysics*, pages 1–24, 1972.

[50] H. R. Wilson and J. D. Cowan. A Mathematical Theory of the Functional Dynamics of Cortical and Thalamic Nervous Tissue. *Kybernetik*, 13:55–80, 1973.

[51] J. Touboul, G. Hermann, and O. Faugeras. Noise-Induced Behaviors in Neural Mean Field Dynamics. *SIAM Journal on Applied Dynamical Systems*, 11(1):49–81, 2012.

[52] A. L. Hodgkin and A. F. Huxley. A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve. *The Journal of Physiology*, 117(4):500–544, August 1952.

[53] S.A. Kauffman. *At Home in the Universe: The Search for Laws of Self-Organization and Complexity*. Oxford paperbacks. Oxford University Press Australia & New Zealand, 1995.

[54] O. Sporns, D. Chialvo, M. Kaiser, and C. Hilgetag. Organization, Development and Function of Complex Brain Networks. *Trends in Cognitive Sciences*, 8(9):418–425, September 2004.

[55] S. C. Ponten, A. Daffertshofer, A. Hillebrand, and C. J. Stam. The Relationship Between Structural and Functional Connectivity: Graph Theoretical Analysis of an EEG Neural Mass Model. *NeuroImage*, 52(3):985–994, 2010.

[56] V. Pernice, B. Staude, S. Cardanobile, and S. Rotter. How Structure Determines Correlations in Neuronal Networks. *PLoS Computational Biology*, 7(5):e1002059, May 2011.

[57] M. Koch. An Investigation of Functional and Anatomical Connectivity Using Magnetic Resonance Imaging. *NeuroImage*, 16(1):241–250, May 2002.

[58] S. B. Eickhoff, S. Jbabdi, S. Caspers, A. R. Laird, P. T. Fox, K. Zilles, and T. E. J. Behrens. Anatomical and Functional Connectivity of Cytoarchitectonic Areas within the Human Parietal Operculum. *The Journal of Neuroscience*, 30:6409 – 6421, 2010. We acknowledge funding by the Human Brain Project/Neuroinformatics Research (National Institute of Biomedical Imaging and Bioengineering, National Institute of Neurological Disorders and Stroke, National Institute of Mental Health; to K.Z.), the Human Brain Project (R01-MH074457-01A1; to S. B. E.) and the Helmholz Initiative on Systems-Biology "The Human Brain Model" (to K.Z. and S.B.E.).

[59] J. Cabral, E. Hugues, M. L. Kringelbach, and G. Deco. Modeling the Outcome of Structural Disconnection on Resting-State Functional Connectivity. *NeuroImage*, 62(3):1342–1353, 2012.

[60] C. Joana, M. L. Kringelbach, and G. Deco. Functional Graph Alterations in Schizophrenia: A Result From a Global Anatomic Decoupling? *Pharmacopsychiatry*, 45 Suppl. 1:557–564, May 2012.

[61] B. Cessac. Increase in Complexity in Random Neural Networks. *Journal de Physique I (France)*, 5(3):409–432, 1995.

[62] B. Cessac, O. Mazet, M. Samuelides, and H. Soula. Mean Field Theory for Random Recurrent Spiking Neural Networks. In *NOLTA '05 (Non Linear Theory and its Applications)*, pages 18–21, Brugge, Belgium, 2005.

[63] M. Samuelides and B. Cessac. Random Recurrent Neural Networks Dynamics. *European Physical Journal Special Topics*, 142(1):89–122, 2007.

[64] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul. Mean-Field Description and Propagation of Chaos in Networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. *The Journal of Mathematical Neuroscience*, 2(1):10, May 2012. This work was partially supported by the ERC grant #227747 NerVi, the FACETSITN Marie-Curie Initial Training Network #237955 and the IP project BrainScaleS #269921.

[65] J. Touboul. Propagation of Chaos in Neural Fields. Submitted to The Annals of Applied Probability, 2011.

[66] G. Piccinini and A. Scarantino. Computation vs. Information Processing: Why Their Difference Matters to Cognitive Science. *Studies in History and Philosophy of Science Part A*, 41(3):237– 246, 2010.

[67] M. Tegmark. Why the Brain is Probably Not a Quantum Computer. *Information Sciences*, 128(3-4):155–179, 2000.

[68] L. M. Ricciardi and H. Umezawa. Brain and Physics of Many-Body Problems. *Kybernetik*, 4:44–48, 1967.

[69] M. Jibu and K. Yasue. *Quantum Brain Dynamics and Consciousness: An Introduction*. Advances in Consciousness Research Series. John Benjamins, 1995.

[70] S. Hameroff and R. Penrose. Orchestrated Reduction of Quantum Coherence in Brain Microtubules: A Model for Consciousness. *Mathematics and Computers in Simulation*, 40(3-4):453–480, 1996.

[71] S. R. Hameroff and R. Penrose. Conscious Events as Orchestrated Space-Time Selections. *Journal of Consciousness Studies*, 3(1):36–53, 1996.

[72] S. Hagan, S. R. Hameroff, and J. A. Tuszyński. Quantum Computation in Brain Microtubules: Decoherence and Biological Feasibility. *Physical Review E*, 65:061901+, June 2002.

[73] M. Sarovar, A. Ishizaki, G. R. Fleming, and K. B. Whaley. Quantum Entanglement in Photosynthetic Light-Harvesting Complexes. *Nature Physics*, 6(6):462–467, June 2010.

[74] E. M. Gauger, E. Rieper, J. J. L. Morton, S. C. Benjamin, and V. Vedral. Sustained Quantum Coherence and Entanglement in the Avian Compass. *Physical Review Letters*, 106:040503, January 2011.

[75] P.-O. Löwdin. Quantum Genetics and the Aperiodic Solid: Some Aspects on the Biological Problems of Heredity, Mutations, Aging, and Tumors in View of the Quantum Theory of the DNA Molecule. volume 2 of *Advances in Quantum Chemistry*, pages 213–360. Academic Press, 1966.

[76] N. Lambert, Y.-N. Chen, Y.-C. Cheng, C.-M. Li, G.-Y. Chen, and F. Nori. Quantum Biology. *Nature Physics*, 9(1):10–18, January 2013.

[77] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland. Quantum Zeno Effect. *Physical Review A*, 41:2295–2300, March 1990.

[78] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000.

[79] P. W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Journal on Computing*, 26(5):1484–1509, October 1997.

[80] P. W. Shor. Why Haven't More Quantum Algorithms Been Found? *Journal of the ACM*, 50(1):87–90, January 2003.

[81] C. Koch and K. Hepp. Quantum Mechanics in the Brain. *Nature*, 440(7084):611, March 2006.

[82] R. Herken. *The Universal Turing Machine: A Half-Century Survey*. Springer-Verlag, 1995.

[83] B. J. Copeland. The Church-Turing Thesis. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, 2002.

[84] M. Mohri. Finite-State Transducers in Language and Speech Processing. *Computational Linguistics*, 23(2):269–311, June 1997.

[85] L. A. Rubel. The Extended Analog Computer. *Advances in Applied Mathematics*, 14(1):39–50, March 1993.

[86] C. E. Shannon. Mathematical Theory of the Differential Analyzer. *Journal of Mathematics and Physics*, 20:337–354, 1941.

[87] J. W. Mills. Polymer Processors. In *2nd International Symposium on Non-Silicon Computation*, 2000.

[88] J. W. Mills. The Nature of the Extended Analog Computer. *Physica D: Nonlinear Phenomena*, 237(9):1235–1256.

[89] A. Adamatzky. From Reaction-Diffusion to Physarum Computing. *Natural Computing*, 8:431–447, 2009.

[90] A. Adamatzky and M. Komosinski. *Artificial Life Models in Hardware*. Springer London, 2009.

[91] J. W. Backus. Can Programming Be Liberated From the von Neumann Style? A Functional Style and its Algebra of Programs. *Communications of the ACM*, 21(8):613–641, August 1978.

[92] S. A. McKee. Reflections on the Memory Wall. In *CF '04: Proceedings of the 1st Conference on Computing Frontiers*, New York, NY, USA, 2004. ACM Press.

[93] R. R. Schaller. Moore's Law: Past, Present, and Future. *IEEE Spectrum*, 34(6):52–59, June 1997.

[94] A. Adamatzky, B. De Lacy Costello, and T. Asai. *Reaction-Diffusion Computers*. Elsevier Science, 2005.

[95] N. G. Rambidi, T. O.-O. Kuular, and E. E. Makhaeva. Information-Processing Capabilities of Chemical Reaction-Diffusion Systems. 1: Belousov-Zhabotinsky Media in Hydrogel Matrices and on Solid Supports. *Advanced Materials for Optics and Electronics*, 8(4):163–171, 1998.

[96] N. G. Rambidi and D. Yakovenchuk. Information-Processing Capabilities of Chemical Reaction-Diffusion Systems. 2: Finding Paths in a Labyrinth Based on Reaction-Diffusion Media. *Advanced Materials for Optics and Electronics*, 9(1):27–34, 1999.

[97] A. Adamatzky. *Collision-Based Computing*. Springer London, 2002.

[98] W. Maass. Networks of Spiking Neurons: the Third Generation of Neural Network Models. *Transactions of the Society for Computer Simulation International*, 14(4):1659–1671, December 1997.

[99] W. S. McCulloch and W. Pitts. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 5(4):115–133, December 1943.

[100] B. J. MacLennan. A Review of Analog Computing. Technical Report UT-CS-07-601, Department of Electrical Engineering & Computer Science, University of Tennessee, Knoxville, 2007.

[101] H. T. Siegelmann. Computation Beyond the Turing Limit. *Science*, 268(5210):545–548, April 1995.

[102] J. Cabessa and H. T. Siegelmann. Evolving Recurrent Neural Networks are Super-Turing. In *International Joint Conference on Neural Networks*, pages 3200–3206. IEEE, 2011.

[103] W. Maass and E. D. Sontag. Analog Neural Nets with Gaussian or Other Common Noise Distributions Cannot Recognize Arbitrary Regular Languages. *Neural Computation*, 11(3):771–782, 1999.

[104] L. Blum, F. Cucker, M. Shub, and S. Smale. Complexity and Real Computation: A Manifesto. *International Journal of Bifurcation and Chaos*, 6:3–26, 1996.

[105] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation*. Springer, October 1997.

[106] S. Aaronson. Guest Column: NP-Complete Problems and Physical Reality. *SIGACT News*, 36(1):30–52, March 2005.

[107] J. D. Hamkins. Infinite Time Turing Machines. *Minds and Machines*, 12(4):521–539, November 2002.

[108] S. Franklin and M. Garzon. *Neural Computability*, volume 1 of *Progress in Neural Networks*, pages 127–145. Ablex, 1991.

[109] S. Franklin and M. Garzon. *Computation on Graphs*, volume 2 of *Progress in Neural Networks*, chapter 13. Ablex, 1994.

[110] J. R. Searle. Minds, Brains, and Programs. *Behavioral and Brain Sciences*, 3(3):417–457, 1980.

[111] S. Harnad. The Symbol Grounding Problem. *Physica D: Nonlinear Phenomena*, 42(1-3):335–346, June 1990.

[112] L. Steels. The Symbol Grounding Problem Has Been Solved. So What's Next? *Symbols, Embodiment and Meaning. Academic Press, New Haven*, 2008.

[113] M. R. Quillian. Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities. *Behavioral Science*, 12(5):410–430, September 1967.

[114] A. M. Collins and E. F. Loftus. A Spreading-Activation Theory of Semantic Processing. *Psychological Review*, 82(6):407–428, November 1975.

[115] S.-I. Amari. Neural Theory of Association and Concept-Formation. *Biological Cybernetics*, 26:175–185, 1977.

[116] H. Cohen and C. Lefebvre. *Handbook of Categorization in Cognitive Science*. Elsevier Science, 2005.

[117] E. Tulving. Memory and Consciousness. *Canadian Psychology*, 25:1 – 12, 1985.

[118] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27, 1948.

[119] J. Lizier, M. Prokopenko, and A. Zomaya. The Information Dynamics of Phase Transitions in Random Boolean Networks. In S. Bullock, J. Noble, R. Watson, and M. A. Bedau, editors, *Artificial Life XI: Proceedings of the 11th International Conference on the Simulation and Synthesis of Living Systems*, pages 374–381. MIT Press, Cambridge, MA, 2008.

[120] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya. Detecting Non-Trivial Computation in Complex Dynamics. In Almeida, Luis M. Rocha, Ernesto Costa, Inman Harvey, and António Coutinho, editors, *Proceedings of the 9th European Conference on Artificial Life (ECAL 2007)*, volume 4648 of *Lecture Notes in Computer Science*, pages 895–904, Berlin / Heidelberg, 2007. Springer.

[121] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya. Local Measures of Information Storage in Complex Distributed Computation. *Information Sciences*, 208:39–54, November 2012.

[122] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya. Information Modification and Particle Collisions in Distributed Computation. *Chaos*, 20(3):037109+, 2010.

[123] H. White. The Entropy of a Continuous Distribution. *The Bulletin of Mathematical Biophysics*, 27(1):135–143, 1965.

[124] C. Vignat and J.-F. Bercher. Analysis of Signals in the Fisher-Shannon Information Plane. *Physics Letters A*, 312(1-2):27–33, 2003.

[125] A. Pouget, P. Dayan, and R. Zemel. Information Processing with Population Codes. *Nature Reviews Neuroscience*, 1:125–132, 2000.

[126] E. T. Rolls and G. Deco. *The Noisy Brain: Stochastic Dynamics as a Principle of Brain Function*. Oxford University Press, USA, 1 edition, March 2010.

[127] Y. Bar-Hillel and R. Carnap. Semantic Information. *British Journal for the Philosophy of Science*, 4:147–157, 1953.

[128] R. Carnap and Y. Bar-Hillel. *An Outline of a Theory of Semantic Information*. Technical report. Research Laboratory of Electronics, MIT, 1952.

[129] L. Floridi. Outline of a Theory of Strongly Semantic Information. *Minds and Machines*, 14:197–221, 2004.

[130] S. D'Alfonso. On Quantifying Semantic Information. *Information*, 2(1):61–101, 2011.

[131] Z. W. Pylyshyn. *Computation and Cognition: Toward a Foundation for Cognitive Science*. MIT, Cambridge, MA, USA, 1986.

[132] A. Newell and H. A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. *Communications of the ACM*, 19(3):113–126, March 1976.

[133] H. Putnam. *Brains and Behavior*, pages 211–235. Analytical Philosophy, Second Series. Oxford: Basil Blackwell, 1963.

[134] J. A. Fodor. *The Language of Thought*. The Language & Thought Series. Harvard University Press, 1975.

[135] S. Horst. The Computational Theory of Mind. In *Stanford Encyclopedia of Philosophy*. 2005.

[136] P. Smolensky. On the Proper Treatment of Connectionism. *Behavioral and Brain Sciences*, 11:1–23, February 1988.

[137] G. Cohen. Hemispheric Differences in Serial Versus Parallel Processing. *Journal of Experimental Psychology*, 97(3):349–356, 1973.

[138] J. M. Polich. Hemispheric Differences for Visual Search: Serial vs Parallel Processing Revisited. *Neuropsychologia*, 20(3):297–307, 1982.

[139] A. W. Ellis, R. Ferreira, P. Cathles-Hagan, K. Holt, L. Jarvis, and L. Barca. Word Learning and the Cerebral Hemispheres: From Serial to Parallel Processing of Written Words. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536):3675–3696, December 2009.

[140] P. Smolensky. Tensor Product Variable Binding and the Representation of Symbolic Structures in Connectionist Systems. *Artificial Intelligence*, 46(1-2):159–216, November 1990.

[141] P. beim Graben and R. Potthast. Inverse Problems in Dynamic Cognitive Modeling. *Chaos*, 19(1):015103, 2009.

[142] C. R. Gallistel and A. P. King. *Memory and the Computational Brain: Why Cognitive Science will Transform Neuroscience*. Blackwell/Maryland Lectures in Language and Cognition. Wiley, 2009.

[143] D. Crevier. *AI: The Tumultuous History of the Search for Artificial Intelligence*. Basic Books, 1993.

[144] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall Series in Artificial Intelligence. Prentice Hall, 2010.

[145] R. Battiti. Using Mutual Information for Selecting Features in Supervised Neural Net Learning. *IEEE Transactions on Neural Networks*, 5(4):537–550, July 1994.

[146] J. C. Principe, D. Xu, Q. Zhao, and J. W. Fisher, III. Learning from Examples with Information Theoretic Criteria. *Journal of VLSI Signal Processing Systems*, 26(1/2):61–77, August 2000.

[147] J. C. Principe. *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*. Springer Publishing Company, Incorporated, 1st edition, 2010.

[148] S. B. Laughlin, R. R. de Ruyter van Steveninck, and J. C. Anderson. The Metabolic Cost of Neural Information. *Nature Neuroscience*, 1(1):36–41, May 1998.

[149] F. E. Theunissen and A. Borst. Information Theory and Neural Coding. *Nature Neuroscience*, 2:947–957, 1999.

[150] N. Brenner, S. P. Strong, R. P. Koberle, W. P. Bialek, and R. R. de Ruyter van Steveninck. Synergy in a Neural Code. *Neural Computation*, 12(7):1531–1552, July 2000.

[151] S. Panzeri. The Role of Spike Timing in the Coding of Stimulus Location in Rat Somatosensory Cortex. *Neuron*, 29(3):769–777, March 2001.

[152] A. Mazzoni, S. Panzeri, N. K. Logothetis, and N. Brunel. Encoding of Naturalistic Stimuli by Local Field Potential Spectra in Networks of Excitatory and Inhibitory Neurons. *PLoS Computational Biology*, 4(12), December 2008.

[153] M. A. Montemurro, M. J. Rasch, Y. Murayama, N. K. Logothetis, and S. Panzeri. Phase-of-Firing Coding of Natural Visual Stimuli in Primary Visual Cortex. *Current Biology*, 18(5):375–380, March 2008.

[154] C. Kayser, M. A. Montemurro, N. K. Logothetis, and S. Panzeri. Spike-Phase Coding Boosts and Stabilizes Information Carried by Spatial and Temporal Spike Patterns. *Neuron*, 61(4):597–608, February 2009.

[155] A. Buehlmann and G. Deco. Optimal Information Transfer in the Cortex Through Synchronization. *PLoS Computational Biology*, 6(9):e1000934, September 2010.

[156] F. Attneave. Some Informational Aspects of Visual Perception. *Psychological Review*, 61(3):183–193, May 1954.

[157] P. Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Computational Neuroscience Series. Mit Press, 2005.

[158] L. F. Abbott and P. Dayan. The Effect of Correlated Variability on the Accuracy of a Population Code. *Neural Computation*, 11(1):91–101, January 1999.

[159] D. Barber. *Bayesian Reasoning and Machine Learning*. Cambridge University Press, 2011.

[160] D. C. Knill and A. Pouget. The Bayesian Brain: The Role of Uncertainty in Neural Coding and Computation. *Trends in Neurosciences*, 27(12):712–719, 2004.

[161] D. J. Higham. An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Review*, 43(3):525–546, March 2001.

[162] H. P. McKean. A Class of Markov Processes Associated with Nonlinear Parabolic Equations. *Proceedings of the National Academy of Sciences*, 56(6):1907–1911, 1966.

[163] H. P. McKean. Propagation of Chaos for a Class of Non-Linear Parabolic Equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations,*

Session 7, Catholic University, 1967), pages 41–57. Air Force Office of Scientific Research, Arlington, Va., 1967.

[164] A. A. Vlasov. On Vibration Properties of Electron Gas. *Journal of Experimental and Theoretical Physics*, 8(3):291+, 1938. (In Russian).

[165] R. FitzHugh. Impulses and Physiological States in Theoretical Models of Nerve Membrane. *Biophysical Journal*, 1(6):445–466, July 1961.

[166] J. Nagumo, S. Arimoto, and S. Yoshizawa. An Active Pulse Transmission Line Simulating Nerve Axon. *Proceedings of the Institute of Radio Engineers*, 50(10):2061–2070, 1962.

[167] G. Wainrib. *Randomness in Neurons: A Multiscale Probabilistic Analysis*. PhD thesis, 2010.

[168] A. Destexhe, Z. F. Mainen, and T. J. Sejnowski. Synthesis of Models for Excitable Membranes, Synaptic Transmission and Neuromodulation Using a Common Kinetic Formalism. *Journal of Computational Neuroscience*, 1(3):195–230, 1994.

[169] G. B. Ermentrout and D. H. Terman. *Mathematical Foundations of Neuroscience*. Applications of Mathematics. Springer, 2010.

[170] A. Destexhe. Fluctuating Synaptic Conductances Recreate In Vivo-Like Activity in Neocortical Neurons. *Neuroscience*, 107(1):13–24, November 2001.

[171] A. Destexhe, M. Badoual, Z. Piwkowska, T. Bal, and M. Rudolph. A Novel Method for Characterizing Synaptic Noise in Cortical Neurons. *Neurocomputing*, 58-60:191–196, 2004.

[172] M. Rudolph, J. G. Pelletier, D. Paré, and A. Destexhe. Estimation of Synaptic Conductances and Their Variances From Intracellular Recordings of Neocortical Neurons In Vivo. *Neurocomputing*, 58-60:387–392, 2004.

[173] W. E. Schiesser. *The Numerical Method of Lines: Integration of Partial Differential Equations*. Academic Press, 1991.

[174] W. E. Schiesser and G. W. Griffiths. *A Compendium of Partial Differential Equation Models: Method of Lines Analysis with Matlab*. Cambridge University Press, 1st edition, March 2009.

[175] I. R. Khan and R. Ohba. Closed-Form Expressions for the Finite Difference Approximations of First and Higher Derivatives Based on Taylor Series. *Journal of Computational and Applied Mathematics*, 107(2):179–193, 1999.

[176] J. Baladron Pezoa, D. Fasoli, and O. Faugeras. Three Applications of GPU Computing in Neuroscience. *Computing in Science and Engineering*, 14(3):40–47, May 2012.

[177] J. Baladron. *Exploring the Neural Codes with Parallel Hardware*. PhD thesis, 2013.

[178] A. S. Ecker, P. Berens, G. Keliris, M. Bethge, N. K. Logothetis, and A. S. Tolias. Decorrelated Neuronal Firing in Cortical Microcircuits. *Science (New York, N.Y.)*, 327(5965):584–587, January 2010.

[179] J. Kremkow, L. Perrinet, G. Masson, and A. Aertsen. Functional Consequences of Correlated Excitatory and Inhibitory Conductances in Cortical Networks. *Journal of Computational Neuroscience*, 28(3):579–594, June 2010.

[180] F. Montani, R. A. A. Ince, R. Senatore, E. Arabzadeh, M. E. Diamond, and S. Panzeri. The Impact of High-Order Interactions on the Rate of Synchronous Discharge and Information Transmission in Somatosensory Cortex. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1901), July 2009.

[181] H. Sompolinsky, H. Yoon, K. Kang, and M. Shamir. Population Coding in Neuronal Systems with Correlated Noise. *Physical Review E*, 64:51904, 2002.

[182] M. T. Wiechert, B. Judkewitz, H. Riecke, and R. W. Friedrich. Mechanisms of Pattern Decorrelation by Recurrent Neuronal Circuits. *Nature Neuroscience*, 13(8):1003–1010, August 2010.

[183] T. Tetzlaff, M. Helias, G. T. Einevoll, and M. Diesmann. Decorrelation of Neural-Network Activity by Inhibitory Feedback. *PLoS Computational Biology*, 8(8):e1002596, August 2012.

[184] O. D. Faugeras, J. D. Touboul, and B. Cessac. A Constructive Mean-Field Analysis of Multi Population Neural Networks with Random Synaptic Weights and Stochastic Inputs. *Frontiers in Computational Neuroscience*, 3(1), 2008.

[185] C. Viroli. Finite Mixtures of Matrix Normal Distributions for Classifying Three-Way Data. *Statistics and Computing*, 21(4):511–522, 2011.

[186] W. Magnus. On the Exponential Solution of Differential Equations for a Linear Operator. *Communications on Pure and Applied Mathematics*, 7(4):649–673, 1954.

[187] L. Isserlis. On a Formula for the Product-Moment Coefficient of any Order of a Normal Frequency Distribution in any Number of Variables. *Biometrika*, 12(1/2):134–139, 1918.

[188] S.-L. Lee and Y.-N. Yeh. On Eigenvalues and Eigenvectors of Graphs. *Journal of Mathematical Chemistry*, 12:121–135, 1993.

[189] S. U. Pillai, T. Suel, and S. Cha. The Perron-Frobenius Theorem: Some of its Applications. *Signal Processing Magazine, IEEE*, 22(2):62–75, 2005.

[190] O. Sporns. Small-World Connectivity, Motif Composition, and Complexity of Fractal Neuronal Connections. *Biosystems*, 85(1):55–64, July 2006.

[191] G. Tononi, G. M. Edelman, and O. Sporns. Complexity and Coherency: Integrating Information in the Brain. *Trends in Cognitive Sciences*, 2(12):474–484, December 1998.

[192] J. C. Spall. Monte Carlo-Based Computation of the Fisher Information Matrix in Nonstandard Settings. In *American Control Conference*, volume 5, pages 3797–3802, 2003.

[193] S. Das, J. C. Spall, and R. Ghanem. An Efficient Calculation of Fisher Information Matrix: Monte Carlo Approach Using Prior Information. In *46th IEEE Conference on Decision and Control*, pages 963–968, 2007.

[194] J. C. Spall. Improved Methods for Monte Carlo Estimation of the Fisher Information Matrix. In *American Control Conference*, pages 2395–2400, 2008.

[195] J. C. Spall. On Monte Carlo Methods for Estimating the Fisher Information Matrix in Difficult Problems. In *Conference on Information Sciences and Systems*, pages 741–746. IEEE, 2009.

[196] S. Das, J. C. Spall, and R. Ghanem. Efficient Monte Carlo Computation of Fisher Information Matrix Using Prior Information. *Computational Statistics & Data Analysis*, 54(2):272–289, February 2010.

[197] T. K. Moon. The Expectation-Maximization Algorithm. *Signal Processing Magazine, IEEE*, 13(6):47–60, 1996.

[198] W. K. Hastings. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57(1):97–109, 1970.

[199] T. Schreiber. Measuring Information Transfer. *Physical Review Letters*, 85(2):461–464, July 2000.

[200] A. Kaiser and T. Schreiber. Information Transfer in Continuous Processes. *Physica D: Nonlinear Phenomena*, 166:43 – 62, 2002.

[201] X. S. Liang and R. Kleeman. Information Transfer between Dynamical System Components. *Physical Review Letters*, 95:244101, December 2005.

[202] A. A. Minai and R. D. Williams. Original Contribution: On the Derivatives of the Sigmoid. *Neural Networks*, 6(6):845–853, June 1993.

[203] L. Carlitz. Eulerian Numbers and Polynomials. *Mathematics Magazine*, 32(5):247–260, 1959.

[204] S. J. Miller. An Identity for Sums of Polylogarithm Functions. *Integers: Electronic Journal Of Combinatorial Number Theory*, 8, 2008.

[205] E. Y. Deeba and D. M. Rodriguez. Stirling's Series and Bernoulli Numbers. *The American Mathematical Monthly*, 98(5):423–426, 1991.

[206] K. Adegoke and O. Layeni. The Higher Derivatives of the Inverse Tangent Function and Rapidly Convergent BBP-Type Formulas for Pi. *Applied Mathematics E-Notes*, 10:70–75, 2010.